

Prostorová analýza voleb
úvodní hodina a
seminář 1

Klíčové záležitosti

- K dosažení známky alespoň E je potřeba alespoň 31 bodů
- 30 bodů seminární práce
- 10 bodů odevzdání všech úkolů
- 10 bodů průběžný test (v poslední hodině)

- „střídání“ přednášek a seminářů
- Na seminářích není vedena docházka, důležité je plnit úkoly
- Na seminářích jsou zadávány úkoly
- Odevzdání vždy den před následujícím seminářem

Organizace seminářů

- Program v sylabu se může změnit v orientační
 - Pokud nebude v předstihu uvedeno jinak, literaturu k seminářům čtěte PŘED seminářem
- Dnes: data, příprava dat
- Příště: popisné statistiky, korelace
- Přes příště: vytváření mapy
- Nakonec: regresní analýza
- Úplně nakonec: „prostorové“ metody/
shrnutí&opakování ???

Semináře

- K čemu jsou:
 - Jak zjistit z agregovaných dat užitečné informace o volebním chování
 - Prakticky využitelné zejména pro práci s volbami, ke kterým se nedělají průzkumy veřejného mínění
 - Rozvoj dovedností v oblasti statistických metod
 - Počítáme s tím, že něco umíte z podzimního kurzu
 - Seznámení s „kartografickými“ metodami
 - Zpracování analýz pro seminární práci
 - OBSAH SEMINÁŘŮ ANI LITERATURY URČENÉ K SEMINÁŘŮM NEBUDE SOUČÁSTÍ ZKOUŠKY
 - Osvojení si dovedností a znalostí bude otestováno seminární prací
 - Literaturu rozhodně čtěte

Seminární práce

- Cíl: popsat a vysvětlit rozdíly v podpoře kandidátů v senátních volbách
- Jednotlivé kroky
 - Popis volební podpory pomocí map
 - Popis volební podpory pomocí deskriptivních statistik
 - Popis vývoje volební podpory pomocí analýzy souvislosti s předchozími volbami
 - Vysvětlení rozdílů v podpoře kandidátů v obcích pomocí regresní analýzy
- PRO ZPRACOVÁNÍ PRÁCE BUDOU DŮLEŽITÉ JAK INFORMACE Z TOHOTO KURZU, TAK Z PODZIMNÍHO KURZU O KVANTITATIVNÍCH METODÁCH

Seminární práce

- Jednotlivé kroky budou konkrétněji zadány na dalších seminářích
- Finální práce bude kompilací úkolů + úvod, závěr, literatura, propojení jednotlivých částí vysvětlujícími pasážemi
- V hodnocení bude také zohledněno, jakým způsobem byly vyřešeny výtky sdělené v hodnocení úkolů

úkoly

- Za účast na seminářích a za odevzdání všech úkolů získáte 10 bodů
- Neodevzdání některého z úkolů = 0 bodů
- Úkoly samotné nebudou nijak hodnoceny, jen vám přinesou zpětnou vazbu
 - Odevzdaný úkol, který neřeší zadání, bude považován za neodevzdaný
- Neúčast na semináři je možná, ale přístup k materiálům k danému semináři budou mít jen zúčastnění
- Termín odevzdání bude sdělen vždy se zadáním úkolu (obvykle půjde o pondělní půlnoc před středečním seminářem)
- Zadání úkolu bude zveřejněno vždy na konci semináře
- Odevzdání úkolů vede k vyšší šanci vytvoření kvalitní seminární práce (a tedy k vyšší šanci na dobré bodové hodnocení)

Nějaké otázky?

Jak nachystat data?

