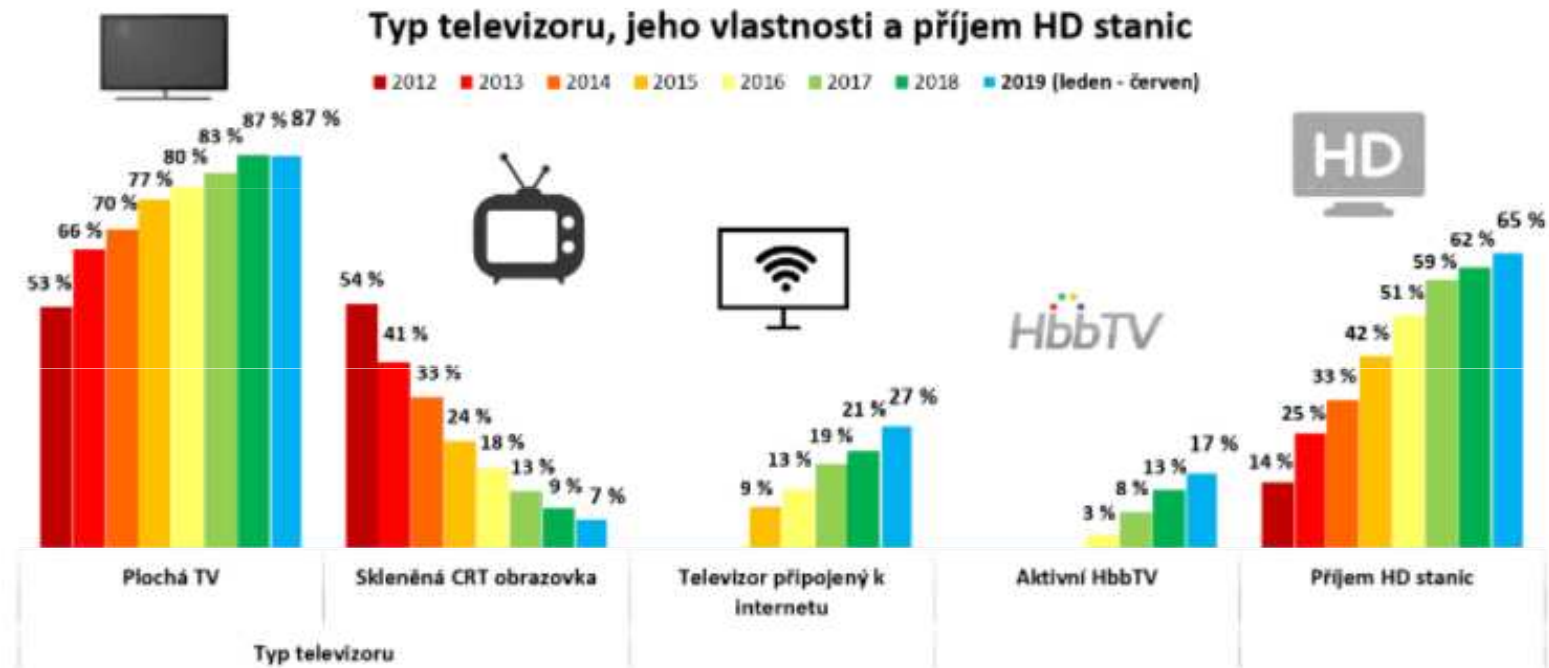


MUNI
FSS

Bivariační analýza

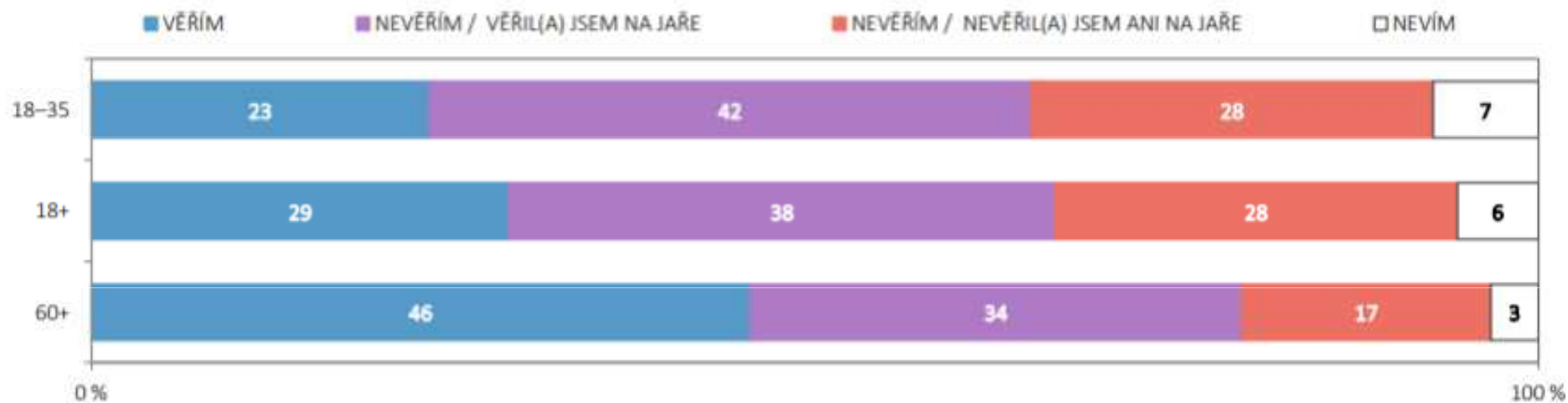
VÝVOJ TV VYBAVENOSTI ČESKÝCH DOMÁCNOSTÍ

93 % televizních domácností má doma funkční televizor



Zdroj: ATO – Nielsen Admosphere. Závěreční a kontinuální výzkum 2012–2019. Základ: všechny domácnosti 10

Důvěra ve vládu v souvislosti s informacemi o covid-19 u široké veřejnosti a mladých Čechů



Covid-19 a šířitelé konspiračních teorií. Průzkum postojů české veřejnosti. 2021. Praha: STEM. Dostupné z: https://www.stem.cz/wp-content/uploads/2021/04/STEM_COVID-19-a-siritele-konspiraci-CZ-verejne.pdf

Tab. 1: Podíly uživatelů online sociálních sítí v populaci 18+ podle pohlaví, věku a vzdělání.

	2019/2020	2020
Pohlaví		
muži	68 %	73 %
ženy	76 %	84 %
Věk		
18-24	95 %	97 %
25-34	89 %	93 %
35-44	80 %	84 %
45-54	69 %	75 %
55-64	60 %	70 %
65 a více	54 %	65 %
Vzdělání		
základní	80 %	87 %
středoškolské bez maturity	72 %	79 %
středoškolské s maturitou	71 %	77 %
vysokoškolské	70 %	77 %

- Zdroj: MACKOVÁ, Alena, Martina NOVOTNÁ, Klára PROCHÁZKOVÁ, Jakub MACEK a Lenka HRBKOVÁ. Češi na sítích, důvěra a polarizace v době pandemie. 2021. 20 s. doi:10.13140/RG.2.2.20950.78401/1.

Jak si zobrazit
výsledky podle
vybrané proměnné/
pro specifickou
skupinu lidí?

