**Na začátku druhého semestru se statistickou analýzou dat doporučujeme osvěžit si základní postupy popisné analýzy dat, jež budou předpokladem k pokračování a k možnosti osvojit si pokročilejší techniky.**

Ujistěte se, že rutinně a bez problému dokážete rozhodnout o povaze analyzovaných proměnných, jejich statusu a smyslu v analýze (úroveň měření, kategorizovaná či spojitá povaha, co je závisle proměnná, nezávisle proměnná, kontrolní proměnná, jaký model vztahu lze očekávat).

Rozložení hodnot jedné proměnné – ujistěte se, že víte, jak získat základní přehled o charakteristikách rozložení hodnot proměnných, které jsou spojité či kategorizované (procedury frequencies, descriptives, explore).

Hledání souvislostí mezi proměnnými – připomeňte si, jak postupovat, chcete-li zjistit, zda jedna věc statisticky souvisí s druhou, ověřte si, zda máte jistotu v rozhodování, jakým způsobem analyzovat různé kombinace proměnných (procedury crosstabs pro kategorizované proměnné, means pro kombinaci kategorizované a spojité proměnné, korelace pro spojité proměnné).

Připomeňte si logiku korelací a asociací jako snahy vyjádřit statistický vztah mezi proměnnými jedním koeficientem – co koneficienty znamenají, kdy se hodí a jaké jsou jejich limity. V mnoha případech lze tento přístup aplikovat paralelně s klasickou popisnou analýzou prostřednictvím crosstabs, či means, jindy jde o jedinou možnost.

Ujistěte se, že máte představu o tom, jak pokračovat v rozpracování vztahu mezi dvěma proměnnými, když chceme zjistit, jak do něj vstupuje další faktor (např. intervenující proměnná). Jak konkrétně rozšířit analýzu v crosstabs a means o další vrstvu, jak zpracovat parciální korelaci.

Ujistěte se, že zvládáte základní operace s daty – vytvoření nových proměnných formou součtového indexu, rekódování, logické operace, aritmetické operace…

**K procvičení můžete použít následující příklady (tato cvičení nejsou povinná, neodevzdávají se, nehodnotí se, je čistě na vás, jestli si je chcete projít):**

**Prozkoumejte charakteristiky prvorodiček v datech ČSÚ z roku 1993 – 2012 PRACUJTE SE SOUBOREM „CSU\_rodice.sav“**

[*https://is.muni.cz/auth/el/fss/podzim2023/SOCb1019/um/datove\_soubory\_a\_dotazniky/CSU\_rodice.sav*](https://is.muni.cz/auth/el/fss/podzim2023/SOCb1019/um/datove_soubory_a_dotazniky/CSU_rodice.sav)

*věk matky* [d18] podle *nejvyšší vzdělání matky* [d14]

1. V jakém věku měly ženy děti v daném období?
2. Existují podle vzdělání odlišnosti? Která vzdělanostní skupina má děti nejpozději a v jakém věku průměrně?
3. Ukažte vhodný graf, v němž budou patrné nejen střední hodnoty, ale také rozdělení věku při porodu prvního dítěte u různých vzdělanostních skupin
4. Jak se rozdíly ve věku prvorodiček podle vzdělání vyvíjely mezi lety 1993 – 2012?

**V sociologii rodiny se často hovoří o vzdělanostní homogamii – tendenci vybírat si partnera vzdělanostně podobného. Nalezneme stopu těchto tendencí i u párů, které jsou v datech ČSÚ zachyceny jako rodiče?** Použijte proměnné *nejvyšší vzdělání otce* [d13] a *nejvyšší vzdělání matky* [d14].

1. Jak je v souboru rozloženo vzdělání u mužů a u žen, která skupina je nejpočetnější?
2. Jaký podíl z vysokoškolaček má partnera s vysokoškolským vzděláním?
3. Jaká kombinace vzdělání je mezi danými páry nejčastější?
4. Vyjádřete těsnost vztahu mezi vzděláním otce a matky příslušným koeficientem a zhodnoťte, zda tendence k vzdělanostní homogamii v našich datech existuje, nebo ne.
5. Vytvořte tři věkové kategorie žen podle *věk matky* [d18] a to konkrétně do 20 let 21-30 a 31+ Ukažte rozložení nové proměnné
6. Srovnejte, jak vypadá vzdělanostní homogamie v rámci těchto věkových kategorií.

**PRACUJTE SE SOUBOREM „bydleni\_Brno.sav“**

<https://is.muni.cz/auth/el/fss/podzim2023/SOCb1019/um/datove_soubory_a_dotazniky/bydleni_Brno.sav>

a) Vypočtěte proměnnou, která udává počet obyvatel na jednu místnost v bytě, tedy podíl

**[b4.kat]** *b4.kat - počet osob bydlících v bytě / domě* a **[B6]** *b6 : b6. Kolik obytných místností (pokojů) má váš byt / dům (mimo samostatnou kuchyň,koupelnu, záchod, předsíň, spíž, sklep a ostatní příslušenství)?*

popište výsledek – kolik mají respondenti průměrně k dispozici místností? Lze říci, že spíše lidé žijí v bytech, kde je více místností než lidí?

b) Souvisí tato proměnná se spokojeností s bydlením? **[INDEXSPOK\_INDEX]** Popište jak.

c) intervenuje nějak věk do této korelace? **[AGE]** *age : R7. Kolik máte let?* Vysvětlete.

d) Popište v jak velikých bytech lidé v Brně bydlí. Použijte proměnnou **[B7]** *b7 : b7. Jaká je celková podlahová plocha Vašeho bytu / domu? Uveďte v metrech čtverečních.*

Uveďte relevantní statistiky a případně vhodné grafy.

e) Jaká je největší velikost bytu, která se ve vzorku vyskytuje.

f) Jaký rozměr podlahové plochy odděluje 10 % respondentů s největšími byty?