

Jak jsem psal v e-mailu, současná podivná doba nám dává příležitost zkoušet věci, na které bychom za normálních okolností ani nepomysleli. A tak jsem se rozhodl vyzkoušet jeden nápad a pokusil jsem se vytvořit značně netypický výukový materiál.

Inspirací k tomu mi byl formát gamebooků. *(Pro ty, kdo se snad s tímto termínem nikdy nesešli se jedná o knižní příběh, který se však různě větví na základě rozhodnutí čtenáře. Na konci každé sekce popisující příběh jste postaveni před volbu*

z několika možností, a tato volba určí pokračování příběhu.) Proto zde na každém slidu je jedno, nebo několik tlačítek. Některá vás pouze pousnou dál po cestě (ta ve tvaru šipek) a některá vás pošlou na různé cesty, které vedou k různým cílům. Kterou cestou se vydáte, záleží na vás. Vy budete mít možnost zvolit si, jak se hlavní postava příběhu zachová. Danou možnost zvolíte kliknutím na příslušné tlačítko a to vás přesune na odpovídající místo. Snad jen dodám, že to, co na vás na které cestě čeká nebudete vědět, dokud se jí nevydáte. A na různých místech najdete různé nápovědy. Dobře si jich všimněte. Někdy záleží na drobných rozdílech.

Každé zakončení je ohodnoceno různým počtem hvězdiček. Celkem je jich devět.

P.S.: To, jak k tomu přistoupíte je samozřejmě na vás. Obejít se dá vždycky všechno a nemůžu vám v tom nijak zabránit. Nicméně doporučuji to alespoň napoprvé zkusit striktně podle instrukcí.

Nyní začněte kliknutím na tlačítko „Úvod“

Úvod

V tomto příběhu se stanete kdysi úspěšným vědcem, který ale v současné době zažívá těžké období. Dlouhou dobu nezaznamenal žádný akademický úspěch a jeho kariéra visí na vlásku. Dostane však příležitost, která by mohla zvrátit cestu vstříc konci akademické práce. To, jak to celé dopadne se však bude odvíjet přímo od vašich rozhodnutí.

Je tam cesta, jak si zachránit kariéru a dál se intenzivně věnovat vědeckému poznání a dokonce je možné, že se vám podaří přijít s převratným zjištěním, které vás vynese na vrchol popularity, kam přece patříte! A jestliže si dáte tu práci, odejdete obohaceni o řadu nových informací.

Nicméně, na těchto cestách je řada slepých odboček, které všechny vedou na jediné místo. K inzerátům pracovních nabídek. Před takovým osudem se lze na poslední chvíli bránit dvěma způsoby:

- a) zvrátit své poslední rozhodnutí a zkusit to jinak
- b) někdy nezbyvá než se vrátit na úplný začátek a odstartovat celý výzkum znovu

Doufám, že vás to bude bavit a že to pro vás bude přínosné. Přeji vám na této cestě řadu těch nejlepších rozhodnutí!

A teď už je na čase opravdu začít:

**Zachránit jednu
vědeckou kariéru!**

„Abraháme!“

„No tak, Abraháme, stůjte!“

Naléhavý hlas vás vytrhnul z myšlenek.

„*Žádná škoda. Stejně to je samý neúspěch,*“ řeknete si v duchu.

Vaše práce v poslední době sklízí spíše posměch než ocenění a dokonce se ve zdech Akademie začíná mluvit o konci jedné kdysi zářné kariéry. Ohlédnete se a vidíte, jak za vámi po chodbě spěchá kolega. Jeho jméno si najednou nevybavujete.

„*Ale copak na tom záleží, stejně už tady dlouho nebudu.*“

„Abraháme, konečně jsem vás našel,“ začne onen kolega, když vás dožene, „máte hned jít za ředitelem.“

„Jasně, hned tam jdu, díky.“

„*No jasně, už je to tady. Bylo nám ctí, ale už tu pro vás nemáme místo.*“

V doprovodu podobných myšlenek dojdete až před dveře ředitelovy kanceláře. Tam se zastavíte. Upravíte si kravatu a uhladíte vlasy na pravé straně. Spíš ze zvyku, než že by se tím skutečně něco změnilo. Až potom se odhodláte zaklepat.

„Dobrý den, Abraháme,“ uvítá vás ředitel, když vstoupíte. „Pojďte se prosím posadit. Asi tušíte, proč jsem si vás nechal zavolat. Vaše práce nám za posledních několik měsíců nepřinesla žádná pozitiva. Akademie se spíše stává terčem posměchu a o vás se mluví jako o člověku, jehož vědecká sláva dávno zapadala prachem. Dělán to nerad, ale jsem nucen v tomto ohledu něco podniknout.“

„*Děkujeme, na shledanou!*“

„Dostali jsme návrh na velký výzkum. Trochu to spěchá, ale vše důležité je v téhle modré složce. Chci, abyste se

Vzít modrou složku a odejít do své pracovny

„Ne, pane řediteli, nechci být dál vysmíván a dál kazit pověst Akademie. Raději odejdu se vztyčenou hlavou.“

Sedíte za stolem ve své pracovně. Modrou složku máte položenou před sebou a zíráte na ni, jako byste mohl prohlédnout skrz její desky. Otevřít jste se jí zatím neodhodlal.

„Takže poslední možnost zachránit si kariéru. No dobrá, tak se do toho pustíme.“

Po otevření složky zjistíte, že se jedná o plán výzkumu na velkém množství lidí napříč populací celého Státu. Zdá se, že celý výzkum je rozdělen do několika částí, přičemž do vašich rukou se dostala pouze jedna z nich. Samotný sběr dat zajišťuje Institut, jehož kontaktní adresu máte ve složce uvedenu. Vaší povinností je však na tuto adresu doručit materiály, pomocí kterých bude Institut data sbírat.

Na další straně dokumentace je specifikován účel vašeho výzkumu. Klade si za cíl zodpovědět na tři otázky:

a) Zda se lidé různých věkových kategorií liší ve schopnosti rozpoznávat spolehlivé mediální zprávy od nespolehlivých.

V této souvislosti složka obsahuje také kvalitně zpracovanou rešerši literatury, ze které vyplývá, že v současné době vědecké důkazy napovídají, že co do schopnosti rozpoznávat spolehlivost mediálních zpráv lze populaci rozdělit do tří věkových skupin:

15 – 22 let, 23 – 60 let a 60 a více let. Tyto tři skupiny obyvatel by se měly od sebe vzájemně lišit. Populace do 15 let věku úmyslně není zahrnuta, neboť se předpokládá, že není dostatečně zasažena mediálním životem.

Je zde také přiložen již hotový testový sešit, obsahující různé zprávy, u kterých má respondent rozhodovat zda jsou či nejsou spolehlivé. Vyšší dosažený skóre indikuje lepší schopnost rozpoznávat spolehlivé zprávy.

b) Zda se u lidí liší jejich ochota k veřejnému vystupování v závislosti na výši jejich čistého příjmu.

Otočit na další stranu

Teorie předpokládá, že lidé s vyššími příjmy budou ochotnější k vystupování na veřejnosti, neboť jejich bohatství jim dává pocit důležitosti a také schopnosti ovlivňovat mínění davu. I k této výzkumné otázce je přiložen dotazník hodnotící ochotu respondenta k veřejnému vystupování. Vyšší skór zde indikuje vyšší ochotu k takovému vystupování.

c) Zda se liší muži a ženy v úrovni projevované agrese vůči vládě Státu.

I k této výzkumné otázce je přiložen dotazník, hodnotící projevovanou agresivitu vůči vládě. Čím vyšší je dosažený skór, tím více agresivity daný respondent projevuje.

„Zdá se, že pro první fázi je tady už všechno připravené,“ říkáte si. „Stačí tyhle připravené materiály vzít a poslat je na adresu Institutu. Oni se už postarají o zbytek.“

**Vzít připravené dotazníky a
odeslat je do Institutu**

**Shrnout si, co je obsahem
výzkumu**

Znovu projít modrou složku

Nadšený z toho, jak kvalitní materiály byly obsaženy v modré složce, a že vám neznámý autor všech těch metod ušetřil spoustu práce, vkládáte Dotazník rozpoznávání spolehlivosti zpráv, Dotazník ochoty k veřejnému vystupování a Dotazník projevované agresivity vůči vládě do obálky a podle instrukcí odesíláte na uvedenou adresu.

Teď nezbyvá než čekat, až institut dodá požadovaná data.

Za nějaký čas skutečně obdržíte datový soubor. Když ho otevřete v počítači, uvidíte následující tabulku:

ID	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	33	41	42
002	25	49	32
003	2	0	36
004	9	19	24
005	7	15	33
...			

Takto tabulka pokračuje až po ID 800, nicméně je vám jasné, že v tomto datovém souboru chybí zásadní informace. Z této tabulky nelze testovat jedinou z výzkumných otázek! Kontaktujte Institut s tím, že zapomněli doplnit anamnestická data.

Institut však odvětví, že nasbírali data z obdržených materiálů tak, jak bylo dohodnuto. Žádná další data pro vás k dispozici nemají a nelze je již doplnit, protože sběr dat byl ukončen.

Nezbývá, než řediteli oznámit neúspěch zapříčiněný vlastní roztržitostí. Jste nucen vyklidit kancelář a v tichosti

Zvrátit poslední volbu

uální pracovní

Znovu zahájit výzkum

„Pomalou, pomalu, musím krotit svoje nadšení, to vede k chybám. Dobře. To znamená, že mám tři úkoly. Srovnat věkové skupiny, co do hrubého skóru Testu spolehlivosti zpráv, posoudit korelace mezi výší platu a hrubým skórem Dotazníku ochoty a srovnat muže a ženy v hrubých skórech Dotazníku agresivity. Fajn. Jaká data k tomu teda potřebuju? Věk, Test zpráv, Výši platu, Dotazník ochoty, Pohlaví a Dotazník agresivity. Dotazníky mám hotové. Takže stačí vymyslet jakým způsobem sbírat a kódovat ty ostatní proměnné. Potřebuju prostě vytvořit Anamnestický dotazník.

**Vrhnout se na vytváření
Anamnestického dotazníku**

**Znovu prolistovat materiály v
modré složce**

Rozhodl jste se, že je lepší postupovat rozvážně. Znovu usedáte za stůl a podrobně zkoumáte obsah modré složky. Základní instrukce k výzkumu.

„Jasně, to jsem četl.“

Test spolehlivosti zpráv, Dotazník ochoty a dotazník agresivity. Dále nacházíte rozpracovaný Anamnestický dotazník. Obsahuje několik tištěných položek a u každé z nich je rukou napsána poznámka.

Pohlaví: **1 - muži, 2 - ženy**

Věk:

Příjem: **zaokrouhlovat na stovky**

Kraj: **1 - 12 podle čísla kraje**

Velikost bydliště: **1 - do 500 obyvatel; 2 - 501 - 2 000; 3 - 2 001 - 10 000; 4 - 10 001 - 50 000;**

50 001 - 100 000; 5 - 100 001 - 500 000; 6 - nad 500 000 obyvatel

Vzdělání: **1 - základní; 2 - výuční; 3 - maturita; 4 - nižší vysoké; 5 - vyšší vysoké; 6 - doktorské; 7 - postdoktorské**

Politická aktivita: **0 - ne; 1 - ano**

Otočit na další stranu 

„Hm, v tomhle návrhu je toho víc, než by bylo potřeba, ale více dat je vždycky lepších než méně dat. Bude lepší to zachovat, ostatně, můžu případně zkusit ještě něco nad rámec těch tří zadaných otázek. Jediné, co v tom Anamnestickém chybí je kódování věku. V zásadě existují dvě varianty, jak k tomu přistoupit. První možnost je nechat všechny vyplnit svůj věk v letech. Snadné a rychlé. Ale trochu problematické. Občasné Státu jsou v tomto ohledu dost hákliví a neradi takové informace sdělují. Ale vzhledem k anonymitě dotazování by to snad ani nemusel být problém. Ač, to vlastně nevím, jakým způsobem to v Institutu sbírají. No, je to trochu riskantní. Navíc nějaký šťoura to může napadnou z hlediska etiky. Abych se tomu vyhnul, můžu se přidržit teorie a nechat každého vybrat, do kterého ze tří intervalů spadá. To znamená 15 – 22; 23 – 60; 60+“

Podstoupit drobné riziko a zvolit variantu konkrétních věků

Přidržit se doporučení teorie a zvolit variantu intervalů.

Rozhodl jste se pro variantu konkrétních věků. To znamená, že finální verze vašeho Anamnestického dotazníku je strukturovaná následujícím způsobem:

Pohlaví: 1 – muži, 2 – ženy

Věk: udávaný v letech

Příjem: zaokrouhlovat na stovky

Kraj: 1 – 12 podle čísla kraje

Velikost bydliště: 1 – do 500 obyvatel; 2 – 501 – 2 000; 3 – 2 001 – 10 000; 4 – 10 001 – 50 000; 50 001- 100 000; 5 – 100 001 – 500 000; 6 – nad 500 000 obyvatel

Vzdělání: 1 – základní; 2 – výuční; 3 – maturita; 4 – nižší vysoké; 5 – vyšší vysoké; 6 – doktorské; 7 – postdoktorské

Politická aktivita: 0 – ne; 1 – ano

K tomu přikládáte Test spolehlivosti zpráv, Dotazník ochoty a Dotazník agresivity.

„V tomto formátu bych snad měl mít všechny potřebné informace. A i nějaké navíc.“

**Zkompletovat sadu materiálů
a odeslat na adresu Institutu**

**Nespěchat s odesláním.
Přemýšlet, zda by nebylo
užitečné ještě něco doplnit.**

Následujících několik dnů přemýšlíte, co ještě přidat k Anamnestickému dotazníku. Přemýšlíte, k čemu je celý výzkum vlastně určený a co by ještě bylo dobré k tomu přidat.

„Jedna věc je splnit zadání, ale když už je tady taková příležitost, mělo by se toho pořádně využít. Přece jenom, bude to výzkum na velkém množství lidí a když přijdu navíc s nějakým zajímavým zjištěním, mohl bych napravit pošramocené jméno Akademie. Beztak za to můžu já. A co být troškař, třeba se zase stanu slavným vědcem!“

Chodíte ode zdi ke zdi a hlava se vám plní různými nápady, které ale vzápětí zase zahazujete. Nic z toho nemá potenciál

k převratnému objevu, který potřebujete. Nakonec začnete uvažovat, jestli má cenu se do toho vlastně pouštět.

**Vzdát snahu o rozšíření
výzkumu**

Dát své kreativě ještě šanci

Vidina převratného úspěchu je tak silná, že se vám nechce hned tak vzdát. Dalších několik dní chodíte po pracovně ode zdi ke zdi. Nakonec se vám v hlavně přece jenom zrodí nápad, který stojí za pokus.

„Když k tomu přidám ještě proměnou týkající se Akademie samotné, můžou z toho být docela zajímavé výsledky. Bud' se ukáže, že lidé vnímají Akademii jako užitečnou a prospěšnou organizaci, což určitě rádo uslyší naše PR oddělení a nebo se ukáže, že jsme vnímání jako zbytečný žrout peněz občanů našeho Státu. To zas bude důležité pro ředitele, protože takové smýšlení je pro nás ohrožující a budeme s tím muset něco dělat. Tak jako tak to pro nás bude užitečné. K tomu, abych něco takového zjistil ale potřebuju nějakou metodu, která zjistí vztah běžného občana k Akademii.“

Hned se do svého záměru pustíte a začnete skládat metodu, která má odhalit vztah běžného člověka k Akademii. A tak ponořen do této práce trávíte den za dnem. Až jednoho odpoledne vás ze zadumání vytrhne zaklepání na dveře. Za nimi stojí onen kolega, kterého jste onehdá potkal na chodbě.

„Abraháme, máš jít za ředitelem.“

„Jo, hned tam budu,“ odpovíte a zamíříte ven z pracovny.

„Abraháme,“ uvítá vás ředitel, „přišel mi dopis z Institutu. Poslouchejte:

Vážená Akademie, k dnešnímu dni jste nedodali materiály ke sběru dat v Celostátním víceúčelovém výzkumu. Musíme vám tedy s politováním oznámit, že vaše účast v tomto šetření je tímto zrušena. Víte, co to znamená, Abraháme? Znamená to, že jste zklamal důvěru, kterou jsem ve vás vložil!“

Jste nucen vyklidit kancelář a v tichosti opustit Akademii. Doma usedáte k počítači a začínáte procházet aktuální

Zvrátit poslední volbu

Znovu zahájit výzkum

Je čas dát všechno do obálky a poslat to Institutu. Nezbyvá než čekat, až data nasbírají a pošlou vám je k analýze. Chvilé klidu je zasloužená. Zatím jste odvedli dobrou práci.