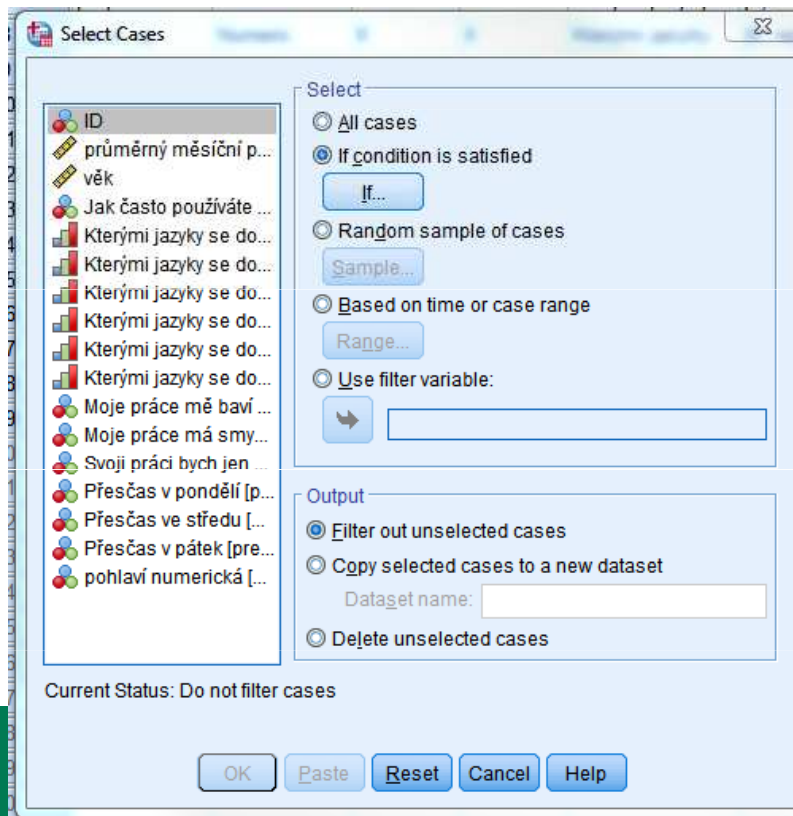
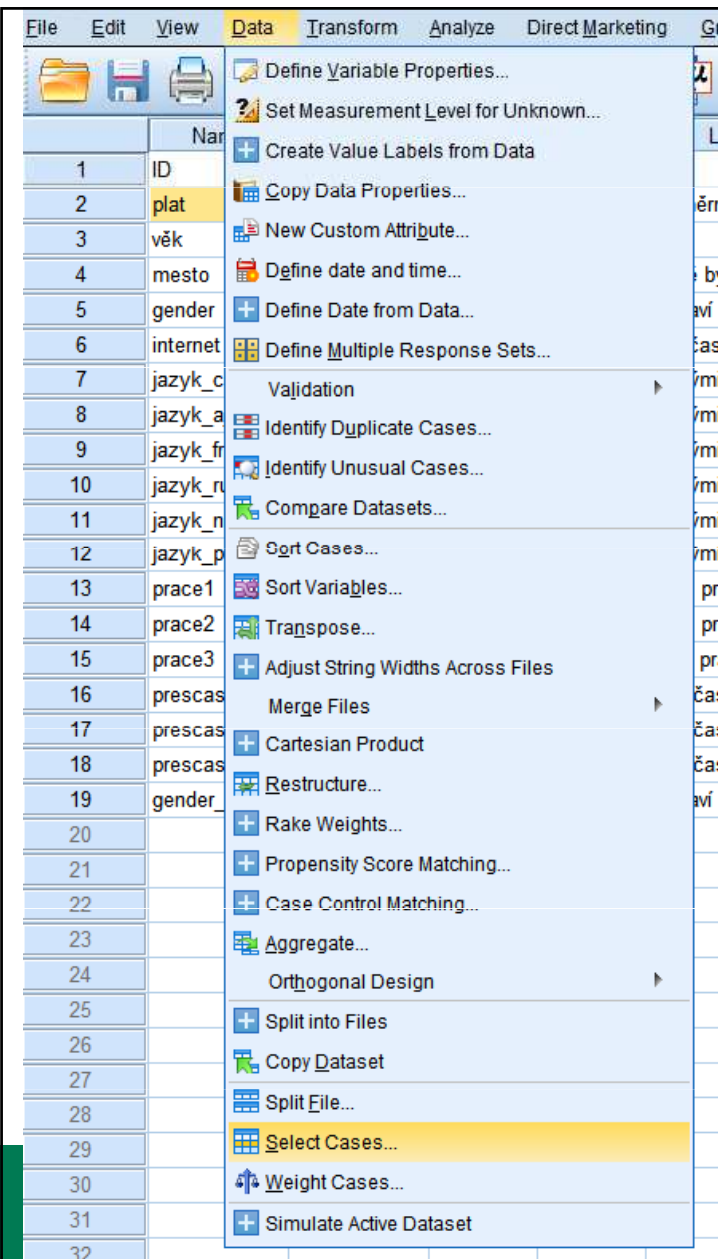
- tři možnosti:
 - 1) SELECT CASES
 - 2) SPLIT FILE
