

# Grafy

POLb1139 Statistické myšlení v sociálních vědách

# Úvod

- Navazujeme na:
  - Četnosti
  - Medián, průměr
- Bude důležité pro:
  - Normální rozdělení
  - Crosstaby
  - korelace

# Důležitost typu proměnné

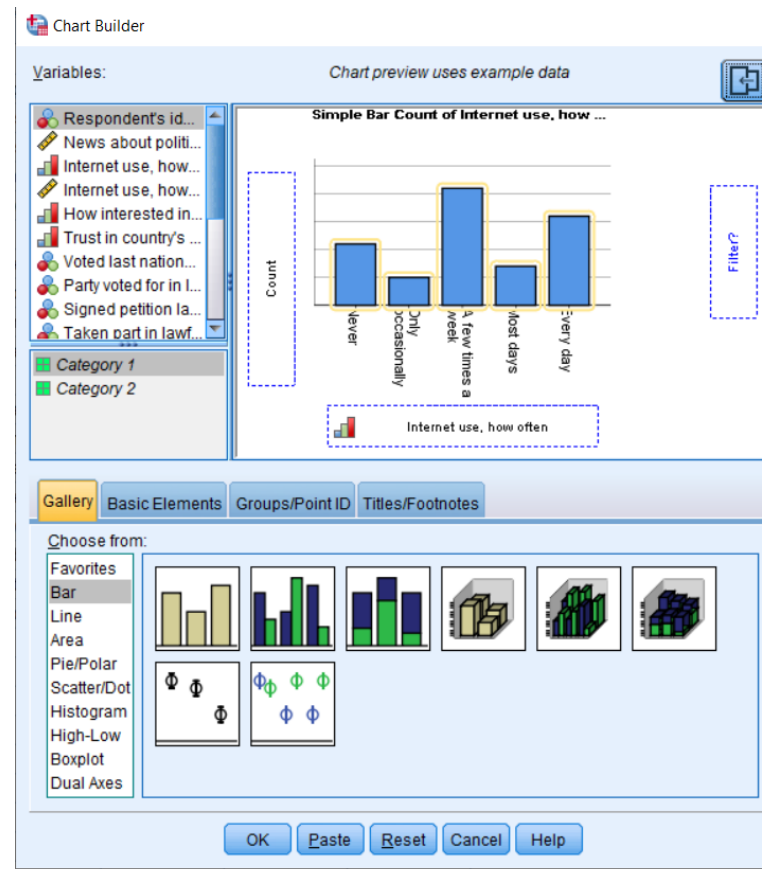
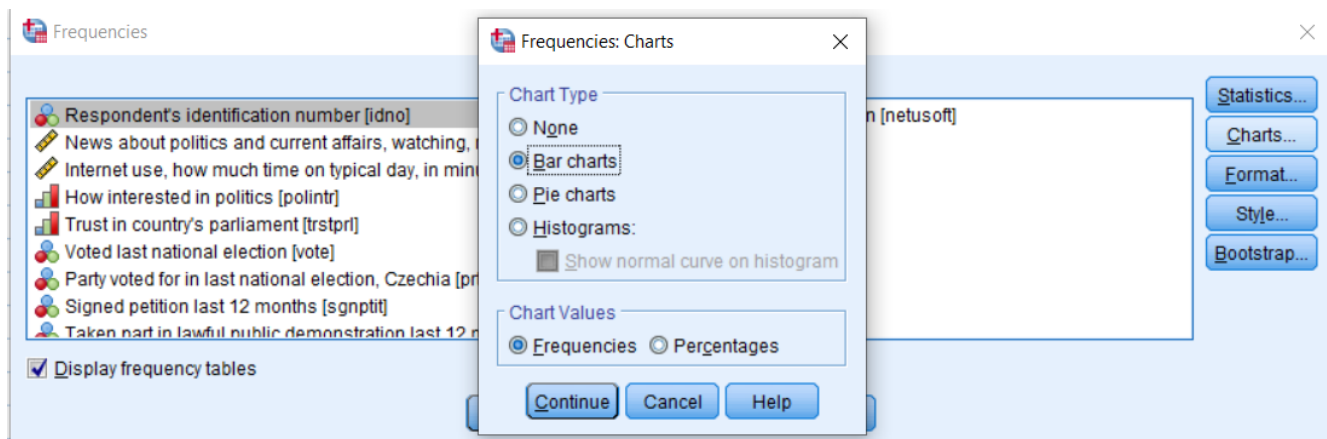
- Různé vizualizace pro různé typy proměnných
- Kategorické – srovnání četností kategorií – sloupcový graf (bar chart)
  - zlomky – koláčový graf (pie charts)
- kardinální – distribuce (histogram, box plot)

# Vizualizace

- Základní pravidlo – nekomplikovat si grafy (život)
- Cíl – vizualizovat distribuci hodnot
- Plně postačí jednoduché sloupcové grafy (bar charts)
- Je zbytečné přidávat různé prvky typu 3D, kombinovat barvy, používat pro efekt koláčové grafy (pie charts) atd.

# Kde se v SPSS dělají grafy

- Analyze → Descriptive Statistics → Frequencies → Charts
- Graphs → Chart Builder → Bar...
- Graphs → Legacy dialogs → Bar...

