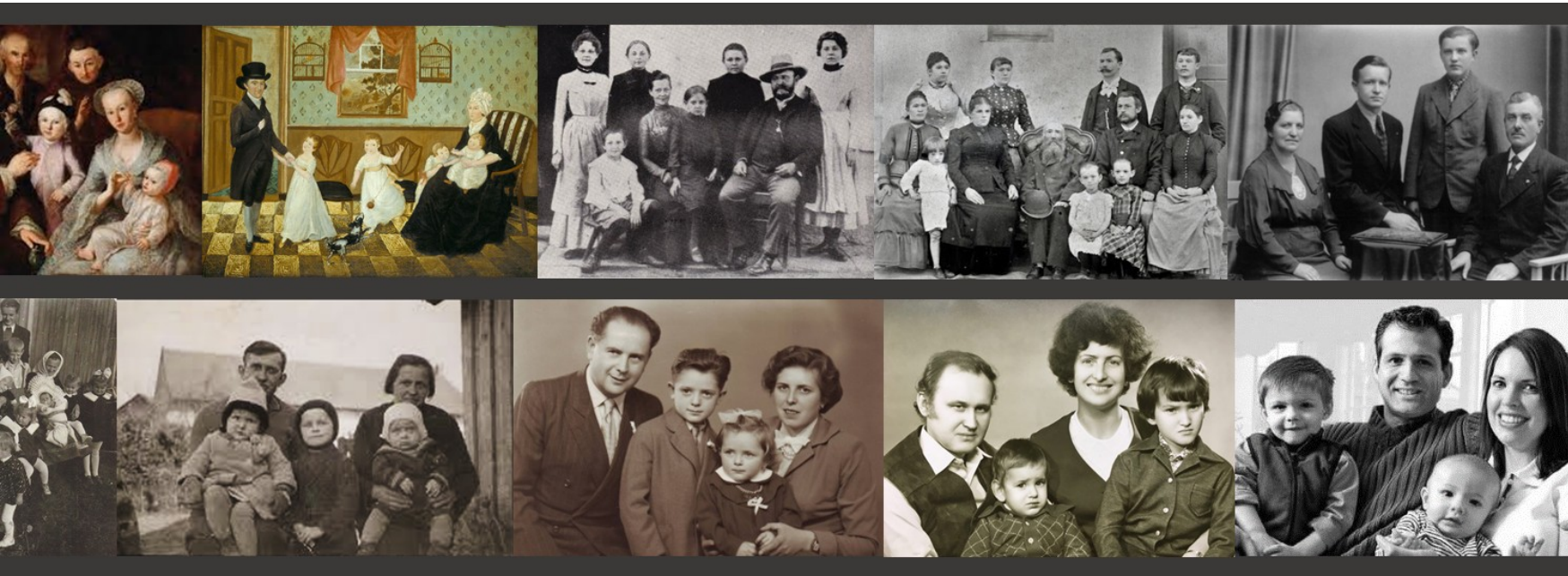


# Výběry a typy redukce

**Dříve, než se začneme bavit o výběru vzorku z populace, připomeňme si, že každý výzkum provádí vlastně mnoho typů výběrů (redukce) z širšího spektra:**