 - 3) BIVARIAČNÍ
ANALÝZA

Select cases

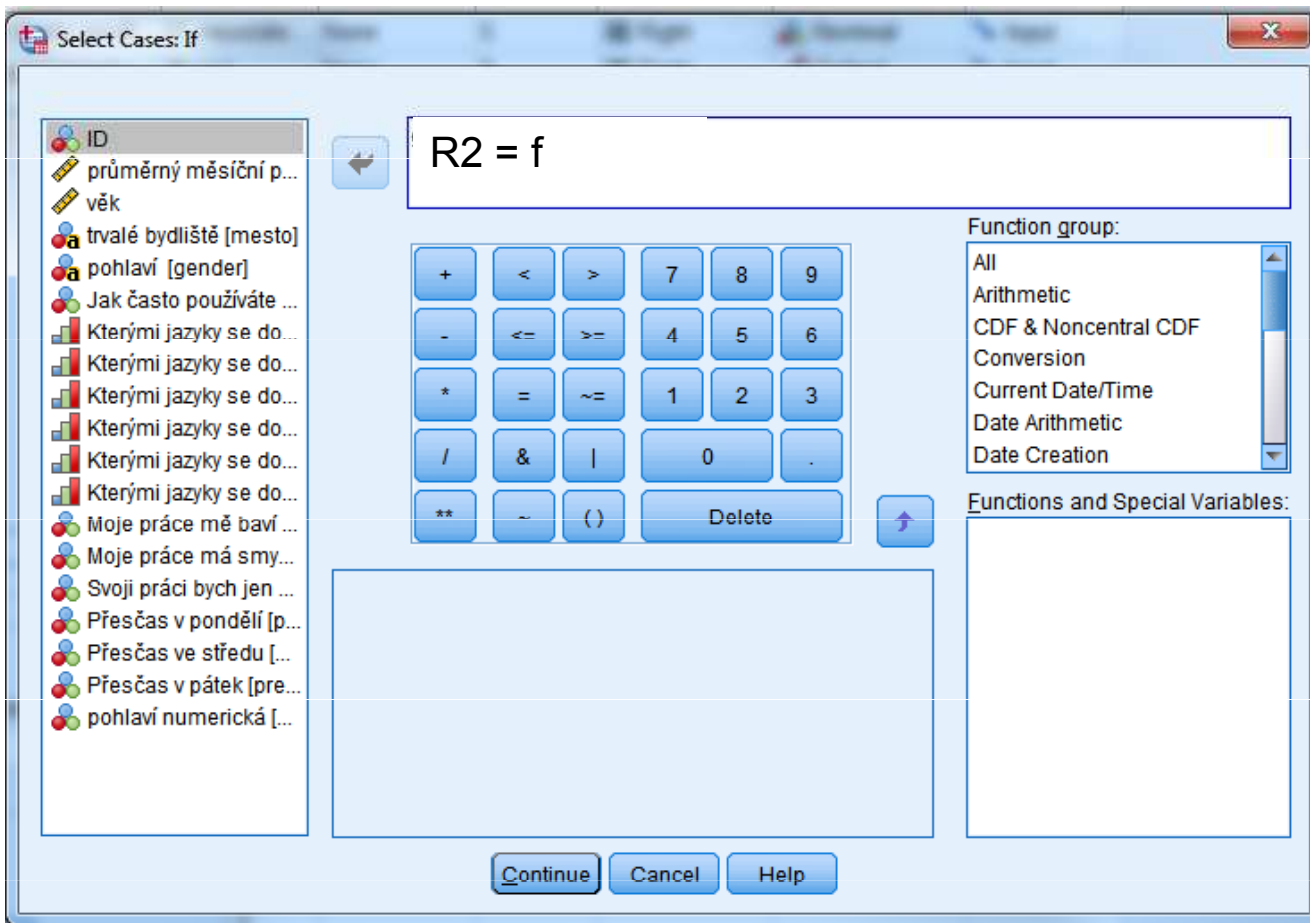
- 1. Data → Select cases
- 2. If condition is satisfied