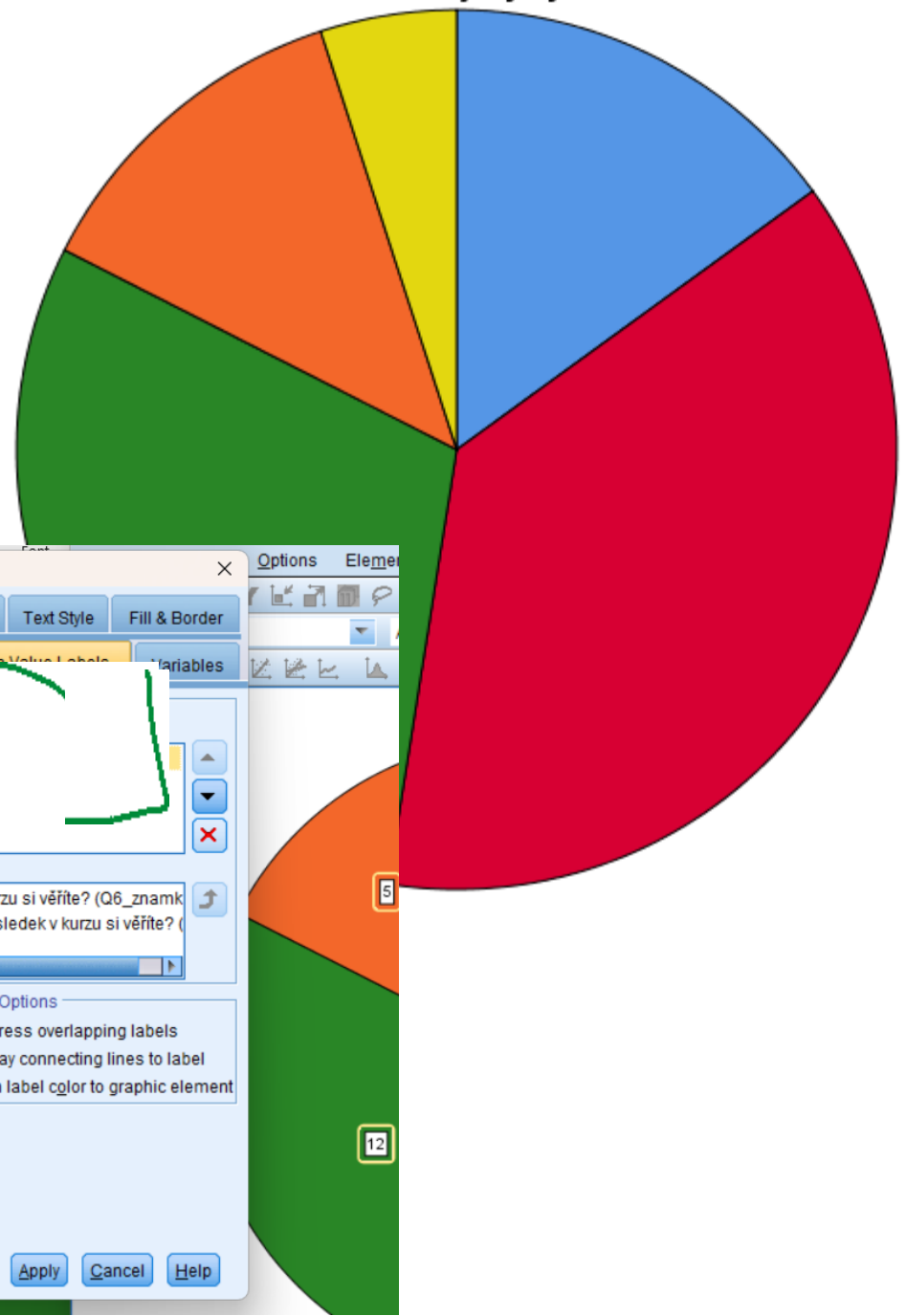
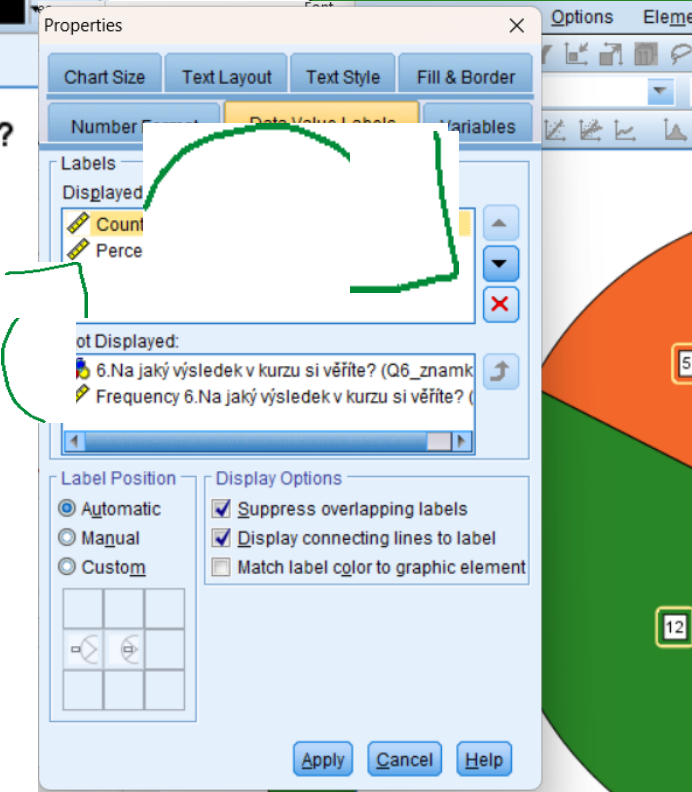
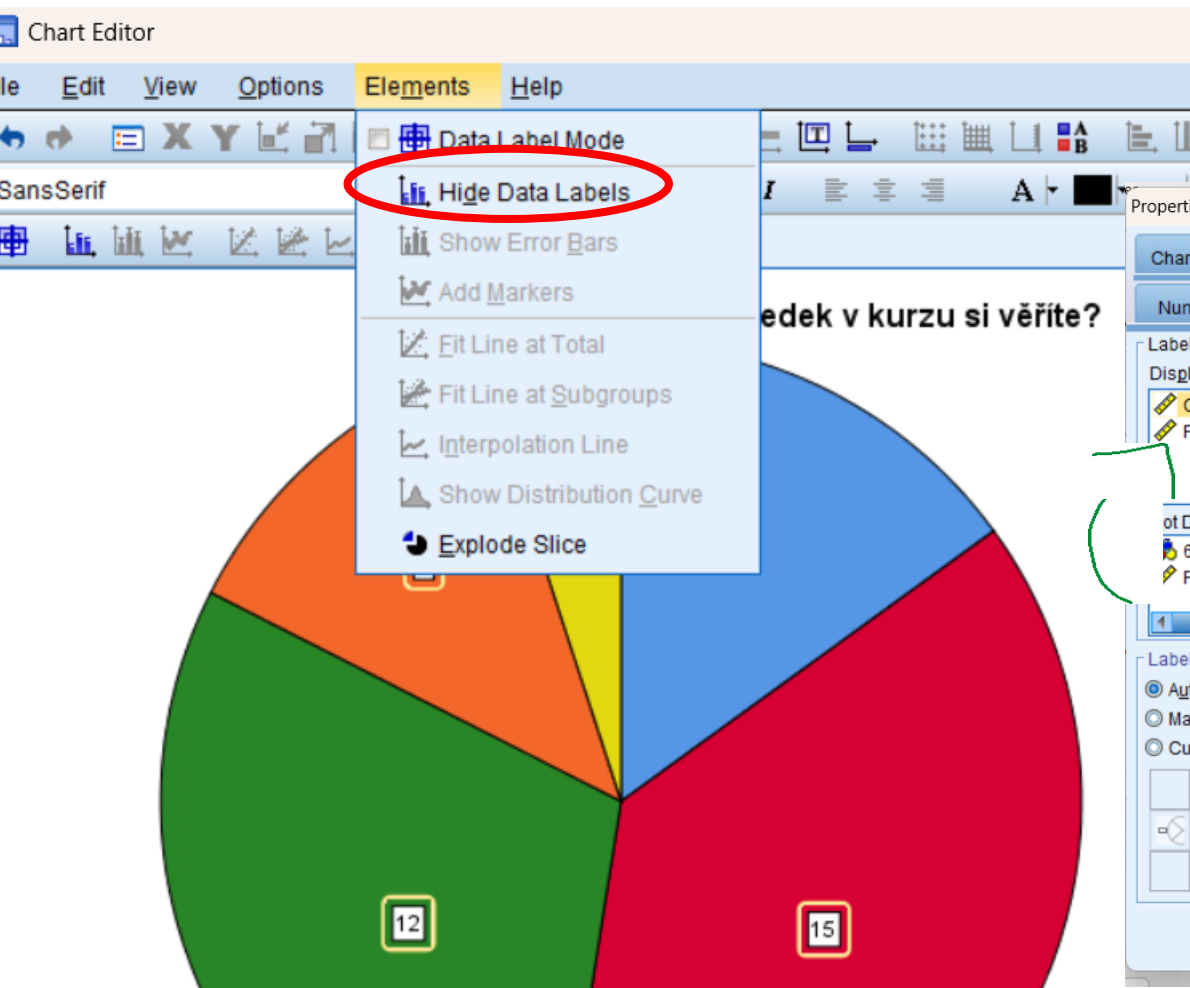


