

Manuál pro analýzu dat v softwaru SPSS

Software SPSS je produkt IBM (<http://www.ibm.com/analytics/us/en/technology/spss/>). SPSS je dostupné v rámci MU z <https://inet.muni.cz/auth/login> (login stejný jako do www.is.muni.cz, seznam dostupných softwarů lze najít v oddílu Provozní služby -> Software).

Načtení datového souboru

File -> Open -> Data -> vybrat datový soubor -> pokud načítáme excelovský soubor, nezapomenout změnit Files of type na Excel (*.xls, *.xlsx, *.xlsm) -> Open -> ve Worksheet vybrat požadovaný list excelovského souboru -> nechat zatržené Read variable names from first row of data -> OK

Uložení datového souboru

File -> Save as -> zadat název souboru -> Save

1. Vizualizace dat

Vytváření grafů pomocí záložky Graphs.

Koláčový graf

Graphs -> Legacy Dialogs -> Pie -> Summaries for groups of cases -> Define -> Define Slices by: zvolit proměnnou (např. Gender) -> OK

Pozn.: Po dvojím kliknutí na graf se nám ukáže okno Chart Editoru, kde lze přidat název grafu (Options -> Title), měnit barvu či ohraničení grafu a upravovat či nastavovat další parametry

Sloupcový graf (na ose y počty lidí)

Graphs -> Legacy Dialogs -> Bar -> Nechat zatržené Simple a Summaries for groups of cases -> Define -> Category Axis: zvolit proměnnou (např. Group) -> OK

Sloupcový graf (na ose y procenta)

Graphs -> Legacy Dialogs -> Bar -> Nechat zatržené Simple a Summaries for groups of cases -> Define -> Category Axis: zvolit proměnnou (např. Group) -> místo N of cases zatrhnout % of cases -> OK

Histogram (na ose y počty lidí)

Graphs -> Legacy Dialogs -> Histogram -> Variable: zvolit proměnnou (např. Age) -> OK

Pozn.: Dále lze zatrhnout Display normal curve a lze vykreslit histogramy podle kategoriální proměnné tím, že se kategoriální proměnná přetáhne do Rows nebo Columns

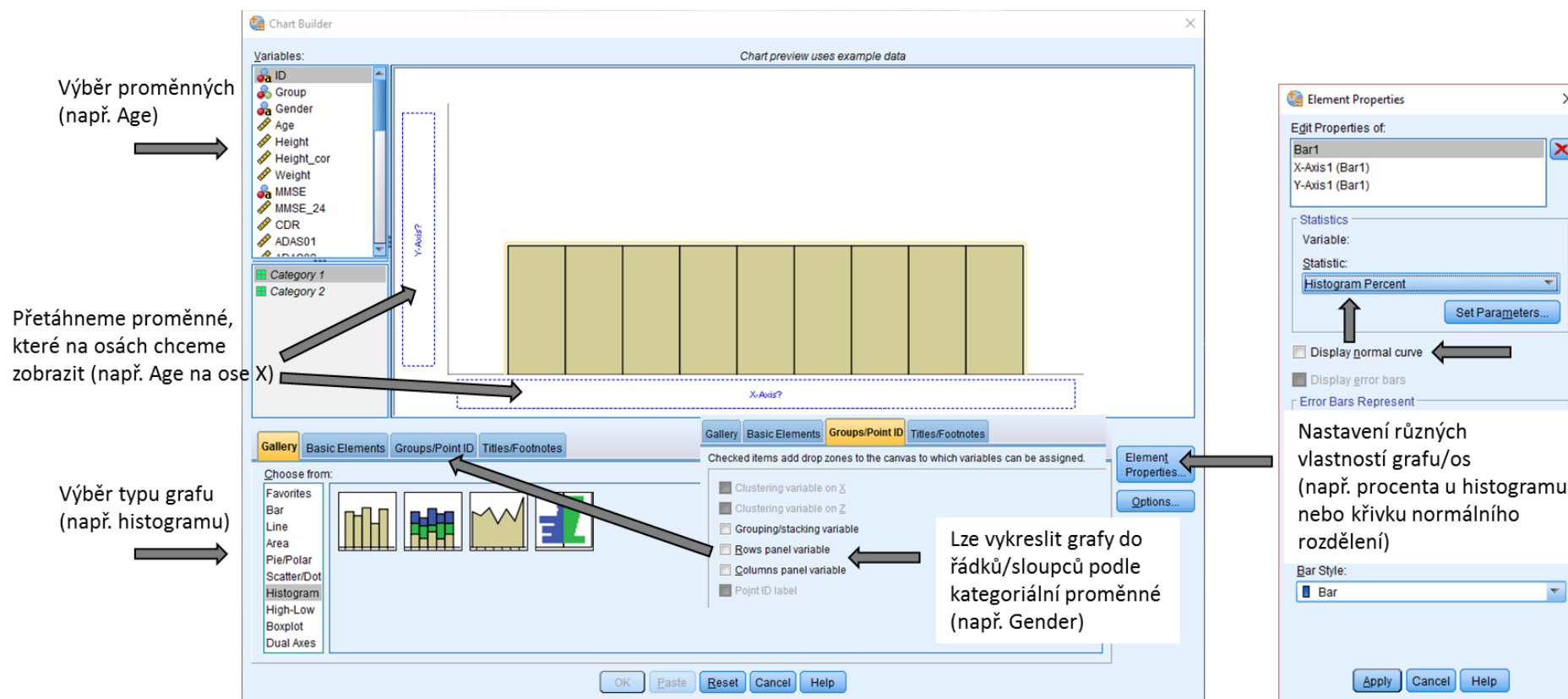
Histogram (na ose y procenta)

Graphs -> Chart Builder -> OK -> Gallery: Histogram -> přetáhnout požadovaný graf do okna výše -> požadovanou proměnnou (např. Age) přetáhnout do pole X-Axis? pod grafem (viz. Obr. 1) -> Element Properties -> Statistic: Histogram Percent -> případně zaškrtnout Display normal curve -> Apply -> Close -> OK

Pozn.: Dále lze v Chart Builder vykreslit histogramy podle kategoriální proměnné kliknutím na záložku Groups/Point ID, zatrhnutím Rows panel variable nebo Columns panel variable a dále přetáhnutím vybrané kategoriální proměnné do políčka Panel?