Co to dělá:

- Vybere z našeho celé vzorku takové respondenty, které splňují naše nastavené podmínky
- Každé další příkazy (statistiky), které SPSS zadáme, pak provádí pouze na vybrané sub-skupině



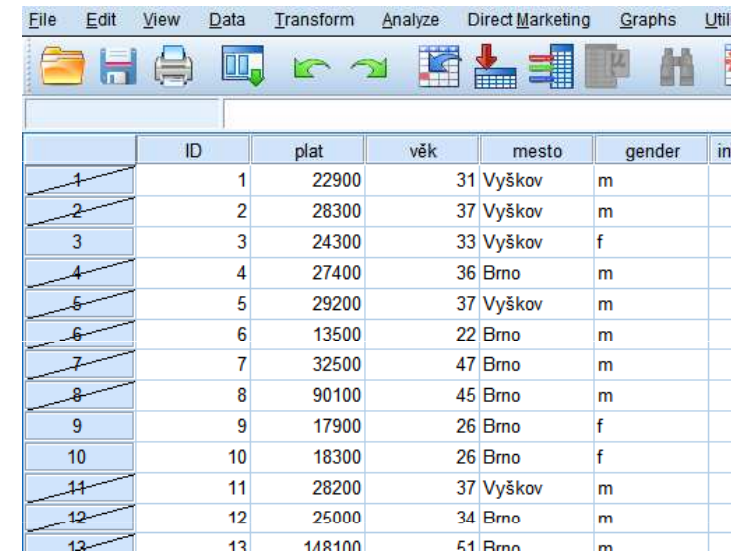
Select cases



- Chceme zobrazit četnosti pro kategorii žen
- 1. naklikneme proměnnou proměnnou R2 pomocí šipky
- 2. specifikujeme podmínku
- Vek_kat je nominální
- $R2 = f$
- 3. continue a ok
- 4. následně znovu vyjedeme četnosti (Analyze → Descriptive Statistics → Frequencies), které se nyní zobrazí jen pro ženy

Select cases

- V datasetu v „Data view“ vidíme nepoužité respondenty (jsou vyškrtnutí)