# Koláčový graf

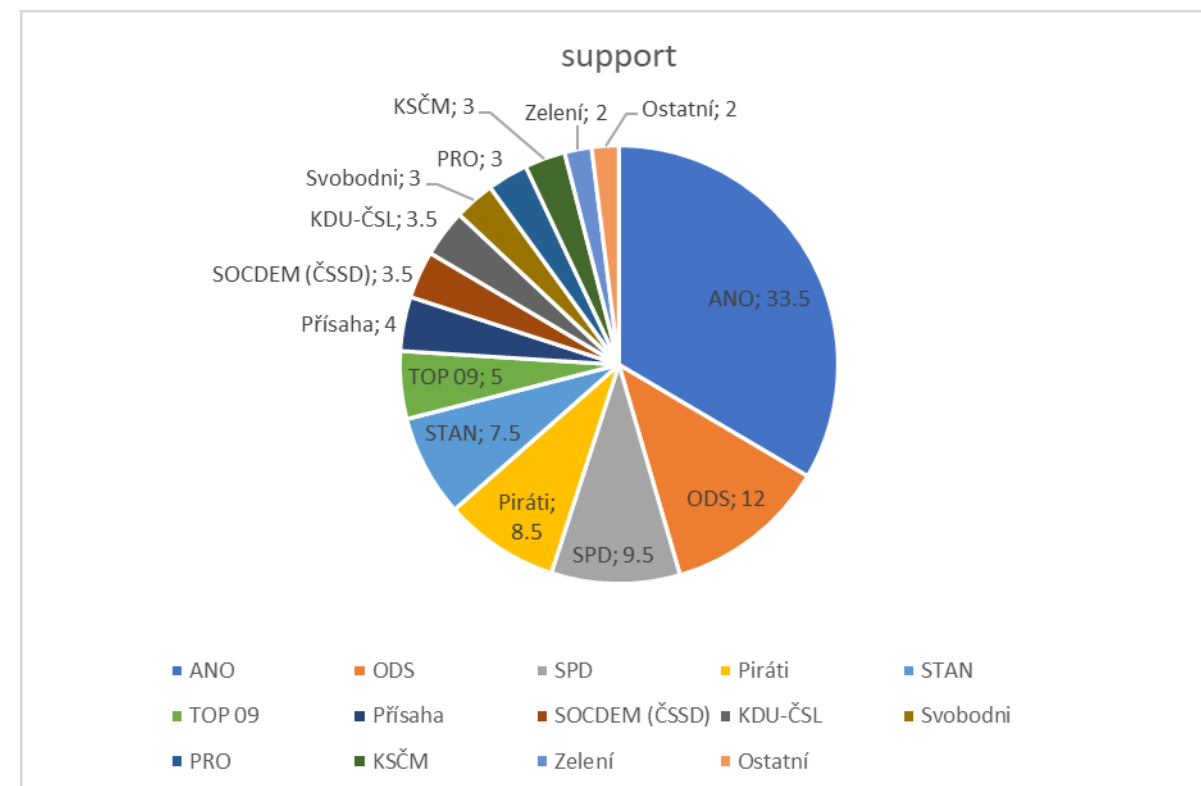
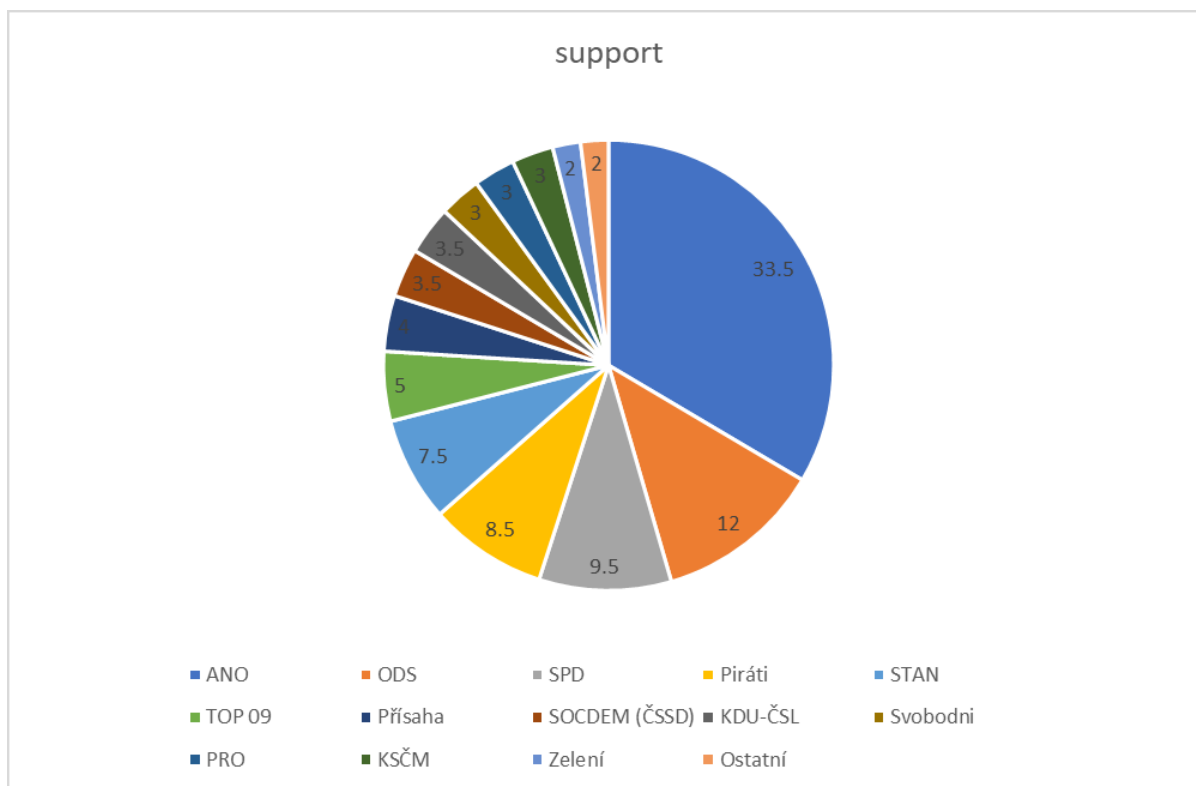
- Jen pokud je počet kategorií malý (max okolo 7)
- Není příliš informativní pokud jsou počty malé nebo podobné
- Analyze → Descriptive Statistics → Frequencies → Charts → pie chart

- Dvojklikem na graf lze otevřít okno pro editaci

6.Na jaký výsledek v kurzu si věříte?