Data – volební výsledky

- 2 možnosti jejich zisku
 - Volby.cz
 - Otevřená data
 - (žádost na ČSÚ – v případě dat, která nejsou v otevřených datech)

 - Data za obce
 - S okrsky je těžká práce
 - Je známé jen aktuální vymezení okrsků
 - <https://volby.tmapy.cz/>
 - [https://geoportal.cuzk.cz/\(S\(b01ot40jeylb2gpdapgd3mtk\)\)/Default.aspx?lng=CZ&mode=TextMeta&side=dsady_RUIAN_vse&metadataID=CZ-00025712-CUZK_SERIES-MD_RUIAN-STATY-SHP](https://geoportal.cuzk.cz/(S(b01ot40jeylb2gpdapgd3mtk))/Default.aspx?lng=CZ&mode=TextMeta&side=dsady_RUIAN_vse&metadataID=CZ-00025712-CUZK_SERIES-MD_RUIAN-STATY-SHP)

Volby.cz

- Otravná manuální práce
- Možnost obejít automatizovaným stahováním (tzv. webscrape)
- Skript lze napsat v VBA/pythonu (excel) nebo v R
 - To se tady učit nebudeme
 - Pokud chcete stahovat data z volby.cz nebo z jiných serverů se systematicky uspořádanými tabulkami, pak se tato schopnost velmi hodí
 - <http://analystcave.com/web-scraping-tutorial/>
 - <http://analystcave.com/excel-tools/excel-scrape-html-add/>
 - <https://www.promptcloud.com/blog/how-to-use-excel-to-scrape-websites>
 - Kurz Volební data management

Otevřená data

- Součást serveru volby.cz
- <http://volby.cz/opendata/opendata.htm>
- Obsahuje systematizovanou informaci o kandidujících subjektech (**registry**) a volebních výsledcích (**okrsková data**)
- Stažení zazipované složky

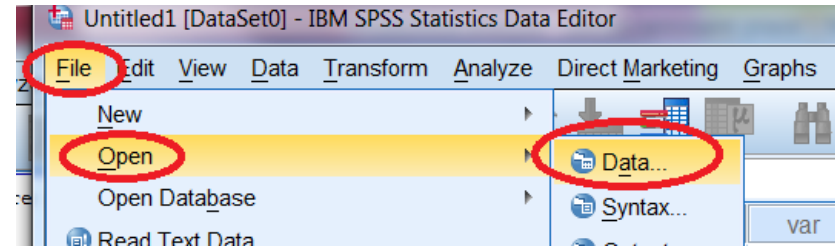
Úprava excelového souboru

- Vyfiltrovat svůj obvod
 - Data – filtr – obvod – číslo obvodu
- Vyfiltrovat druhé kolo
 - Data – filtr – kolo – 2 – ctrl+a – ctrl+c – ctrl+n – ctrl+v
- Souhrn dle obce
 - Data – souhrn – u změny ve sloupci OBEC; použít funkci součet; označit všechny položky od VOL_SEZNAM, kromě KC – ok
- Výsledek do nového sešitu
 - ctrl+a – ctrl+c – ctrl+n – klepnout pravým do první buňky a vložit jako hodnoty
- Upravit nový sešit
 - Vložit sloupec za sloupec OBEC – nový sloupec pojmenovat „Celkem“ – rozdělit sloupec obec (data – text do sloupců – oddělovač – mezera – ok) – seřadit dle celkem – smazat data v řádcích, které neobsahují slovo celkem - smazat prázdné sloupce, sloupce kde jsou jen nuly a sloupec celkem
 - Přejmenujte sloupce HLASY01, ..., HLASYX podle jmen uvedených v registrech
- ctrl+s – název: obvodN_II
 - Před uložením přejmenovat také list
 - Název sešitu krátký, bez mezer a diakritiky