Ona chvíle odpočinku ale netrvá zase příliš dlouho. Za nějaký čas obdržíte datový soubor. Když ho otevřete v počítači, uvidíte následující tabulku:

ID	Věk	Pohlaví	Vzdělán í	Kraj	Velikost bydliště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

Tímto způsobem tabulka pokračuje až po ID 800.

„Výborně, a teď je čas se pustit do analýz.“

Začít analyzovat data

Je na čase pustit se do další práce. Zkoumáte, jak souvisí věk se schopností posuzovat spolehlivost mediálních zpráv. V tuto chvíli je čas přidržet se známé teorie a ze všeho nejdřív sloučit respondenty do skupin podle věku tak, jak bylo uvedeno v rešerši.

Vytvoříte si tedy novou proměnnou „Věk_Int“, která nabývá hodnoty 1 pro lidi ve věku 15 – 22 let, hodnoty 2 pro lidi mezi 23 a 60 lety a hodnoty 3 pro lidi starší 60 let. Tyto tři skupiny pak porovnáte v hrubých skórech Dotazníku mediálních zpráv.

Výsledky lze shrnout asi takto:

Všechny tři skupiny se od sebe vzájemně liší statisticky významně.

„Což vzhledem k velikosti vzorku není nijak překvapivé.“

Věcný rozdíl mezi skupinou od 15 let a skupinou 60+ je zhruba půl směrodatné odchylky, přičemž skupina od 15 dosahuje vyššího skóru.

Skupina 23 – 60 let dosahuje v Testu schopnosti ze všech tří skupin nejlepších výsledků. Věcný rozdíl mezi skupinou od 15 let a skupinou 23 – 60 se rovná jedné směrodatné odchylce. Rozdíl mezi skupinou 23 – 60 a 60+ je tedy jedna a půl směrodatné odchylky.

Tuto otázku lze tedy uzavřít tak, že skupina od 15 let je o trochu lepší v hodnocení spolehlivosti mediálních zpráv, než skupina 60+. Skupina 23 – 60 je pak výrazněji silnější než obě zbývající skupiny.

Zpracovat tyto výsledky do článku

Blíže výsledky prozkoumat

Vaše vědecká zvědavost vám nedá a rozhodnete se detailněji prozkoumat výsledky. Proto si v dalším kroku zobrazíte rozložení proměnné Věk_Int pomocí scatterplotu. Rozložení grafu napovídá, že by v daném vzorku mohly existovat nikoli tři, ale čtyři skupiny. Proto si zobrazíte ještě jeden scatterplot, tentokrát z proměnné Věk. Analýzou jednotlivých případů se pak ukáže, že se vzorek skutečně rozdělil do čtyř nezávislých skupin.

Zůstala zachována skupina 15 – 22 let a skupina 60+, ale prostřední skupina se navíc rozdělila na dvě. Vznikla tak skupina 23 – 45 let a skupina 46 – 60 let.

Vytvoříte si další proměnnou „Věk_Int2“, která nabývá hodnoty 1 pro interval 15 – 22let, 2 pro interval 23 – 45, 3 pro interval 46 – 60 let a 4 pro skupinu 60+. Tyto čtyři skupiny pak znovu porovnáte v hrubých skórech Testu důvěryhodnosti mediálních zpráv.

Ukázalo se, že v takovém rozložení jsou výsledky poněkud jiné.

Skupina 4 dosahuje nejhorších výsledků. Skupiny 1 a 3, které se vzájemně neliší, dosahují o půl směrodatné odchylky lepších výsledků než skupina 3. Skupina 2 dosahuje zdaleka nejlepších výsledků o dvě a půl směrodatné odchylky lepších než skupina 4.

Z těchto výsledků tedy vyplývá, že lidé starší 60 let dokáží jen velmi špatně rozlišovat spolehlivé a nespolehlivé zprávy. O něco lépe jsou na tom v tomto směru lidé mezi 15 a 22 a mezi 46 a 60 lety. Zdaleka nejlepší v tomto jsou pak lidé mezi 23 a 45 lety.

Zpracovat tyto výsledky do článku

První ze tří částí máte úspěšně provedenou a sepsanou. Dokonce jste přišel s výsledky, které do značné míry korigují dosavadní předpoklady. To lze považovat za velký úspěch. Dále se tedy věnujete zpracování těchto informací do článku, který hodláte na konci výzkumu publikovat. Po napsání příslušné části textu je čas přistoupit k dalším analýzám.

Znovu jste sedl k datovému souboru. Je na čase pokračovat druhou částí.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

Dalším krokem je prozkoumání výzkumné otázky b). Tedy korelační vztah čistého příjmu s ochotou respondentů vystupovat na veřejnosti, která je reprezentována hrubým skórem v Dotazníku ochoty k veřejnému vystupování. V první fázi analýzy se bohužel ukázalo, že proměnná Čistý příjem je značné množství chybějících hodnot. Dále jste také zjistil, že tímto neduhem netrpí pouze proměnná Čistý příjem, ale také ostatní proměnné, se kterými jste doposud nepracoval. Tento problém je nutné nějak řešit.

„Hm, musím nějak naložit s těmi chybějícími hodnotami. V zásadě existují dva běžné způsoby, jak s tím naložit. Buď se zabavím všech respondentů, od kterých nemám kompletní data a nebo chybějící data v Čistém příjmu nahradím průměrnými hodnotami.“

Jak tak v zamyšlení pochodujete ode zdi ke zdi po pracovně, padne vám zrak na knihu, na kterou jste dávno zapomněli. Její hřbet nese název *„Pokročilé postupy řešení statistických problémů.“*

Vymazat všechny nekompletní záznamy

Nahlédnout do knihy *Pokročilých postupů*

Doplnit chybějící hodnoty příjmů průměrnými hodnotami

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, ale vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

**Zpracovat tyto výsledky do
článku**

Pokročilé postupy řešení statistických problémů

Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p -hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu

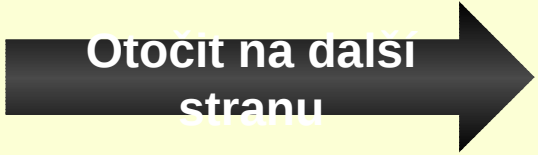


Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat **„Hodnoty chybějící z definice“**. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých respondentů

Otočit na další

stranu



Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami i o zmíněných softwarech (Acock, 2005)

[Přejít na seznam literatury](#)

Zavřít knihu

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.

Vratit se k tomuto odkazu

- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.

Vratit se k tomuto odkazu

- Acock, A. C. (2005). Working with missing values. ***Journal of Marriage and Family*** 67(4), 1012-1028.

Vratit se k tomuto odkazu

Četbou knihy jste si ověřil své domněnky a načerpal novou inspiraci pro řešení nastalého problému. Teď je potřeba toto úsilí zužitkovat. Je potřeba posunout výzkum zase kousek dál a rozhodnout se, jak tedy vyřešíte problém

s chybějícími hodnotami.

Vzhledem k tomu, že pro takzvané moderní postupy nemáte zařízení s dostatečnou výpočetní kapacitou, musíte vybrat to nejvhodnější z tradičních postupů.

Použít Mean substitution

Použít Pairwise deletion

Použít Indicator Adjustment

Použít Case deletion

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, ale vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po použití Pairwise delation z této konkrétní analýzy vypadlo 25% výzkumného vzorku. Vzhledem k jeho rozsahu si analýza i v tomto rozsahu zachovává svou statistickou sílu. Odhalila silnou pozitivní korelaci mezi výší čistého příjmu a ochotou k veřejným vystoupením. Tento výsledek podporuje dosavadní mínění teorie, že bohatší lidé jsou ochotnější k veřejnému vystupování. Z dostupných dat však nelze zjistit nic o tom, jaký mechanismus tuto větší ochotu způsobuje.

„To by chtělo navazující kvalitativní studii, která by se bohatých lidí ptala na důvody, proč jsou ochotní vystupovat na veřejnosti a snažila se odhalit nějaký mechanismus. Hm. Tohle je dobrá věta do diskuse. To si musím někam poznačit.“

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po aplikování Indicator adjustment postupu a doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost korelace čistého příjmu a hrubého skóru Dotazníku, nicméně tato korelace se ukázala být velmi nízká. Díky indikátoru víte, že průměrný plat byl doplněn do 25% případů. Momentálně se vám tedy podařilo vyvrátit mylný předpoklad teorie, je však otázka, jak moc oněch 25% doplněných hodnot tento výsledek zkresluje.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu. V prvním případě jste objevil mezeru v teorii a v druhém případě jste teorii potvrdil. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo korigovat jednu nepřesnost v teorii a v druhém případě, pokud připustíme korekci síly korelace, která je zkreslená 25% doplněných průměrných hodnot, tak by ta korelace mohla být dostatečně silná, aby to stálo za zmínku. V podstatě to teorii podporuje. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydliště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo najít dvě trhliny v teoretických předpokladech. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste se sešli k datovému souboru.

ID	Věk	Pohlaví	Vzdělán í	Kraj	Velikost bydliště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

**Pokračovat další výzkumnou
otázkou**

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

Použít výsledky včetně outlierů

Vyřadit outliery a provést analýzy znovu

**Nahlédnout do knihy
*„Pokročilé postupy řešení statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.

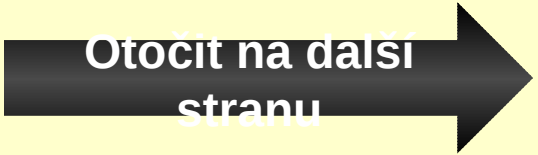
První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outlierem jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v tomto interpretativní zodpovědnosti

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

[Vrátit se k tomuto odkazu](#)

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet ještě jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

Zabývat se výzkumnou **otázkou**
interakce

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak [Přejít na seznam literatury](#) individuální analýze i problémům plynoucím z plošných řešení.

To je jen plošné shrnutí toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermana Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci interpretativní zodpovědnosti

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

Zabývat se výzkumnou otázkou
interakce

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outlierem jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otociť na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v tomto interpretativní zodpovědnosti

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý. Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů. Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

První ze tří částí máte úspěšně provedenou a sepsanou. Podařilo se vám potvrdit to, co předpokládá teorie. Dále se tedy věnujete zapracování těchto informací do článku, který hodláte na konci výzkumu publikovat. Po napsání příslušné části textu je čas přistoupit k dalším analýzám. Znovu jste sedl k datovému souboru. Je na čase pokračovat druhou částí.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

Dalším krokem je prozkoumání výzkumné otázky b). Tedy korelační vztah čistého příjmu s ochotou respondentů vystupovat na veřejnosti, která je reprezentována hrubým skórem v Dotazníku ochoty k veřejnému vystupování.

V první fázi analýzy se bohužel ukázalo, že proměnná Čistý příjem je značné množství chybějících hodnot. Dále jste také zjistil, že tímto neduhem netrpí pouze proměnná Čistý příjem, ale také ostatní proměnné, se kterými jste doposud nepracoval. Tento problém je nutné nějak řešit.

„Hm, musím nějak naložit s těmi chybějícími hodnotami. V zásadě existují dva běžné způsoby, jak s tím naložit. Buď se zabavím všech respondentů, od kterých nemám kompletní data a nebo chybějící data v Čistém příjmu nahradím průměrnými hodnotami.“

Jak tak v zamyšlení pochodujete ode zdi ke zdi po pracovně, padne vám zrak na knihu, na kterou jste dávno zapomněli. Její hřbet nese název *„Pokročilé postupy řešení statistických problémů.“*

Vymazat všechny nekompletní záznamy

Doplnit chybějící hodnoty příjmů průměrnými hodnotami

Nahlédnout do knihy *Pokročilých postupů*

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané Státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

**Zpracovat tyto výsledky do
článku**

Pokročilé postupy řešení statistických problémů

Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p -hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu



Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat „**Hodnoty chybí z definice**“. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých respondentů chybí uvedení

Otáčít na další
stranu



Tento postup zachovává co největší množství respondentů a tedy co největší statistickou sílu, je zde však problém, že vzorek pro zpracování každé jednotlivé analýzy je pak odlišný a nelze tudíž jednotlivé kroky slučovat. V tomto případě by tedy nebylo možné studovat interakce věku a asertivity z hlediska prožívaného wellbeingu. Tento postup by vyžadoval Case delation.

Mean substitution je postup, který nahrazuje chybějící hodnoty průměrem vzorku v dané proměnné. Jedná se o logický postup, jestliže lze předpokládat, že daná proměnná ve vzorku odpovídá normálnímu rozložení. V takovém případě, neznáme-li hodnotu dané proměnné u konkrétní osoby, nejpravděpodobnější je, že tato hodnota se pohybuje blízko kolem průměru. Tento postup je však problematický, pokud v konkrétním vzorku je velké množství chybějících hodnot. Jestliže ve vzorku chybí například 30% hodnot, pak nahrazení všech těchto hodnot průměrem značně redukuje rozptyl ve vzorku, a tím pádem podhodnocuje hodnotu korelace. Dalším problémem je necitlivost takového postupu vůči extrémním hodnotám. Jestliže mi u respondenta chybí jeho měsíční příjem a tento člověk je buďto milionář nebo se naopak potácí na hranici chudoby, pak v obou těchto případech je nahrazení této chybějící hodnoty průměrem velmi nešťastné a povede ke zkreslení výsledků. Pomocí Mean substitution není také vhodné doplňovat z definice chybějící hodnoty.

Indicator adjustment je lehká úprava strategie doplňování průměru. V tomto případě se do datového souboru přidá proměnná, tzv. indikátor, který nabývá hodnoty 0 v případech, kdy cílová hodnota chybí a 1 v případech, kdy je cílová hodnota pozorována. Následně jsou chybějící hodnoty doplněny průměrem vzorku. Přítomnost indikátoru zde udržuje informaci o počtu doplněných hodnot a tím pádem umožňuje rozsahu zkreslení. Více o tomto postupu (Cohen 1983) a následně [Přejít na seznam literatury](#) [Otočit na další stránu](#)

Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami viz kniha *Missing Data Imputation* (Acocck, 2005).

[Přejít na seznam literatury](#)

[Zavřít knihu](#)

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Acock, A. C. (2005). Working with missing values. ***Journl of Marriage and family*** 67(4), 1012-1028.

Vraťit se k tomuto odkazu

Četbou knihy jste si ověřil své domněnky a načerpal novou inspiraci pro řešení nastalého problému. Teď je potřeba toto úsilí zužitkovat. Je potřeba posunout výzkum zase kousek dál a rozhodnout se, jak tedy vyřešíte problém s chybějícími hodnotami. Vzhledem k tomu, že pro takzvané moderní postupy nemáte zařízení s dostatečnou výpočetní kapacitou, musíte vybrat to nejvhodnější z tradičních postupů.

Použít Case delation

Použít Pairwise delation

Použít Mean substitution

Použít Indicator Adjustment

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané Státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po použití Pairwise delation z této konkrétní analýzy vypadlo 25% výzkumného vzorku. Vzhledem k jeho rozsahu si analýza i v tomto rozsahu zachovává svou statistickou sílu. Odhalila silnou pozitivní korelaci mezi výší čistého příjmu a ochotou k veřejným vystoupením. Tento výsledek podporuje dosavadní mínění teorie, že bohatší lidé jsou ochotnější k veřejnému vystupování. Z dostupných dat však nelze zjistit nic o tom, jaký mechanismus tuto větší ochotu způsobuje.