Krabicový graf

Graphs -> Legacy Dialogs -> Boxplot -> Nechat zatržené Simple a zatrhnout Summaries of separate variables -> Define -> Boxes Represent: zvolit proměnnou (např. Age) -> OK



Obr. 1: Dialogové okno Chart editoru pro vytvoření grafu

2. Příprava dat pro analýzu

Nastavení formátu u MMSE na numerickou proměnnou

1. způsob: Přejít na Variable View a u vybrané proměnné ve sloupci Type rozkliknout tlačítko se třemi tečkami -> zatrhnout Numeric -> OK -> u vybrané proměnné přejít do sloupce Measure -> vybrat Scale
2. způsob: Data -> Define Variable Properties -> Variables to Scan: zvolit proměnnou (MMSE) -> Continue -> Type: vybrat Numeric -> Measurement Level: vybrat Scale -> OK

Nastavení formátu u scan_date na datum

1. způsob: Přejít na Variable View a kliknout u vybrané proměnné do sloupce Type -> zatrhnout Date -> v pravém okně z nabídky vybrat formát (např. dd-mmm-yyyy)
2. způsob: Data -> Define Variable Properties -> Variables to Scan: zvolit proměnnou (scan_date) -> Continue -> Type: vybrat Date a formát (např. dd-mmm-yyyy) -> Measurement Level: vybrat např. Scale nebo Ordinal -> OK

Úprava počtu desetinných míst u Weight

Přejít na Variable View a kliknout a nastavit Decimals např. na hodnotu 1

Identifikace duplikací

Data -> Identify Duplicate Cases -> Define matching cases by: vybrat proměnnou (ID) -> Sort within matching groups by: vybrat proměnnou (např. scan_date) -> Sort: Ascending -> nechat zaškrtnuté Indicator of primary cases a Last case in each group is primary -> lze ponechat název Name: PrimaryLast

Pozn.: Výsledkem je proměnná PrimaryLast, která primární záznamy označuje hodnotou 1 a duplikátní hodnotou 0. Je patrné, že jako primární je definován záznam s nejnovějším datem. Pokud chceme, aby byl jako primární záznam označen ten s nejstarším datem, tak zaškrtneme First case in each group is primary -> lze ponechat název Name: PrimaryFirst

Odstranění duplikací:

Data -> Select cases -> zaškrtnout Use filter variable -> do okénka přetáhnout proměnnou identifikující primární a duplikátní záznamy (např. PrimaryFirst) -> zaškrtnout Delete unselected cases -> OK

Odstranění chybějících a chybných hodnot

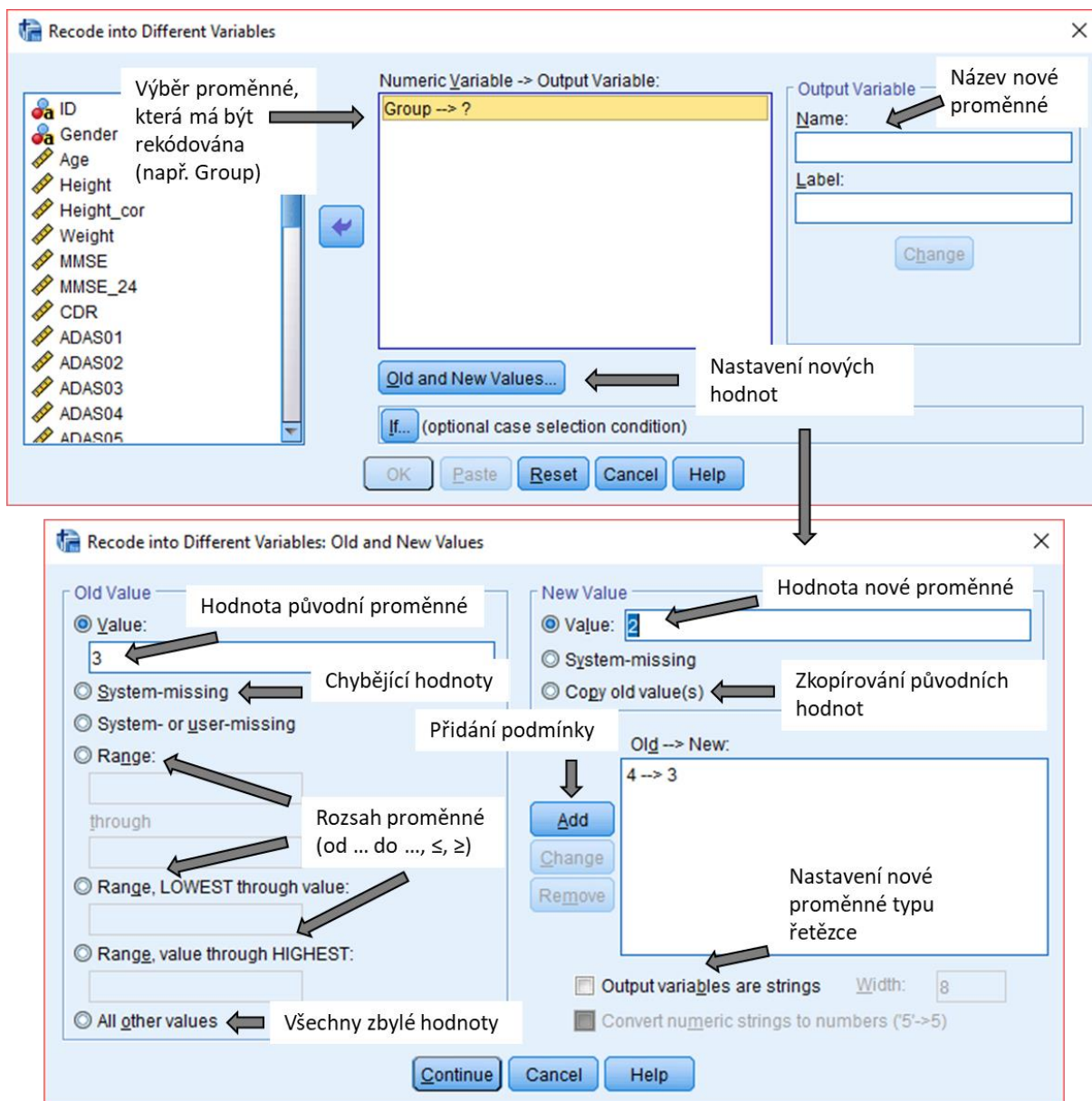
Data -> Select cases -> zatrhnout If condition is satisfied -> kliknout na tlačítko If -> do okénka zapsat podmínky (např. Age < 110 & MISSING(Age)=0 & MISSING(Weight)=0) -> Continue -> zaškrtnout Delete unselected cases -> OK

