

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 = test (v poslední hodině, podrobnosti sdělí doc. Pink)

- „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 čtete PŘED seminářem
- Dnes: příprava dat
- 7. 3. Kontrola dat a popisné statistiky
- 21. 3. Jak se liší volební podpora stran v prostoru a čase?
- 28. 3. Jak se dělá mapa? Práce s daty a práce s GISem.
- 18. 4. Proč se podpora stran v různých místech liší?
- 2. 5. „prostorové“ metody/ shrnutí & opakování
- 7. 5. rezerva

# 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
- **Odevzdání 20. 5.**, v případě nedostatku bodů je možné práci přepracovat. Přepracované práce mají deadline 20. 6.

# 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
- Termíny odevzdání v sylabu (vždy jde úterní půlnoc před čtvrtěčným seminářem)
- Zadání úkolu bude zveřejněno vždy na konci semináře
- Při zpracování úkolů můžete spolupracovat
  
- Odevzdání úkolů vede k vyšší šanci vytvoření kvalitní seminární práce (a tedy k vyšší šanci na vyšší 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>
- [https://volby.cz/opendata/se2018/se2018\\_opendata.htm](https://volby.cz/opendata/se2018/se2018_opendata.htm)
- Obsahuje systematizovanou informaci o kandidujících subjektech (**registry**) a volebních výsledcích (**okrsková data**)
- Stažení zazipované složky

# Postup po senátní volby

# Úprava excelového souboru

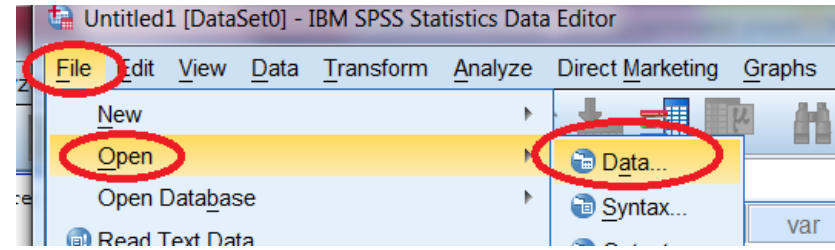
- Vyfiltrovat svůj obvod
  - Data – filtr – obvod – číslo obvodu
- Vyfiltrovat druhé kolo
  - Data – filtr – kolo – 2 – ctrl+a – ctrl+c – ctrl+n – ctrl+v
- Souhrn dle obce
  - Data – souhrn – u změny ve sloupci OBEC; použít funkci součet; označit všechny položky od VOL\_SEZNAM, kromě KC – ok
- Výsledek 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

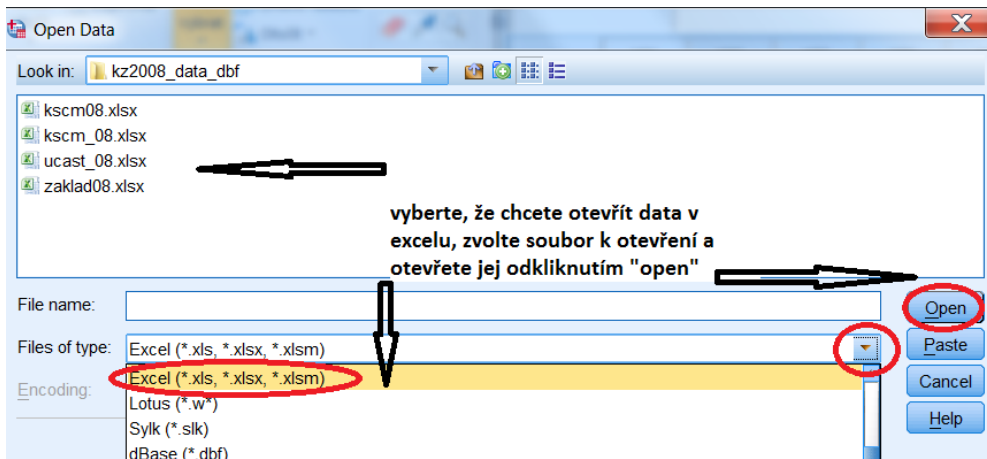


# Spojení 1. a 2. kola a přidání dalších dat

- 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.