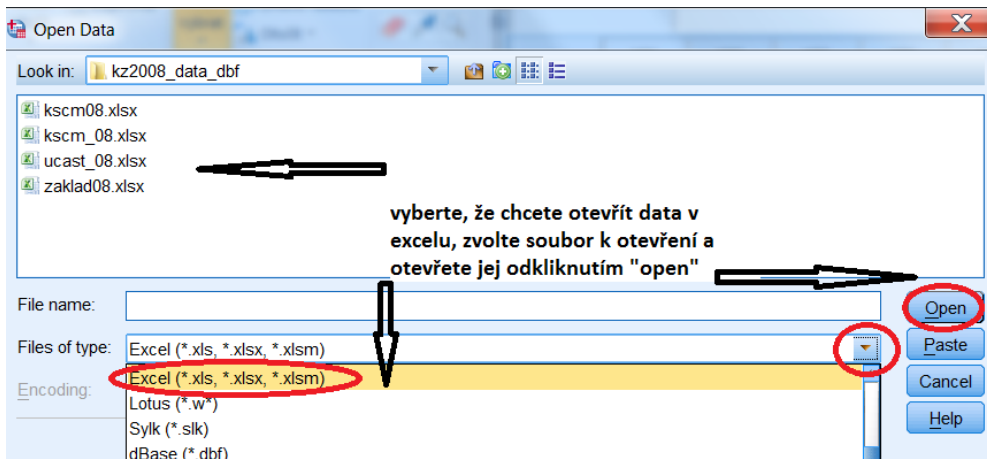
- V původním sešitu označit a smazat data za 2. kolo
- Zrušit filtr za kolo
- Upravit data za 1.kolo
- Souhrn dle obce
 - Data – souhrn – u změny ve sloupci OBEC; použít funkci součet; označit všechny položky od VOL_SEZNAM, kromě KC – ok
- Výsledek do nového sešitu
 - ctrl+a – ctrl+c – ctrl+n – klepnout pravým do první buňky a vložit jako hodnoty
- Upravit nový sešit
 - Vložit sloupec za sloupec OBEC – nový sloupec pojmenovat „Celkem“ – rozdělit sloupec obec (data – text do sloupců – oddělovač – mezera – ok) – seřadit dle celkem smazat prázdné sloupce, sloupce kde jsou jen nuly a sloupec celkem
 - Přejmenujte sloupce HLASY01, ...,HLASYX podle jmen uvedených v registrech
 - Za název v každém sloupci přidejte „_2“
- ctrl+s – název: obvodN_II
 - Před uložením přejmenovat také list

Spojení 1. a 2. kola a přidání dalších dat

- SPSS
- Přidání dat
 - Z předchozích voleb
 - Ze sčítání lidu
 - Z jiných zdrojů (MPSV, MŠMT, cokoli,...)



- Otevřete soubory, které chcete propojit, v SPSS



- Po otevření tabulek začněte se spojováním.

- Nejprve je nutné seřadit data v každé spojované tabulce podle sloupce OBEC

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

1 : OBEC 505927

	OBEC			BALKY	ODEVZ_OBAL	F
1	5059			15770	15526	
2	5061		1343	486	484	
3	5062		3564	1291	1277	
4	5064		1100	370	369	
5	5067		3497	1046	1041	
6	5067		2754	1141	1128	
7	5071		859	291	291	
8	5072		992	462	456	
9	5072		4389	1602	1595	
10	5073		1048	272	268	
11	507504	1	2664	765	753	
12	507580	1	5545	1480	1454	

Kliknout pravým a zvolit "sort ascending"

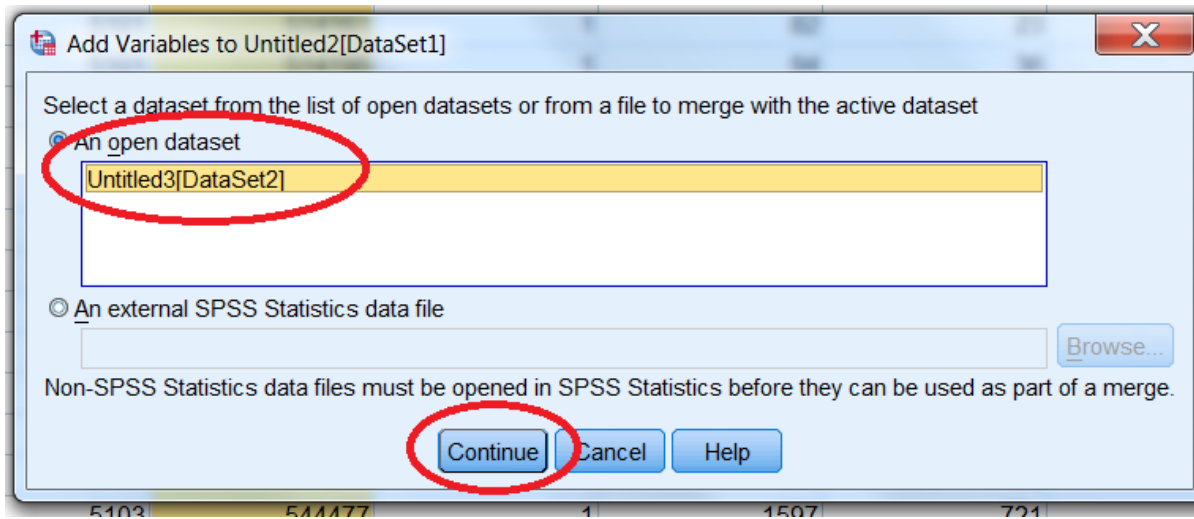
Spojení dat

- Spojování začnete od 1. kola posledních voleb (v úkolu nejlépe od účasti v krajských volbách)