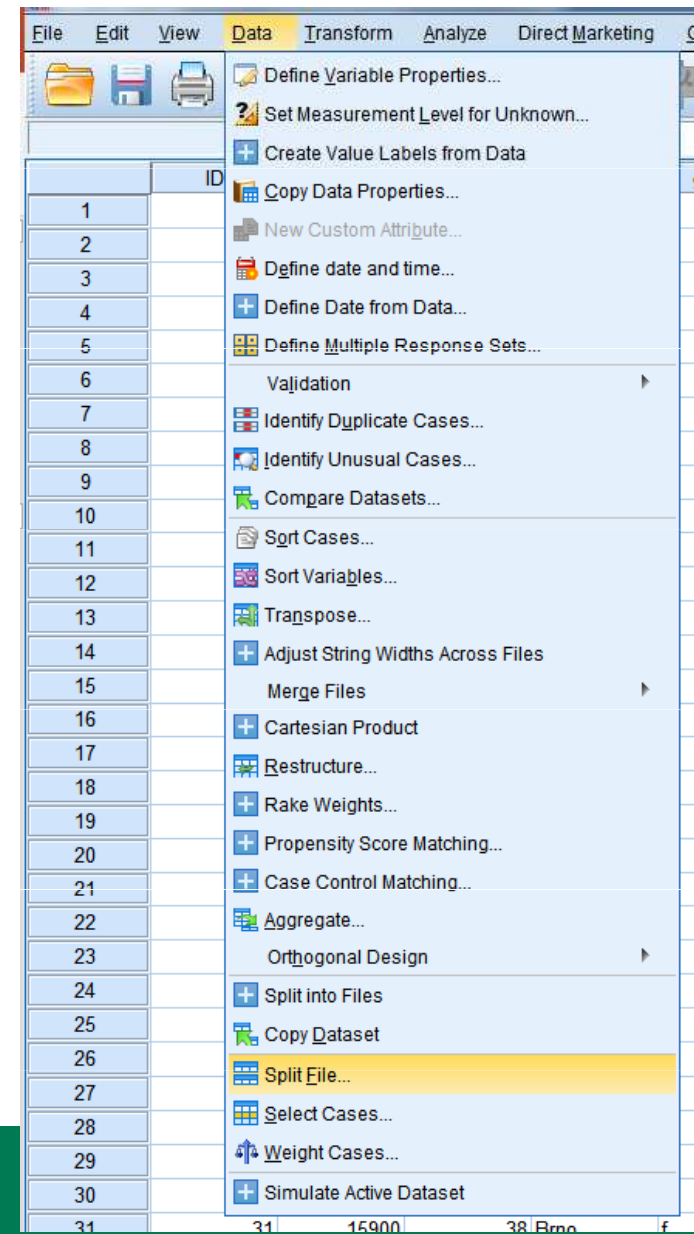
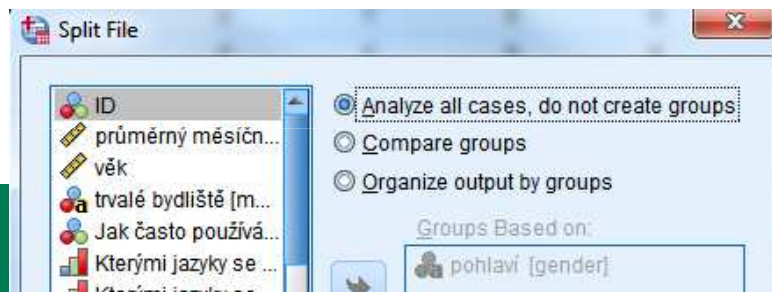
	ID	plat	věk	mesto	gender	in
1	1	22900	31	Vyškov	m	
2	2	28300	37	Vyškov	m	
3	3	24300	33	Vyškov	f	
4	4	27400	36	Brno	m	
5	5	29200	37	Vyškov	m	
6	6	13500	22	Brno	m	
7	7	32500	47	Brno	m	
8	8	90100	45	Brno	m	
9	9	17900	26	Brno	f	
10	10	18300	26	Brno	f	
11	11	28200	37	Vyškov	m	
12	12	25000	34	Brno	m	
13	13	148100	51	Brno	m	

- Pokud chceme opět pracovat s celým vzorkem, podmínku zrušíme přes příkaz
- Data → Select cases → All cases

Split file

- 1. Data → Split file
- 2. Compare groups
- 3. Vložíme proměnnou, podle které se má output rozdělit
 - Tj. Pokud podle pohlaví, tak R2
- Co to dělá:
 - Rozdělí vzorek podle zadané proměnné
 - V outputu pak vyjede výsledek pro sub-skupiny zvlášť