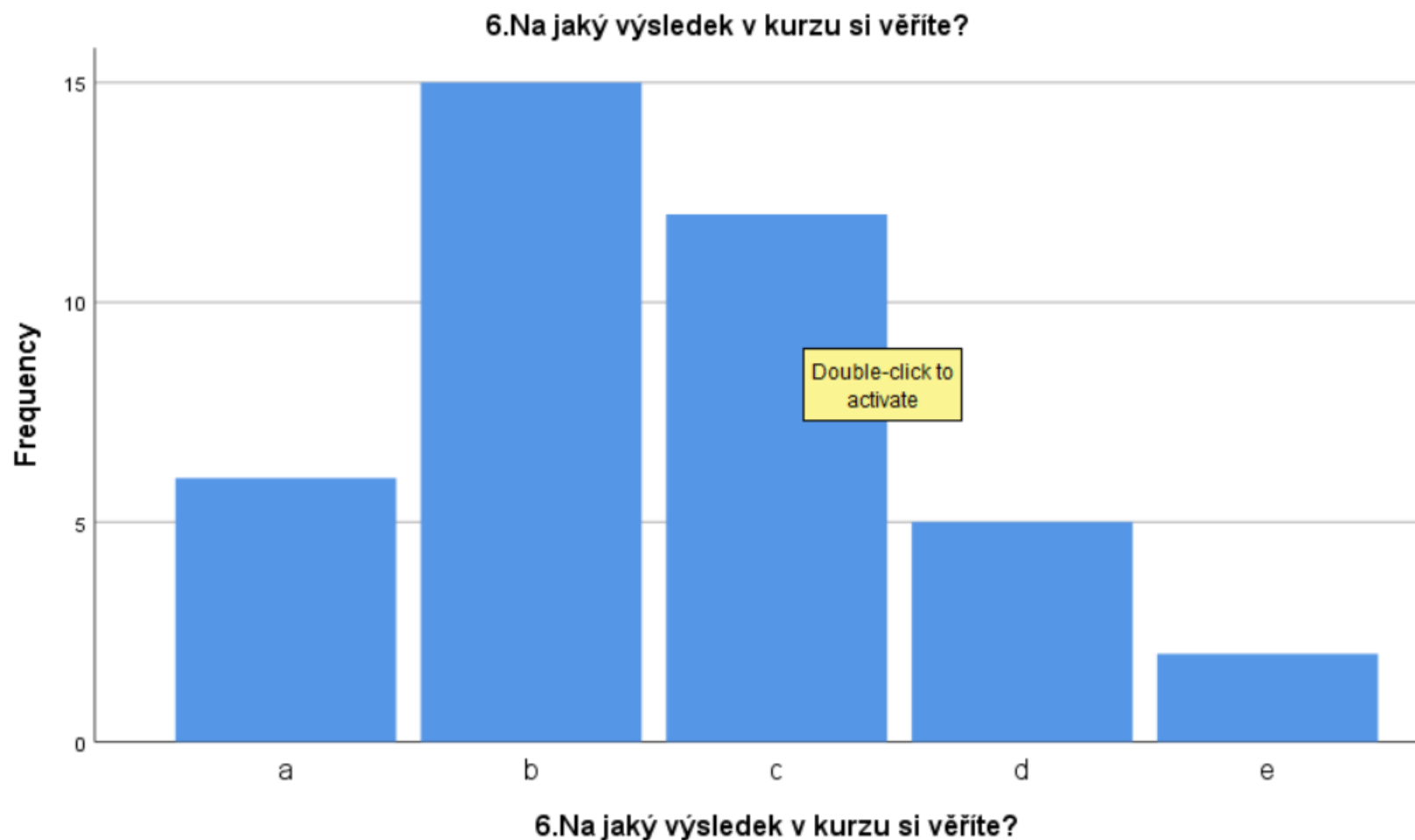
# Problém v případě mnoha kategorií





# Sloupcový graf

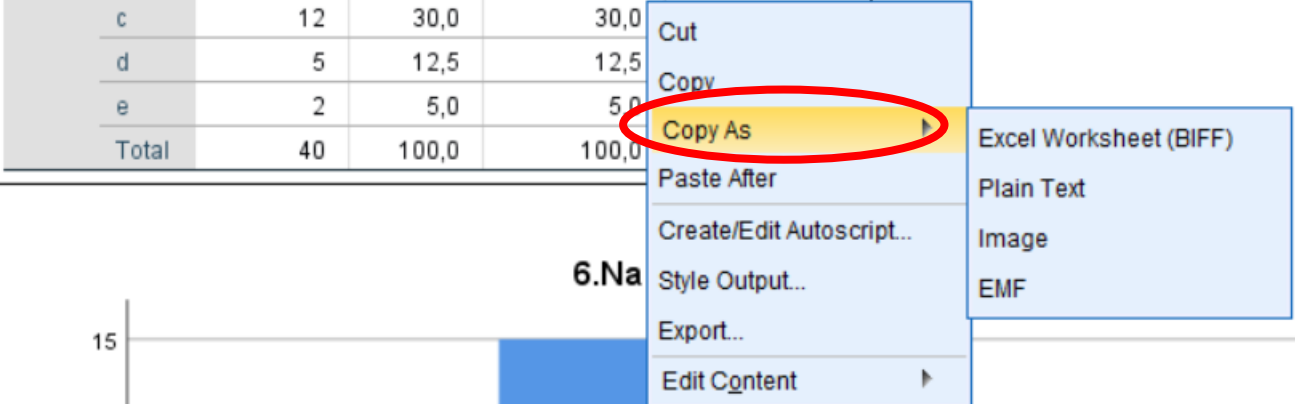
- Ukazuje četnosti každé z kategorií
- Neukazuje procenta
- Řešení -> export do excelu



# Možnosti exportu

- Export konkrétní tabulky nebo grafu
  - Kliknutím pravým tlačítkem myši na věc v outputu
  - Copy as -> excel worksheet pro tabulky  
->image nebo microsoft object pro grafy

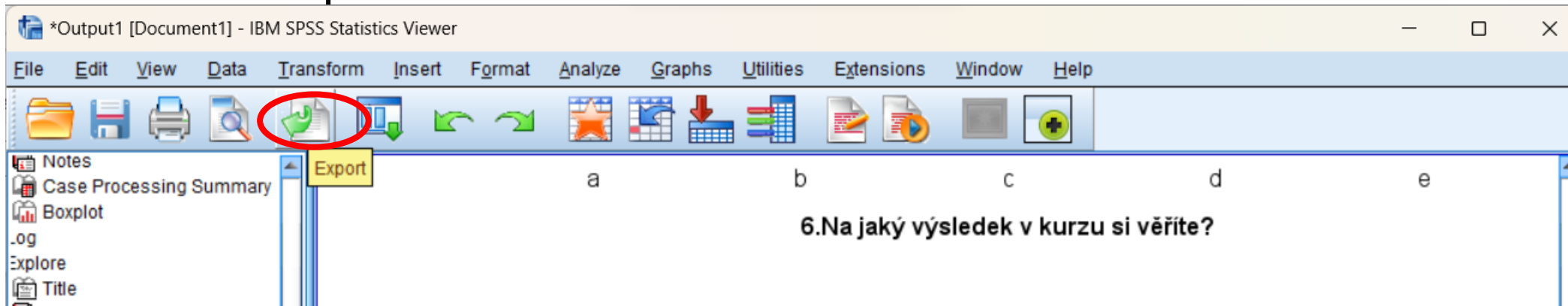
6.Na jaký výsledek v kurzu si věříte?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	a	6	15,0	15,0	15,0
	b	15	37,5	37,5	52,5
	c	12	30,0	30,0	
	d	5	12,5	12,5	
	e	2	5,0	5,0	
Total		40	100,0	100,0	



6.Na jaký výsledek v kurzu si věříte?

15

- Export celého outputu



\*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Export

6.Na jaký výsledek v kurzu si věříte?