Rekódování proměnné Gender, aby obsahovala pouze hodnoty F a M

1. způsob – ručně: Najít si řádek s hodnotou FF a hodnotu FF přepsat na F
2. způsob – vytvořením nové proměnné: Transform -> Recode into Different Variables -> do okénka Input Variable přetáhnout proměnnou Gender -> Name: zadat název nové proměnné (např. Gender_rek) -> Change -> kliknout na tlačítko Old and New Values -> v oddílu Old Value zadat Value: FF -> zatrhnout Output variables are strings -> v oddílu New Value zadat Value: F -> Add -> v oddílu Old Value zatrhnout All other values -> v oddílu New Value zatrhnout Copy old values -> Add -> Continue -> OK
3. způsob – pozměněním původní proměnné: Transform -> Recode into Same Variables -> do okna Variables přetáhnout proměnnou Gender -> Old and New Values -> v sekci Old Value vepsat Value: FF -> v sekci New Value vepsat Value: F -> Add -> Continue -> OK

Rekódování proměnné Group, aby obsahovala pouze hodnoty 1 (CN), 2 (MCI) a 3 (AD)

Transform -> Recode into Different Variables -> do okénka Input Variable přetáhnout proměnnou Group -> Name: zadat název nové proměnné (např. Group_3kat) -> Change -> Old and New Values -> v oddílu Old Value zadat Value: 4 -> v oddílu New Value zadat Value: 3 -> Add -> v oddílu Old Value zadat 3 -> v oddílu New Value zadat 2 -> Add -> v oddílu Old Value zatrhnout All other values -> v oddílu New Value zatrhnout Copy old values -> Add -> Continue -> OK



Obr. 2: Ukázka rekódování proměnné do jiné proměnné

Vytvoření textových popisků u kvalitativní proměnné

Přejít do Variable View -> kliknout u příslušné proměnné (Group_3kat) do okénka Value a rozkliknout tlačítko se třemi tečkami -> Value: 1, Label: CN -> Add -> Value: 2, Label: MCI -> Add -> Value: 3, Label: AD -> Add -> OK

Pozn. 1: Při zadávání textových popisků nepíšeme uvozovky, doplní se samy.

Pozn. 2: Hodnoty a testové popisky lze přepínat v hlavním panelu pomocí tlačítka.



3. Popisná sumarizace dat

Popisná sumarizace dat pomocí Analyze -> Descriptive Statistics

Popisná sumarizace kvalitativních dat – frekvenční tabulka

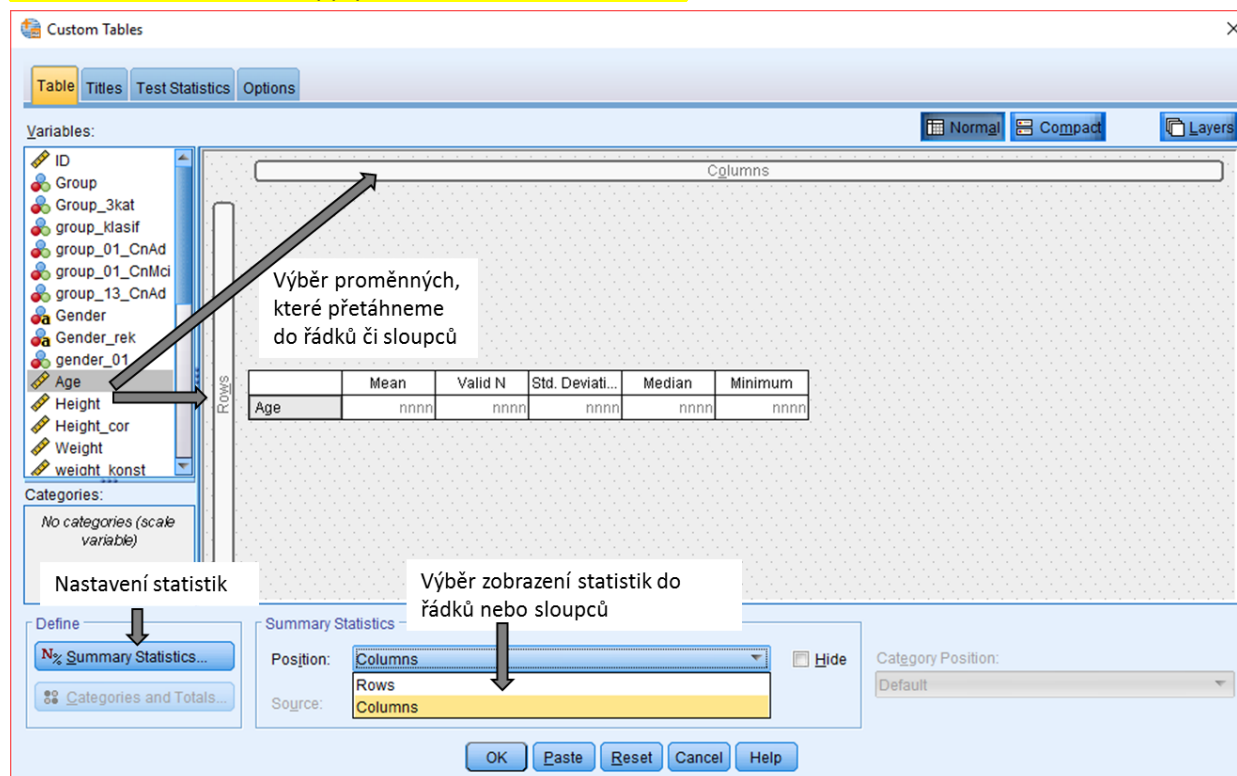
1. způsob: Analyze -> Descriptive Statistics -> Frequencies -> Variables: zvolit proměnnou (např. Group) -> OK
2. způsob: Kliknout pravým tlačítkem na název proměnné (např. Group) v Data View nebo Variable View -> Descriptive Statistics