„To by chtělo navazující kvalitativní studii, která by se bohatých lidí ptala na důvody, proč jsou ochotní vystupovat na veřejnosti a snažila se odhalit nějaký mechanismus. Hm. Tohle je dobrá věta do diskuse. To si musím někam poznačit.“

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedli korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po aplikování Indicator adjustment postupu a doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost korelace čistého příjmu a hrubého skóru Dotazníku ochoty k vystupování, nicméně tato korelace se ukázala být velmi nízká. Díky indikátoru víte, že průměrný plat byl doplněn do 25% případů. Momentálně se vám tedy podařilo vyvrátit mylný předpoklad teorie, je však otázka, jak moc oněch 25% doplněných hodnot tento výsledek zkresluje.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu. V prvním případě jste mezeru v teorii a v druhém případě jste teorii potvrdil. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydliště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo korigovat jednu nepřesnost v teorii a v druhém případě, pokud připustíme korekci síly korelace, která je zkreslená 25% doplněných průměrných hodnot, tak by ta korelace mohla být dostatečně silná, aby to stálo za zmínku. V podstatě to teorii podporuje. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydliště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo najít dvě trhliny v teoretických předpokladech. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

Použít výsledky včetně outlierů

Vyřadit outliery a provést analýzy znovu

**Nahlédnout do knihy
*„Pokročilé postupy řešení statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrní a interpretovatelní zodpovědnosti.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý. Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**


**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.

První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermana Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrní a interpretovatelní zodpovědnosti.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý. Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů. Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Rozhodl jste se pro variantu intervalů. To znamená, že finální verze vašeho Anamnestického dotazníku je strukturovaná následujícím způsobem:

Pohlaví: 1 – muži, 2 – ženy

Věk: 1 – 15 – 22 let; 2 – 23 – 60 let; 3 – 60+ let

Příjem: zaokrouhlovat na stovky

Kraj: 1 – 12 podle čísla kraje

Velikost bydliště: 1 – do 500 obyvatel; 2 – 501 – 2 000; 3 – 2 001 – 10 000; 4 – 10 001 – 50 000; 50 001- 100 000; 5 – 100 001 – 500 000; 6 – nad 500 000 obyvatel

Vzdělání: 1 – základní; 2 – výuční; 3 – maturita; 4 – nižší vysoké; 5 – vyšší vysoké; 6 – doktorské; 7 – postdoktorské

Politická aktivita: 0 – ne; 1 – ano

K tomu přikládáte Test spolehlivosti zpráv, Dotazník ochoty a Dotazník agresivity.

„V tomto formátu bych snad měl mít všechny potřebné informace. A i nějaké navíc.“

**Zkompletovat sadu
materiálů a odeslat na
adresu Institutu**

**Nespěchat s odesláním.
Přemýšlet, zda by nebylo
užitečné ještě něco doplnit.**

Následujících několik dnů přemýšlíte, co ještě přidat k Anamnestickému dotazníku. Přemýšlíte, k čemu je celý výzkum vlastně určený a co by ještě bylo dobré k tomu přidat.

„Jedna věc je splnit zadání, ale když už je tady taková příležitost, mělo by se toho pořádně využít. Přece jenom, bude to výzkum na velkém množství lidí a když přijdu navíc s nějakým zajímavým zjištěním, mohl bych napravit pošramocené jméno Akademie. Beztak za to můžu já. A co být troškař, třeba se zase stanu slavným vědcem!“

Chodíte ode zdi ke zdi a hlava se vám plní různými nápady, které ale vzápětí zase zahazujete. Nic z toho nemá potenciál k převratnému objevu, který potřebujete. Nakonec začnete uvažovat, jestli má cenu se do toho vlastně pouštět.

**Vzdát snahu o rozšíření
výzkumu**

Dát své kreativitě ještě šanci

Vidina převratného úspěchu je tak silná, že se vám nechce hned tak vzdát. Dalších několik dní chodíte po pracovně ode zdi ke zdi. Nakonec se vám v hlavně přece jenom zrodí nápad, který stojí za pokus.

„Když k tomu přidám ještě proměnou týkající se Akademie samotné, můžou z toho být docela zajímavé výsledky. Bud' se ukáže, že lidé vnímají Akademii jako užitečnou a prospěšnou organizaci, což určitě rádo uslyší naše PR oddělení a nebo se ukáže, že jsme vnímáni jako zbytečný žrout peněz občanů našeho Státu. To zas bude důležité pro ředitele, protože takové smýšlení je pro nás ohrožující a budeme s tím muset něco dělat. Tak jako tak to pro nás bude užitečné. K tomu, abych něco takového zjistil ale potřebuju nějakou metodu, která zjistí vztah běžného občana k Akademii.“

Hned se do svého záměru pustíte a začnete skládat metodu, která má odhalit vztah běžného člověka k Akademii. A tak ponořen do této práce trávíte den za dnem. Až jednoho odpoledne vás ze zadumání vytrhne zaklepání na dveře. Za nimi stojí onen kolega, kterého jste onehdá potkal na chodbě.

„Abraháme, máš jít za ředitelem.“

„Jo, hned tam budu,“ odpovíte a zamíříte ven z pracovny.

„Abraháme,“ uvítá vás ředitel, „přišel mi dopis z Institutu. Poslouchejte:

Vážená Akademie, k dnešnímu dni jste nedodali materiály ke sběru dat v Celostátním víceúčelovém výzkumu. Musíme vám tedy s politováním oznámit, že vaše účast v tomto šetření je tímto zrušena. Víte, co to znamená, Abraháme? Znamená to, že jste zklamal důvěru, kterou jsem ve vás vložil!“

Jste nucen vyklidit kancelář a v tichosti opustit Akademii. Doma usedáte k počítači a začínáte procházet aktuální

Zvrátit poslední volbu

Znovu zahájit výzkum

Je čas dát všechno do obálky a poslat to Institutu. Nezbývá než čekat, až data nasbírají a pošlou vám je k analýze. Chvilé klidu je zasloužená. Zatím jste odvedli dobrou práci.

Ona chvíle odpočinku ale netrvá zase příliš dlouho. Za nějaký čas obdržíte datový soubor. Když ho otevřete v počítači, uvidíte následující tabulku:

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	1	2	1	4	1	8	NA	28	5	0
002	3	2	3	12	6	46	21 400	5	7	1
003	2	2	3	6	2	45	20 200	28	3	1
004	2	2	6	3	6	26	134 700	16	5	1
005	3	2	1	4	5	12	NA	19	46	1
...										

Tímto způsobem tabulka pokračuje až po ID 800.

„Výborně, a teď je čas se pustit do analýz.“

Začít analyzovat data

Je na čase pustit se do další práce. Zkoumáte, jak souvisí věk se schopností posuzovat spolehlivost mediálních zpráv. V tuto chvíli je čas přidržet se známé teorie a ze všeho nejdřív sloučit respondenty do skupin podle věku tak, jak bylo uvedeno v rešerši.

Vytvoříte si tedy novou proměnnou „Věk_Int“, která nabývá hodnoty 1 pro lidi ve věku 15 – 22 let, hodnoty 2 pro lidi mezi 23 a 60 lety a hodnoty 3 pro lidi starší 60 let. Tyto tři skupiny pak porovnáte v hrubých skórech Dotazníku mediálních zpráv.

Výsledky lze shrnout asi takto:

Všechny tři skupiny se od sebe vzájemně liší statisticky významně.

„Což vzhledem k velikosti vzorku není nijak překvapivé.“

Věcný rozdíl mezi skupinou od 15 let a skupinou 60+ je zhruba půl směrodatné odchylky, přičemž skupina od 15 dosahuje vyššího skóru.

Skupina 23 – 60 let dosahuje v Testu schopnosti ze všech tří skupin nejlepších výsledků. Věcný rozdíl mezi skupinou od 15 let a skupinou 23 – 60 se rovná jedné směrodatné odchylce. Rozdíl mezi skupinou 23 – 60 a 60+ je tedy jedna a půl směrodatné odchylky.

Tuto otázku lze tedy uzavřít tak, že skupina od 15 let je o trochu lepší v hodnocení spolehlivosti mediálních zpráv, než skupina 60+. Skupina 23 – 60 je pak výrazněji silnější než obě zbývající skupiny.

Zpracovat tyto výsledky do článku

Bližší výsledky prozkoumat

Vaše vědecká zvědavost vám nedá a rozhodnete se detailněji prozkoumat výsledky. Proto si v dalším kroku zobrazíte rozložení Věku pomocí scatterplotu. Rozložení grafu napovídá, že by v daném vzorku mohly existovat nikoli tři, ale čtyři skupiny. Zjistíte, že dvě z těchto čtyřech skupin jsou tvořeny výhradně lidmi z věkové skupiny tři. Protože je ale vzhledem ke struktuře vašich dat nejste schopni rozlišit, nelze nijak získat další informace.

**Zpracovat tyto výsledky do
článku**

První ze tří částí máte úspěšně provedenou a sepsanou. Podařilo se vám potvrdit to, co předpokládá teorie. „Sakra, v důsledku vlastního špatného rozhodnutí jsem se připravil o další informace. Kdybych se nerozhodl sbírat věk pomocí intervalů, mohlo z toho vzejít ještě něco zajímavého. No co, z článku, se to nikdo nemusí dozvědět. Podám to jako úspěšné potvrzení dosavadní teorie.“

Dále se tedy věnujete zapracování těchto informací do článku, který hodláte na konci výzkumu publikovat. Po napsání příslušné části textu je čas přistoupit k dalším analýzám.

Znovu jste sedl k datovému souboru. Je na čase pokračovat druhou částí.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	1	2	1	4	1	8	NA	28	5	0
002	3	2	3	12	6	46	21 400	5	7	1
003	2	2	3	6	2	45	20 200	28	3	1
004	2	2	6	3	6	26	134 700	16	5	1
005	3	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

Dalším krokem je prozkoumání výzkumné otázky b). Tedy korelační vztah čistého příjmu s ochotou respondentů vystupovat na veřejnosti, která je reprezentována hrubým skórem v Dotazníku ochoty k veřejnému vystupování. V první fázi analýzy se bohužel ukázalo, že proměnná Čistý příjem je značné množství chybějících hodnot. Dále jste také zjistil, že tímto neduhem netrpí pouze proměnná Čistý příjem, ale také ostatní proměnné, se kterými jste doposud nepracoval. Tento problém je nutné nějak řešit.

„Hm, musím nějak naložit s těmi chybějícími hodnotami. V zásadě existují dva běžné způsoby, jak s tím naložit. Bud' se zabavím všech respondentů, od kterých nemám kompletní data a nebo chybějící data v Čistém příjmu nahradím průměrnými hodnotami.“

Jak tak v zamyšlení pochodujete ode zdi ke zdi po pracovně, padne vám zrak na knihu, na kterou jste dávno zapomněli. Její hřbet nese název *„Pokročilé postupy řešení statistických problémů.“*

Vymazat všechny nekompletní záznamy

Doplnit chybějící hodnoty příjmů průměrnými hodnotami

Nahlédnout do knihy *Pokročilých postupů*

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, ale vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

**Zpracovat tyto výsledky do
článku**

Pokročilé postupy řešení statistických problémů

Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p -hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu



Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat **„Hodnoty chybějící z definice“**. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých respondentů chybí uvedení

Otáčít na další
stranu



Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami i o zmíněných softwarech (Acock, 2005).

Zavřít knihu

Přejít na seznam
literatury

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Acock, A. C. (2005). Working with missing values. ***Journl of Marriage and family*** 67(4), 1012-1028.

Vraťit se k tomuto odkazu

Četbou knihy jste si ověřil své domněnky a načerpal novou inspiraci pro řešení nastalého problému. Teď je potřeba toto úsilí zužitkovat. Je potřeba posunout výzkum zase kousek dál a rozhodnout se, jak tedy vyřešíte problém s chybějícími hodnotami.

Vzhledem k tomu, že pro takzvané moderní postupy nemáte zařízení s dostatečnou výpočetní kapacitou, musíte vybrat to nejvhodnější z tradičních postupů.

Použít Case delation

Použít Pairwise delation

Použít Mean substitution

Použít Indicator Adjustment

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, ale vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po použití Pairwise delation z této konkrétní analýzy vypadlo 25% výzkumného vzorku. Vzhledem k jeho rozsahu si analýza i v tomto rozsahu zachovává svou statistickou sílu. Odhalila silnou pozitivní korelaci mezi výší čistého příjmu a ochotou k veřejným vystoupením. Tento výsledek podporuje dosavadní mínění teorie, že bohatší lidé jsou ochotnější k veřejnému vystupování. Z dostupných dat však nelze zjistit nic o tom, jaký mechanismus tuto větší ochotu způsobuje.

„To by chtělo navazující kvalitativní studii, která by se bohatých lidí ptala na důvody, proč jsou ochotní vystupovat na veřejnosti a snažila se odhalit nějaký mechanismus. Hm. Tohle je dobrá věta do diskuse. To si musím někde poznačit.“

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po aplikování Indicator adjustment postupu a doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost korelace čistého příjmu a hrubého skóru Dotazníku, nicméně tato korelace se ukázala být velmi nízká. Díky indikátoru víte, že průměrný plat byl doplněn do 25% případů. Momentálně se vám tedy podařilo vyvrátit mylný předpoklad teorie, je však otázka, jak moc oněch 25% doplněných hodnot tento výsledek zkresluje.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu. V prvním případě jste podpořil prozatimní teoretické předpoklady, i když s jistou pochybností a v druhém případě jste teorii potvrdil. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydl. bydliště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	1	2	1	4	1	8	NA	28	5	0
002	3	2	3	12	6	46	21 400	5	7	1
003	2	2	3	6	2	45	20 200	28	3	1
004	2	2	6	3	6	26	134 700	16	5	1
005	3	2	1	4	5	12	NA	19	46	1
...										

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. V prvním případě jste podpořil prozatimní teoretické předpoklady, i když s jistou pochybností a v druhém případě, pokud připustíme korekci síly korelace, která je zkreslená 25% doplněných průměrných hodnot, tak by ta korelace mohla být dostatečně silná, aby to stálo za zmínku. V podstatě to teorii podporuje. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	1	2	1	4	1	8	NA	28	5	0
002	3	2	3	12	6	46	21 400	5	7	1
003	2	2	3	6	2	45	20 200	28	3	1
004	2	2	6	3	6	26	134 700	16	5	1
005	3	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. V prvním případě jste podpořil prozatimní teoretické předpoklady, i když s jistou pochybností a v druhém případě vaše zjištění teorii spíše nabourávají. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	1	2	1	4	1	8	NA	28	5	0
002	3	2	3	12	6	46	21 400	5	7	1
003	2	2	3	6	2	45	20 200	28	3	1
004	2	2	6	3	6	26	134 700	16	5	1
005	3	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. O'Boe, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrní a interpretovat je s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

**Zpracovat tyto výsledky do
článku**

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věci. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci interpretativní zodpovědnosti

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý. Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů. Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

Použít výsledky včetně outlierů

Vyřadit outliery a provést analýzy znovu

**Nahlédnout do knihy
*„Pokročilé postupy řešení statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermana Aguinise, Ryana K. Gottfredsona a Harryho J. O'Boe, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

První ze tří částí máte úspěšně provedenou a sepsanou. Podařilo se vám potvrdit to, co předpokládá teorie. Dále se tedy věnujete zapracování těchto informací do článku, který hodláte na konci výzkumu publikovat. Po napsání příslušné části textu je čas přistoupit k dalším analýzám.

Znovu jste sedl k datovému souboru. Je na čase pokračovat druhou částí.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	1	2	1	4	1	8	NA	28	5	0
002	3	2	3	12	6	46	21 400	5	7	1
003	2	2	3	6	2	45	20 200	28	3	1
004	2	2	6	3	6	26	134 700	16	5	1
005	3	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

Dalším krokem je prozkoumání výzkumné otázky b). Tedy korelační vztah čistého příjmu s ochotou respondentů vystupovat na veřejnosti, která je reprezentována hrubým skórem v Dotazníku ochoty k veřejnému vystupování. V první fázi analýzy se bohužel ukázalo, že proměnná Čistý příjem je značné množství chybějících hodnot. Dále jste také zjistil, že tímto neduhem netrpí pouze proměnná Čistý příjem, ale také ostatní proměnné, se kterými jste doposud nepracoval. Tento problém je nutné nějak řešit.