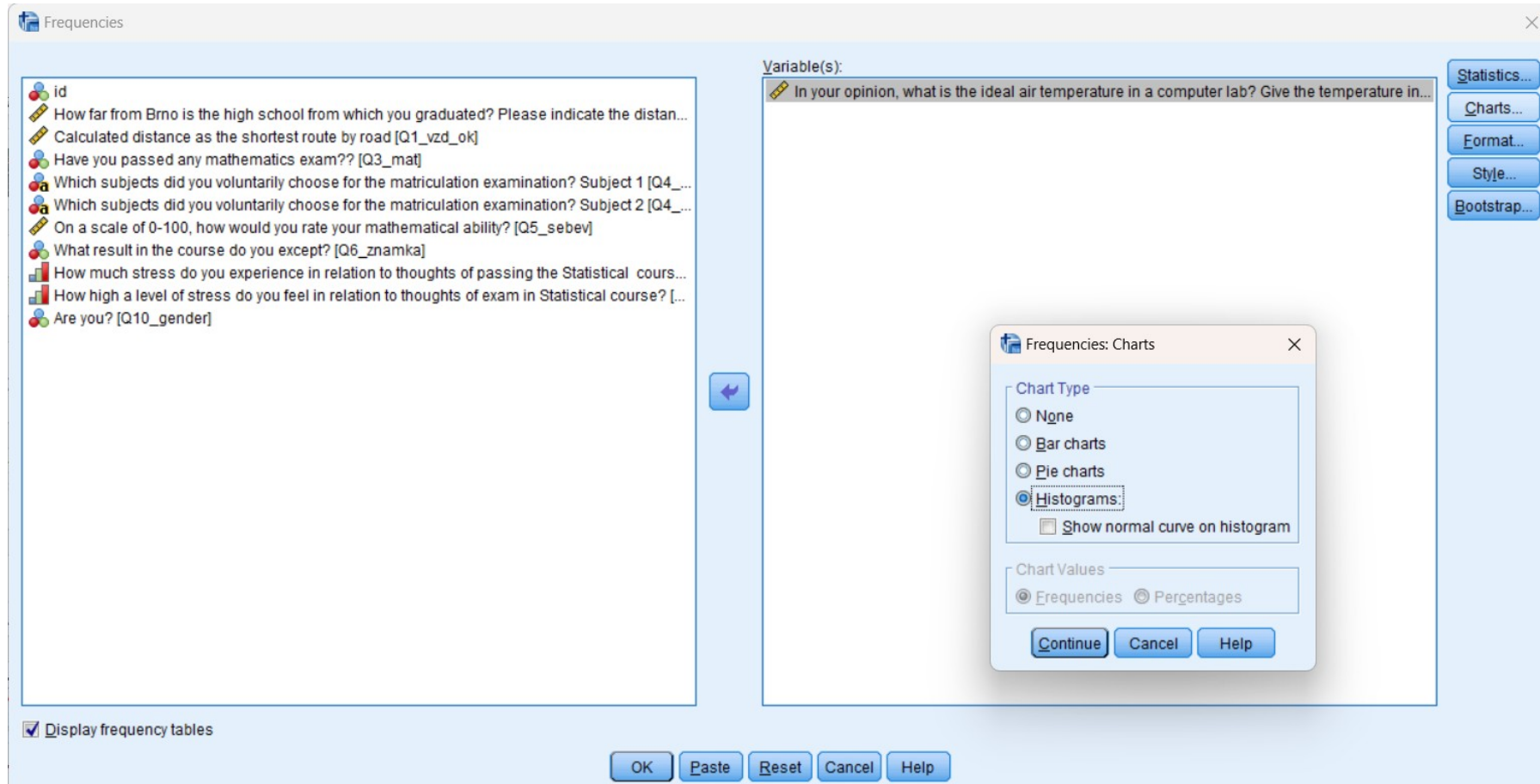
a b c d e

- Používejte export do excelu, v něm je snazší tabulky upravit, až po úpravě v excelu je kopírujte do wordu

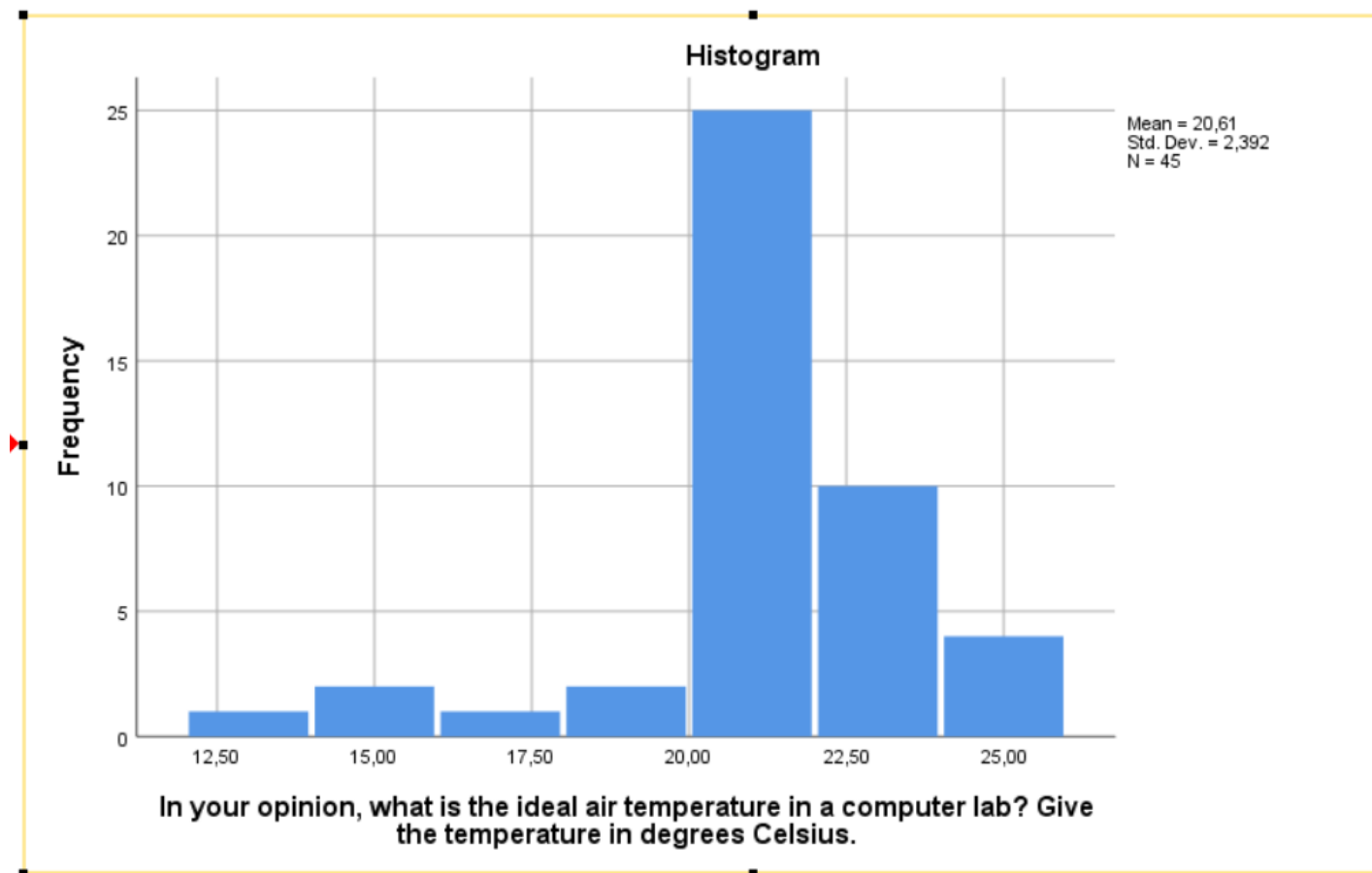
# Kardinální proměnné

- Intervalové a poměrové (SPSS nerozlišuje – obě jsou *scale*)
- Numerické kódy (zpravidla) odpovídají reálným pozorovaným hodnotám
- Více možností jednorozměrné analýzy oproti nominálním a ordinálním proměnným
- Vizualizace – stejná pravidla