Popisná sumarizace kvantitativních dat

1. způsob: Analyze -> Descriptive Statistics -> Descriptives -> Variables: zvolit proměnnou (např. Age) -> OK

Pozn.: Automaticky se vypisuje průměr, směrodatná odchylka, minimum a maximum, lze vybrat též rozptyl, rozsah či standardní chybu, pokud bychom vypsali medián, lze využít následující tři postupy.

2. způsob: Analyze -> Tables -> Custom Tables -> OK -> vybrané proměnné přetáhnout do pravého okna (do Rows nebo Columns) -> Summary Statistics -> zvolit např. Valid N, Mean, Std. Deviation, Median, Minimum, Maximum -> Apply to Selection -> Close -> OK



Obr. 3: Popisná sumarizace pomocí Custom Tables

3. způsob: Analyze -> Descriptive Statistics -> Explore -> Dependent List: vybrat proměnné -> zahrnout Statistics -> OK
4. způsob: Kliknout pravým tlačítkem na název proměnné (např. Age) v Data View nebo Variable View -> Descriptive Statistics

Popisná sumarizace – zapnutí filtru (vyfiltrování posledních 20 pacientů)

Data -> Select Cases -> zahrnout Based on time or case range -> Range -> vyplnit First Case: 813, Last Case: 833 -> Continue -> zahrnout Filter out unselected cases -> OK

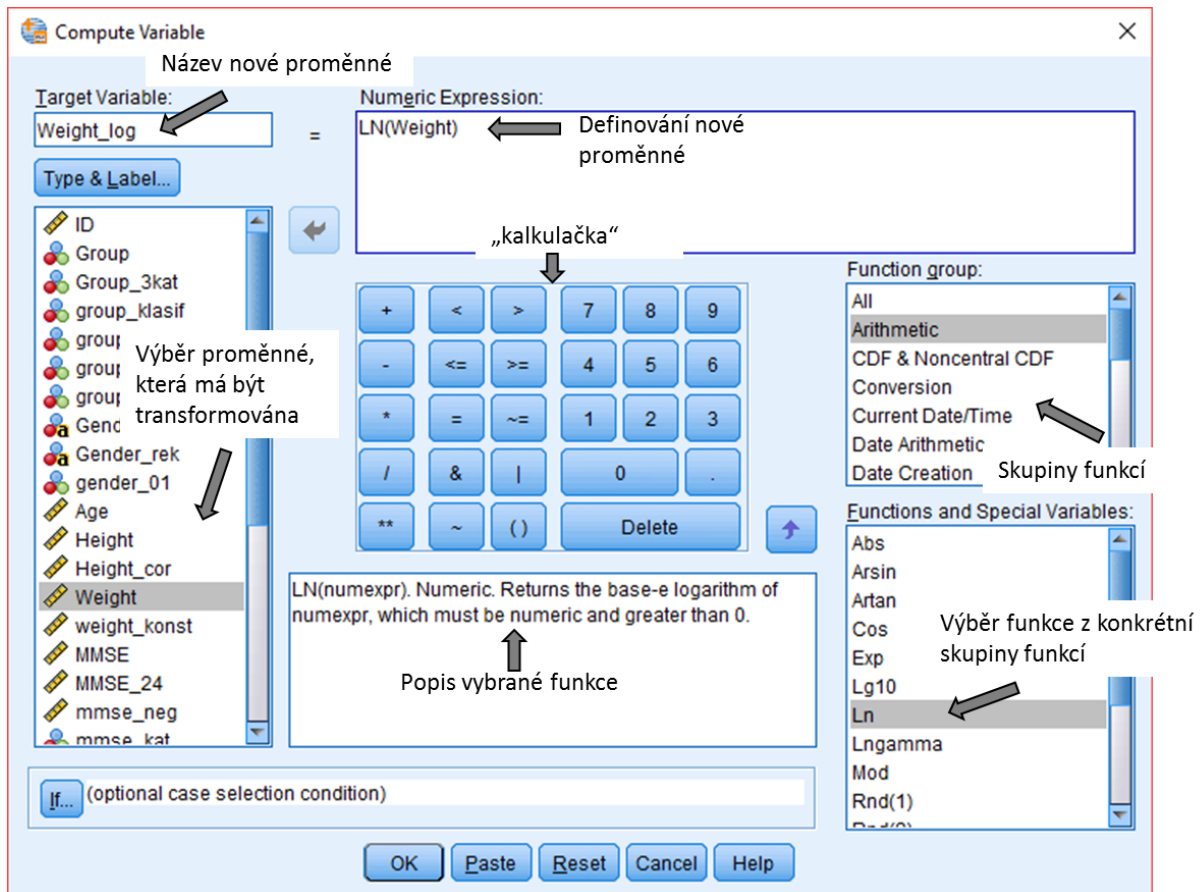
Analyze -> Descriptive Statistics -> Descriptives -> Variables: zvolit proměnnou (např. Age) -> OK