1 : OBEC

	EK	VOL_SEZNAM	VYD_OBALKY	ODE
1	1	107	44	
2	1	117	42	
3	1	309	139	
4	1	597	250	
5	1	334	138	
6	1	129	58	
7	2	147	41	
8	1	244	160	
9	1	180	108	
10	1	1597	721	

- Vyberte



Add Variables from DataSet2

Excluded Variables:
KOLO(*)

New Active Dataset:
VYD_ OBEC
VYD_ OBALKY(*)
ODEVZ_ OBAL(*)
PL_ HL_ CELK(*)
Papuga(*)
Stiborský(*)
Skařupa(*)
Vlček(*)
Opálka(*)

Key Variables:
OBEC

**proměnnou OBEC
přesuňte přes toto
okno do pole Key
Variables**

Zaškrkněte obě pole

Match cases on key variables
 Cases are sorted in order of key variables in both datasets

Both files provide cases
 Non-active dataset is keyed table
 Active dataset is keyed table

Indicate case source as variable: source01

(*)=Active dataset
(+)=DataSet2

OK Paste Reset Cancel Help

509736

1

1196

486

481

415

Sčítání lidu

- https://www.czso.cz/csu/czso/otevrena_data_pro_vysledky_scitani_lidu_domu_a_bytu_2011_slodb_2011
- **Výběr údajů ze SLDB 2011 za obyvatelstvo, domy a byty, domácnosti a vyjížděku**
 - Popis dat
 - Obyvatelstvo (excel v csv. složce)
 - Po otevření: data – text do sloupců – oddělovač: čárka
 - Vyfiltrování obcí: data – filtr – typuz_naz: obec
 - Vykopírování do nového sešitu
 - Přejmenování sloupců (pomocí souboru popis dat)
 - Zjednodušení a zkrácení názvů! (jinak problémy v SPSS)
 - Smazání údajů jen za ženy/muže
 - **Výběr sloupců závisí na teoretických předpokladech**

Sčítání lidu 2021

- Vloženo v isu
- Všechna data dostupná zde:
- <https://www.czso.cz/csu/czso/vysledky-scitani-2021-otevrena-data>

- Spojení stejným způsobem jako výše
- Export do excelu
- Smazání obcí „navíc“
- Vypočtení procent

Nezaměstnanost

- <https://data.mpsv.cz/web/data/vizualizace5>
- Data neobsahují identifikační kód
- Stažení dat za každý okres, do kterého zasahuje senátní obvod
- Připsání identifikace okresu do sloupce tabulky
- Seřazení dle okresu a obce
- Připsání kódu ze struktury území
(https://www.czso.cz/csu/czso/i_zakladni_uzemni_i_ciselniky_na_uzemi_cr_a_klasifikace_cz_nuts)
- Pozor na Prahu (někdy v datech je, někdy není)

Úkol

- Vytvořit datovou matici obsahující:
 - Podporu kandidátů v senátních volbách 2010 a 2016
 - Podporu stran ve volbách do Poslanecké sněmovny v roce 2017 a 2021 (pozor na názvy! – za každý název umístěte něco jako „_PS17“)
 - Podporu stran v krajských volbách 2016 a 2020
 - Spočítané rozdíly mezi volbami (viz video „uprava vzorcu“)
- Odevzdání do půlnoci 13. 3. 2022.

Četba k příštímu semináři

- Voda, Petr. 2015. *Jaká je role postkomunismu? Volební geografie České a Rakouské republiky v letech 1990-2013*. Brno: CDK. Str. 56-62.
- Doporučená:
- Mareš, Rabušic, Soukup: 75-122