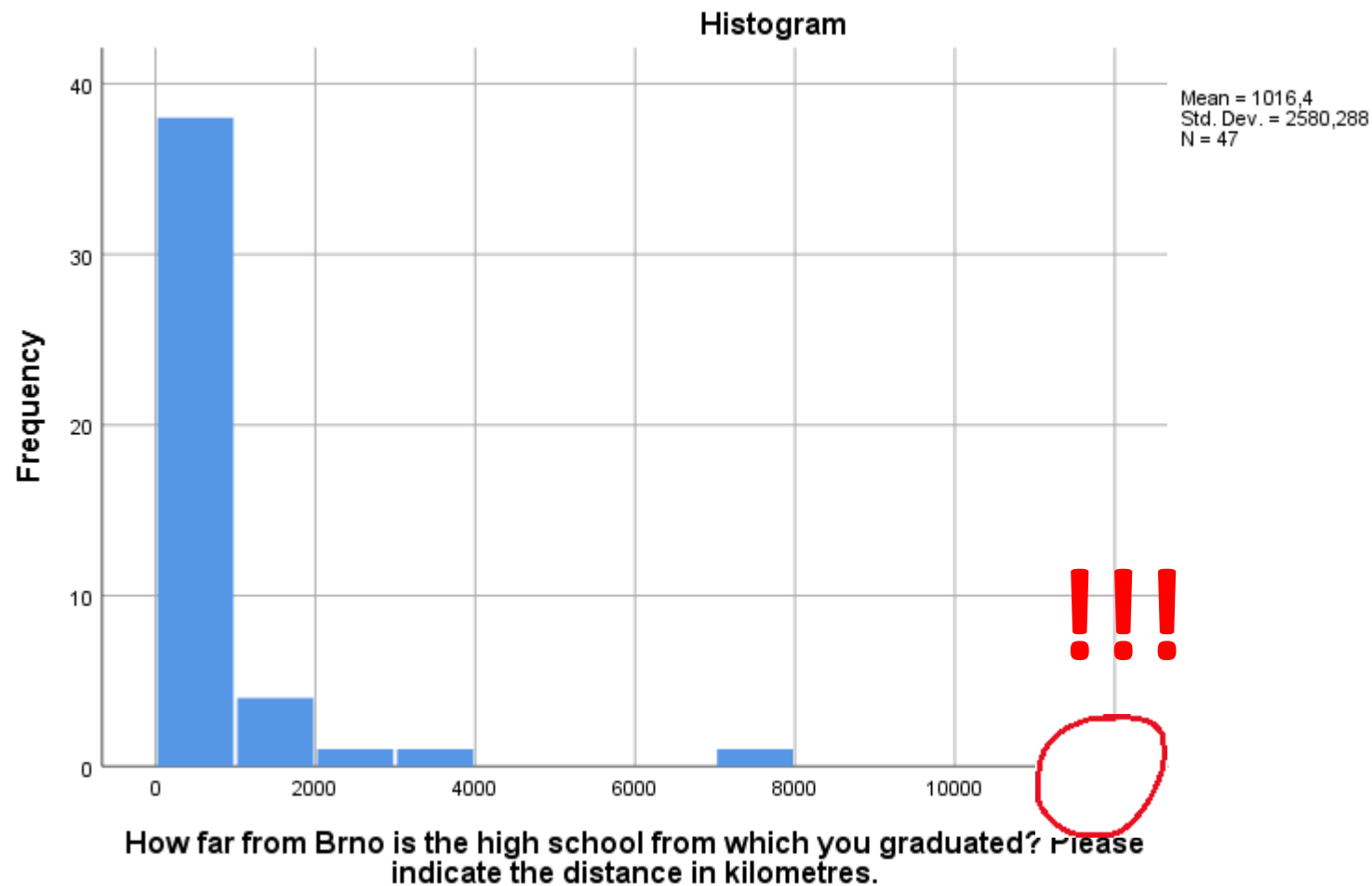
# Analyze → Descriptive Statistics → Frequencies → Charts

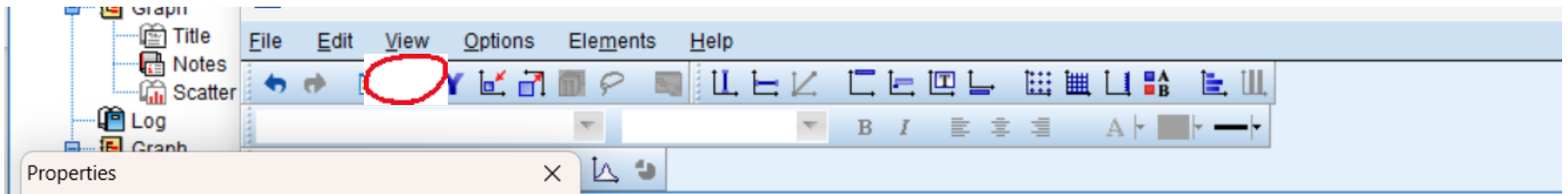


# Proměnná teplota v místnosti



# Proměnná vzdálenost





Properties

Number Format   Grid Lines   Variables

Chart Size   Lines   Labels & Ticks

Range

	Auto	Custom	Data
Minimum	<input checked="" type="checkbox"/>	0	0
Maximum	<input checked="" type="checkbox"/>	100000	21543
Major Increment	<input checked="" type="checkbox"/>		
Origin	<input checked="" type="checkbox"/>	0	