**Čas** - pro jaký čas platí výsledky našeho výzkumu? Pro jaké životní období? Pro jakou kohortu?



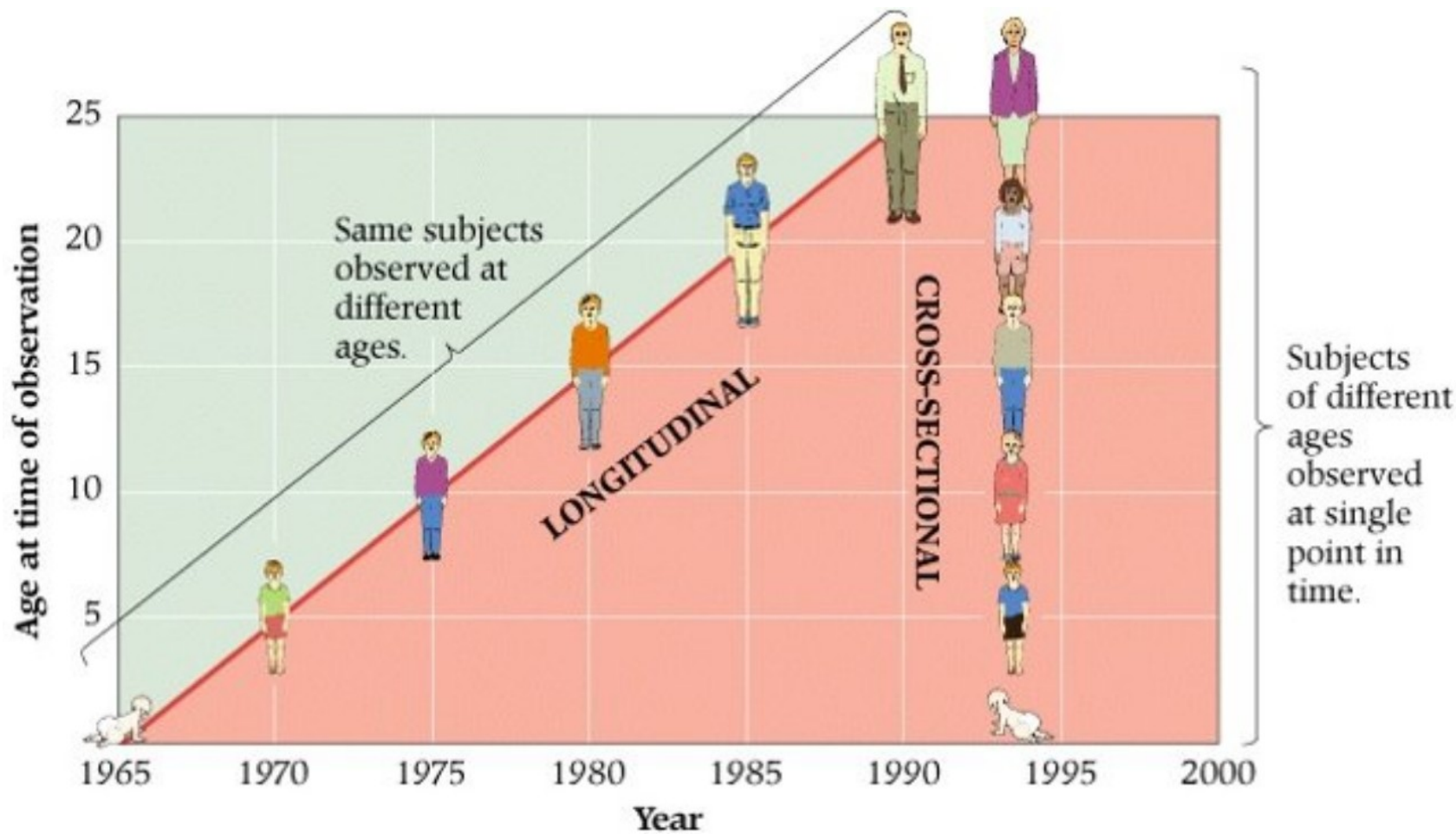
průřezové strategie

kohortní studie

studie trendů

panelové studie

+  
longitudinální výzkum



# Časová dimenze dat

**průřezové strategie**

**kohortní studie**

**studie trendů**

**panelová studie**

**+**

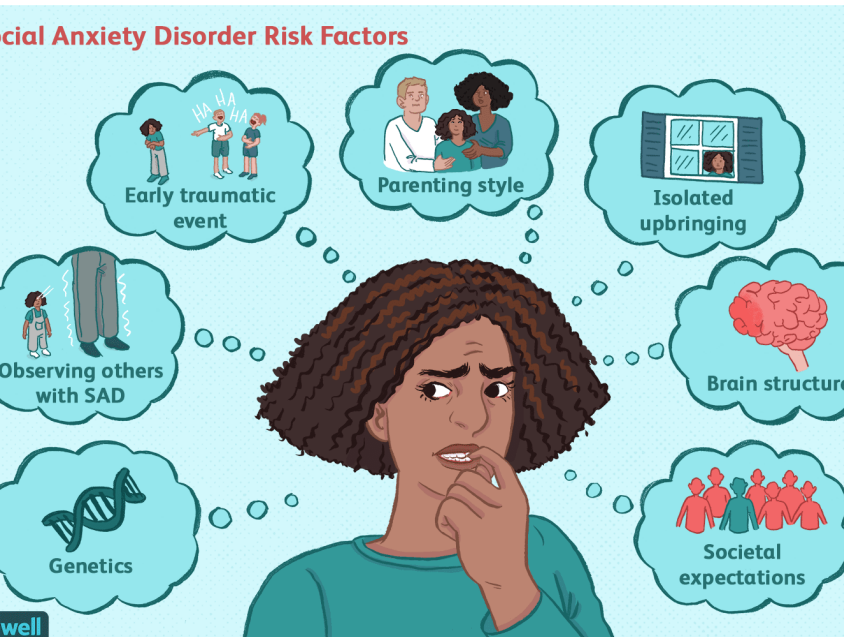
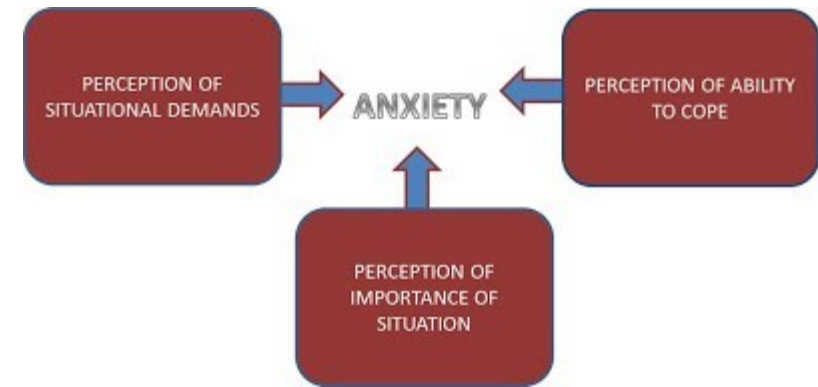
**longitudinální výzkum**

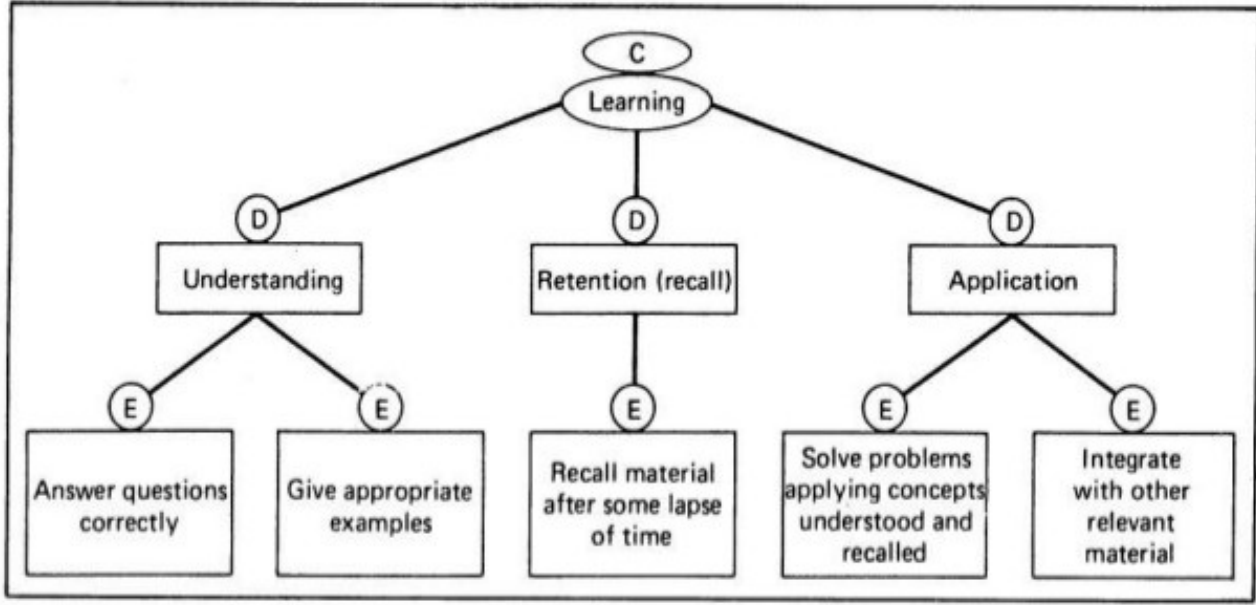
# Výběry a typy redukce

Dříve, než se začneme bavit o výběru vzorku z populace, připomeňme si, že každý výzkum provádí vlastně mnoho typů výběrů (redukce) z širšího spektra:






## Komplexnost systému

- počet sledovaných jevů a souvislostí mezi nimi  
liši se kvali/kvanti přístup – vždy však redukuje (v kvali. méně)





**COMPLEXITY CAN GROW QUICKLY**

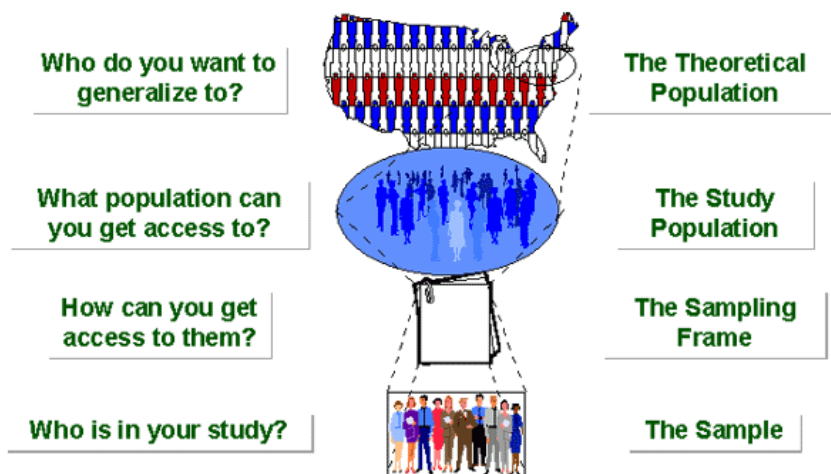
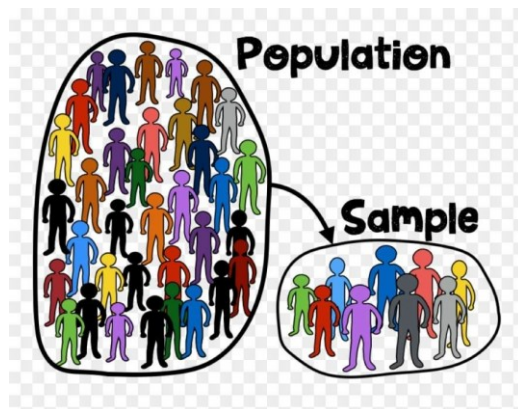
# of Elements	2	3	4	6	8
1 to 1 Relationships					
	1	3	6	15	28
Permutations (1,2), (2,1)	2	6	24	720	40,320

