**Příprava, validace a základní popis dat**

***Datové soubory:***

* **RA\_BASELINE.sav** (základní charakteristiky pacientů s revmatoidní artritidou v době zahájení léčby biologickými preparáty)
* **RA\_VISITS.sav** (parametry aktivity onemocnění měřené opakovaně v průběhu biologické léčby)
* **RA\_STOP.sav** (záznamy o všech přerušení/ukončení léčby)

***Cíle:***

1. Vyhodnocení poklesu aktivity onemocnění (DAS28) po jednom roce biologické léčby.
2. Vyhodnocení setrvání na léčbě.

***Úkoly (syntax celého postupu práce si ukládejte):***

1. **Načtěte** všechny tři datové soubory a vhodně si je pojmenujte (*File -> Open -> Data*).

*Pozn.:* většinu drobných úprav (např. právě pojmenování datasetu) je jednodušší provádět až v rámci vygenerovaného kódu.

1. Připravte **datový soubor** pro vyhodnocení daných cílů.
   1. Ze souboru RA\_VISITS zjistěte hodnoty parametrů aktivity nemoci po 12 měsících léčby (*Data -> Select Cases*)
   2. Ze souboru RA\_STOP určete datum prvního přerušení nebo ukončení léčby (*Data -> Aggregate*)
   3. Do souboru RA\_BASELINE doplňte data odvozená v předchozích bodech (*Data -> Merge Files -> Add Variables*)

*Pozn. 1:* před spojováním datových souboru je vhodné mít data v obou tabulkách seřazena podle klíčů (*Data -> Sort Cases*).

*Pozn.* 2: čtěte a upravujte si syntax - některé části automaticky generovaného kódu jsou redundantní (např. při *Select Cases*).

*Pozn.* 3: využijte příkazů „*Rename variables XXX=YYY“* a „*Delete variables XXX“* ke zpřehlednění datového souboru.

1. Vypočítejte další potřebné **odvozené proměnné**.
   1. Věk při zahájení léčby (*Transform -> Compute Variable*)
   2. Věk při diagnóze RA (*Transform -> Compute Variable*)
   3. Délka trvání nemoci při zahájení léčby (*Transform -> Compute Variable*)
   4. DAS28 při zahájení léčby a po roce léčby (*Transform -> Compute Variable*)
   5. Změna DAS28 po roce léčby (*Transform -> Compute Variable)*
   6. Aktivita nemoci dle DAS28 při zahájení léčby a po roce léčby: remise/nízká/střední/vysoká (*Transform -> Recode into Different Variables*)
   7. Doba setrvání na léčbě, přičemž datum uzavření dat je 27. 11. 2017 (*Transform -> Compute Variable*)
   8. Indikátor cenzorování pro ukončení léčby (*Transform -> Compute Variable)*

*Pozn. 1:* veškeré textové argumenty musí být zapsány v uvozovkách.

*Pozn. 2*: kód, který už znáte, pište rovnou.

*Pozn. 3:* vzorec pro výpočet DAS28 a hranice pro stanovení aktivity onemocnění jsou uvedeny na konci dokumentu (potřebnou funkci pro jakékoli výpočty včetně vysvětlivek lze zpravidla najít v kategorizované nabídce menu *Transform -> Compute Variable*).

*Pozn. 4:* většinu *recode* funkcí lze nahradit pomocí *Transform -> Compute Variable* s využitím podmínek.

1. Definujte pro oba cíle **validní kohortu** pacientů (pomocí binárních indikátorů 1/missing).
   1. Najděte a prověřte chybné nebo duplicitní záznamy (*Data -> Identify Duplicate Cases*)
   2. Validujte hodnoty jednotlivých proměnných (*Data -> Sort Cases*)
   3. Kohorta pro cíl 1): diagnóza RA po 16 letech věku, vyplněnost pohlaví, věku, trvání nemoci, DAS28 při zahájení a po roce léčby (*Transform -> Compute Variable)*
   4. Kohorta pro cíl 2): diagnóza RA po 16 letech věku, vyplněnost pohlaví, věku, trvání nemoci, DAS28 při zahájení léčby (*Transform -> Compute Variable)*

*Pozn. 1:* filtry definující různé překrývající se kohorty pacientů je vhodné vyjádřit formou indikátorů *1/missing* případně *1/0*.

*Pozn. 2*: pro rychlou kontrolu rozsahu a validity hodnot zpravidla stačí seřazení nebo základní deskripce dané proměnné v nabídce pod pravým kliknutím myši.

1. Navrhněte a proveďte **základní popisnou analýzu** vstupních charakteristik a srovnejte tři kohorty (*Analyze -> Custom Tables*).
   1. všichni pacienti
   2. kohorta pro cíl 1)
   3. kohorta pro cíl 2)

*Otázka:* proč je důležité srovnat analyzované kohorty a jaké závěry ze srovnání plynou?

1. Navrhněte a proveďte **vyhodnocení změny aktivity nemoci** (DAS28) po roce léčby (*Analyze -> Custom Tables*).

*Pozn.:* k průběžné deskriptivní analýze lze využít všechny různé možnosti SPSS v záložkách *Analyze* a *Graphs* (pro tvorbu finálních výstupů je zpravidla vhodné kombinovat výpočetní schopnosti SPSS s grafickými možnostmi jiných programů, např. MS Office).

*Otázka:* lze na základě výsledků konstatovat nějaký závěr ohledně účinnosti léčby?

1. Vykreslete **křivku setrvání na léčbě** (*Analyze -> Survival -> Kaplan-Meier*).