„Hm, musím nějak naložit s těmi chybějícími hodnotami. V zásadě existují dva běžné způsoby, jak s tím naložit. Buď se zabavím všech respondentů, od kterých nemám kompletní data a nebo chybějící data v Čistém příjmu nahradím průměrnými hodnotami.“

Jak tak v zamyšlení pochodujete ode zdi ke zdi po pracovně, padne vám zrak na knihu, na kterou jste dávno zapomněli. Její hřbet nese název *„Pokročilé postupy řešení statistických problémů.“*

Vymazat všechny nekompletní záznamy

Doplnit chybějící hodnoty příjmů průměrnými hodnotami

Nahlédnout do knihy *Pokročilých postupů*

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané Státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

**Zpracovat tyto výsledky do
článku**

Pokročilé postupy řešení statistických problémů

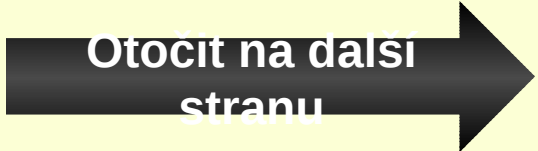
Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p-hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu

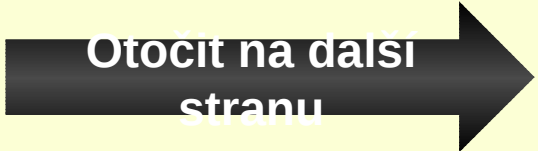


Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat „**Hodnoty chybějící z definice**“. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



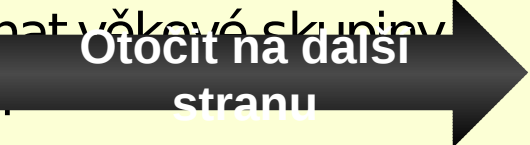
Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých respondentů chybí uvedení

Otáčet na další
stranu



Tento postup zachovává co největší množství respondentů a tedy co největší statistickou sílu, je zde však problém, že vzorek pro zpracování každé jednotlivé analýzy je pak odlišný a nelze tudíž jednotlivé kroky slučovat. V tomto případě by tedy nebylo možné studovat interakce věku a asertivity z hlediska prožívaného wellbeingu. Tento postup by vyžadoval Case delation.

Mean substitution je postup, který nahrazuje chybějící hodnoty průměrem vzorku v dané proměnné. Jedná se o logický postup, jestliže lze předpokládat, že daná proměnná ve vzorku odpovídá normálnímu rozložení. V takovém případě, neznáme-li hodnotu dané proměnné u konkrétní osoby, nejpravděpodobnější je, že tato hodnota se pohybuje blízko kolem průměru. Tento postup je však problematický, pokud v konkrétním vzorku je velké množství chybějících hodnot. Jestliže ve vzorku chybí například 30% hodnot, pak nahrazení všech těchto hodnot průměrem značně redukuje rozptyl ve vzorku, a tím pádem podhodnocuje hodnotu korelace. Dalším problémem je necitlivost takového postupu vůči extrémním hodnotám. Jestliže mi u respondenta chybí jeho měsíční příjem a tento člověk je buďto milionář nebo se naopak potácí na hranici chudoby, pak v obou těchto případech je nahrazení této chybějící hodnoty průměrem velmi nešťastné a povede ke zkreslení výsledků. Pomocí Mean substitution není také vhodné doplňovat z definice chybějící hodnoty.

Indicator adjustment je lehká úprava strategie doplňování průměru. V tomto případě se do datového souboru přidá proměnná, tzv. indikátor, který nabývá hodnoty 0 v případech, kdy cílová hodnota chybí a 1 v případech, kdy je cílová hodnota pozorována. Následně jsou chybějící hodnoty doplněny průměrem vzorku. Přítomnost indikátoru zde udržuje informaci o počtu doplněných hodnot a tím pádem možného rozsahu zkreslení. Více o tomto postupu (Cohen 1983) a následně [Otáčit na další stranu](#)

Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami i o zmíněných softwarech (Acock, 2005).

[Přejít na seznam literatury](#)

Zavřít knihu

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.
[Vratit se k tomuto odkazu](#)
- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.
[Vratit se k tomuto odkazu](#)
- Acock, A. C. (2005). Working with missing values. ***Journal of Marriage and Family*** 67(4), 1012-1028.
[Vratit se k tomuto odkazu](#)

Četbou knihy jste si ověřil své domněnky a načerpal novou inspiraci pro řešení nastalého problému. Teď je potřeba toto úsilí zužitkovat. Je potřeba posunout výzkum zase kousek dál a rozhodnout se, jak tedy vyřešíte problém s chybějícími hodnotami. Vzhledem k tomu, že pro takzvané moderní postupy nemáte zařízení s dostatečnou výpočetní kapacitou, musíte vybrat to nejvhodnější z tradičních postupů.

Použít Case delation

Použít Pairwise delation

Použít Mean substitution

Použít Indicator Adjustment

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tři set osob. Takto redukováný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané Státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po použití Pairwise delation z této konkrétní analýzy vypadlo 25% výzkumného vzorku. Vzhledem k jeho rozsahu si analýza i v tomto rozsahu zachovává svou statistickou sílu. Odhalila silnou pozitivní korelaci mezi výší čistého příjmu a ochotou k veřejným vystoupením. Tento výsledek podporuje dosavadní mínění teorie, že bohatší lidé jsou ochotnější k veřejnému vystupování. Z dostupných dat však nelze zjistit nic o tom, jaký mechanismus tuto větší ochotu způsobuje.

„To by chtělo navazující kvalitativní studii, která by se bohatých lidí ptala na důvody, proč jsou ochotní vystupovat na veřejnosti a snažila se odhalit nějaký mechanismus. Hm. Tohle je dobrá věta do diskuse. To si musím někde poznačit.“

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po aplikování Indicator adjustment postupu a doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost korelace čistého příjmu a hrubého skóru Dotazníku ochoty k vystupování, nicméně tato korelace se ukázala být velmi nízká. Díky indikátoru víte, že průměrný plat byl doplněn do 25% případů. Momentálně se vám tedy podařilo vyvrátit mylný předpoklad teorie, je však otázka, jak moc oněch 25% doplněných hodnot tento výsledek zkresluje.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedli korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu. V prvním případě jste mezeru v teorii a v druhém případě jste teorii potvrdil. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo korigovat jednu nepřesnost v teorii a v druhém případě, pokud připustíme korekci síly korelace, která je zkreslená 25% doplněných průměrných hodnot, tak by ta korelace mohla být dostatečně silná, aby to stálo za zmínku. V podstatě to teorii podporuje. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Vzdělán í	Kraj	Velikost bydliště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo najít dvě trhliny v teoretických předpokladech. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste se sešli k datovému souboru.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nhlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“*a**

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

Zabývat se výzkumnou **otázkou**
interakce

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. O'Boe, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

Použít výsledky včetně outlierů

Vyřadit outliery a provést analýzy znovu

**Nahlédnout do knihy
*„Pokročilé postupy řešení statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outlierem jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otociť na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v tomto interpretativní zodpovědnosti

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Vzhledem k rozsahu dat přemýšlíte, že byste se ještě zabýval vlivem interakce politické aktivity a věku na míru agresivity

**Zabývat se výzkumnou
otázkou interakce**

Publikovat

Zpracováním těchto dat byl článek dokončen!

Rád byste se ještě zabýval doplňkovým výzkumem, abyste maximalizoval užitek z obdržených dat, nicméně podrobná analýza outlierů si vyžádala příliš mnoho času.

Publikovat

„Fajn, potřebuju vytvořit Anamnestický dotazník, který mi bude pokrývat potřebné výzkumné otázky. Takže k výzkumné otázce a) potřebuju věk, k výzkumné otázce b) potřebuju čistý příjem a k výzkumné otázce c) potřebuju znát pohlaví.

Co se týče kódování, tak nejjednodušší je pohlaví. Tam není co řešit. 1 – muži, 2 – ženy. Čistý příjem se u nás, až na nějaké extrémy, běžně pohybuje mezi 10 000 Mincí a 50 000 Mincí. Při takovém rozpětí hodnot nemá význam řešit každou jednotku. Nemůžu to ale zase zkrouhnout moc. Kdybych to rozlišoval příliš hrubě, nebylo by možné udělat korelační analýzu. Proto by bylo nejspíš užitečnější zaokrouhlovat příjem na stovky Mincí. Ovšem větší problém bude s tím, jak kódovat věk. V zásadě existují dvě varianty, jak k tomu přistoupit. První možnost je nechat všechny vyplnit svůj věk v letech. Snadné a rychlé. Ale trochu problematické. Občasné státní jsou v tomto ohledu dost háklivé a neradi takové informace sdělují. Ale vzhledem k anonymitě dotazování by to snad ani nemusel být problém. Ač, to vlastně nevím, jakým způsobem to v Institutu sbírají. No, je to trochu riskantní. Navíc nějaký šťoura to může napadnou z hlediska etiky. Abych se tomu vyhnul, můžu se přidržet teorie a nechat každého vybrat, do kterého ze tří intervalů spadá. To znamená 15 – 22; 23 – 60; 60+. Ten samý problém vlastně vyvstává i u příjmu, pokud zvolím jako interval, opět narazím na problém s korelační analýzou. Takže tam nemá

Podstoupit drobné riziko a zvolit variantu konkrétních věků

Přidržet se doporučení teorie a zvolit variantu intervalů.

Rozhodl jste se pro variantu konkrétních věků. To znamená, že finální verze vašeho Anamnestického dotazníku je strukturovaná následujícím způsobem:

Pohlaví: 1 – muži, 2 – ženy

Věk: udávaný v letech

Příjem: zaokrouhlovat na stovky

K tomu přikládáte Test spolehlivosti zpráv, Dotazník ochoty a Dotazník agresivity.

„V tomto formátu bych snad měl mít všechny potřebné informace.“

**Zkompletovat sadu
materiálů a odeslat na
adresu Institutu**

**Nespěchat s odesláním.
Přemýšlet, zda by nebylo
užitečné ještě něco doplnit.**

Následujících několik dnů přemýšlíte, co ještě přidat k Anamnestickému dotazníku. Přemýšlíte, k čemu je celý výzkum vlastně určený a co by ještě bylo dobré k tomu přidat.

„Jedna věc je splnit zadání, ale když už je tady taková příležitost, mělo by se toho pořádně využít. Přece jenom, bude to výzkum na velkém množství lidí a když přijdu navíc s nějakým zajímavým zjištěním, mohl bych napravit pošramocené jméno Akademie. Beztak za to můžu já. A co být troškař, třeba se zase stanu slavným vědcem!“

Chodíte ode zdi ke zdi a hlava se vám plní různými nápady, které ale vzápětí zase zahazujete. Nic z toho nemá potenciál k převratnému objevu, který potřebujete. Nakonec začnete uvažovat, jestli má cenu se do toho vlastně pouštět.

**Vzdát snahu o rozšíření
výzkumu**

Dát své kreativitě ještě šanci

Vidina převratného úspěchu je tak silná, že se vám nechce hned tak vzdát. Dalších několik dní chodíte po pracovně ode zdi ke zdi. Nakonec se vám v hlavně přece jenom zrodí nápad, který stojí za pokus.

„Když k tomu přidám ještě proměnou týkající se Akademie samotné, můžou z toho být docela zajímavé výsledky. Bud' se ukáže, že lidé vnímají Akademii jako užitečnou a prospěšnou organizaci, což určitě rádo uslyší naše PR oddělení a nebo se ukáže, že jsme vnímání jako zbytečný žrout peněz občanů našeho Státu. To zas bude důležité pro ředitele, protože takové smýšlení je pro nás ohrožující a budeme s tím muset něco dělat. Tak jako tak to pro nás bude užitečné. K tomu, abych něco takového zjistil ale potřebuju nějakou metodu, která zjistí vztah běžného občana k Akademii.“

Hned se do svého záměru pustíte a začnete skládat metodu, která má odhalit vztah běžného člověka k Akademii. A tak ponořen do této práce trávíte den za dnem. Až jednoho odpoledne vás ze zadumání vytrhne zaklepání na dveře. Za nimi stojí onen kolega, kterého jste onehdá potkal na chodbě.

„Abraháme, máš jít za ředitelem.“

„Jo, hned tam budu,“ odpovíte a zamíříte ven z pracovny.

„Abraháme,“ uvítá vás ředitel, „přišel mi dopis z Institutu. Poslouchejte:

Vážená Akademie, k dnešnímu dni jste nedodali materiály ke sběru dat v Celostátním víceúčelovém výzkumu. Musíme vám tedy s politováním oznámit, že vaše účast v tomto šetření je tímto zrušena. Víte, co to znamená, Abraháme? Znamená to, že jste zklamal důvěru, kterou jsem ve vás vložil!“

Jste nucen vyklidit kancelář a v tichosti opustit Akademii. Doma usedáte k počítači a začínáte procházet aktuální

Zvrátit poslední volbu

Znovu zahájit výzkum

Je čas dát všechno do obálky a poslat to Institutu. Nezbyvá než čekat, až data nasbírají a pošlou vám je k analýze. Chvilé klidu je zasloužená. Zatím jste odvedli dobrou práci.

Ona chvíle odpočinku ale netrvá zase příliš dlouho. Za nějaký čas obdržíte datový soubor. Když ho otevřete v počítači, uvidíte následující tabulku:

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	15	2	NA	8	28	5
002	62	2	21 400	46	5	7
003	35	2	20 200	45	28	3
004	46	2	134 700	26	16	5
005	67	2	NA	12	19	46
...						

Tímto způsobem tabulka pokračuje až po ID 800.

„Výborně, a teď je čas se pustit do analýz.“

Začít analyzovat data

Je na čase pustit se do další práce. Zkoumáte, jak souvisí věk se schopností posuzovat spolehlivost mediálních zpráv. V tuto chvíli je čas přidržet se známé teorie a ze všeho nejdřív sloučit respondenty do skupin podle věku tak, jak bylo uvedeno v rešerši.

Vytvoříte si tedy novou proměnnou „Věk_Int“, která nabývá hodnoty 1 pro lidi ve věku 15 – 22 let, hodnoty 2 pro lidi mezi 23 a 60 lety a hodnoty 3 pro lidi starší 60 let. Tyto tři skupiny pak porovnáte v hrubých skórech Dotazníku mediálních zpráv.

Výsledky lze shrnout asi takto:

Všechny tři skupiny se od sebe vzájemně liší statisticky významně.

„Což vzhledem k velikosti vzorku není nijak překvapivé.“

Věcný rozdíl mezi skupinou od 15 let a skupinou 60+ je zhruba půl směrodatné odchylky, přičemž skupina od 15 dosahuje vyššího skóru.

Skupina 23 – 60 let dosahuje v Testu schopnosti ze všech tří skupin nejlepších výsledků. Věcný rozdíl mezi skupinou od 15 let a skupinou 23 – 60 se rovná jedné směrodatné odchylce. Rozdíl mezi skupinou 23 – 60 a 60+ je tedy jedna a půl směrodatné odchylky.

Tuto otázku lze tedy uzavřít tak, že skupina od 15 let je o trochu lepší v hodnocení spolehlivosti mediálních zpráv, než skupina 60+. Skupina 23 – 60 je pak výrazněji silnější než obě zbývající skupiny.

Zpracovat tyto výsledky do článku

Bližší výsledky prozkoumat

Vaše vědecká zvědavost vám nedá a rozhodnete se detailněji prozkoumat výsledky. Proto si v dalším kroku zobrazíte rozložení proměnné Věk_Int pomocí scatterplotu. Rozložení grafu napovídá, že by v daném vzorku mohly existovat nikoli tři, ale čtyři skupiny. Proto si zobrazíte ještě jeden scatterplot, tentokrát z proměnné Věk. Analýzou jednotlivých případů se pak ukáže, že se vzorek skutečně rozdělil do čtyř nezávislých skupin.

Zůstala zachována skupina 15 – 22 let a skupina 60+, ale prostřední skupina se navíc rozdělila na dvě. Vznikla tak skupina 23 – 45 let a skupina 46 – 60 let.

Vytvoříte si další proměnnou „Věk_Int2“, která nabývá hodnoty 1 pro interval 15 – 22let, 2 pro interval 23 – 45, 3 pro interval 46 – 60 let a 4 pro skupinu 60+. Tyto čtyři skupiny pak znovu porovnáte v hrubých skórech Testu důvěryhodnosti mediálních zpráv.

Ukázalo se, že v takovém rozložení jsou výsledky poněkud jiné.

Skupina 4 dosahuje nejhorších výsledků. Skupiny 1 a 3, které se vzájemně neliší, dosahují o půl směrodatné odchylky lepších výsledků než skupina 3. Skupina 2 dosahuje zdaleka nejlepších výsledků o dvě a půl směrodatné odchylky lepších než skupina 4.