# Výběry a typy redukce

Dříve, než se začneme bavit o výběru vzorku z populace, připomeňme si, že každý výzkum provádí vlastně mnoho typů výběrů (redukce) z širšího spektra:

## Rozsah zkoumaných jednotek

- Počet případů (jednotek), které jsme schopni zkoumat v kvalitativním výzkumu omezen intenzitou interakce (nejsme schopni analyzovat stovky rozhovorů) v kvantitativním výzkumu rozpočtem (technicky vzato jsme schopni analyzovat celé populace)





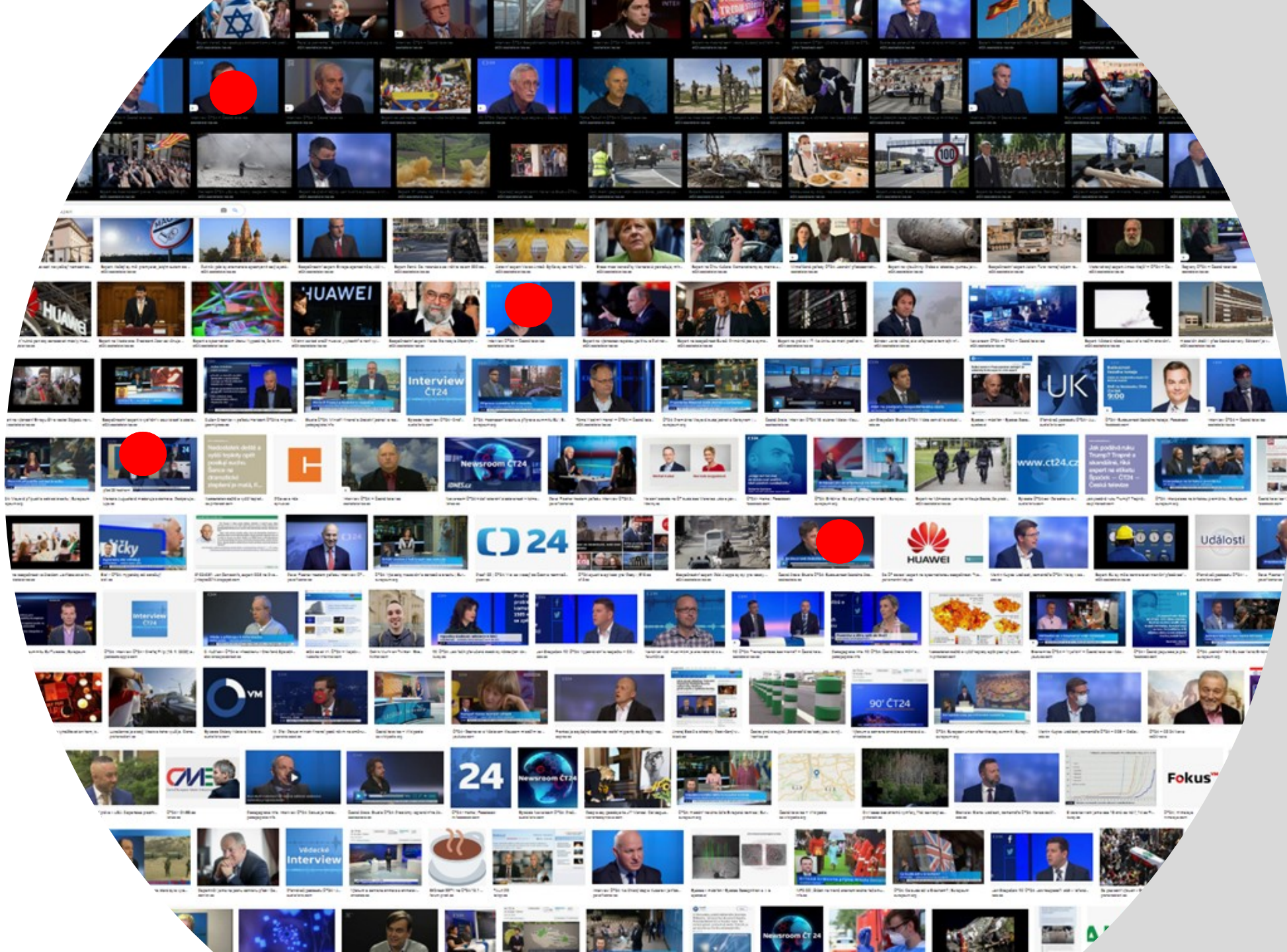
ANNO MDCCXLII

nn  
mm

nn  
kk  
ii  
hh  
gg  
ff  
ee  
dd  
cc  
bb

nn  
kk  
ii  
hh  
gg  
ff  
ee  
dd  
cc  
bb

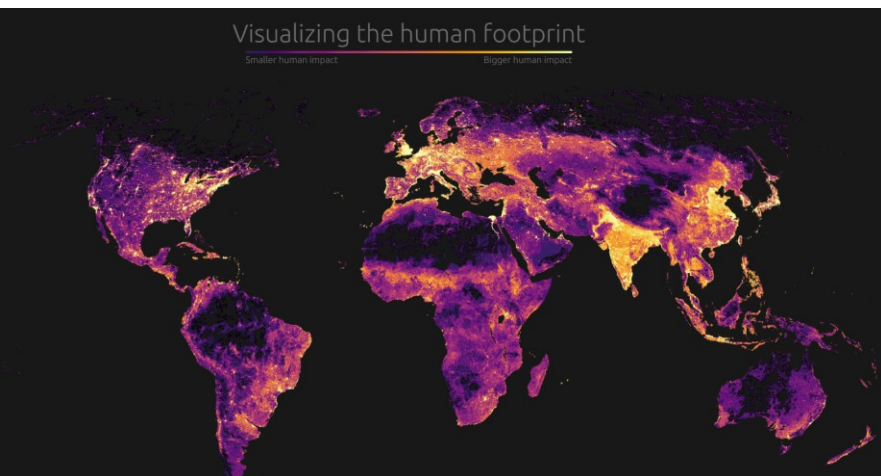
nn  
kk  
ii  
hh  
gg  
ff  
ee  
dd  
cc  
bb





