

### Cvičení 3.: Výpočet číselných charakteristik jednorozměrného souboru

**Úkol 1.:** U 100 náhodně vybraných domácností byl zjišťován způsob zásobování bramborami (znak X, varianty 1 = vlastní sklep, 2 = jinde, 3 = nákup) a bydliště (znak Y, varianty 1 = velké město, 2 = malé město, 3 = vesnice).

způsob zásobování	bydliště		
	velké město	malé město	vesnice
vlastní sklep	13	15	14
jinde	11	7	2
nákup	19	9	10

Pro oba znaky určete modus.

**Návod:** Otevřeme datový soubor brambory.sta se třemi proměnnými X, Y, četnost a devíti případy. Proměnná X obsahuje varianty způsobu zásobování, proměnná Y varianty bydliště a proměnná četnost obsahuje odpovídající simultánní absolutní četnosti dvojic variant  $(x_{[j]}, y_{[k]})$ ,  $j = 1, 2, 3$ ,  $k = 1, 2, 3$ .

Výpočet modu: Statistika – Základní statistiky/tabulky – Popisné statistiky – OK – klikneme na tlačítko se závažím – zaškrtneme Stav zapnuto, vybereme proměnnou vah četnost – OK – Proměnné X, Y – OK – Detailní výsledky – zaškrtneme Modus.

Proměnná	Popisné statistiky (brambory)	
	Modus	Četnost modu
X	1,000000	42
Y	1,000000	43

Proměnná X má modus 1, tj. nejvíce domácností skladuje brambory ve vlastním sklepě a proměnná Y má také modus 1, tj. nejvíce domácností bydlí ve velkém městě.

**Úkol 2.:** Otevřete datový soubor znamky.sta.

Pro známky z matematiky a angličtiny vypočítejte medián, dolní a horní kvartil, kvartilovou odchylku, a to pro všechny studenty, pak zvlášť pro muže a zvlášť pro ženy.

**Návod:**

Statistika – Základní statistiky/tabulky – Popisné statistiky – OK – Proměnné X, Y – OK – Detailní výsledky - zaškrtneme Medián, Dolní & horní kvartily, Kvartil. rozpětí – Výpočet.

Proměnná	Popisné statistiky (znamky)			
	Medián	Spodní kvartil	Horní kvartil	Kvartilové rozpětí
X	2,500000	1,000000	4,000000	3,000000
Y	3,000000	2,000000	3,500000	1,500000

Interpretace např. dolního kvartilu známek z angličtiny: aspoň čtvrtina studentů má z angličtiny nejhůř za dvě.

Počítáme-li tyto charakteristiky pro ženy (resp. pro muže), použijeme filtr: tlačítko Select Cases – Zapnout filtr – včetně případů – některé, vybrané pomocí výrazu  $Z=0$  (resp.  $Z=1$ ).

Pro ženy:

Proměnná	Popisné statistiky (znamky.sta) Zhrnout podmínku: z=0			
	Medián	Dolní kvartil	Horní kvartil	Kvartilové rozpětí
X	1,500000	1,000000	3,000000	2,000000
Y	2,000000	1,000000	4,000000	3,000000

Pro muže:

Proměnná	Popisné statistiky (znamky.sta) Zhrnout podmínku: z=1			
	Medián	Dolní kvartil	Horní kvartil	Kvartilové rozpětí
X	4,000000	2,000000	4,000000	2,000000
Y	3,000000	3,000000	3,000000	0,000000

Vidíme, že v matematice i angličtině ženy dosahují lepších výsledků. Zajímavé jsou výsledky z angličtiny pro muže: zde splývá medián s oběma kvartily.

**Úkol 3.:** Otevřete datový soubor ocel.sta.

Pro mez plasticity a mez pevnosti vypočítejte aritmetický průměr, směrodatnou odchylku, rozptyl a koeficient variace. Výsledky porovnejte s údaji ve skriptech Popisná statistika (viz str. 30).

**Návod:**

Statistiky – Základní statistiky/tabulky – Popisné statistiky – OK – Proměnné X, Y – OK – Detailní výsledky - zaškrtneme Průměr, Směrodat. odchylka, Rozptyl, Variační koeficient – Výsledky.

Proměnná	Popisné statistiky (ocel.sta)			
	Průměr	Rozptyl	Sm.odch.	Koef.prom.
X	95,8833	1070,240	32,71453	34,11910
Y	114,4000	1075,125	32,78911	28,66181

Vysvětlení: Rozptyl a směrodatná odchylka vyjdou ve STATISTICE jinak než ve skriptech, protože STATISTICA ve vzorci pro výpočet rozptylu nepoužívá  $1/n$ , ale  $1/(n-1)$ . Koeficient variace (v tabulce označený jako Koef. prom.) je udán v procentech.

Pokud bychom chtěli získat stejné výsledky jako ve skriptech, museli bychom přepočítat získaný rozptyl, směrodatnou odchylku a koeficient variace: k výstupní tabulce přidáme tři nové proměnné, které vložíme za proměnnou v4. V první proměnné bude přepočítaný rozptyl, ve druhé přepočítaná směrodatná odchylka a ve třetí přepočítaný koeficient variace. Do Dlouhého jména první proměnné napíšeme  $=\sqrt{2*59/60}$ , do Dlouhého jména druhé proměnné napíšeme  $=\sqrt{v5}$  a do Dlouhého jména třetí proměnné napíšeme  $=100*v6/v1$ . Dostaneme tabulku:

Proměnná	Popisné statistiky (ocel.sta)						
	Průměr	Rozptyl	Sm.odch.	Koef.prom.	NProm1	NProm2	NProm3
X	95,8833	1070,240	32,71453	34,11910	1052,403	32,44076	33,83358
Y	114,4000	1075,125	32,78911	28,66181	1057,207	32,51471	28,42195

## Úkol k samostatnému řešení

Máme k dispozici údaje o rozměrech lebek staroegyptské populace. Jedná se o 216 mužů a 109 žen.

Znak X ... největší délka mozkovny v mm (tj. přímá vzdálenost kranio-metrických bodů glabella a opisthocranion)

Znak Y ... největší šířka mozkovny v mm (tj. přímá vzdálenost kranio-metrických bodů euryon dx a euryon sin)

Znak Z ... pohlaví osoby (1 ... muž, 0 ... žena)

Údaje jsou uloženy v souboru lebky.sta.

a) Pro největší délku a největší šířku mozkovny mužů vypočtete aritmetický průměr, směrodatnou odchylku, rozptyl, koeficient variace, šikmost a špičatost.

b) Pro největší délku a největší šířku mozkovny žen vypočtete aritmetický průměr, směrodatnou odchylku, rozptyl, koeficient variace, šikmost a špičatost.

Výsledky:

Ad a)

Proměnná	Popisné statistiky (lebky.sta)						
	Zhrnout podmínku: z=1						
	N platných	Průměr	Rozptyl	Sm. odch.	Koef.	Šikmost	Špičatost
X	216	182,032	40,578	6,370	3,499	-0,055	-0,451
Y	216	137,185	23,169	4,813	3,509	0,085	-0,249

Ad b)

Proměnná	Popisné statistiky (lebky.sta)						
	Zhrnout podmínku: z=0						
	N platných	Průměr	Rozptyl	Sm. odch.	Koef.	Šikmost	Špičatost
X	109	174,532	38,677	6,219	3,563	-0,039	-0,163
Y	109	134,147	22,052	4,696	3,501	0,030	0,567