ZRUŠENÍ →



M U N I
F S S

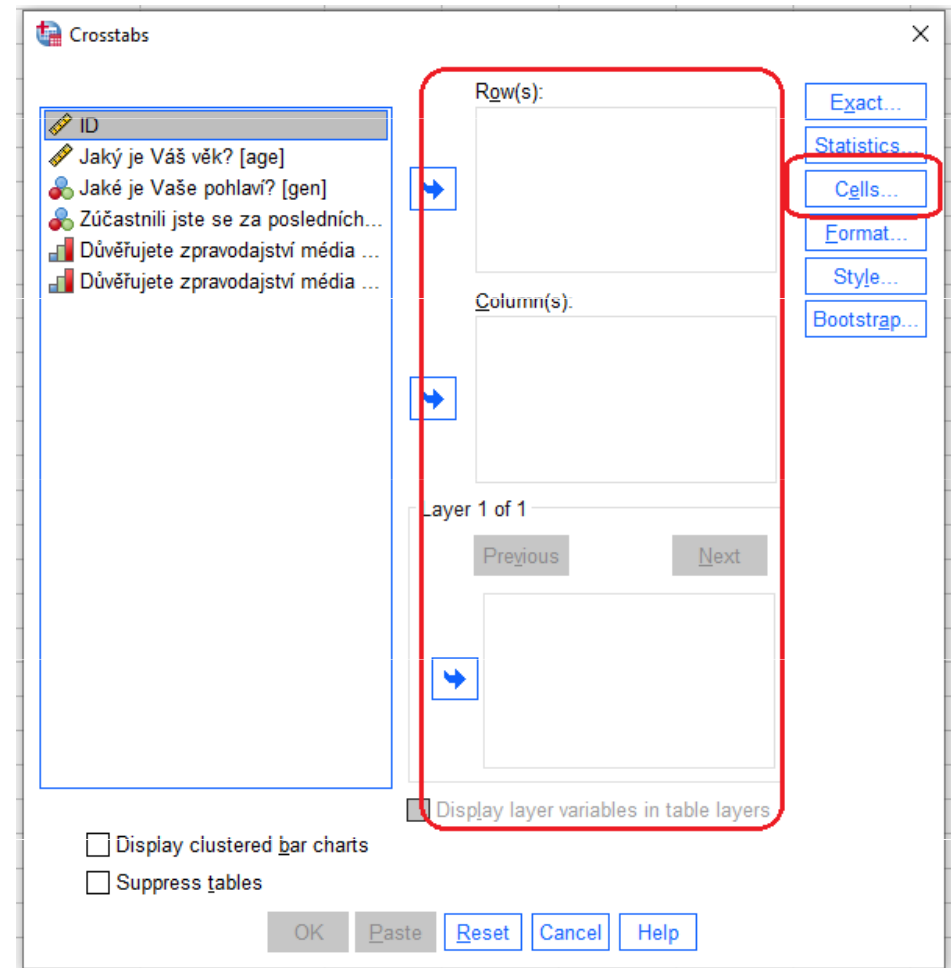
Základy bivariace

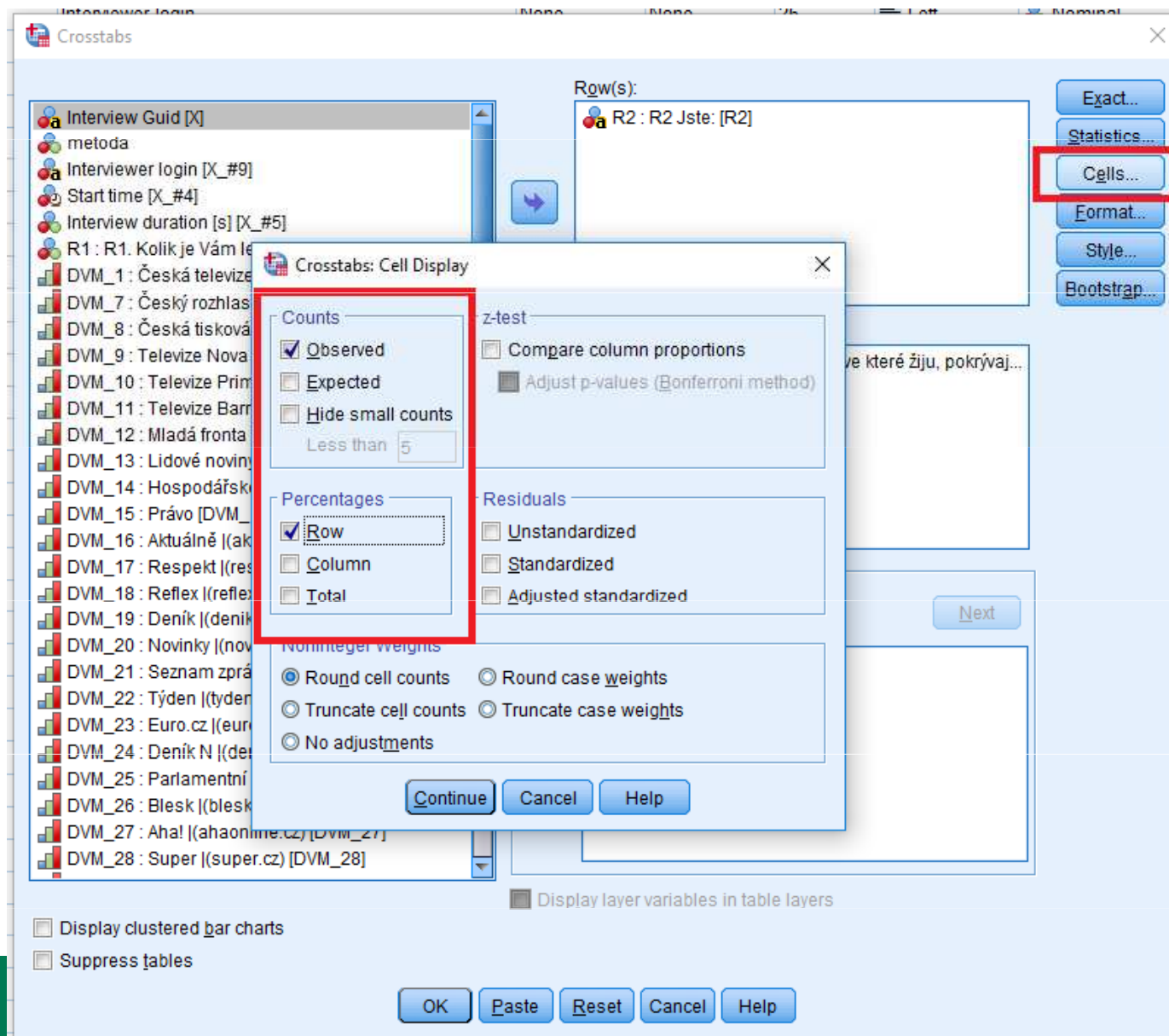
Kontingenční tabulka (crosstab)