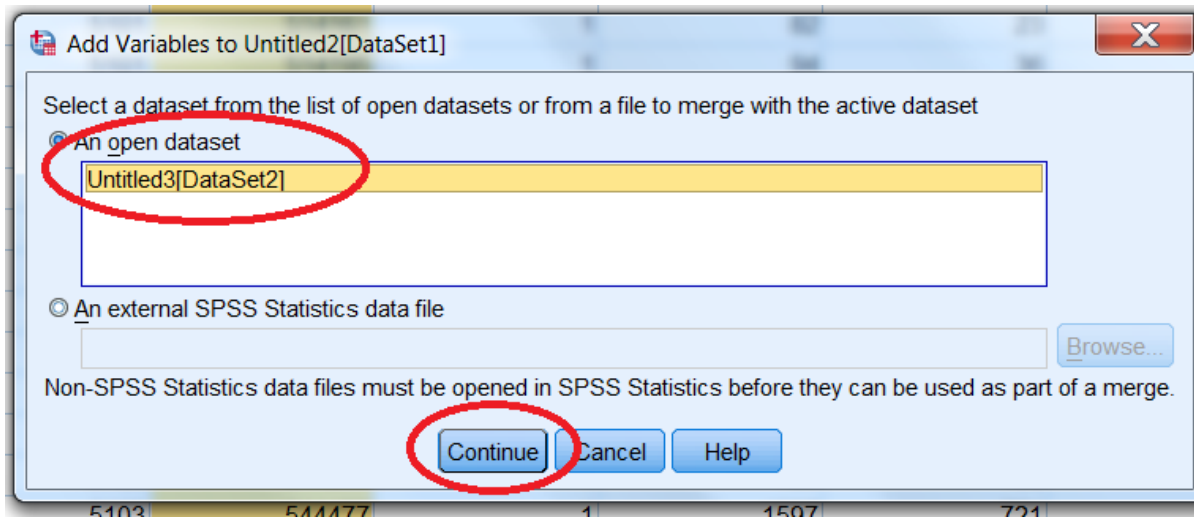
# Spojení dat

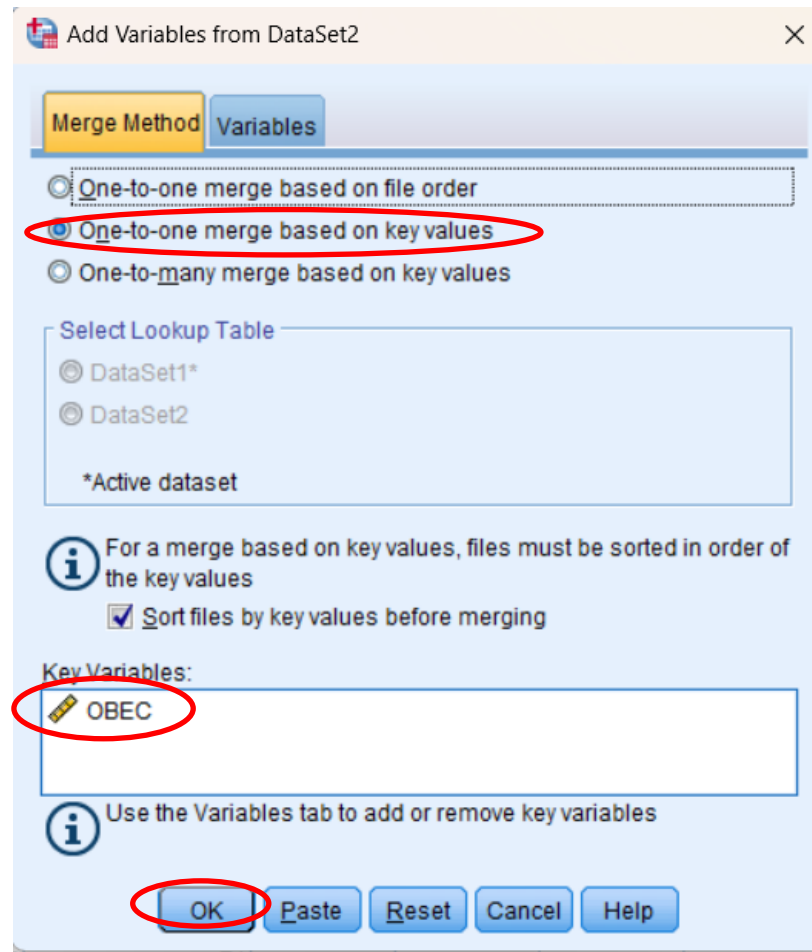
- Spojování začněte od 1. kola posledních voleb (v úkolu nejlépe od účasti v krajských volbách)