Pozn.: Nezapomenout poté vypnout filtr Data -> Select Cases -> All cases -> OK

4. Transformace dat

Logaritmická transformace

Transform -> Compute Variable -> Target Variable: *název nové proměnné (např. Weight_log)* -> Function group: Arithmetic -> Functions and Special Variables: *vybrat Ln a přetáhnout do okna Numeric Expression* -> *do argumentu funkce vložíme vybranou proměnnou (např. Weight)*



Obr. 4: Ukázka transformace proměnné na novou proměnnou pomocí Compute Variable

Standardizace dat

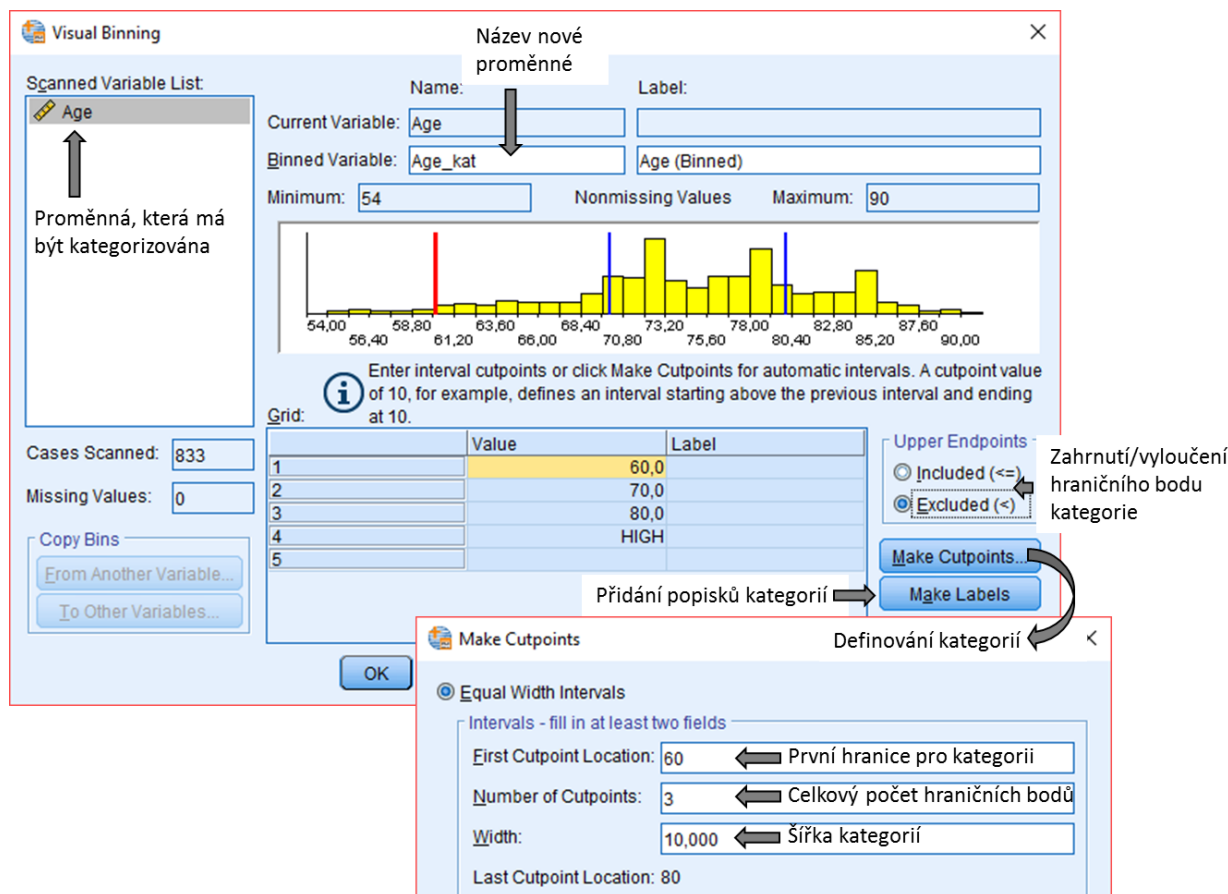
Analyze -> Descriptive Statistics -> Descriptives -> Variables: *vybrat proměnné* -> *zatrhnout Save standardized values as variables* -> OK

Centrování dat

Transform -> Compute Variable -> Target Variable: *zadat jméno nové proměnné (např. Height_centr)* -> Numeric Expression: *vybrat proměnnou, kterou chceme centrovat a přidat minus průměrná hodnota (např. Height-175.97)*

Kategorizace

Transform -> Visual Binning -> Variables to Bin: *vybrat proměnnou (např. Age)* -> Continue -> Binned Variable: *název nové proměnné (např. Age_kat)* -> Make Cutpoints -> *vyplnit First Cutpoint Location: 60, Number of Cutpoints: 3, Width: 10* -> Apply -> Upper Endpoints: *zatrhnout Excluded (<)* -> *kliknout na Make Labels* -> OK



Obr. 5: Ukázka kategorizace proměnné pomocí Visual Binning

5. Intervaly spolehlivosti

Výpočet intervalu spolehlivosti a střední chyby průměru (standard error)

Analyze -> Tables -> Custom Tables -> OK -> do okna vpravo přetáhnout vybrané proměnné (buď do Rows, nebo Columns) -> Summary Statistics -> zvolit Mean, Lower CL for Mean, Upper CL for Mean, Standard Error of Mean -> Apply to Selection -> Close -> OK

6. Statistické testy pro kvantitativní data – parametrické testy

Jednovýběrový t-test

Analyze -> Compare Means -> One-Sample T test -> Test Variables: zvolit proměnnou (např. Hippocampus volume (mm3)) -> Test Value: vyplnit např. 6575 -> OK

Párový t-test

Analyze -> Compare Means -> Paired-Samples T Test -> Paired Variables: přesunout vybrané proměnné (např. Hippocampus_volume (mm3) jako Variable1 a Hippocampus_volume_24 (mm3) jako Variable2) -> OK