Z těchto výsledků tedy vyplývá, že lidé starší 60 let dokáží jen velmi špatně rozlišovat spolehlivé a nespolehlivé zprávy. O něco lépe jsou na tom v tomto směru lidé mezi 15 a 22 a mezi 46 a 60 lety. Zdaleka nejlepší v tomto jsou pak lidé mezi 23 a 45 lety.

Zpracovat tyto výsledky do článku

První ze tří částí máte úspěšně provedenou a sepsanou. Dokonce jste přišel s výsledky, které do značné míry korigují dosavadní předpoklady. To lze považovat za velký úspěch. Dále se tedy věnujete zpracování těchto informací do článku, který hodláte na konci výzkumu publikovat. Po napsání příslušné části textu je čas přistoupit k dalším analýzám.

Znovu jste sedl k datovému souboru. Je na čase pokračovat druhou částí.

ID	Věk	Pohlaví	Vzdělání	Kraj	Velikost bydlíště	Schopnost rozlišování zpráv	Čistý příjem	Ochota k veřejnému vystupování	Agresivita vůči vládě	Politická aktivita
001	15	2	1	4	1	8	NA	28	5	0
002	62	2	3	12	6	46	21 400	5	7	1
003	35	2	3	6	2	45	20 200	28	3	1
004	46	2	6	3	6	26	134 700	16	5	1
005	67	2	1	4	5	12	NA	19	46	1
...										

Pokračovat další výzkumnou otázkou

Dalším krokem je prozkoumání výzkumné otázky b). Tedy korelační vztah čistého příjmu s ochotou respondentů vystupovat na veřejnosti, která je reprezentována hrubým skórem v Dotazníku ochoty k veřejnému vystupování. V první fázi analýzy se bohužel ukázalo, že proměnná Čistý příjem je značné množství chybějících hodnot. Dále jste také zjistil, že tímto neduhem netrpí pouze proměnná Čistý příjem, ale také ostatní proměnné, se kterými jste doposud nepracoval. Tento problém je nutné nějak řešit.

„Hm, musím nějak naložit s těmi chybějícími hodnotami. V zásadě existují dva běžné způsoby, jak s tím naložit. Buď se zabavím všech respondentů, od kterých nemám kompletní data a nebo chybějící data v Čistém příjmu nahradím průměrnými hodnotami.“

Jak tak v zamyšlení pochodujete ode zdi ke zdi po pracovně, padne vám zrak na knihu, na kterou jste dávno zapomněli. Její hřbet nese název *„Pokročilé postupy řešení statistických problémů.“*

Vymazat všechny nekompletní záznamy

Doplnit chybějící hodnoty příjmů průměrnými hodnotami

Nahlédnout do knihy *Pokročilých postupů*

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukováný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, ale vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

**Zpracovat tyto výsledky do
článku**

Pokročilé postupy řešení statistických problémů

Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p -hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu

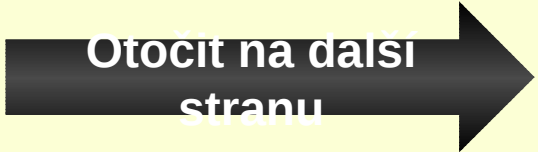


Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat „**Hodnoty chybějící z definice**“. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých respondentů

Otočit na další
stranu



Tento postup zachovává co největší množství respondentů a tedy co největší statistickou sílu, je zde však problém, že vzorek pro zpracování každé jednotlivé analýzy je pak odlišný a nelze tudíž jednotlivé kroky slučovat. V tomto případě by tedy nebylo možné studovat interakce věku a asertivity z hlediska prožívaného wellbeingu. Tento postup by vyžadoval Case delation.

Mean substitution je postup, který nahrazuje chybějící hodnoty průměrem vzorku v dané proměnné. Jedná se o logický postup, jestliže lze předpokládat, že daná proměnná ve vzorku odpovídá normálnímu rozložení. V takovém případě, neznáme-li hodnotu dané proměnné u konkrétní osoby, nejpravděpodobnější je, že tato hodnota se pohybuje blízko kolem průměru. Tento postup je však problematický, pokud v konkrétním vzorku je velké množství chybějících hodnot. Jestliže ve vzorku chybí například 30% hodnot, pak nahrazení všech těchto hodnot průměrem značně redukuje rozptyl ve vzorku, a tím pádem podhodnocuje hodnotu korelace. Dalším problémem je necitlivost takového postupu vůči extrémním hodnotám. Jestliže mi u respondenta chybí jeho měsíční příjem a tento člověk je buďto milionář nebo se naopak potácí na hranici chudoby, pak v obou těchto případech je nahrazení této chybějící hodnoty průměrem velmi nešťastné a povede ke zkreslení výsledků. Pomocí Mean substitution není také vhodné doplňovat z definice chybějící hodnoty.

Indicator adjustment je lehká úprava strategie doplňování průměru. V tomto případě se do datového souboru přidá proměnná, tzv. indikátor, který nabývá hodnoty 0 v případech, kdy cílová hodnota chybí a 1 v případech, kdy je cílová hodnota pozorována. Následně jsou chybějící hodnoty doplněny průměrem vzorku. Přítomnost indikátoru zde udržuje informaci o počtu doplněných hodnot a tím pádem možného rozsahu zkreslení. Více o tomto postupu (Cohen 1983) a následně

[Přejít na seznam literatury](#)

[Otočit na další stranu](#)

Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami i o zmíněných softwarech (Acock, 2005).

[Přejít na seznam literatury](#)

Zavřít knihu

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Acock, A. C. (2005). Working with missing values. ***Journal of Marriage and Family*** 67(4), 1012-1028.

Vraťit se k tomuto odkazu

Četbou knihy jste si ověřil své domněnky a načerpal novou inspiraci pro řešení nastalého problému. Teď je potřeba toto úsilí zužitkovat. Je potřeba posunout výzkum zase kousek dál a rozhodnout se, jak tedy vyřešíte problém s chybějícími hodnotami. Vzhledem k tomu, že pro takzvané moderní postupy nemáte zařízení s dostatečnou výpočetní kapacitou, musíte vybrat to nejvhodnější z tradičních postupů.

Použít Case delation

Použít Pairwise delation

Použít Mean substitution

Použít Indicator Adjustment

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tři set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, ale vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po použití Pairwise delation z této konkrétní analýzy vypadlo 25% výzkumného vzorku. Vzhledem k jeho rozsahu si analýza i v tomto rozsahu zachovává svou statistickou sílu. Odhalila silnou pozitivní korelaci mezi výší čistého příjmu a ochotou k veřejným vystoupením. Tento výsledek podporuje dosavadní mínění teorie, že bohatší lidé jsou ochotnější k veřejnému vystupování. Z dostupných dat však nelze zjistit nic o tom, jaký mechanismus tuto větší ochotu způsobuje.

„To by chtělo navazující kvalitativní studii, která by se bohatých lidí ptala na důvody, proč jsou ochotní vystupovat na veřejnosti a snažila se odhalit nějaký mechanismus. Hm. Tohle je dobrá věta do diskuse. To si musím někde poznačit.“

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po aplikování Indicator adjustment postupu a doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost korelace čistého příjmu a hrubého skóru Dotazníku, nicméně tato korelace se ukázala být velmi nízká. Díky indikátoru víte, že průměrný plat byl doplněn do 25% případů. Momentálně se vám tedy podařilo vyvrátit mylný předpoklad teorie, je však otázka, jak moc oněch 25% doplněných hodnot tento výsledek zkresluje.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu. V prvním případě jste objevil mezeru v teorii a v druhém případě jste teorii potvrdil. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	15	2	NA	8	28	5
002	62	2	21 400	46	5	7
003	35	2	20 200	45	28	3
004	46	2	134 700	26	16	5
005	67	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo korigovat jednu nepřesnost v teorii a v druhém případě, pokud připustíme korekci síly korelace, která je zkreslená 25% doplněných průměrných hodnot, tak by ta korelace mohla být dostatečně silná, aby to stálo za zmínku. V podstatě to teorii podporuje. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	15	2	NA	8	28	5
002	62	2	21 400	46	5	7
003	35	2	20 200	45	28	3
004	46	2	134 700	26	16	5
005	67	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo najít dvě trhliny v teoretických předpokladech. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	15	2	NA	8	28	5
002	62	2	21 400	46	5	7
003	35	2	20 200	45	28	3
004	46	2	134 700	26	16	5
005	67	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

Použít výsledky včetně outlierů

Vyřadit outliery a provést analýzy znovu

**Nahlédnout do knihy
*„Pokročilé postupy řešení statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

Použít výsledky včetně outlierů

Vyřadit outliery a provést analýzy znovu

**Nahlédnout do knihy
*„Pokročilé postupy řešení statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. O'Learyho, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů jejich detekce a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrní a interpretovatelní zodpovědnosti.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrní a interpretovatelní zodpovědnosti.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

První ze tří částí máte úspěšně provedenou a sepsanou. Podařilo se vám potvrdit to, co předpokládá teorie. Dále se tedy věnujete zpracování těchto informací do článku, který hodláte na konci výzkumu publikovat. Po napsání příslušné části textu je čas přistoupit k dalším analýzám.

Znovu jste sedl k datovému souboru. Je na čase pokračovat druhou částí.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	15	2	NA	8	28	5
002	62	2	21 400	46	5	7
003	35	2	20 200	45	28	3
004	46	2	134 700	26	16	5
005	67	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

Dalším krokem je prozkoumání výzkumné otázky b). Tedy korelační vztah čistého příjmu s ochotou respondentů vystupovat na veřejnosti, která je reprezentována hrubým skórem v Dotazníku ochoty k veřejnému vystupování. V první fázi analýzy se bohužel ukázalo, že proměnná Čistý příjem je značné množství chybějících hodnot. Dále jste také zjistil, že tímto neduhem netrpí pouze proměnná Čistý příjem, ale také ostatní proměnné, se kterými jste doposud nepracoval. Tento problém je nutné nějak řešit.

„Hm, musím nějak naložit s těmi chybějícími hodnotami. V zásadě existují dva běžné způsoby, jak s tím naložit. Buď se zabavím všech respondentů, od kterých nemám kompletní data a nebo chybějící data v Čistém příjmu nahradím průměrnými hodnotami.“

Jak tak v zamyšlení pochodujete ode zdi ke zdi po pracovně, padne vám zrak na knihu, na kterou jste dávno zapomněli. Její hřbet nese název *„Pokročilé postupy řešení statistických problémů.“*

Vymazat všechny nekompletní záznamy

Doplnit chybějící hodnoty příjmů průměrnými hodnotami

Nahlédnout do knihy *Pokročilých postupů*

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukováný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané Státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

**Zpracovat tyto výsledky do
článku**

Pokročilé postupy řešení statistických problémů

Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p -hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu

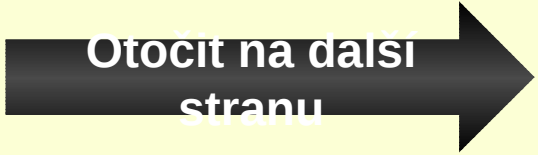


Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat „**Hodnoty chybějící z definice**“. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých respondentů

Otočit na další
stranu



Tento postup zachovává co největší množství respondentů a tedy co největší statistickou sílu, je zde však problém, že vzorek pro zpracování každé jednotlivé analýzy je pak odlišný a nelze tudíž jednotlivé kroky slučovat. V tomto případě by tedy nebylo možné studovat interakce věku a asertivity z hlediska prožívaného wellbeingu. Tento postup by vyžadoval Case delation.

Mean substitution je postup, který nahrazuje chybějící hodnoty průměrem vzorku v dané proměnné. Jedná se o logický postup, jestliže lze předpokládat, že daná proměnná ve vzorku odpovídá normálnímu rozložení. V takovém případě, neznáme-li hodnotu dané proměnné u konkrétní osoby, nejpravděpodobnější je, že tato hodnota se pohybuje blízko kolem průměru. Tento postup je však problematický, pokud v konkrétním vzorku je velké množství chybějících hodnot. Jestliže ve vzorku chybí například 30% hodnot, pak nahrazení všech těchto hodnot průměrem značně redukuje rozptyl ve vzorku, a tím pádem podhodnocuje hodnotu korelace. Dalším problémem je necitlivost takového postupu vůči extrémním hodnotám. Jestliže mi u respondenta chybí jeho měsíční příjem a tento člověk je buďto milionář nebo se naopak potácí na hranici chudoby, pak v obou těchto případech je nahrazení této chybějící hodnoty průměrem velmi nešťastné a povede ke zkreslení výsledků. Pomocí Mean substitution není také vhodné doplňovat z definice chybějící hodnoty.

Indicator adjustment je lehká úprava strategie doplňování průměru. V tomto případě se do datového souboru přidá proměnná, tzv. indikátor, který nabývá hodnoty 0 v případech, kdy cílová hodnota chybí a 1 v případech, kdy je cílová hodnota pozorována. Následně jsou chybějící hodnoty doplněny průměrem vzorku. Přítomnost indikátoru zde udržuje informaci o počtu doplněných hodnot a tím pádem umožňuje rozsahu zkreslení. Více o tomto postupu například [Prejít na seznam literatury](#) (3) [Otočit na další stranu](#)

Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami i o zmíněných softwarech (Acock, 2005).

[Přejít na seznam literatury](#)

Zavřít knihu

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Acock, A. C. (2005). Working with missing values. ***Journd of Marriage and family*** 67(4), 1012-1028.

Vraťit se k tomuto odkazu

Četbou knihy jste si ověřil své domněnky a načerpal novou inspiraci pro řešení nastalého problému. Teď je potřeba toto úsilí zužitkovat. Je potřeba posunout výzkum zase kousek dál a rozhodnout se, jak tedy vyřešíte problém s chybějícími hodnotami. Vzhledem k tomu, že pro takzvané moderní postupy nemáte zařízení s dostatečnou výpočetní kapacitou, musíte vybrat to nejvhodnější z tradičních postupů.

Použít Case delation

Použít Mean substitution

Použít Pairwise delation

Použít Indicator Adjustment

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukováný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané Státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po použití Pairwise delation z této konkrétní analýzy vypadlo 25% výzkumného vzorku. Vzhledem k jeho rozsahu si analýza i v tomto rozsahu zachovává svou statistickou sílu. Odhalila silnou pozitivní korelaci mezi výší čistého příjmu a ochotou k veřejným vystoupením. Tento výsledek podporuje dosavadní mínění teorie, že bohatší lidé jsou ochotnější k veřejnému vystupování. Z dostupných dat však nelze zjistit nic o tom, jaký mechanismus tuto větší ochotu způsobuje.

„To by chtělo navazující kvalitativní studii, která by se bohatých lidí ptala na důvody, proč jsou ochotní vystupovat na veřejnosti a snažila se odhalit nějaký mechanismus. Hm. Tohle je dobrá věta do diskuse. To si musím někde poznačit.“

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po aplikování Indicator adjustment postupu a doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost korelace čistého příjmu a hrubého skóru Dotazníku ochoty k vystupování, nicméně tato korelace se ukázala být velmi nízká. Díky indikátoru víte, že průměrný plat byl doplněn do 25% případů. Momentálně se vám tedy podařilo vyvrátit mylný předpoklad teorie, je však otázka, jak moc oněch 25% doplněných hodnot tento výsledek zkresluje.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu. V prvním případě jste mezeru v teorii a v druhém případě jste teorii potvrdil. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	15	2	NA	8	28	5
002	62	2	21 400	46	5	7
003	35	2	20 200	45	28	3
004	46	2	134 700	26	16	5
005	67	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo korigovat jednu nepřesnost v teorii a v druhém případě, pokud připustíme korekci síly korelace, která je zkreslená 25% doplněných průměrných hodnot, tak by ta korelace mohla být dostatečně silná, aby to stálo za zmínku. V podstatě to teorii podporuje. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	15	2	NA	8	28	5
002	62	2	21 400	46	5	7
003	35	2	20 200	45	28	3
004	46	2	134 700	26	16	5
005	67	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo najít dvě trhliny v teoretických předpokladech. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	15	2	NA	8	28	5
002	62	2	21 400	46	5	7
003	35	2	20 200	45	28	3
004	46	2	134 700	26	16	5
005	67	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
„Pokročilé postupy řešení
statistických problémů“**

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermana Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrně zodpovědní.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný poustup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**


**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.

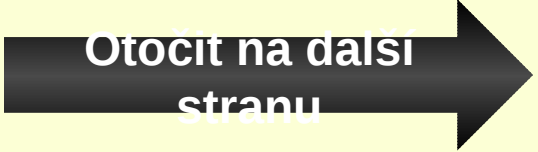
První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outlierem jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermana Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrní a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

Zpracovat tyto výsledky do článku

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

[Vrátit se k tomuto odkazu](#)

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Rozhodl jste se pro variantu intervalů. To znamená, že finální verze vašeho Anamnestického dotazníku je strukturovaná následujícím způsobem:

Pohlaví: 1 – muži, 2 – ženy

Věk: 1 – 15 – 22 let; 2 – 23 – 60 let; 3 – 60+ let

Příjem: zaokrouhlovat na stovky

K tomu přikládáte Test spolehlivosti zpráv, Dotazník ochoty a Dotazník agresivity.