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





# Postup po krajské volby

# Úprava excelového souboru

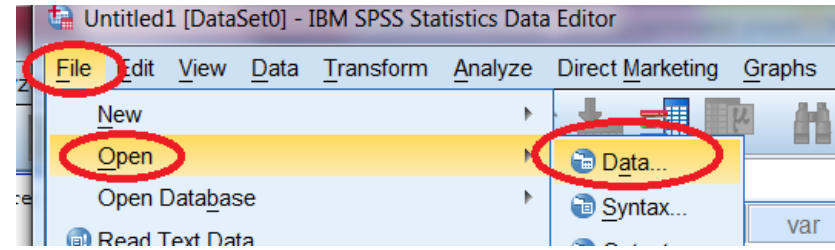
- Složka obsahuje 2 soubory kzt6 (účast) a kzt6p (hlasy pro strany)
- Vyfiltrovat svůj kraj z tabulky kzt6
  - Data – filtr – okres – vybrat čísla okresů
  - Číselník okresů dle krajů:  
<https://www.czso.cz/documents/10180/25385875/13315241+420213p3.pdf/295ebb5e-1e4e-4d50-bd80-379c5527b911?version=1.0>
- ctrl+a → ctrl+c → ctrl+n → ctrl+v
- Souhrn dle obce
  - Data – souhrn – u změny ve sloupci OBEC; použít funkci součet; označit položky VOL\_SEZNAM, VYD\_OBALKY, ODEVZ\_OBAL a PL\_HL\_CELK
- 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
- Uložit pod jménem *jmenokraje\_ucast20*

- Otevřít sešit kzt6p
- Smazat zbytečné sloupce TYP\_FORM OPRAVA CHYBA OKRSEK KC\_1 KC\_2, KC\_3, KC\_4, KC\_sum a hlasy01-konec dat
- Pomocí data – filtr vybrat okresy patřící do kraje
- Vykopírovat vybraná data do nového sešitu
  - Parlamentní strany + ostatní strany nad 5 %
- Vybrat stranu (číslo dle webové prezentace ve volby.cz)
  - Data – filtr – kstrana
  - ctrl+a → ctrl+c → ctrl+n → ctrl+v
- Souhrn dle obce
  - Data – souhrn – u změny ve sloupci OBEC; použít funkci součet; označit položku poc\_hlasy
- 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 a kstrana
- Přejmenovat sloupec poc\_hlasy na strana20
- Uložit sešit pod názvem strana20
- Opakovat pro všechny strany nad 5% a strany které mají zástupce v poslanecké sněmovně (ANO, ODS, KDU, TOP, Piráti, STAN, SPD)

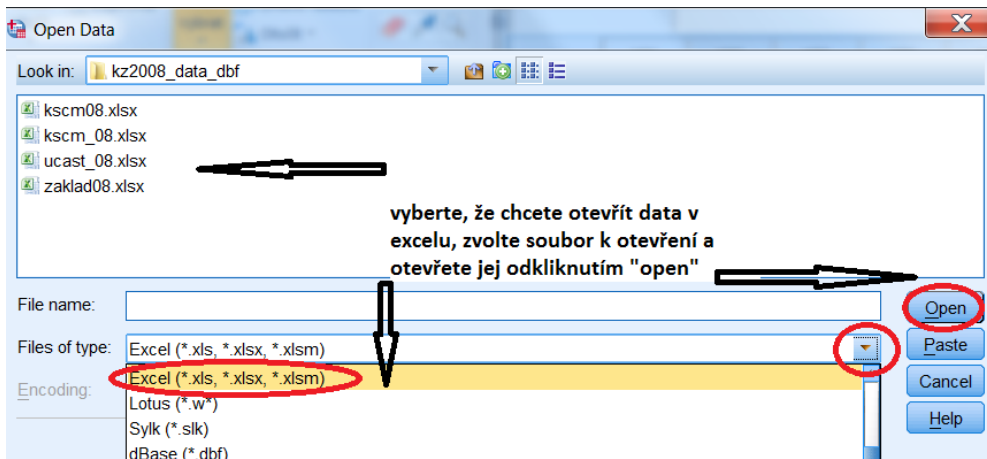


# Spojení účasti a výsledků a přidání dalších dat

- 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.

# Spojení dat

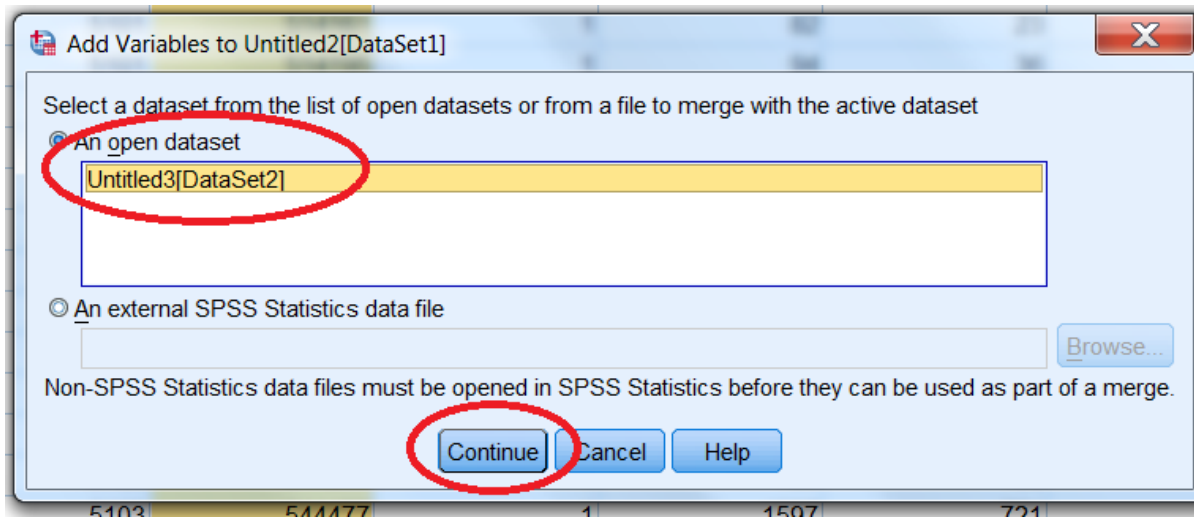
- Spojování začnete od účasti v krajských volbách)