Dvouvýběrový t-test

Analyze -> Compare Means -> Independent-Samples T Test -> Test Variables: zvolit proměnnou (např. Putamen_volume (mm3)) -> Grouping Variable: zvolit proměnnou (např. Gender_rek) -> Define Groups -> doplnit Group 1: F a Group 2: M -> Continue -> OK

Analýza rozptylu (ANOVA) a post-hoc testy

Analyze -> Compare Means -> One-Way ANOVA -> Dependent List: *zvolit proměnnou (např. Hippocampus_volume (mm3))* -> Factor: *zvolit proměnnou (např. Group_3kat)* -> Options -> *zatrhnout* Homogeneity of variance test *a* Brown-Forsythe -> Continue -> Post Hoc -> *zatrhnout* Tukey (v případě vyrovnaných počtů subjektů ve skupinách) *nebo* Scheffe (pro stejné i nestejně počty subjektů ve skupinách) -> Continue -> OK

7. Statistické testy pro kvantitativní data – neparametrické testy

Wilcoxonův test – jednovýběrový

Analyze -> Nonparametric Tests -> One Sample -> *na záložce* Objective *zatrhnout* Customize analysis -> *na záložce* Fields *do pole* Test Fields *přesunout proměnnou, kterou chceme testovat (např. MMSE)* -> *na záložce* Settings *zatrhnout* Customize tests -> *zatrhnout* Compare median to hypothesized -> Hypothesized median: *doplnit 27,5* -> Run

Wilcoxonův test – párový

1. způsob: Analyze -> Nonparametric Tests -> Related Samples -> *na záložce* Objective *zatrhnout* Customize analysis -> *na záložce* Fields *do pole* Test Fields *přesunout proměnné, které chceme testovat (např. MMSE a MMSE_24)* -> *na záložce* Settings *zatrhnout* Customize tests -> Wilcoxon matched-pair signed-rank (2 samples) -> Run

Pozn.: Chceme-li jinou neparametrickou obdobu párového t-testu, můžeme zvolit znaménkový test zaškrtnutím Sign Test (2 samples)

2. způsob: Analyze -> Nonparametric Tests -> Legacy Dialogs -> 2 Related Samples -> Test Pairs: *vložit proměnné (např. MMSE a MMSE_24)* -> Test Type: Wilcoxon (*případně* Sign) -> OK

Mannův-Whitneyův test

1. způsob: Analyze -> Nonparametric Tests -> Independent Samples -> *na záložce* Objective *zatrhnout* Customize analysis -> *na záložce* Fields *do pole* Test Fields *přesunout proměnnou, kterou chceme testovat (např. Hippocampus_volume (mm3)) a do pole Groups přesunout grupovací proměnnou (např. Gender_rek)* -> *na záložce* Settings *zatrhnout* Customize tests -> Mann-Whitney U (2 samples) -> Run

2. způsob: Analyze -> Nonparametric Tests -> Legacy Dialogs -> 2 Independent Samples -> Test Variable List: *vybrat proměnné (např. Hippocampus_volume (mm3))* -> Grouping Variable: *vybrat grupovací proměnnou, která musí být typu numeric (např. Gender_rek2)* -> Define Groups -> *doplnit* Group 1: 1, Group 2: 2 -> Continue -> OK

Pozn.: textovou proměnnou lze změnit na proměnnou s číselnými kategoriemi pomocí Transform -> Automatic Recode -> vybrat proměnnou (např. Gender_rek) -> New Name: vyplnit název nové proměnné (např. Gender_rek2) -> Add New Name -> OK

Kruskalův-Wallisův test

1. způsob: Analyze -> Nonparametric Tests -> Independent Samples -> *na záložce* Objective *zatrhnout* Customize analysis -> *na záložce* Fields *do pole* Test Fields *přesunout proměnnou, kterou chceme testovat (např. Hippocampus_volume (mm3)) a do pole Groups přesunout grupovací proměnnou (např. Group_3kat)* -> *na záložce* Settings *zatrhnout* Customize tests -> Kruskal-Wallis 1-way ANOVA (k samples) -> Run

2. způsob: Analyze -> Nonparametric Tests -> Legacy Dialogs -> K Independent Samples -> Test Variable List: *vybrat proměnné (např. Hippocampus_volume (mm3))* -> Grouping Variable: *vybrat grupovací proměnnou (např. Group_3kat)* -> Define Range: *doplnit rozsah grupovací proměnné* Minimum: 1, Maximum: 3 -> Continue -> OK

Pozn.: Post hoc testy ke Kruskal-Wallis ANOVě počítané 2. způsobem v SPSS neexistují, je možné to řešit sadou Mannových-Whitneyových testů (manuálně je pak nutno provést Bonferroniho korekci)

8. Ověření předpokladů statistických testů

Vykreslení Q-Q grafu pro jednotlivé skupiny

1. způsob: Analyze -> Descriptive Statistics -> Q-Q Plots -> Variables: *zvolit proměnnou (např. Hippocampus_volume (mm3))* -> Test Distribution: Normal -> OK

Pozn.: Nelze pomocí tohoto způsobu nastavit grupovací proměnnou.