„V tomto formátu bych snad měl mít všechny potřebné informace.“

**Zkompletovat sadu
materiálů a odeslat na
adresu Institutu**

**Nespěchat s odesláním.
Přemýšlet, zda by nebylo
užitečné ještě něco doplnit.**

Následujících několik dnů přemýšlíte, co ještě přidat k Anamnestickému dotazníku. Přemýšlíte, k čemu je celý výzkum vlastně určený a co by ještě bylo dobré k tomu přidat.

„Jedna věc je splnit zadání, ale když už je tady taková příležitost, mělo by se toho pořádně využít. Přece jenom, bude to výzkum na velkém množství lidí a když přijdu navíc s nějakým zajímavým zjištěním, mohl bych napravit pošramocené jméno Akademie. Beztak za to můžu já. A co být troškař, třeba se zase stanu slavným vědcem!“

Chodíte ode zdi ke zdi a hlava se vám plní různými nápady, které ale vzápětí zase zahazujete. Nic z toho nemá potenciál

k převratnému objevu, který potřebujete. Nakonec začnete uvažovat, jestli má cenu se do toho vlastně pouštět.

**Vzdát snahu o rozšíření
výzkumu**

Dát své kreativě ještě šanci

Vidina převratného úspěchu je tak silná, že se vám nechce hned tak vzdát. Dalších několik dní chodíte po pracovně ode zdi ke zdi. Nakonec se vám v hlavně přece jenom zrodí nápad, který stojí za pokus.

„Když k tomu přidám ještě proměnou týkající se Akademie samotné, můžou z toho být docela zajímavé výsledky. Bud' se ukáže, že lidé vnímají Akademii jako užitečnou a prospěšnou organizaci, což určitě rádo uslyší naše PR oddělení a nebo se ukáže, že jsme vnímání jako zbytečný žrout peněz občanů našeho Státu. To zas bude důležité pro ředitele, protože takové smýšlení je pro nás ohrožující a budeme s tím muset něco dělat. Tak jako tak to pro nás bude užitečné. K tomu, abych něco takového zjistil ale potřebuju nějakou metodu, která zjistí vztah běžného občana k Akademii.“

Hned se do svého záměru pustíte a začnete skládat metodu, která má odhalit vztah běžného člověka k Akademii. A tak ponořen do této práce trávíte den za dnem. Až jednoho odpoledne vás ze zadumání vytrhne zaklepání na dveře. Za nimi stojí onen kolega, kterého jste onehdá potkal na chodbě.

„Abraháme, máš jít za ředitelem.“

„Jo, hned tam budu,“ odpovíte a zamíříte ven z pracovny.

„Abraháme,“ uvítá vás ředitel, „přišel mi dopis z Institutu. Poslouchejte:

Vážená Akademie, k dnešnímu dni jste nedodali materiály ke sběru dat v Celostátním víceúčelovém výzkumu. Musíme vám tedy s politováním oznámit, že vaše účast v tomto šetření je tímto zrušena. Víte, co to znamená, Abraháme? Znamená to, že jste zklamal důvěru, kterou jsem ve vás vložil!“

Jste nucen vyklidit kancelář a v tichosti opustit Akademii. Doma usedáte k počítači a začínáte procházet aktuální

Zvrátit poslední volbu

Znovu zahájit výzkum

Je čas dát všechno do obálky a poslat to Institutu. Nezbyvá než čekat, až data nasbírají a pošlou vám je k analýze. Chvilé klidu je zasloužená. Zatím jste odvedli dobrou práci.

Ona chvíle odpočinku ale netrvá zase příliš dlouho. Za nějaký čas obdržíte datový soubor. Když ho otevřete v počítači, uvidíte následující tabulku:

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

Tímto způsobem tabulka pokračuje až po ID 800.

„Výborně, a teď je čas se pustit do analýz.“

Začít analyzovat data

Je na čase pustit se do další práce. Zkoumáte, jak souvisí věk se schopností posuzovat spolehlivost mediálních zpráv. V tuto chvíli je čas přidržet se známé teorie a ze všeho nejdřív sloučit respondenty do skupin podle věku tak, jak bylo uvedeno v rešerši.

Vytvoříte si tedy novou proměnnou „Věk_Int“, která nabývá hodnoty 1 pro lidi ve věku 15 – 22 let, hodnoty 2 pro lidi mezi 23 a 60 lety a hodnoty 3 pro lidi starší 60 let. Tyto tři skupiny pak porovnáte v hrubých skórech Dotazníku mediálních zpráv.

Výsledky lze shrnout asi takto:

Všechny tři skupiny se od sebe vzájemně liší statisticky významně.

„Což vzhledem k velikosti vzorku není nijak překvapivé.“

Věcný rozdíl mezi skupinou od 15 let a skupinou 60+ je zhruba půl směrodatné odchylky, přičemž skupina od 15 dosahuje vyššího skóru.

Skupina 23 – 60 let dosahuje v Testu schopnosti ze všech tří skupin nejlepších výsledků. Věcný rozdíl mezi skupinou od 15 let a skupinou 23 – 60 se rovná jedné směrodatné odchylce. Rozdíl mezi skupinou 23 – 60 a 60+ je tedy jedna a půl směrodatné odchylky.

Tuto otázku lze tedy uzavřít tak, že skupina od 15 let je o trochu lepší v hodnocení spolehlivosti mediálních zpráv, než skupina 60+. Skupina 23 – 60 je pak výrazněji silnější než obě zbývající skupiny.

Zpracovat tyto výsledky do článku

Bližší výsledky prozkoumat

Vaše vědecká zvědavost vám nedá a rozhodnete se detailněji prozkoumat výsledky. Proto si v dalším kroku zobrazíte rozložení Věku pomocí scatterplotu. Rozložení grafu napovídá, že by v daném vzorku mohly existovat nikoli tři, ale čtyři skupiny. Zjistíte, že dvě z těchto čtyřech skupin jsou tvořeny výhradně lidmi z věkové skupiny tři. Protože je ale vzhledem ke struktuře vašich dat nejste schopni rozlišit, nelze nijak získat další informace.

**Zpracovat tyto výsledky do
článku**

První ze tří částí máte úspěšně provedenou a sepsanou. Podařilo se vám potvrdit to, co předpokládá teorie. „Sakra, v důsledku vlastního špatného rozhodnutí jsem se připravil o další informace. Kdybych se nerozhodl sbírat věk pomocí intervalů, mohlo z toho vzejít ještě něco zajímavého. No co, z článku, se to nikdo nemusí dozvědět. Podám to jako úspěšné potvrzení dosavadní teorie.“

Dále se tedy věnujete zapracování těchto informací do článku, který hodláte na konci výzkumu publikovat. Po napsání příslušné části textu je čas přistoupit k dalším analýzám.

Znovu jste sedl k datovému souboru. Je na čase pokračovat druhou částí.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

Pokračovat další výzkumnou otázkou

Dalším krokem je prozkoumání výzkumné otázky b). Tedy korelační vztah čistého příjmu s ochotou respondentů vystupovat na veřejnosti, která je reprezentována hrubým skórem v Dotazníku ochoty k veřejnému vystupování. V první fázi analýzy se bohužel ukázalo, že proměnná Čistý příjem je značné množství chybějících hodnot. Dále jste také zjistil, že tímto neduhem netrpí pouze proměnná Čistý příjem, ale také ostatní proměnné, se kterými jste doposud nepracoval. Tento problém je nutné nějak řešit.

„Hm, musím nějak naložit s těmi chybějícími hodnotami. V zásadě existují dva běžné způsoby, jak s tím naložit. Buď se zabavím všech respondentů, od kterých nemám kompletní data a nebo chybějící data v Čistém příjmu nahradím průměrnými hodnotami.“

Jak tak v zamyšlení pochodujete ode zdi ke zdi po pracovně, padne vám zrak na knihu, na kterou jste dávno zapomněli. Její hřbet nese název *„Pokročilé postupy řešení statistických problémů.“*

Vymazat všechny nekompletní záznamy

Doplnit chybějící hodnoty příjmů průměrnými hodnotami

Nahlédnout do knihy *Pokročilých postupů*

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, ale vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

**Zpracovat tyto výsledky do
článku**

Pokročilé postupy řešení statistických problémů

Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p -hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu

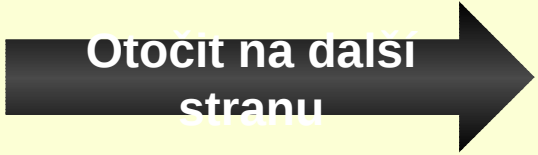


Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat „**Hodnoty chybějící z definice**“. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých

Otočit na další
stranu



Tento postup zachovává co největší množství respondentů a tedy co největší statistickou sílu, je zde však problém, že vzorek pro zpracování každé jednotlivé analýzy je pak odlišný a nelze tudíž jednotlivé kroky slučovat. V tomto případě by tedy nebylo možné studovat interakce věku a asertivity z hlediska prožívaného wellbeingu. Tento postup by vyžadoval Case delation.

Mean substitution je postup, který nahrazuje chybějící hodnoty průměrem vzorku v dané proměnné. Jedná se o logický postup, jestliže lze předpokládat, že daná proměnná ve vzorku odpovídá normálnímu rozložení. V takovém případě, neznáme-li hodnotu dané proměnné u konkrétní osoby, nejpravděpodobnější je, že tato hodnota se pohybuje blízko kolem průměru. Tento postup je však problematický, pokud v konkrétním vzorku je velké množství chybějících hodnot. Jestliže ve vzorku chybí například 30% hodnot, pak nahrazení všech těchto hodnot průměrem značně redukuje rozptyl ve vzorku, a tím pádem podhodnocuje hodnotu korelace. Dalším problémem je necitlivost takového postupu vůči extrémním hodnotám. Jestliže mi u respondenta chybí jeho měsíční příjem a tento člověk je buďto milionář nebo se naopak potácí na hranici chudoby, pak v obou těchto případech je nahrazení této chybějící hodnoty průměrem velmi nešťastné a povede ke zkreslení výsledků. Pomocí Mean substitution není také vhodné doplňovat z definice chybějící hodnoty.

Indicator adjustment je lehká úprava strategie doplňování průměru. V tomto případě se do datového souboru přidá proměnná, tzv. indikátor, který nabývá hodnoty 0 v případech, kdy cílová hodnota chybí a 1 v případech, kdy je cílová hodnota pozorována. Následně jsou chybějící hodnoty doplněny průměrem vzorku. Přítomnost indikátoru zde udržuje informaci o počtu doplněných hodnot a tím pádem možného rozsahu zkreslení. Více o tomto postupu například [Prejít na seznam literatury](#) (3) [Otočit na další stranu](#)

Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami i o zmíněných softwarech (Acock, 2005).

[Přejít na seznam literatury](#)

Zavřít knihu

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.

Vraťte se k tomuto odkazu

- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.

Vraťte se k tomuto odkazu

- Acock, A. C. (2005). Working with missing values. ***Journal of Management and Family Business*** (4), 1012-1028.

Vraťte se k tomuto odkazu

Četbou knihy jste si ověřil své domněnky a načerpal novou inspiraci pro řešení nastalého problému. Teď je potřeba toto úsilí zužitkovat. Je potřeba posunout výzkum zase kousek dál a rozhodnout se, jak tedy vyřešíte problém s chybějícími hodnotami. Vzhledem k tomu, že pro takzvané moderní postupy nemáte zařízení s dostatečnou výpočetní kapacitou, musíte vybrat to nejvhodnější z tradičních postupů.

Použít Case delation

Použít Pairwise delation

Použít Mean substitution

Použít Indicator Adjustment

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tři set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, ale vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po použití Pairwise delation z této konkrétní analýzy vypadlo 25% výzkumného vzorku. Vzhledem k jeho rozsahu si analýza i v tomto rozsahu zachovává svou statistickou sílu. Odhalila silnou pozitivní korelaci mezi výší čistého příjmu a ochotou k veřejným vystoupením. Tento výsledek podporuje dosavadní mínění teorie, že bohatší lidé jsou ochotnější k veřejnému vystupování. Z dostupných dat však nelze zjistit nic o tom, jaký mechanismus tuto větší ochotu způsobuje.

„To by chtělo navazující kvalitativní studii, která by se bohatých lidí ptala na důvody, proč jsou ochotní vystupovat na veřejnosti a snažila se odhalit nějaký mechanismus. Hm. Tohle je dobrá věta do diskuse. To si musím někde poznačit.“

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po aplikování Indicator adjustment postupu a doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost korelace čistého příjmu a hrubého skóru Dotazníku, nicméně tato korelace se ukázala být velmi nízká. Díky indikátoru víte, že průměrný plat byl doplněn do 25% případů. Momentálně se vám tedy podařilo vyvrátit mylný předpoklad teorie, je však otázka, jak moc oněch 25% doplněných hodnot tento výsledek zkresluje.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu. V prvním případě jste podpořil prozatimní teoretické předpoklady, i když s jistou pochybností a v druhém případě jste teorii potvrdil. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. V prvním případě jste podpořil prozatimní teoretické předpoklady, i když s jistou pochybností a v druhém případě, pokud připustíme korekci síly korelace, která je zkreslená 25% doplněných průměrných hodnot, tak by ta korelace mohla být dostatečně silná, aby to stálo za zmínku. V podstatě to teorii podporuje. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. V prvním případě jste podpořil prozatimní teoretické předpoklady, i když s jistou pochybností a v druhém případě vaše zjištění teorii spíše nabourávají. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermana Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrný a interpretovat s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

První ze tří částí máte úspěšně provedenou a sepsanou. Podařilo se vám potvrdit to, co předpokládá teorie. Dále se tedy věnujete zpracování těchto informací do článku, který hodláte na konci výzkumu publikovat. Po napsání příslušné části textu je čas přistoupit k dalším analýzám.

Znovu jste sedl k datovému souboru. Je na čase pokračovat druhou částí.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

Dalším krokem je prozkoumání výzkumné otázky b). Tedy korelační vztah čistého příjmu s ochotou respondentů vystupovat na veřejnosti, která je reprezentována hrubým skórem v Dotazníku ochoty k veřejnému vystupování. V první fázi analýzy se bohužel ukázalo, že proměnná Čistý příjem je značné množství chybějících hodnot. Dále jste také zjistil, že tímto neduhem netrpí pouze proměnná Čistý příjem, ale také ostatní proměnné, se kterými jste doposud nepracoval. Tento problém je nutné nějak řešit.

„Hm, musím nějak naložit s těmi chybějícími hodnotami. V zásadě existují dva běžné způsoby, jak s tím naložit. Buď se zabavím všech respondentů, od kterých nemám kompletní data a nebo chybějící data v Čistém příjmu nahradím průměrnými hodnotami.“

Jak tak v zamyšlení pochodujete ode zdi ke zdi po pracovně, padne vám zrak na knihu, na kterou jste dávno zapomněli. Její hřbet nese název *„Pokročilé postupy řešení statistických problémů.“*

Vymazat všechny nekompletní záznamy

Doplnit chybějící hodnoty příjmů průměrnými hodnotami

Nahlédnout do knihy *Pokročilých postupů*

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukovaný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané Státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

**Zpracovat tyto výsledky do
článku**

Pokročilé postupy řešení statistických problémů

Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p -hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu

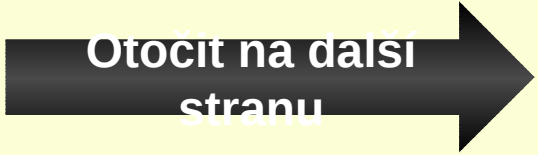


Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat „**Hodnoty chybějící z definice**“. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých

Otočit na další
stranu



Tento postup zachovává co největší množství respondentů a tedy co největší statistickou sílu, je zde však problém, že vzorek pro zpracování každé jednotlivé analýzy je pak odlišný a nelze tudíž jednotlivé kroky slučovat. V tomto případě by tedy nebylo možné studovat interakce věku a asertivity z hlediska prožívaného wellbeingu. Tento postup by vyžadoval Case delation.