- Zjistíme, jak na tom jsme s rozložením absolutních i relativních četností jedné (nebo více) proměnné s ohledem na další jednu (nebo více) proměnou
- Základ: Analyze – Descr. Stats. – Crosstab...
- Alternativně můžeme sestavit tabulku v Analyze – Tables – Custom tables
- Tímto způsobem pracujeme pouze s **kategorickými proměnnými**
 - Pokud máme například proměnnou měřící přesný věk v letech, tak si ji musíme rekódovat do věkových kategorií (viz minulá hodina)

Crosstabs

- Co nás zajímá?
 - **Rows** – řádky tabulky. Sem vložíme všechny položky, které chceme mít v řadách.
 - **Columns** – sloupce tabulky. Sem vložíme položky, které chceme mít ve sloupcích.
 - **Layer** – vrstva. Pokud chceme mít širší tabulku, která obsahuje více než 2 položky.
 - **Cells** – buňky. Zde si můžeme zvolit, co vše má tabulka obsahovat (tj, zda chceme výsledky v % nebo absolutních četnostech, např.).



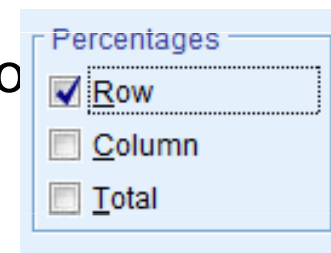


Counts – absolutní četnosti

- Poměrně málo používané, říkají nám málo o kontextu (Je 250 hodně? Málo?)
- Hodí se nám při kontrole kvality dat – na některou otázku měla odpovídat jen určitá část respondentů (podle pohlaví). Bylo tomu skutečně tak?

Relativní četnosti (%) – řádky, sloupce nebo total?

- Záleží na našem záměru a na tom, co dává smysl.
- „Kolik % lidí, kteří rozhodně souhlasí s tím, že příjmy jejich domácnosti pokrývají všechny jejich náklady zcela důvěřují Hospodářským novinám?”
- „Kolik % respondentů, kteří naprosto nedůvěřují HN absolutně nesouhlasí s tím, že jim příjmy vystačí na pokrytí jejich potřeb?”
- „Kolik % vzorku tvoří lidé, kteří cítí nejvyšší míru finančního bezpečí a zároveň považují HN za zcela důvěryhodné médium?”
- Málokdy vybereme více možností pro % - není to přehledné
- Výběr záleží na plánovaný způsob interpretace - více k tomuto následující lekci



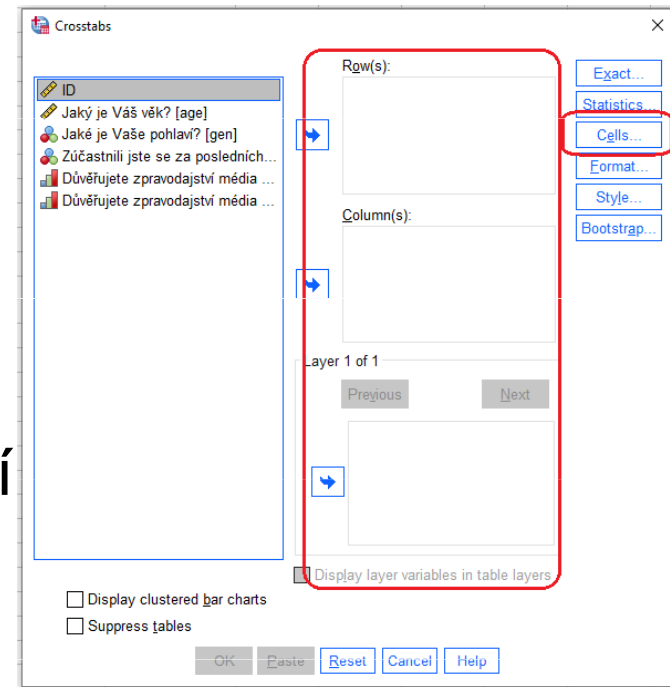
Jak s tím do grafu?