Display line at origin

Type

Scale

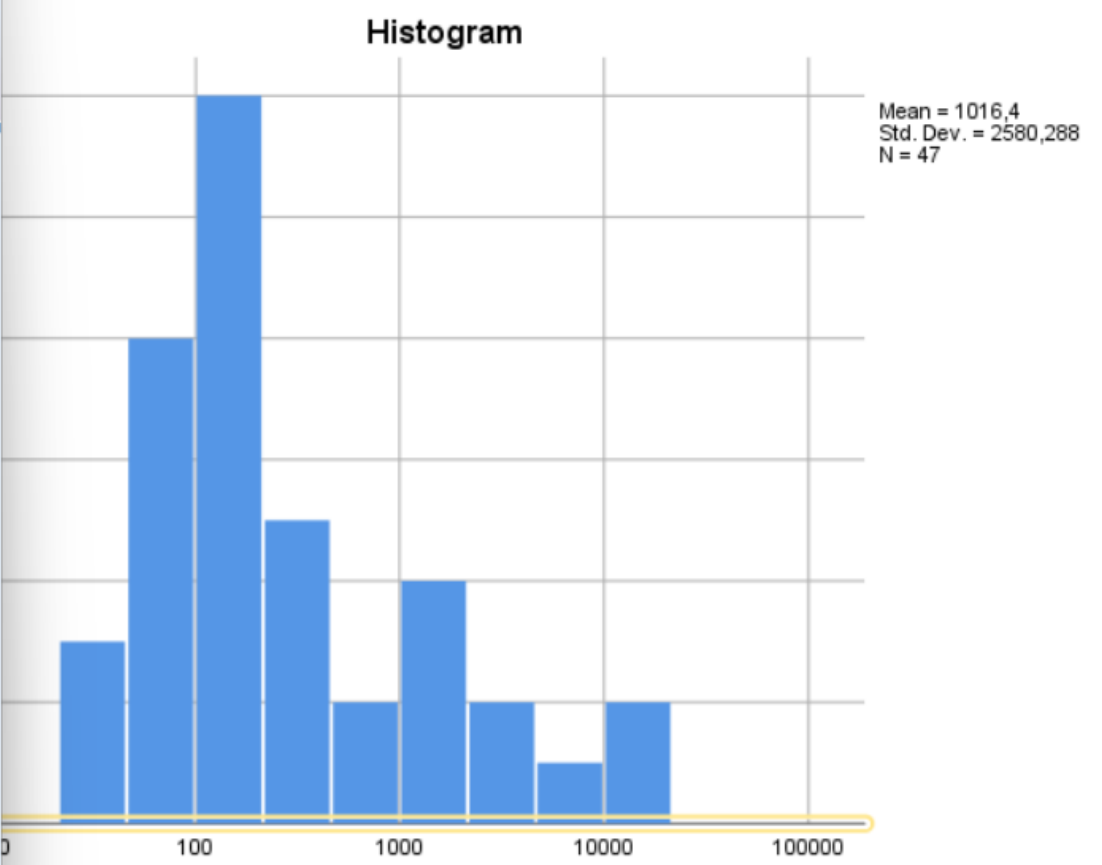
Safe

Base: 10

Exponent: 0,5

Lower margin (%): 5   Upper margin (%): 5

Apply   Close   Help



How is the high school from which you graduated? Please indicate the distance in kilometres.

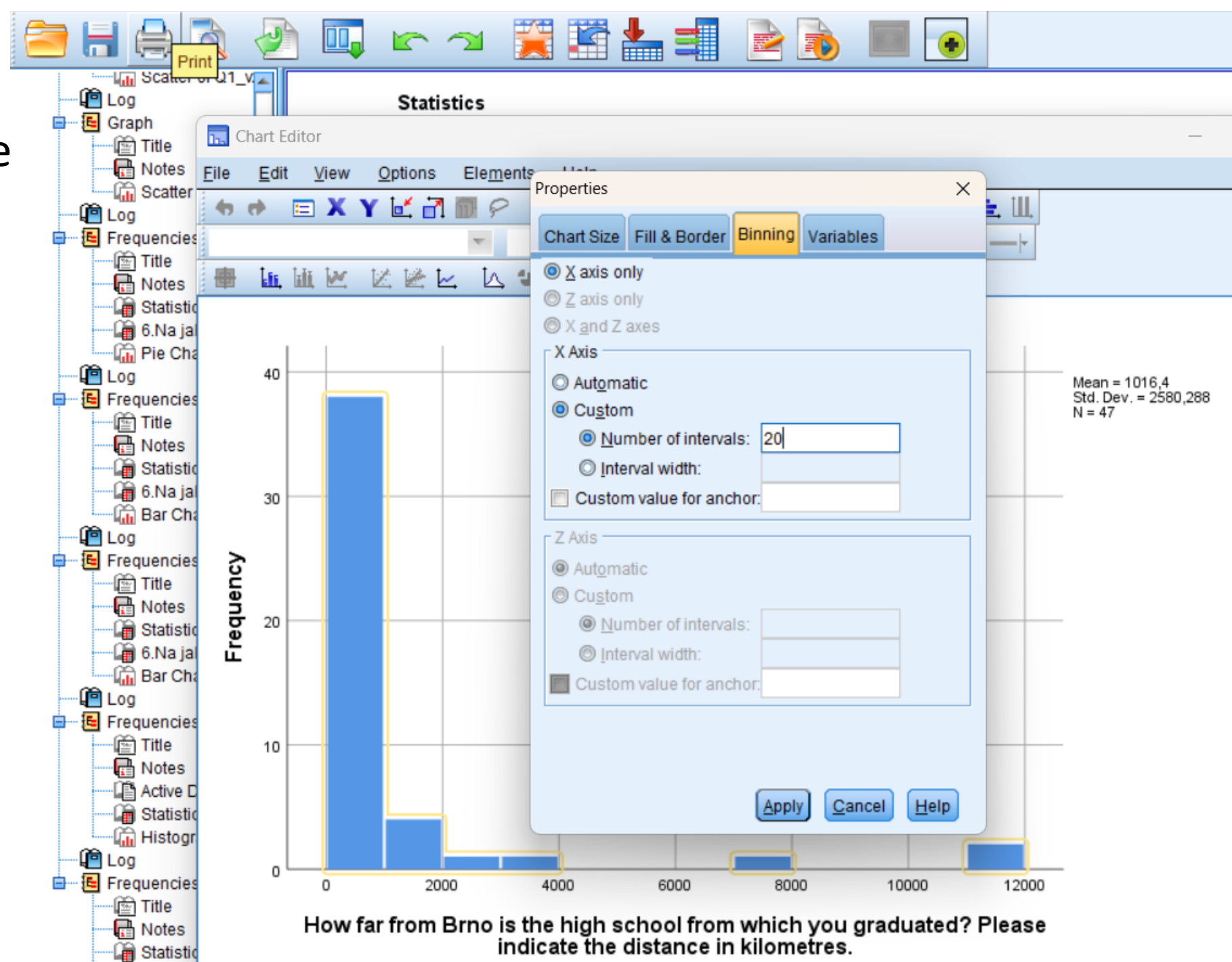
# Logaritmovaná osa

- Řešení pro situaci kdy máme velké množství nízkých hodnot a několik případů s velmi vysokými hodnotami
  - Typická situace: populace
  - Důležité také pro scatter plot
- Problem: nefunguje pro 0 a záporná čísla

Logaritmus vyjadřuje hodnotu kterou je třeba umocnit základ (zde 10), aby vzniklo dané číslo (např. pro 100 je to 2, protože  $10^2$  je 100)



- Úprava sloupců
- Dvojklik na některý ze sloupců
- Binning
- Buďto zvýšení počtu kategorií
- Nebo změna šířky intervalu