Mean substitution je postup, který nahrazuje chybějící hodnoty průměrem vzorku v dané proměnné. Jedná se o logický postup, jestliže lze předpokládat, že daná proměnná ve vzorku odpovídá normálnímu rozložení. V takovém případě, neznáme-li hodnotu dané proměnné u konkrétní osoby, nejpravděpodobnější je, že tato hodnota se pohybuje blízko kolem průměru. Tento postup je však problematický, pokud v konkrétním vzorku je velké množství chybějících hodnot. Jestliže ve vzorku chybí například 30% hodnot, pak nahrazení všech těchto hodnot průměrem značně redukuje rozptyl ve vzorku, a tím pádem podhodnocuje hodnotu korelace. Dalším problémem je necitlivost takového postupu vůči extrémním hodnotám. Jestliže mi u respondenta chybí jeho měsíční příjem a tento člověk je buďto milionář nebo se naopak potácí na hranici chudoby, pak v obou těchto případech je nahrazení této chybějící hodnoty průměrem velmi nešťastné a povede ke zkreslení výsledků. Pomocí Mean substitution není také vhodné doplňovat z definice chybějící hodnoty.

Indicator adjustment je lehká úprava strategie doplňování průměru. V tomto případě se do datového souboru přidá proměnná, tzv. indikátor, který nabývá hodnoty 0 v případech, kdy cílová hodnota chybí a 1 v případech, kdy je cílová hodnota pozorována. Následně jsou chybějící hodnoty doplněny průměrem vzorku. Přítomnost indikátoru zde udržuje informaci o počtu doplněných hodnot a tím pádem umožňuje rozsahu zkreslení. Více o tomto postupu například [Prejit na seznam literatury](#) [Otočit na další stranu](#)

Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami i o zmíněných softwarech (Acock, 2005).

[Přejít na seznam literatury](#)

Zavřít knihu

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.

Vraťit se k tomuto odkazu

- Acock, A. C. (2005). Working with missing values. ***Journd of Marriage and family*** 67(4), 1012-1028.

Vraťit se k tomuto odkazu

Četbou knihy jste si ověřil své domněnky a načerpal novou inspiraci pro řešení nastalého problému. Teď je potřeba toto úsilí zužitkovat. Je potřeba posunout výzkum zase kousek dál a rozhodnout se, jak tedy vyřešíte problém s chybějícími hodnotami. Vzhledem k tomu, že pro takzvané moderní postupy nemáte zařízení s dostatečnou výpočetní kapacitou, musíte vybrat to nejvhodnější z tradičních postupů.

Použít Pairwise delation

Použít Case delation

Použít Mean substitution

Použít Indicator Adjustment

Po vymazání všech nekompletních záznamů, vám zůstal datový soubor o velikosti sotva tří set osob. Takto redukováný vzorek již pro tak rozsáhlou populaci jako jsou všichni občané Státu není v žádném případě dostačující.

„Sakra,“ řekl jste si, „není možné, aby dvě třetiny lidí nevedly svůj čistý příjem. Ano, ta otázka je trochu ošemetná, ale dvě třetiny?“

Samozřejmě, že to tak nebylo, vymazal jste ze souboru všechny nekompletní záznamy. To znamená i ty respondenty, kteří sice uvedli čistý příjem, ale nevedli některou z jiných proměnných, jako například velikost bydliště nebo politickou aktivitu.

S takto poničeným datovým souborem není možné ve výzkumu dále pokračovat.

Zvrátit poslední volbu

Znovu zahájit výzkum

Po použití Pairwise delation z této konkrétní analýzy vypadlo 25% výzkumného vzorku. Vzhledem k jeho rozsahu si analýza i v tomto rozsahu zachovává svou statistickou sílu. Odhalila silnou pozitivní korelaci mezi výší čistého příjmu a ochotou k veřejným vystoupením. Tento výsledek podporuje dosavadní mínění teorie, že bohatší lidé jsou ochotnější k veřejnému vystupování. Z dostupných dat však nelze zjistit nic o tom, jaký mechanismus tuto větší ochotu způsobuje.

„To by chtělo navazující kvalitativní studii, která by se bohatých lidí ptala na důvody, proč jsou ochotní vystupovat na veřejnosti a snažila se odhalit nějaký mechanismus. Hm. Tohle je dobrá věta do diskuse. To si musím někde poznačit.“

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po aplikování Indicator adjustment postupu a doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost korelace čistého příjmu a hrubého skóru Dotazníku ochoty k vystupování, nicméně tato korelace se ukázala být velmi nízká. Díky indikátoru víte, že průměrný plat byl doplněn do 25% případů. Momentálně se vám tedy podařilo vyvrátit mylný předpoklad teorie, je však otázka, jak moc oněch 25% doplněných hodnot tento výsledek zkresluje.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

Po doplnění chybějících hodnot Čistého příjmu průměrnými hodnotami ze vzorku jste provedl korelační analýzu. Výsledky sice prokázaly statistickou významnost pozitivní korelace čistého příjmu a hrubého skóru Dotazníku ochoty, což by naznačovalo, že bohatší lidé jsou ochotnější k vystupování na veřejnosti, nicméně tato korelace se ukázala být velmi nízká. Věcný význam tohoto vztahu je tedy v podstatě nulový. Opět jste prokázal nedostatky v současné teorii.

Zpracovat tyto výsledky do článku

Zkusit ještě jiný postup

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu. V prvním případě jste mezeru v teorii a v druhém případě jste teorii potvrdil. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

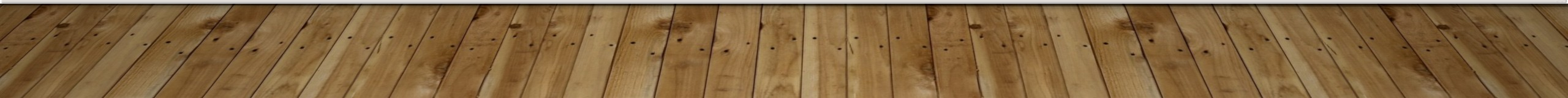
ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo korigovat jednu nepřesnost v teorii a v druhém případě, pokud připustíme korekci síly korelace, která je zkreslená 25% doplněných průměrných hodnot, tak by ta korelace mohla být dostatečně silná, aby to stálo za zmínku. V podstatě to teorii podporuje. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce. Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

Pokračovat další výzkumnou otázkou



I druhou ze tří částí výzkumu máte úspěšně za sebou. Článek uspokojivě narůstá na objemu a zdá se, že přináší zajímavé poznatky. Zatím se vám podařilo najít dvě trhliny v teoretických předpokladech. Začínáte se těšit, až ho publikujete. Teď je čas přistoupit k poslední zadané výzkumné otázce.

Znovu jste sedl k datovému souboru.

ID	Věk	Pohlaví	Čistý příjem	Schopnost rozlišování zpráv	Ochota k veřejnému vystupování	Agresivita vůči vládě
001	1	2	NA	8	28	5
002	3	2	21 400	46	5	7
003	2	2	20 200	45	28	3
004	2	2	134 700	26	16	5
005	3	2	NA	12	19	46
...						

**Pokračovat další výzkumnou
otázkou**

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermana Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrní a interpretovatelní zodpovědnosti.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být interpretativně zodpovědní.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

**Zpracovat tyto výsledky do
článku**

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Třetí otázkou je, zda se liší muži a ženy v úrovni projevované agrese vůči vládě. Analýza neprokázala, že by se muži a ženy v tomto ohledu lišili.

Kromě toho jste dodatečnými analýzami odhalil přítomnost řady outlierů v obou skupinách.

„V tom množství outlierů se může potenciálně skrývat docela problém. No, můžu outliery z analýz vyřadit a zkusit to znovu. Počkat! Někde jsem tu měl tu knihu. Třeba něco vyčtu.“

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

**Nahlédnout do knihy
*„Pokročilé postupy řešení
statistických problémů“***

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outliery jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. O'Learyho, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů jejich detekce a 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v této věci snažím být opatrní a interpretovat je s odpovědností.

Zavřít knihu

Seznam literatury

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.

Vrátit se k tomuto odkazu

Kniha vás ani v tomto případě nezklamala. Díky jejím informacím se rozšířily vaše možnosti, jak z celou záležitostí naložit.

**Použít výsledky včetně
outlierů**

**Vyřadit outliery a provést
analýzy znovu**

Použít obojí výsledky

**Investovat čas a energii do
individuální analýzy**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

**Zpracovat tyto výsledky do
článku**

„No, výsledky jsou výsledky. Navíc, tak velký vzorek by měl být poměrně reprezentativní. Jsou holt více a méně agresivní muži a agresivnější a méně agresivní ženy. V průměru to bude asi jedno.“

Veden takovou úvahou se rozhodnete ušetřit čas na dalších analýzách a vynaložit ho k dokončení článku.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

„Je pravděpodobné, že extrémní hodnoty jsou výsledkem nějakých chyb. Bude lepší je z dalších analýz vyřadit.“

Veden touto myšlenkou provede analýzu znovu na vzorku očištěném od outlierů a skutečně docházíte k jiným výsledkům. V tomto případě se ukázal jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

Varianta včetně outlierů neprokázala žádný rozdíl mezi muži a ženami. Analýza očištěná od outlierů ukázala jak věcně tak statisticky významný rozdíl. Napovídá, že muži jsou výrazně náchylnější k agresivitě vůči vládě Státu. To vlastně potvrzuje, co se tak nějak předpokládalo.

Zpracovat tyto výsledky do článku

Vyzkoušet jiný postup

„Asi bude nejlepší investovat ještě nějaký čas a úsilí do podrobnější analýzy. Ty výsledky zatím nejsou vlastně nijak vypovídající.“

Veden touto myšlenkou se pouštíte do detailního rozboru všech outlierů. Od Institutu jste si vyžádal původní data včetně odpovědí na jednotlivé položky všech metod a upravoval, třídil a interpretoval jste jako divý.

Úsilí se ovšem vyplatilo. Nápravou několika chybně zadaných hodnot a vyřazením několika lidí, jejichž odpovědi byly dostatečně podivné na to, aby se dalo usuzovat na nepochopení, jste docílil toho, že následná analýza prokázala rozdíl mezi muži a ženami, jehož věcná významnost není zanedbatelná.

Navíc se zdá, že prostředí, v jakém se lidé běžně nachází a jejich politická aktivita či neaktivita může radikalizovat či umírňovat agresivní projevy. A to jak u žen tak u mužů.

Tyto analýzy si však vyžádaly takové množství času, že už nezbyl na to, abyste se tím dál zabýval, je nutné dokončit článek.

**Zpracovat tyto výsledky do
článku**

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

Zpracováním těchto dat byl článek dokončen!

Publikovat

„Abraháme, jste si tím jistý? Je vám jasné, že to znamená odchod z Akademie?“

„Ano, vím. Přemýšlel jsem o tom a myslím, že je na čase vyklidit pole.“

„No dobrá, v tom případě mi bylo ctí s vámi pracovat. Doufám, že se potkáme někdy někde jinde.“

Potom, co jste opustil ředitelovu kancelář, opustil jste Akademii a došel domů. Tam jste dlouhou dobu prostě jenom seděl na gauči a přemýšlel co dál. Nakonec jste začal přebírat věci, které jste si přinesl z teď už bývalého zaměstnání. Hrneček na čaj, zarámovaný diplom, nějaké knihy. Strukturální modelování, Kvalitativní analýza a tak dále. Nakonec vám v ruce zůstanou „*Pokročilé postupy řešení statistických problémů*“. Lehnete si na gauči a začnete knihou listovat.

**Číst si v „*Pokročilých
postupech*“**

Vzteky knihu zahodit

Pokročilé postupy řešení statistických problémů

Práce s chybějícími hodnotami

Otočit na další stranu 

Předmluva

V ideálním světě, kde výzkumníkům vždy vše hraje pěkně do karet zůstávají naše p -hodnoty nízké, věcné významnosti vysoké a naše datové soubory vždy bezchybné a kompletní. Bohužel, nežijeme v ideálním výzkumnickém světě, a tak se se všemi těmito problémy musíme v naší výzkumnické praxi neustále potýkat. V této kapitole se zaměříme na poslední zmíněný problém, tedy chybějící hodnoty.

Otocit na další
stranu

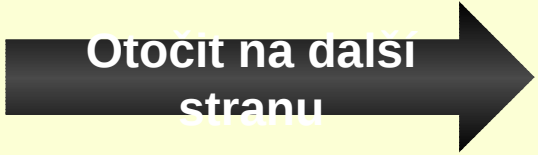


Chybějící hodnoty lze rozdělit do několika typů, přičemž od konkrétního typu se pak odvíjí různé způsoby práce s nimi. Prvním typem je **Missing by definition of the subpopulation**, nebo by se to také dalo nazvat „**Hodnoty chybějící z definice**“. K němu může docházet například u velkých výzkumných souborů, kde se setkávají různé podskupiny (subpopulace), u kterých nás pak zajímají rozdílné věci. Například, pokud se zajímám o prožívání po dobu těhotenství, žen se budu ptát na prožívání vlastního těhotenství. Mužů se pak budu ptát na prožívání těhotenství jejich partnerek. Z toho logicky vyplývá, že u mužů zaznamenám chybějící hodnoty u otázek na vlastní těhotenství a u žen pak chybějící hodnoty u otázek na těhotenství jejich partnerky (v tomto příkladu jsou záměrně vynechána registrovaná partnerství).

Druhým typem chybějících hodnot je **Missing completely at random (MCAR)**. V tomto případě jsou chybějící hodnoty roztroušeny napříč datovým souborem zcela náhodně a nezávisle na jakémkoliv ukazateli. Představíme-li si datový soubor jako matici dat a zvýrazníme v ní chybějící hodnoty, v tomto případě zde nebude možné odhalit jakýkoliv vzorec.

Třetím typem je **Missing at random (MAR)**. Sem spadají takové chybějící hodnoty, u kterých lze pozorovat mechanismus, nicméně ten mechanismus nesouvisí přímo se zjišťovanou proměnnou. Například, když lidé neodpoví na položky týkající se depresivity, nicméně mechanismem, který za tím lze najít není samotná úroveň depresivity, ale například jejich rodinná situace. Tedy, že lidé ve špatné rodinné situaci spíše neodpovídají na otázky o depresivitě než lidé kteří jsou v dobré rodinné situaci.

Otociť na další
stranu



Posledním typem je **Missing not at random (MNAR)**, kdy neochota odpovídat na konkrétní položku přímo souvisí s konceptem, který daná položka zjišťuje. Například když se v datovém souboru ukáže, že lidé trpící depresí neodpovídají na položky ohledně vlastní depresivity.

Způsoby, jak se s chybějícími hodnotami vypořádat lze rozdělit na tzv. **tradiční** a **moderní**. Mezi tradiční způsoby patří Case delation, Pairwise delation, Mean substitution a Indicator Adjustment. Mezi moderní se pak řadí Simple imputation, Multiple imputation nebo Full information maximum likelihood approach.

Case delation je nejběžnější metoda řešení problému chybějících hodnot. Jedná se o vyřazení všech nekompletních záznamů ze vzorku. To může vést ke ztrátě 20% - 50% datového souboru a tím pádem ke značnému omezení statistické síly. U rozsáhlých vzorků však je tento problém do značné míry zanedbatelný. Další komplikací je však typ chybějící hodnoty. Case delation je vhodné použít především u MCAR missingů. Jestliže se však chybějící hodnoty ve vzorku řídí nějakým pravidlem, hrozí, že tento postup dá nakonec vzniknout nereprezentativnímu vzorku (například pokud by lidé žijící na hranici chudoby odmítli uvést svůj měsíční příjem, case delation těchto případů z výzkumu vyřadí tuto část populace).