- Někdy si chceme rozložení dvou proměnných ukázat graficky. Máme 2 možnosti
 - Jedná se o dvě kategorické položky a/nebo chceme jen základní sloupcový graf
 - Zvolíme Display clustered bar charts v kartě Crosstabs
 - Nejedná se o dvě kategorické položky a/nebo chceme jiný druh grafu
 - Zvolíme Graphs – Chart builder



Praktické cvičení

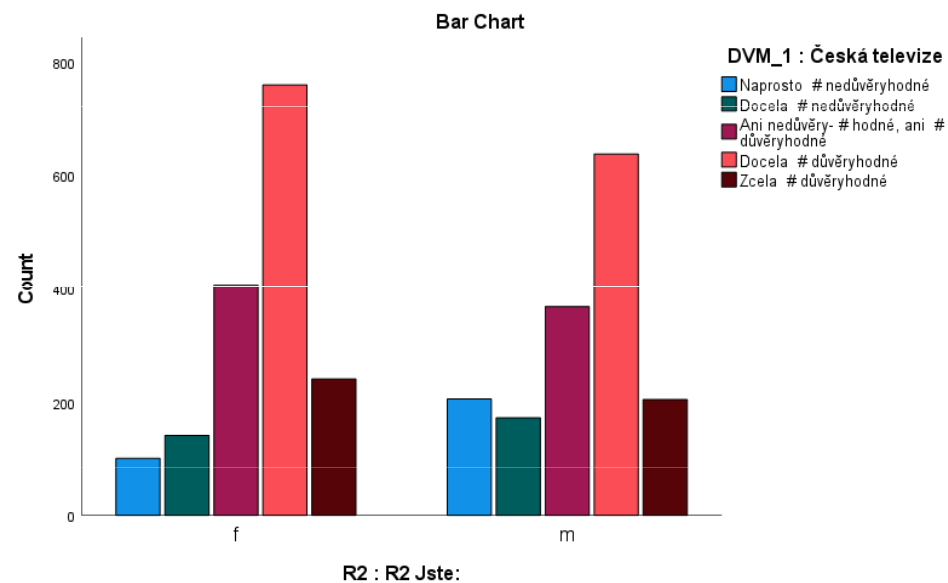
- Stáhněte si datový soubor z minulé hodiny - „transformace_lekce“
- Proveďte bivariaci položky pohlaví (R2) a důvěry ve zpravodajství České televice (DVM_1). Udělejte tak pomocí tabulky i graficky
- Otázka: Kolik % mužů považuje zpravodajství ČT za naprosto nedůvěryhodné?



R2 : R2 Jste: * DVM_1 : Česká televize Crosstabulation

% within R2 : R2 Jste:

		DVM_1 : Česká televize					Total
		Naprosto # nedůvěryhodné	Docela # nedůvěryhodné	Ani nedůvěry- # hodné, ani # důvěryhodné	Docela # důvěryhodné	Zcela # důvěryhodné	
R2 : R2 Jste:	f	6,1%	8,6%	24,7%	46,1%	14,6%	100,0%
	m	12,9%	10,9%	23,2%	40,2%	12,9%	100,0%
Total		9,5%	9,7%	23,9%	43,2%	13,7%	100,0%



Praktické cvičení vol. 2

- Proveďte bivariaci stejných položek, tentokrát ale jen pomocí tabulky.
 - Kolik % z těch, kteří zcela nedůvěřují zpravodajství ČT, jsou ženy?

Praktické cvičení vol. 2

- V předchozí části jsme zadávali procenta jako **Rows**, nyní potřebujeme **Columns**

R2 : R2 Jste: * DVM_1 : Česká televize Crosstabulation

% within DVM_1 : Česká televize

DVM_1 : Česká televize

	Naprosto # nedůvěryhodn é	Docela # nedůvěryhodn é	Ani nedůvěry- # hodné, ani # důvěryhodné	Docela # důvěryhodné	Zcela # důvěryhodné	Total
R2 : R2 Jste: f	33,0%	45,0%	52,6%	54,4%	54,1%	51,0%
m	67,0%	55,0%	47,4%	45,6%	45,9%	49,0%
Total	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Praktické cvičení vol. 3

- Krok 1: Vytvořte si věkové kategorie pomocí kvartilů (transformace podle způsoby z předchozí hodiny)
- Krok 2: Vyberte si jednu z kategorických položek měřící důvěru ke zpravodajství některého média (DVM_XX)
- Krok 3: Proveďte tabulkovou bivariaci této položky a Vaší nové ordinální položky věku (vyberte % podle věku – pokud věk máte v rows, zadejte % podle ROW)
- Krok 4: Zamyslete se nad rozložením této tabulky. Co Vám ukazuje? Co Vás překvapuje?