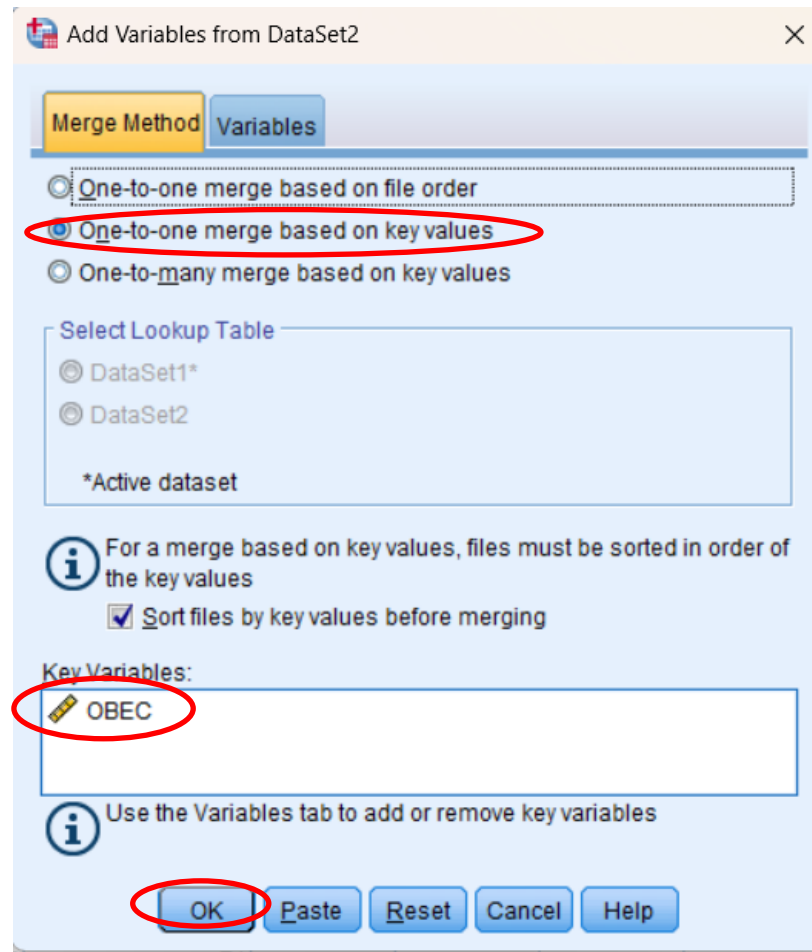
1 : OBEC

	EK	VOL_SEZNAM	VYD_OBALKY	ODE
1				
2				
3				
4				
5	1	107		44
6	1	117		42
7	1	309		139
8	1	597		250
9	1	334		138
10	1	129		58
11	2	147		41
12				
13				
14	1	1597		721

v okně s otevřeným souborem UCAST\_ROK  
v nabídce "Data" zvolte položku "Merge Files"  
a poté "Add Variables"

- Vyberte





# Rozdíly mezi výukovým videem a daty z roku 2020

- 1:40 nyní nejsou data s výsledky rozdělena po krajích. Volební výsledky za všechny obce všech krajů jsou v jediném souboru (Kzt6p)
- 11:34 Není potřeba se dívat na čísla krajů
- 12:35 Nemažte sloupec okres. Použijte jej pro stejné odfiltrování jako v případě účasti

# Sčítání lidu

- [https://www.czso.cz/csu/czso/otevrena\\_data\\_pro\\_vysledky\\_scitani\\_lidu\\_domu\\_a\\_bytu\\_2011\\_slodb\\_2011](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 vyjížďku**
  - 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 celý kraj
- 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](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 2012 a 2018
  - 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 „připrava\_dat\_pro\_mapu“)
  - Všechna data budou v jedné velké tabulce, nikoli v jednotlivých sešitech
  - Pro výsledky relevantních kandidátů v senátních volbách vypočtete procenta (hlasy pro kandidáta/platné hlasy\*100)
- Odevzdání do půlnoci 5. 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