# Výběry a typy redukce

## Další typy redukce: prostor, etnicita, třída...



Google RODINA

Rodina a děti jsou pro Čechy důležitější než vysok...  
novinky.cz

Rodina a Boží vůle I. | Milost.cz  
milost.cz

Co znamená, že rodina drží pohromadě...  
cestykoobe.cz

Rodina je neuvěřití hodnota, kterou společnost...  
christnet.eu

Co dělat, když z nás vznikla sešivaná rodina...  
flowee.cz

Rodina - základ státu  
w300228.savana-hosting.cz

Svatý stolec vyzývá k podpoře a ochraně ro...  
civex.cz

Světlo Vůdce: Rodina  
svetlovsude.blogspot.com

Vztahy podle horoskopu: kdo chce založit rodi...  
prozeny.cz

Rodina sešivaná ze zbytků. Co tě čeká v patch...  
femina.cz

Rodina je muž, žena a děti. Adopce homosexuálů...  
bezpolitickerekretnosti.cz

5 znaků, podle kterých poznáte, že žijete v...  
supermammy.maminka.cz

Rodina a její význam | Zdravě.cz  
vztahy-rodina.zdrave.cz

Rodina | FH Investor  
phinvestor.cz

Rodina - slovní zákeba - Klet nakladatel...  
klett.cz

Velká rodina / Diskuze | ProMa...  
promaminky.cz

Rodina - Aktuálně.cz  
aktualne.cz

Rodina, přátelé, komunita. To vše je důležité pro naše...  
janetman.cz

Klasická česká rodina mizí, ůšla jsou neuprosná. Má...  
reflex.cz

Rodina | Mateřská škola Hlubočky  
mahlubočky.cz

rodina - Mateřská škola Bellova  
mabellova.cz

Královská rodina v roce 2019: 8 nej...  
prozeny.cz

Jak se rodi děti do duhových rodin  
dvatatove.cz

Přání dětem, které nejen dnešní Den rodiny tráví be...  
idnes.cz

Související vyhledávání  
rodina kreslená  
rodina citáty  
rodina obrázky

Královská rodina vyzývá k služnosti na internetu - L...  
idnes.cz

Duggaroví: nejpočetnější americká rodina s 19 dět...  
zeňa.aktualne.cz

Co dělat, když z nás vznikla sešivaná rodina - F...  
flowee.cz

Obrázkový test: Jaci jste a jaká byla vaše rodina? | H...  
horoskopyblesk.cz

Rodina samolepky-samolepky.cz | SKL...  
skl...

šťastná rodina | Milost.cz  
milost.cz

Rodinné focení ceník | Fotograf v Pr...  
fotoprme.cz

Home - Asociace Děte a Rodina, z.s.  
dtearodina.cz

Rodina už není základ státu. Můžeme na to doje...  
zpravy.aktualne.cz

Rodina je, když... | Rádio Junior  
juniorrozhlas.cz

# Nejčastější otázka – kolik lidí potřebuji, aby to bylo reprezentativní?

Tak zaprvé:

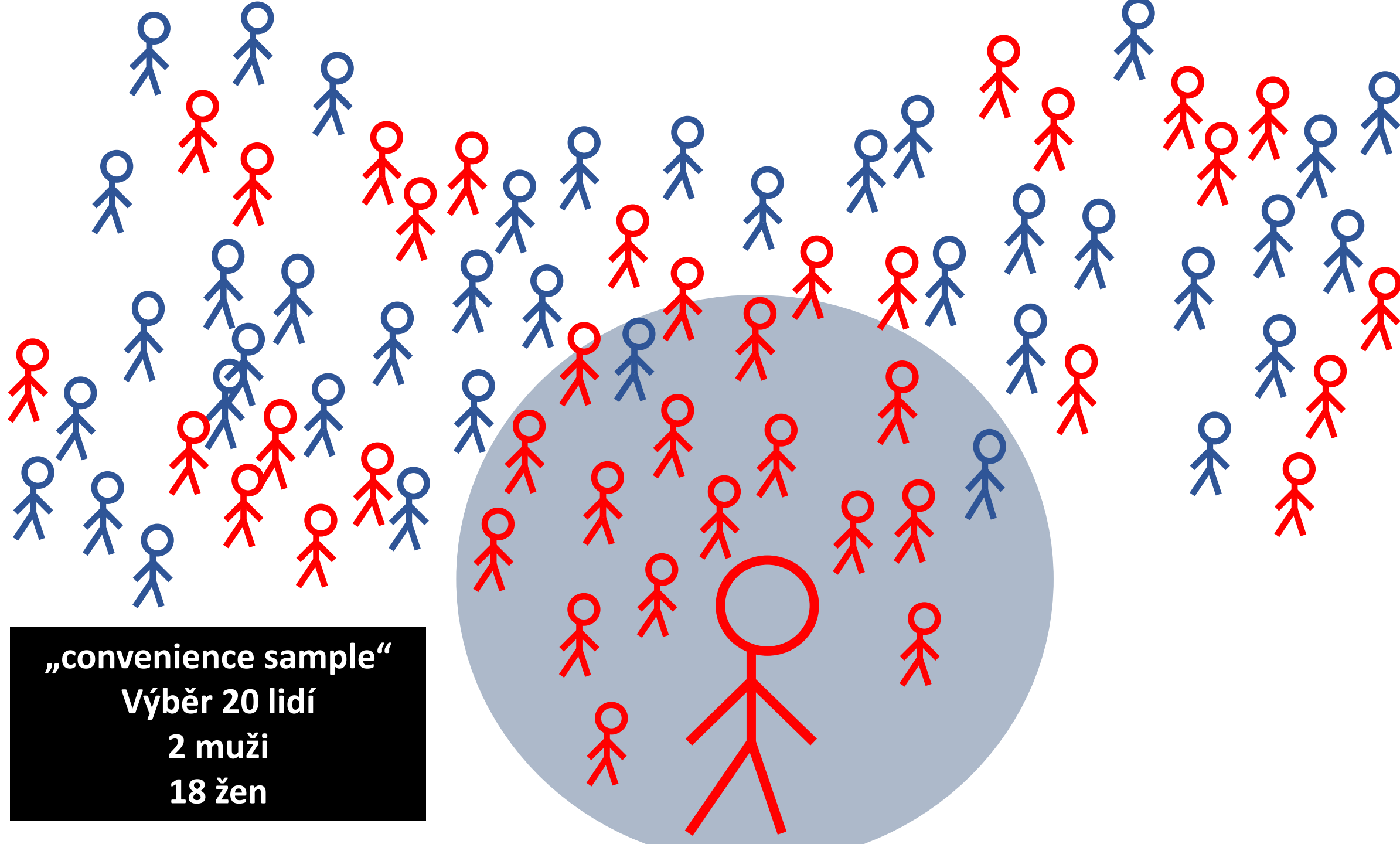
## Počet není to nejdůležitější!

**V kvalitativním výzkumu** pracujeme s jiným pojetím reprezentativity – zde počet případů hraje roli v tzv. teoretickém nasycení (saturaci): Jakmile získáváme informaci, která se začíná opakovat, zapadá do vytyčených kategorií, je výběr případů nasycený. Zde nepoužíváme pravděpodobnostní (náhodné) výběry.

**V kvantitativním výzkumu** je počet případů jednoznačně spjat s rozsahem výběrové chyby jen pokud je výběr náhodný: V náhodném (pravděpodobnostním) výběru klesá výběrová chyba (zkreslení) s velikostí vzorku.

Věnujme tedy pozornost spíše principu výběru, než počtu samotnému!