Pairwise delation je postup, který do zpracování zařazuje všechny, kteří odpověděli na položky, kterých se analýza týká, nehledě na to, že u jiných položek, které v tu chvíli nejsou předmětem analýzy, mají chybějící hodnoty. Například, pokud chci srovnat věkové skupiny z hlediska jejich asertivity, vezmu do analýzy všechny respondenty, u kterých mám k dispozici jak věk tak míru asertivity bez ohledu na to, že u některých z nich nemám k dispozici míru well-beingu. Stejně tak pokud chci srovnat věkové skupiny z hlediska well-beingu, neberu v úvahu, zda u některých respondentů

Otočit na další

stranu



Tento postup zachovává co největší množství respondentů a tedy co největší statistickou sílu, je zde však problém, že vzorek pro zpracování každé jednotlivé analýzy je pak odlišný a nelze tudíž jednotlivé kroky slučovat. V tomto případě by tedy nebylo možné studovat interakce věku a asertivity z hlediska prožívaného wellbeingu. Tento postup by vyžadoval Case delation.

Mean substitution je postup, který nahrazuje chybějící hodnoty průměrem vzorku v dané proměnné. Jedná se o logický postup, jestliže lze předpokládat, že daná proměnná ve vzorku odpovídá normálnímu rozložení. V takovém případě, neznáme-li hodnotu dané proměnné u konkrétní osoby, nejpravděpodobnější je, že tato hodnota se pohybuje blízko kolem průměru. Tento postup je však problematický, pokud v konkrétním vzorku je velké množství chybějících hodnot. Jestliže ve vzorku chybí například 30% hodnot, pak nahrazení všech těchto hodnot průměrem značně redukuje rozptyl ve vzorku, a tím pádem podhodnocuje hodnotu korelace. Dalším problémem je necitlivost takového postupu vůči extrémním hodnotám. Jestliže mi u respondenta chybí jeho měsíční příjem a tento člověk je buďto milionář nebo se naopak potácí na hranici chudoby, pak v obou těchto případech je nahrazení této chybějící hodnoty průměrem velmi nešťastné a povede ke zkreslení výsledků. Pomocí Mean substitution není také vhodné doplňovat z definice chybějící hodnoty.

Indicator adjustment je lehká úprava strategie doplňování průměru. V tomto případě se do datového souboru přidá proměnná, tzv. indikátor, který nabývá hodnoty 0 v případech, kdy cílová hodnota chybí a 1 v případech, kdy je cílová hodnota pozorována. Následně jsou chybějící hodnoty doplněny průměrem vzorku. Přítomnost indikátoru zde udržuje informaci o počtu doplněných hodnot a tím pádem možného rozsahu zkreslení. Více o tomto postupu například (Gelman, Gelman, Su, & Su, 2002)

[Přejít na seznam literatury](#)

[Otočit na další stranu](#)

Shrneme-li jednoduše vlastnosti tradičních přístupů, žádný z nich neposkytuje ideální řešení. Jejich použitelnost vždy závisí na konkrétních okolnostech.

K moderním přístupům, které se snaží některé tyto nedostatky kompenzovat, patří **Simple imputation** postup, který vytvoří nový dataset, kde jsou všechny chybějící hodnoty doplněny metodou maximální pravděpodobnosti. Tento přístup je založen na ve vzorku obsažených vztazích mezi všemi proměnnými do kterého vstupuje určitý stupeň náhodných chyb. Tento postup však podhodnocuje standardní chybu měření a tedy nadhodnocuje přesnost.

Tento problém řeší **Multiple imputation**, postup který vytvoří 5 - 10 datových souborů různým doplněním. Analýza je pak provedena na všech těchto souborech a uchovává všechny hodnoty parametrů i standardních chyb. Pak je možné použít průměrné hodnoty.

Full information maximum likelihood approach nepoužívá doplnění chybějících hodnot, ale využívá všechny dostupné informace ze vzorku, aby metodou maximální věrohodnosti poskytl co nejpřesnější odhady chybějících parametrů.

Tyto pokročilé moderní metody jsou ve srovnání s tradičními přesnější, nicméně výpočetně náročnější. Existují však softwary, které použití těchto přístupů umožňují. Jedná se například o NORM, ale Simple imputation umí třeba i SPSS.

Více o práci s chybějícími hodnotami i o zmíněných softwarech (Acock, 2005).

[Přejít na seznam literatury](#)

[Otocit na další stranu](#)

Pokročilé postupy řešení statistických problémů

Práce s outliery

Otočit na další
stranu



Předmluva

Mluvím-li o outlierech, mám na mysli pozorované hodnoty měření, které se výrazně liší od ostatních hodnot zjištěných ve vzorku. Jestliže Vám, vážený čtenáři, tato definice připadá vágní pak vezte, že je tomu opravdu tak. Na to, co je to vlastně outlier, jak ho poznat a jak pak s takovým záznamem naložit existují desítky různých názorů. A to na každou ze zmíněných tři věcí. Následující stránky pokrývají část možných variant.

Otocit na další
stranu



Přidržíme se předpokladu, že outlier znamená nepřiměřeně vysokou, či nepřiměřeně nízkou hodnotu konkrétního parametru ve srovnání s ostatními hodnotami stejné proměnné napříč datovým souborem. Pro další uvažování je důležité určit, co takovou nepřiměřenou hodnotu parametru způsobilo.


První, nejběžnější a nejlépe řešitelná varianta je, že tato hodnota je způsobena **prostou chybou v přepisu dat**. K tomu může snadno docházet při manuálním přepisu fyzických dat do elektronické podoby. Řešením je samozřejmě kontrola původních dat a korekce takových chybových hodnot.

Druhým typem chyby, která vede k přítomnosti outlierů v datech jsou **chyby na straně respondenta**, které mohou plynout z nepochopení zadání, špatné interpretaci odpověďové škály a tak podobně. V tomto případě nezbyvá, než takové hodnoty vyloučit z analýzy, protože jejich zachování by vedlo ke zkreslení výsledků.

Třetí zdroj outlierů lze shrnout jako **netypické projevy**. Zde již narážíme na všudypřítomný problém „sto lidí sto názorů.“ Netypické projevy samozřejmě zahrnují některé jednotlivce, kteří ve srovnání s ostatními jednoduše dosahují v konkrétní proměnné extrémních hodnot. Je však popisována řada definic outlierů, některé odvozené od konkrétních statistických postupů. Od toho se potom odvíjí jak způsoby identifikace takových případů tak způsoby nakládání s nimi.

Samozřejmě, existuje také přístup, který nenakládá s outlierem jako s problémem, který musí být nutně vyřešen, ale jako se zdrojem zajímavého druhu informací.

Otocit na další
stranu



Takový zdroj může v některých případech mnoho napovědět. Dalo by se říct, že studiem výjimek se nejvíce dovídáme o pravidlech. Tento způsob práce samozřejmě vyžaduje individuální přístup a studium různých příčin u každého jednotlivého respondenta klade vysoké nároky jak na čas, tak na úsilí. Je nicméně pravda, že i korekce chybových dat vyžaduje zvýšené úsilí a individuální přístup ke každé hodnotě, kterou lze považovat za outlier.

Může se zdát, že v případě outlierů zatížených dat máme k dispozici dvě možnosti. První možností je investovat mnoho času a úsilí k individuálnímu posouzení každé takové hodnoty, což pravděpodobně povede k jejich redukci, či k tomu, že se z těchto hodnot dozvíme něco, co z celkového obrazu není patrné. Druhá možnost znamená plošné řešení outlierových hodnot. Zde buď zahrneme takové hodnoty do analýz nebo je z nich vyloučíme.

V prvním případě riskujeme značně zkreslené výsledky, zatímco v druhém případě riskujeme ztrátu důležitých informací a nereprezentativnost takových výsledků. Také riskujeme nařčení ze strany oponentů, že manipulujeme s daty záměrně takovým způsobem, abychom získali výsledky, které se nám, abych tak řekl, hodí do krámu. V každém případě jsou tyto možnosti značně omezující. Proto někteří výzkumníci aplikují třetí možnost.

Tou třetí možností je publikování dvojích výsledků, jak včetně outlierů tak bez nich. Takový postup lze označit za sebeprotektivní, protože přesouvá část zodpovědnosti za interpretaci na čtenáře. Je to však způsob, jak se vyhnout náročné individuální analýze i problémům plynoucím z plošných řešení.

[Přejít na seznam literatury](#) toho, co lze říct o práci s outliery. V tomto ohledu důrazně doporučuji nahlédnout do článku autorů Hermána Aguinise, Ryana K. Gottfredsona a Harryho J. Oo, publikovaného v roce 2013, který shrnuje 14 definic outlierů, 39 způsobů 20 postupů nakládání s nimi. Některé jsou poměrně komplikované, takže se i já v tomto interpretativní zodpovědnosti

Zavřít knihu

Seznam literatury

- Cohen, J ., & Cohen, P. (1983). ***Applied multiple regression and correlation analysis for the behavioral sciences (2nd ed)***. Hillsdale, NJ : Erlbaum.
- Cohen, J ., Cohen, P., West, S., & Aiken, L. (2003). ***Applied multiple regression and correlation analysis for the behavioral sciences (3rd ed)***. Mahwah, NJ : Erlbaum.
- Acock, A. C. (2005). Working with missing values. ***Journal of Marriage and Family* 67(4)**, 1012-1028.
- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. ***Organizational Research Methods* 16(2)**, 270-301.

Vrátit se k tomuto odkazu

Vrátit se k tomuto odkazu

Vrátit se k tomuto odkazu

Vrátit se k tomuto odkazu

„Tohle je mi k ničemu!“

Kniha letí přes pokoj a otevřená zůstane ležet u stěny, do které narazila.

Od té chvíle Abrahám upadl do apatie. Zmizela všechna motivace k práci a k čemukoliv dalšímu. Snad to brzo přejde a najde si jiné místo. Přejme mu, aby se zase dostal k vědecké činnosti, která pro něj je tak důležitá.

Odmítl jste svoji šanci a tak vaše cesta skončila velmi záhy.



Zavřít

**Vyzkoušet jinou
cestu**

„Tohle všechno já vím! Bohužel, mně to nijak Nepomáhá! Sakra, co teď budu dělat?“

Od té chvíle Abrahám upadl do apatie. Zmizela všechna motivace k práci a k čemukoliv dalšímu. Snad to brzo přejde a najde si jiné místo. Přejme mu, aby se zase dostal k vědecké činnosti, která pro něj je tak důležitá.

Odmítl jste svoji šanci a tak vaše cesta skončila velmi záhy, ale alespoň něco jste se dozvěděl.



Zavřít

**Vyzkoušet jinou
cestu**

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste.

Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé dále narůstá. Neprokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste.

Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá.

Neprokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal značný rozruch ve společnosti. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste.

Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá.

Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlady. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Váš výzkum se vypořádá s určitými aspekty těchto důsledků.

Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní si můžete vypořádat s takovými aspekty těchto důsledků

Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá.

Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch.

Mluví

se

o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Nepochybně rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti také nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak také nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.

Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let.

Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytila vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zklamání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.

Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytli ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéry příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

**Vyzkoušet jinou
cestu**

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéry příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády.

Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytli ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste.

Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

**Vyzkoušet jinou
cestu**

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste.

Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste.

Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlady. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nechtěli předpovídat. Nyní se můžete vypořádat s etickými aspekty těchto důsledků.

Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste.

Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme neměli předpokládat. Nyní se musíte rozhodnout s jakými aspekty těchto důsledků.

Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní.

Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlady. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Nepochybně rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéry příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéry příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytila vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Nepochybně rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s politickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlady.

Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy

Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytila vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Neprokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluk které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Nepochybně rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Pokud se budete muset vyrovnat s etickými aspekty těchto důsledků.

Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Nepochybně rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkřížení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy.

Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy.

Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytila vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlady. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jste nemohl předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto

Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Nепrokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv vlády.

Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou

Zavřít

nemohou předpokládat.

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.

Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéry příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytila vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlady.

Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsou nemohli předpovídat. Nyní se musíte vypořádat s etickými aspekty těchto

Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jste nemohli předpovědět. Nyní se musíte vypořádat s etickými aspekty těchto důsledků.

Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jste ne mohli předpokládat. Nyní se musíte vyrovnat s etickými aspekty těchto důsledků.

Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády.

Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládnou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které ne nemohli předpokládat. Nyní se můžete pořádat s etickými aspekty těchto důsledků.

Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády. Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musí vydat s etickými aspekty těchto

Zavřít

**Vyzkoušet jinou
cestu**

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé. Váš zájem o to, jak interakce politické aktivity a věku ovlivňuje míru projevované agresivity proti vládě Státu vyústil v závěr, že politicky aktivní lidé jsou agresivnější než ti politicky neaktivní. Nejvýraznější je to pak u věkové skupiny od dvacet tří let. Vašich zjištění vyplývajících ze studia interakce politické aktivity a věku se chytla vláda. Její způsob řešení této situace ale spočíval v omezování možností politických aktivit a policejní tlak na politicky aktivní jedince v kritickém věku od 23 let. Tato skupina je nyní výrazně omezována, aby neohrožovala pozici a vliv Vlády.

Vy jste se stal mezi lidmi velmi slavným. Vešlo v obecnou známost, že výzkum, který byl následně vládou zneužit a vedl k útlaku obyvatelstva, jste provedl vy.

Z vědeckého hlediska bylo zkoumání této interakce samozřejmě správným rozhodnutím. Nicméně někdy vedou k důsledkům, které jsme nemohli předpokládat. Nyní se musíte vypořádat s etickými aspekty těchto

Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Zájem vyvolalo i potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytli ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéry příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní.

Neprokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní.

Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste.

Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Korekce teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv vyvolala ve vědeckých kruzích rozruch. Řada vědců se touto záležitostí dále zabývá a vaše vědecká popularita roste. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní.

Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Nepochybně rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuvzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Neprokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy.

Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Neprokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Možná, kdybyste byl schopen rozlišit ty dva shluky, které jste viděl na scatterplotu, bylo by to celé trochu jinak. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Neprokázaní rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Potvrzení toho, že výše příjmů koreluje s ochotou lidí k veřejnému vystupování vyvolalo ve vědeckých kruzích zájem. Různé týmy se pustily do vámi navrhované kvalitativní analýzy mechanismu, který se zatím skrývá. Vaše vědecké renomé narůstá. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl.mNepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Neprokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepříspěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Nepříliš průkazné výsledky, které jste poskytl ohledně korelace výše příjmů a ochoty k veřejnému vystupování vaší kariéře příliš neprospěly. Sklidil jste v tomto ohledu spíše kritické ohlasy. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Nепrokázání rozdílů v agresivitě mezi muži a ženami zapadlo mezi ostatními, významnějšími výsledky z poslední doby. Touto částí svého článku jste ke znovuzkříšení své vědecké pověsti nijak nepřispěl.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Nicméně, objevují se i takové hlasy, které tyto výsledky zpochybňují. Argumentují tím, že vyřazením outlierů z analýz jste docílil nadhodnocení významnosti celého závěru.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Vaše výsledky ohledně míry agresivity vůči vládě Státu u mužů a žen vyvolaly ve vědeckých kruzích bouřlivou debatu. Vědci intenzivně porovnávají oba závěry a intenzivně se mezi sebou dohadují, který z nich je věrohodnější a co z toho celého vlastně plyne. Tento rozruch značně přispívá vašemu vědeckému renomé.



Zavřít

Vyzkoušet jinou
cestu

Potvrzení teoretických předpokladů ohledně věkových skupin a jejich schopnosti rozlišovat spolehlivost mediálních zpráv zapadlo ve víru jiných, zajímavějších zjištění. Tímto počinem jste své kariéře příliš neprospěl. Vaše výsledky ohledně korelace výše příjmů a ochoty k veřejnému vystupování nevyvolaly ve vědeckých kruzích příliš pozitivní reakce. Několik výzkumníků se vůči nim ohradilo s poukázáním na to, že vámi zvolený způsob řešení problému chybějících hodnot mohl mít výrazný vliv na podhodnocení věcné významnosti korelace. Vámi poskytnuté výsledky tedy nelze prý považovat za relevantní. Důkaz, že se muži a ženy liší v úrovni projevované agresivity vůči Vládě vyvolal ve společnosti značný rozruch. Mluví se o tom jak ve vědeckých kruzích, tak ve vládních úřadech. Oba tábory zvažují, jaký bude v tomto ohledu další postup. Vaše jméno je samozřejmě v této souvislosti skloňováno velmi často. Vědečtí pracovníci také velmi oceňují množství práce, které jste prokázal individuální analýzou outlierů.



Zavřít

Vyzkoušet jinou
cestu