2. způsob: Analyze -> Descriptive Statistics -> Explore -> Dependent List: *zvolit proměnné (např. Hippocampus_volume (mm3))* -> Factor List: *vybrat grupovací proměnnou (např. Gender_rek)* -> Plots -> *odtrhnout Stem-and-leaf a zatrhnout Normality plots with tests (a případně i Histogram)* -> Continue -> Display: *zatrhnout Plots* -> OK

9. Další užitečné příkazy

Popisná sumarizace kvantitativní proměnné podle kategorií kvalitativní proměnné

Analyze -> Tables -> Custom tables -> OK -> Columns: *přetáhnout kvantitativní proměnnou (např. Hippocampus_volume (mm3))* -> Summary Statistics: *zvolit např. Valid N, Mean, Std. Deviation, Median, Minimum, Maximum* -> Apply to Selection -> Close -> Rows: *přetáhnout kvalitativní proměnnou (např. Gender_rek)* -> Categories and Totals -> *Ize zaškrtnout Total (sumarizace pro všechny kategorie dohromady)* -> Apply -> OK

Pozn.: Pokud chceme zobrazit popisné statistiky do řádků, stačí zvolit Position: Rows

Vytvoření šablony grafů

Upravit si graf do finální podoby v Chart Editor -> File -> Save Chart Template -> zaškrtnout, která nastavení se mají uložit -> Continue -> File name: pojmenovat šablonu -> vybrat adresář, kam se to uloží -> Save

Úprava grafu podle šablony

2x kliknout do grafu -> Chart Editor -> File -> Apply Chart Template -> vybrat styl -> Open

Rozdělení souboru pomocí SPLIT FILE

Data -> Split File... -> zatrhnout Organize output by groups -> zvolit proměnnou (např. Group_3kat) do Groups Based on: -> OK

Zrušení rozdělení souboru

Data -> Split File... -> zatrhnout Analyze all cases, do not create groups -> OK

10. Analýza kontingenčních tabulek

Kontingenční tabulka absolutních četností

Analyze -> Descriptive Statistics -> Crosstabs -> Rows (select variables) (např. *Group_3kat*) -> Columns (select variables) (např. *Age_kat*) -> Cells -> v sekci Counts *zatrhnout* Observed -> Continue -> OK

Kontingenční tabulka procent

Analyze -> Descriptive Statistics -> Crosstabs -> Rows (select variables) (např. *Group_3kat*) -> Columns (select variables) (např. *Age_kat*) -> Cells -> v sekci Percentages *zatrhnout* Row, Columns *nebo* Total -> Continue -> OK

Pokud chceme kontingenční tabulku jenom procent

Analyze -> Descriptive Statistics -> Crosstabs -> Rows (select variables) (např. *Group_3kat*) -> Columns (select variables) (např. *Age_kat*) -> Cells -> v sekci Percentages *zatrhnout* Row, Columns *nebo* Total -> v sekci Counts *nesmí být zatrhnuté nic* -> Continue -> OK

Kontingenční tabulka – očekávané četnosti

Analyze -> Descriptive Statistics -> Crosstabs -> Rows (select variables) (např. *Group_3kat*) -> Columns (select variables) (např. *Age_kat*) -> Cells -> v sekci Counts *zatrhnout* Expected -> Continue -> OK

Pearsonův chí-kvadrát test

Analyze -> Descriptive Statistics -> Crosstabs -> Rows (select variables) (např. *Group_3kat*) -> Columns (select variables) (např. *Age_kat*) -> Statistics -> *zatrhnout* Chi-square -> Continue -> OK

Fisherův exaktní test

Analyze -> Descriptive Statistics -> Crosstabs -> Rows (select variables) (např. *Gender_rek*) -> Columns (select variables) (např. *mmse_kat*) -> Statistics -> *zatrhnout* Chi-square -> Continue -> Exact -> *zatrhnout* Exact -> Continue -> OK

Poznámka: u 2x2 tabulek se vypočítá automaticky při zatržení Chi-square (není třeba zatrhávat Exact).

McNemarův test

Analyze -> Descriptive Statistics -> Crosstabs -> Rows (select variables) (např. *mmse_kat*) -> Columns (select variables) (např. *mmse_24_kat*) -> Statistics -> *zatrhnout* McNemar -> Continue -> OK

Relativní riziko (RR) a poměr šancí (OR)

Analyze -> Descriptive Statistics -> Crosstabs -> Rows (select variables) (např. *Gender_rek*) -> Columns (select variables) (např. *mmse_kat2*) -> Statistics -> *zatrhnout* Risk -> Continue -> OK

!!! Pozor, lze vypočítat pouze pro tabulky 2x2 !!!

Poměr šancí (OR) včetně p-hodnoty

Analyze -> Regression -> Binary Logistic -> vybrat závisle proměnnou a kovariátu -> Categorical (lze zvolit, zda je referenční kategorie Last nebo First; při změně kliknout na Change) -> Options (*zatrhnout* CI for exp(B), aby se vypsaly intervaly spolehlivosti) -> Continue -> OK

11. ROC analýza

- I. Analyze -> ROC Curve -> *zadat* Test Variable (*např.* MMSE) a State Variable (*např.* group_01_CnMci) a jako Value of State Variable *zadat* rizikovou kategorii (*např.* 1) -> na záložce Options lze zvolit, zda „Larger test result indicates more positive test“ nebo „Smaller test result indicates more positive test“ -> Continue
- II. Pokud chceme vypočítat k AUC intervaly spolehlivosti a p-hodnotu: zatrhnout „Standard error and confidence interval“
- III. Pokud chceme získat tabulku se senzitivitou a 1-specifitou: zatrhnout „Coordinate points of the ROC Curve“
- IV. Pro zjištění cut-off (nejlepšího dělicího bodu) je nutné tabulku s 1-specifitou zkopírovat do Excelu, vypočítat specifitu a nalézt nejlepší cut-off a to následovně:
 - a. kliknout levým tlačítkem myši na tabulku -> Copy -> vložit do Excelu
 - b. v Excelu spočítat specifitu (tzn. udělat 1-(sloupeček s 1-Specificity)) a pak spočítat součet senzitivity a specifity -> vybrat řádek s největší hodnotou součtu senzitivity a specifity -> to je nejlepší cut-off
- V. Vytvoření kategorizovaného MMSE skóre s využitím cut-off: Transform -> Recode into Different Variables -> pojmenovat novou proměnnou (*např.* mmse_kat3) v Output Variable a kliknout Change -> Old and New Values -> do Range, LOWEST through value napsat 28,5 -> do New Value (Value) napsat 1 -> Add -> kliknout na All other values -> do New Value (Value) napsat 0 -> Add -> Continue -> OK
- VI. Ověření vypočítané senzitivity a specifity: Analyze -> Descriptive Statistics -> Crosstabs -> do Rows (*např.* mmse_kat3) do Columns (*např.* group_01_CnMci) -> Cells -> na záložce Percentages zatrhnout Column -> Continue -> OK -> senzitivita je tedy procento v levém dolním rohu, specifita je procento v pravém horním rohu – záleží ale na nakódování dat a v Rows musí být výsledek diagnostického testu a v Columns skutečnost