**Je náš výběr náhodný?**



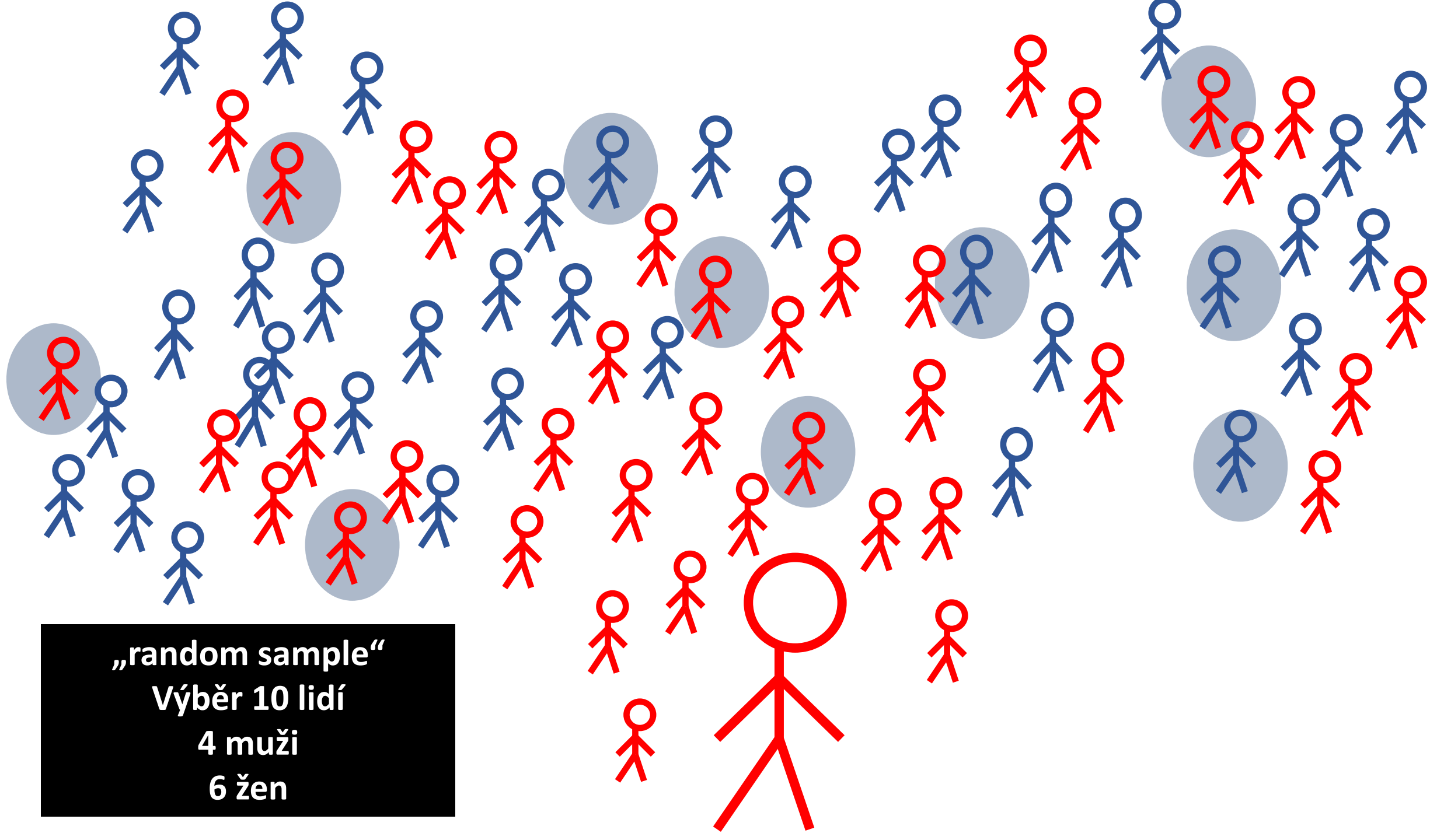
„convenience sample“

Výběr 20 lidí

2 muži

18 žen





„random sample“  
Výběr 10 lidí  
4 muži  
6 žen

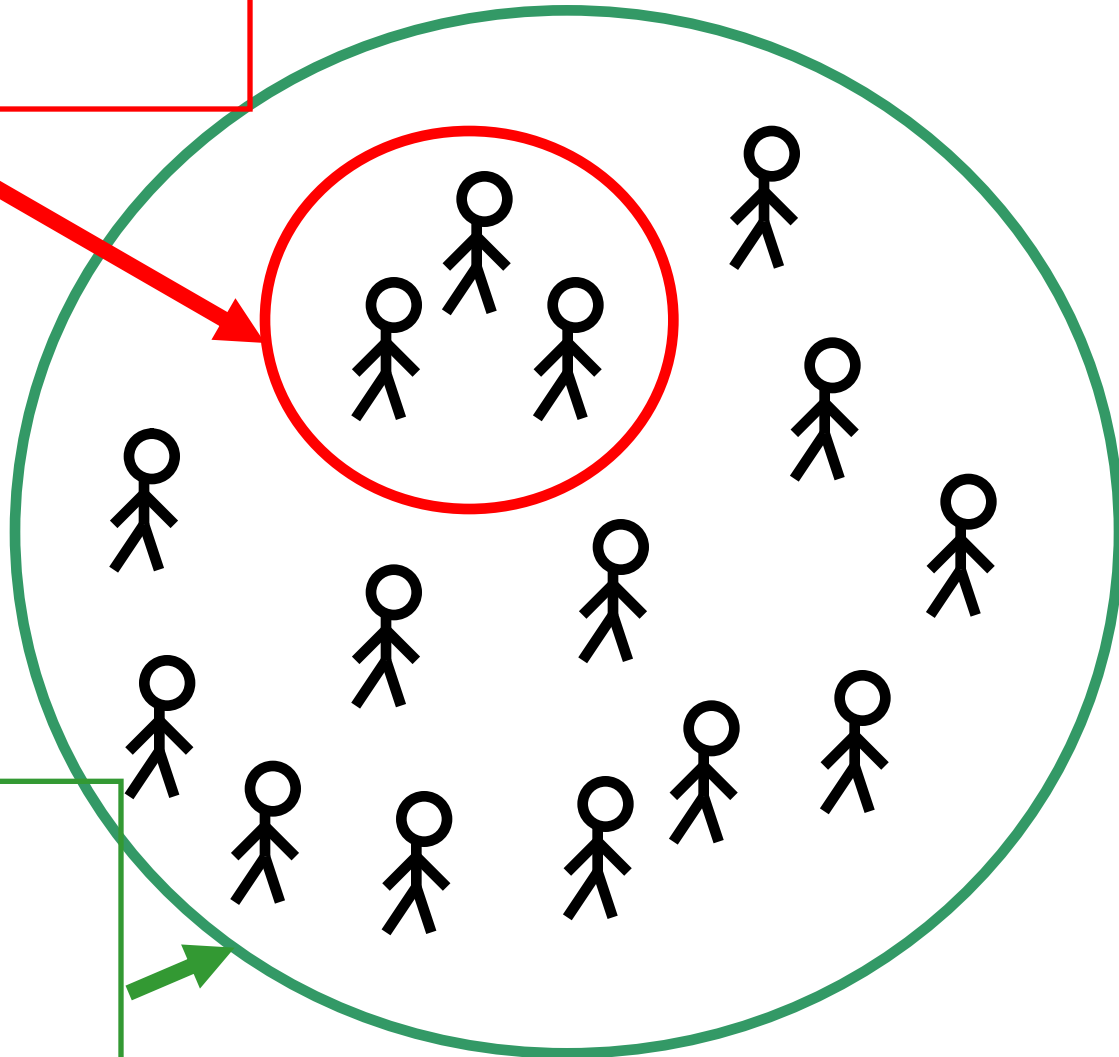


**VZOREK**/ sample/ výběrový soubor  
Výsledky analýz = STATISTIKY

VÝBĚR

INFERENCE

**POPULACE**/population  
základní soubor  
Neznámé vlastnosti =  
PARAMETRY



# ZÁKLADNÍ POJMY

**CÍLOVÁ POPULACE** (též základní soubor, základní populace) – soubor jednotek pro něž předpokládáme platnost našich výsledků

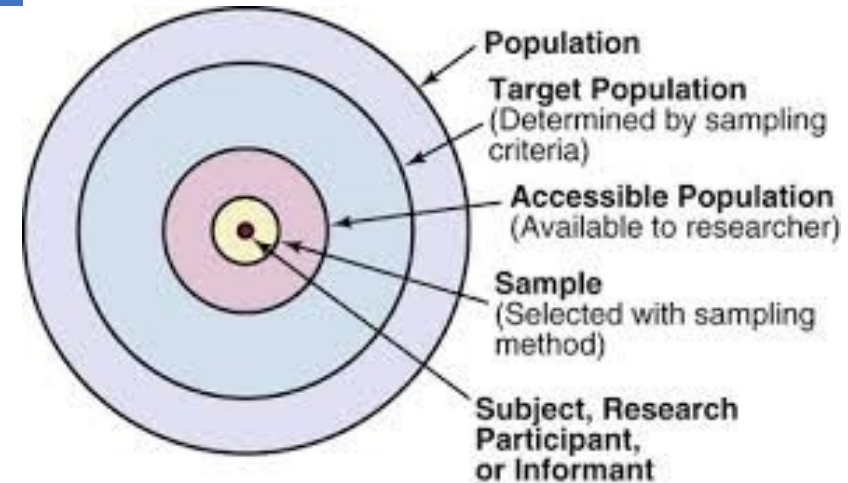
**VÝZKUMNÁ POPULACE** – soubor jednotek teoreticky dostupných pro výběr  
(oba pojmy se v praxi často používají synonymicky)

**VZOREK** (sample, výběrový soubor) – soubor jednotek, které zkoumáme

**OPORA VÝBĚRU** – seznam jednotek, umožňující náhodný výběr  
(např. seznam žáků školy, seznam adres s byty)

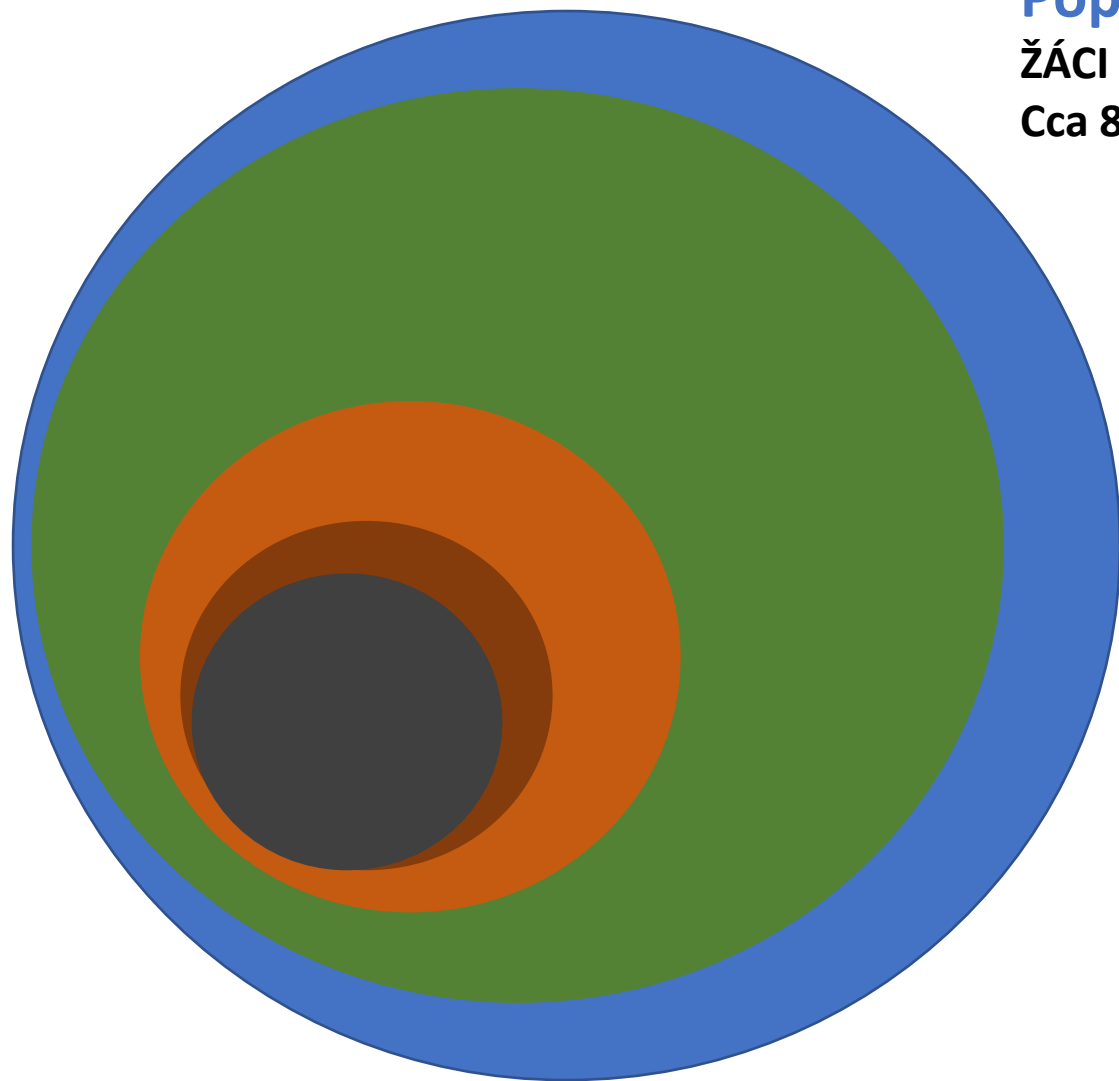
**VÝBĚR** – proces redukce počtu zkoumaných jednotek