12. Korelační analýza

Bodový graf (“Scatterplot”)

Graphs -> Legacy Dialogs -> Scatter/Dot... -> zvolit Simple Scatter -> Define -> zvolit proměnné (*např.* Putamen_volume (mm3) jako X a Amygdala_volume (mm3) jako Y -> OK

Poznámka 1 – přidání regresní přímky: 2x kliknout na graf -> Elements -> Fit line at total (případně v nově vyskočeném okně zrušit zatržení "Attach label to line" a kliknout na Apply)

Poznámka 2 – vykreslení podle skupin: při zadávání přidat kategoriální proměnnou do Set Markers by.

Výpočet Pearsonova korelačního koeficientu

Analyze -> Correlate -> Bivariate -> zvolit proměnné (*např.* Amygdala_volume (mm3) a Putamen_volume (mm3)) -> v sekci Correlation Coefficients zatrhnout Pearson -> OK

Výpočet Pearsonova korelačního koeficientu – 2. způsob

Analyze -> Descriptive Statistics -> Crosstabs -> zvolit proměnné (*např.* Amygdala_volume (mm3) a Putamen_volume (mm3)) -> Statistics -> zatrhnout Correlations -> Continue -> je vhodné zatrhnout Suppress tables -> OK

Poznámka – spočte to zároveň i Spearmanův korelační koeficient

Výpočet Spearmanova korelačního koeficientu

Analyze -> Correlate -> Bivariate -> zvolit proměnné (*např.* MMSE a Hippocampus_volume (mm3)) -> v sekci Correlation Coefficients zatrhnout Spearman -> OK

Výpočet parciální korelace (parametrické)

Analyze -> Correlate -> Partial -> zvolit proměnné (*např.* Pallidum_volume (mm3) a Nucl_caud_volume

(mm3) do Variables a Age do Controlling for) -> OK

Poznámka – neparametrickou parciální korelaci nelze v SPSS vyklikat, musí se použít syntax.

13. Regresní analýza

Lineární regrese a odstranění vlivu kovariát

Analyze -> Regression -> Linear -> zvolit *proměnné* (např. Nucl_caud_volume (mm3) jako Dependent, Age, gender_01, ... jako Independent(s)) -> OK -> vypíše *regresní koeficienty a p-hodnoty*

- Vykreslení Q-Q grafu a histogramu standardizovaných reziduí: na záložce Plots kliknout na *ZRESID a zaškrtnout Histogram -> Continue
- Vykreslení Q-Q grafu a histogramu pro rezidua – 2. způsob: kliknout na Save a v sekci Residuals zaškrtnout Standardized či Unstandardized -> Continue -> OK
Analyze -> Descriptive Statistics -> Explore -> Dependent List (zvolit např. RES_1) -> Plots -> zaškrtnout Normality plots with tests a Histogram -> Continue -> OK
- Vykreslení bodového grafu predikovaných hodnot a reziduí: na záložce Plots zvolit na Y osu *ZRESID a na osu X *ZPRED
- Uložení reziduí a predikovaných hodnot: na záložce Save zaškrtnout při Predicted Values i při Residuals: Standardized nebo Unstandardized podle toho, co nám vyhovuje -> Continue -> OK

14. Analýza přežití

Kaplanův-Meierův odhad funkce přežití pro jeden výběr

Analyze -> Survival -> Kaplan-Meier... -> zvolit *proměnné* Time a Status (např. čas jako Time a úmrtí jako Status) -> kliknout na Define event (do Single value zvolit příslušnou hodnotu, která kóduje událost (většinou hodnota 1)) -> OK

- Vykreslení křivky přežití: kliknout na Options a zatrhnout Survival v sekci Plots
- Vypsání tabulky pro výpočet x-letého přežití: kliknout na Options a zatrhnout Survival table(s) v sekci Statistics
- Vypsání tabulky s mediánem přežití: kliknout na Options a zatrhnout Mean and median survival v sekci Statistics

Kaplanův-Meierův odhad funkce přežití pro více výběrů

Analyze -> Survival -> Kaplan-Meier -> zvolit *proměnné* Time a Status (např. čas jako Time a úmrtí jako Status) -> kliknout na Define event (do Single value zvolit příslušnou hodnotu, která kóduje událost (většinou hodnota 1)) -> zvolit Factor (proměnná podle které chceme rozdělit data do skupin) -> OK

- Vykreslení křivek přežití: kliknout na Options a zatrhnout Survival v sekci Plots
- Výpočet testu na srovnání přežití ve skupinách: kliknout na Compare Factor a zatrhnout Log rank v sekci Test Statistics
- Vypsání tabulky pro výpočet x-letého přežití: kliknout na Options a zatrhnout Survival table(s) v sekci Statistics
- Vypsání tabulky s mediánem přežití: kliknout na Options a zatrhnout Mean and median survival v sekci Statistics

Coxův model proporcionálních rizik

Analyze -> Survival -> Cox Regression -> zvolit *proměnné* Time, Status a Covariates (např. OS_doba_mesice jako Time, úmrtí jako Status a ECOG1 jako Covariates) -> kliknout na Define event (do Single value zvolit příslušnou hodnotu, která kóduje událost (většinou hodnota 1)) -> OK

- Vykreslení křivek přežití: kliknout na Plots a zatrhnout Survival
- Interval spolehlivosti: kliknout na Options a zatrhnout CI for exp(B) v záložce Model Statistics