**Jednotky:** osoby, místa, texty, obrazy, časy, události, kontexty...





## TAKŽE NAPŘÍKLAD:



### Populace:

ŽÁCI PRVNÍCH ROČNÍKŮ ZÁKLADNÍCH ŠKOL V ČR

Cca 80 000

### Dostupná populace:

ČESKY HOVOŘÍCÍ ŽÁCI PRVNÍCH ROČNÍKŮ ZÁKLADNÍCH ŠKOL V ČR

Cca 72 000

### Výběr

NÁHODNÝ VÝBĚR ŽÁKŮ

1500

### Vzorek pro výzkum

ŽÁCI, KTEŘÍ VYPLNILI DOTAZNÍK

900

### Validní data

ŽÁCI, KTEŘÍ SMYSLUPLNĚ

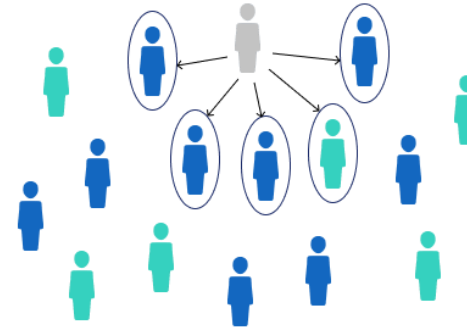
VYPLNILI DOTAZNÍK

750

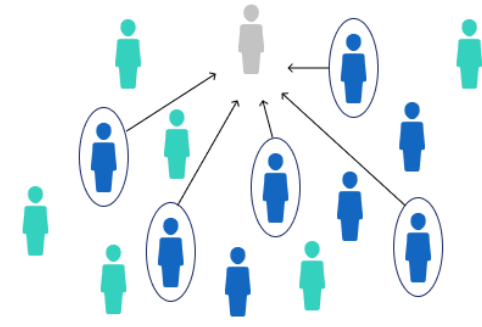
# TYPY VÝBĚRŮ I

<https://www.scribbr.com/methodology/sampling-methods/>

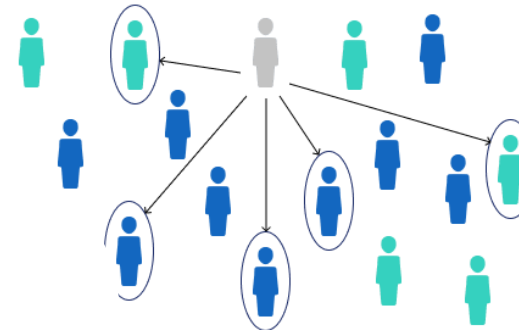
Convenience sample



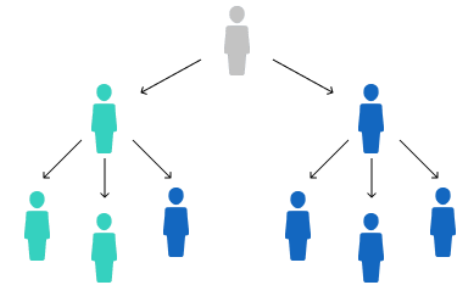
Voluntary response sample



Purposive sample



Snowball sample



Liší se pro kvalitativní a kvantitativní výzkum

U kvalitativního výzkumu je výběr veden logikou reprezentativity pro výzkumný problém



Jednotky reprezentují jednotlivé dimenze problému

Výběr vzorku je veden pravidlem teoretické saturace


Prakticky to znamená nenáhodný výběr vedený snahou po rozmanitosti, odlišných úhlech pohledu, konfrontace různých typů expertního vědění

(expertem není tazatel, ale komunikační partner)

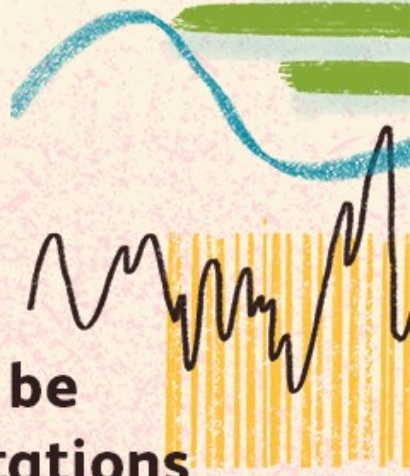

# Common Reasons for Using Samples in Research



It's impractical to survey every member of a population



Allows for inferences about the characteristics of a population



Samples can be helpful representations of the population in question



# TYPY VÝBĚRŮ II

v kvantitativním výzkumu má vzorek reprezentovat populaci  
existuje několik technik výběru:

## **náhodný výběr (pravděpodobnostní)**

dále se rozlišuje:

- prostý náhodný
- systematický náhodný
- stratifikovaný náhodný

## **kvótní výběr**

**anketa (samovýběr)**

## **Složitější typy výběrů**

**– víceúrovňová struktura (reprezentativita v různých úrovních)**

**Např. škola – třída - žák**

**– párované výběry**

**Např. žák - rodič**

## prostý náhodný výběr

**každá jednotka základní populace musí mít stejnou pravděpodobnost dostat se do výběru**

pokud je tato podmínka dodržena, data reprezentují cílovou populaci s chybou, která je odhadnutelná a závislá na velikosti vzorku

**reprezentuje známé i neznámé vlastnosti populace**

je třeba opory výběru – seznam jednotek cílové populace z níž je vybírán vzorek

*(např. seznam obyvatel ČR, seznam domácností)*

## stratifikovaný náhodný výběr

**prostý náhodný výběr je organizačně velmi náročný, proto se náhodný výběr většinou provádí v několika krocích**

*např.*

*výběr z okresů v rámci ČR*

*výběr sídel v rámci okresů*

*výběr škol v rámci sídel*

*výběr žáků v rámci školy*



**stratifikovaný náhodný výběr je reprezentativní, je však třeba vhodně zvolit kroky, aby nedošlo ke zkreslení**

**pro agentury náhodný výběr = stratifikovaný**

## kvótní výběr

**stanovuje kvóty – vyjádření základních parametrů populace, které mají být ve výběru dodrženy**

**kvóty jsou stanoveny na základě sčítání lidu, podle pohybu obyvatelstva, či jiných vyčerpávajících šetření**

**obvykle je použito několik málo charakteristik:**

*např. pohlaví, věk, vzdělání, velikost místa bydliště*

z ortodoxního hlediska neplatí pro kvótní výběr pravděpodobnostní statistika (inferenční statistika) – o tom dále

## srovnání výběrů

Náhodný výběr	Kvótní výběr
+ Kontrolovatelná reprezentativita Kontrola tazatelů	+ Rychlost Pružnost Láce
- Náročnost Nízká návratnost Nutnost opory výběru	- Nekontrolovatelné odchyly od reprezentativnosti Obtížná kontrola tazatelů Velký vliv tazatelů



Přispějte svým názorem a přivydělejte si



Upravte se a zadejte informace, abychom mohli provést vaše **spřávné výzkumy**

Vyplňujte online **dotazníky** na různá témata

Vybírejte si **peníze**, věcné odměny nebo přispějte na **dobrou věc**

> **Vyzkoušet členství v panelu**

**VÝZKUMNÉ „PANELY“ – PROFESIONÁLNÍ RESPONDENTI 😊**

**NEPLÉST S PANELOVÝM ŠETŘENÍM**

Přihlášení


E-mail



Heslo

Přihlásit

Zapomenuté

 Přihlásit se

 **Český národní panel**  
8 167 To se mi líbí

 Dát stránce To se mi líbí  Sdílet

**PRO VÝZKUM**  
NÁZORY MĚNÍ SVĚT

 Mobilní aplikace pro Android

## Rozhodování o konstrukci vzorku

Chci-li dosáhnout reprezentativity s danou mírou chyby, musím splnit dvě podmínky:

1) Pravděpodobnostní výběr (paradoxně pouze u náhodného výběru lze stanovit pravděpodobnost chyby, zatímco u systematických výběrů ne)  
-(o tom jsme mluvili výše – typy výběrů)

(tady je na místě rozhodování do jaké míry jsem schopen docílit pravděpodobnostního výběru, nebo do jaké míry se uchýlit ke kvótnímu)

2) Velikost výběrového souboru (při pravděpodobnostním výběru je dosažení dané míry přesnosti jen otázkou velikosti souboru)  
-(tím se nyní budeme zabývat)

## Jak velký vzorek potřebuji?

Kritéria rozhodování:

Zvolená přesnost

*chyba 3 % = 1000 lidí, chyba 6 % = 300 lidí, atd.*

Nutná třídění do podskupin

*dvojměrná, tříměrná, x-měrná třídění  
typy použitých proměnných a počty kategorií*

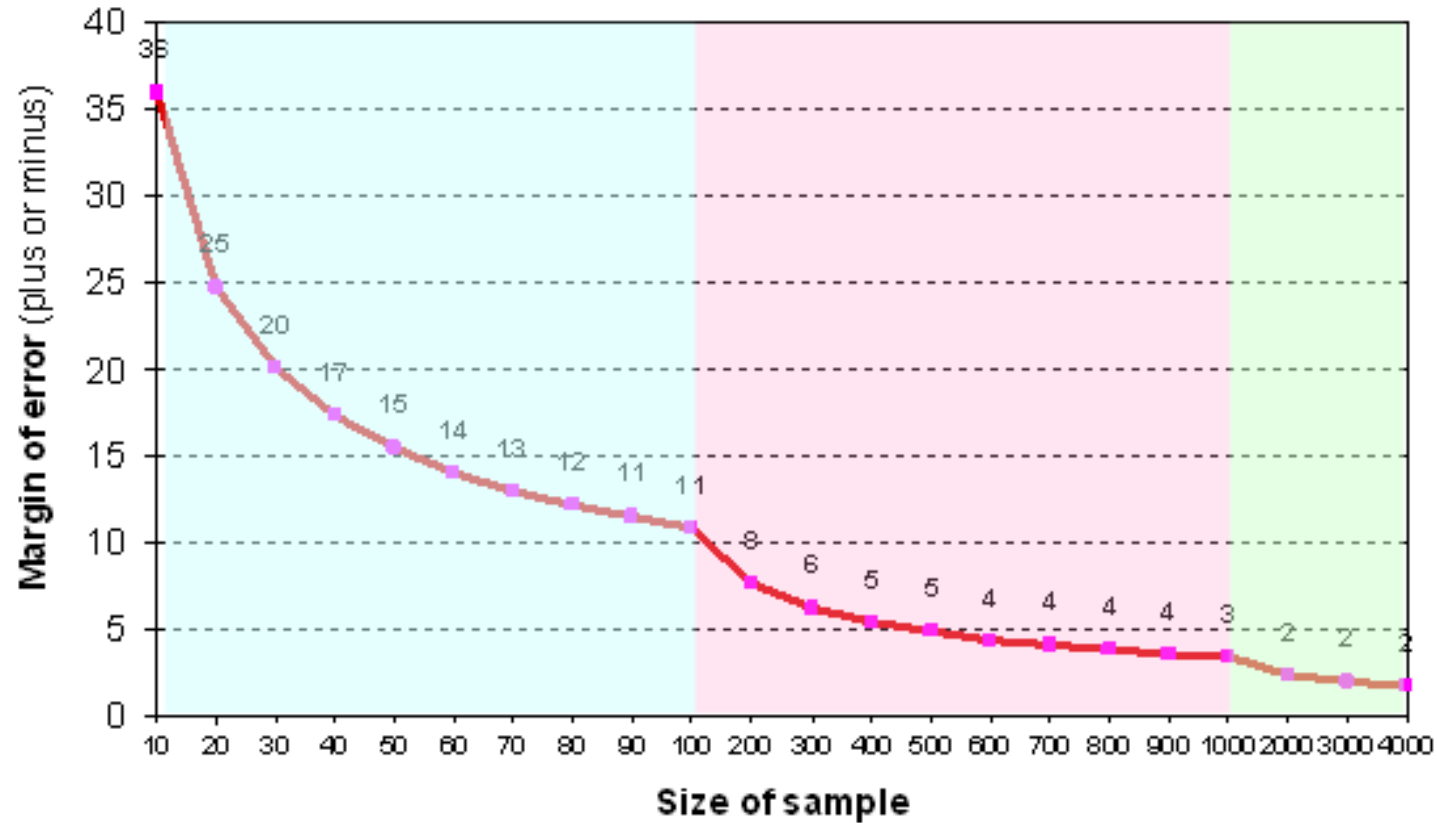
Non-response

Dostupnost a heterogenita cílové populace

Velikost cílové populace (hraje roli pouze u malých populací)

# výběrová chyba

## Sample Size and Sampling Error



# Non-response

**Při plánování vzorku je třeba počítat s neochotou odpovídat**  
2 aspekty problému:

## **Navýšení velikosti o předpokládaný „odpad“**

- Např. on-line survey má návratnost mezi 20 – 40 %, potřebuji-li vzorek 100 lidí, obešlu 500
- Např. longitudinální výzkum má úmrtnost 30 % ročně. Potřebuji-li 100 lidí za 5 let, musím začínat s vzorkem cca 370 lidí

## **Sledování struktury odmítnutí**

Čím vyšší je míra odmítnutí, tím více se výběr blíží anketě  
Je-li odmítnutí rozloženo náhodně, není to takový problém jako když systematicky odmítají konkrétní skupiny -> zkreslení výsledků

- Alespoň sledovat návratnost (mít přehled kolik jednotek jsme oslovili)
- Screening před dotazováním – získáváme základní údaje i o těch, kteří nespolupracují (např. tazatel může zaznamenat pohlaví, odhad věku bydliště...)

**Vedle toho se samozřejmě snažíme zavést motivační prvky!!**

## Důsledky pro analýzu

Pro koho je vzorek reprezentativní?  
Jedná se o vzorek, nebo vyčerpávající výběr?

-> Aplikace inferenční statistiky

**Inference:** usuzování ze vzorku na základní populaci

**Statistická významnost:** pravděpodobnost, že statistiky vypočtené na vzorku platí také pro populaci

Takže například

Potřebuji srovnat školní výkon dětí z rozvedených a intaktních rodin

Zvolená přesnost analýzy pro daný problém:

+ - 5 % (cca 500 případů)

Počet skupin, jež potřebuji srovnávat:

Děti z rozvedených a intaktních rodin (500+500)

+ Kontrolní proměnné – elaborace?

*Např. vzdělání rodin – alespoň třídění nižší/vyšší*

Nižší vzdělání rodičů – rozvedení (500)

Nižší vzdělání rodičů – intaktní (500)

Vyšší vzdělání rodičů – rozvedení (500)

Vyšší vzdělání rodičů – intaktní (500)

Celkem 2000

Non-response/návratnost? (počítejme např. 50 %)

→ Oslovit vzorek 4000 dětí