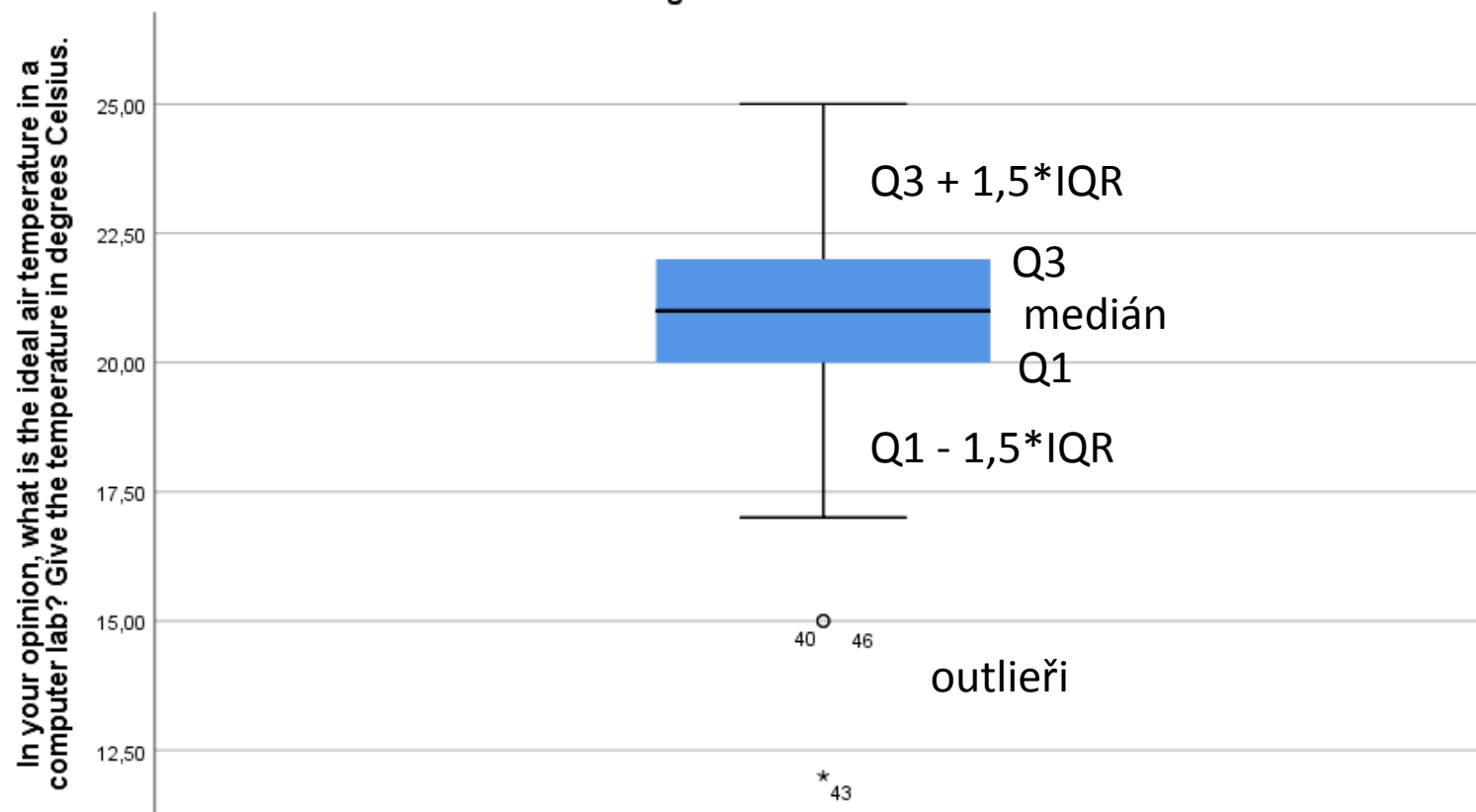


# Odlehlé případy (outliers)

- $< Q1 - 1,5 * IQR$
- $> Q3 + 1,5 * IQR$
  
- Outliers leží za těmito hodnotami
- Vhodné poznat pro určité druhy analýzy (vliv na výsledky)
  
- vizualizace pomocí krabicového grafu (boxplot)

# Graphs -> Chart Builder -> Boxplot

1-D Boxplot of In your opinion, what is the ideal air temperature in a computer lab? Give the temperature in degrees Celsius.



# Grafy pro ilustraci souvislosti

- kategorická x kategorická – složený sloupcový graf (clustered/stacked bar chart)
- kategorická x kardinální – sada boxplotů
- kardinální x kardinální – bodový graf (scatterplot)

STATISTICS Data Editor

Analyze Utilities Extensions Window Help

Chart Builder...  
Graphboard Template Chooser...  
Weibull Plot...  
Compare Subgroups  
Regression Variable Plots

Width	Missing	Columns
2	None	12
4	None	12
4	None	12
1		
10	None	
10	None	
3	{0, bad}...	
3	{1,0, a}...	
1	{1, very low	
1	{1, very low	
11	None	
18	{1, female}	

Line...  
Area...  
Pie...  
High-Low...  
Boxplot...  
Error Bar...  
Population Pyramid...  
Scatter/Dot...  
Histogram...

Bar Charts

Simple  
Clustered  
Stacked

Data in Chart Are

Summaries for groups of cases  
 Summaries of separate variables  
 Values of individual cases

Define Cancel Help

Define Clustered Bar: Summaries for Groups of Cases

Bars Represent

N of cases  % of cases  
 Cum. N  Cum. %  
 Other statistic (e.g., mean)

Variable:

Change Statistic...

Category Axis:

Define Clusters by:

Panel by

Rows:

Nest variables (no empty rows)

Columns:

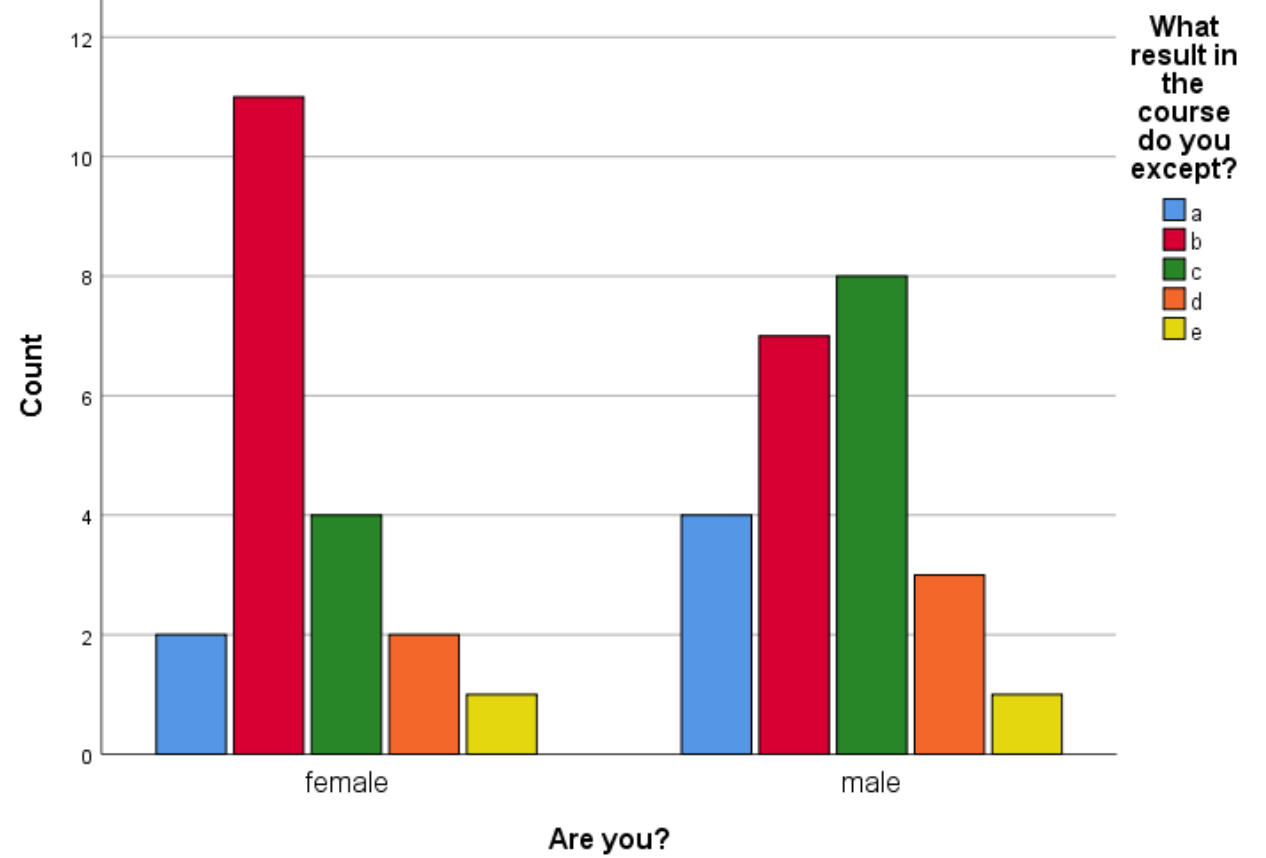
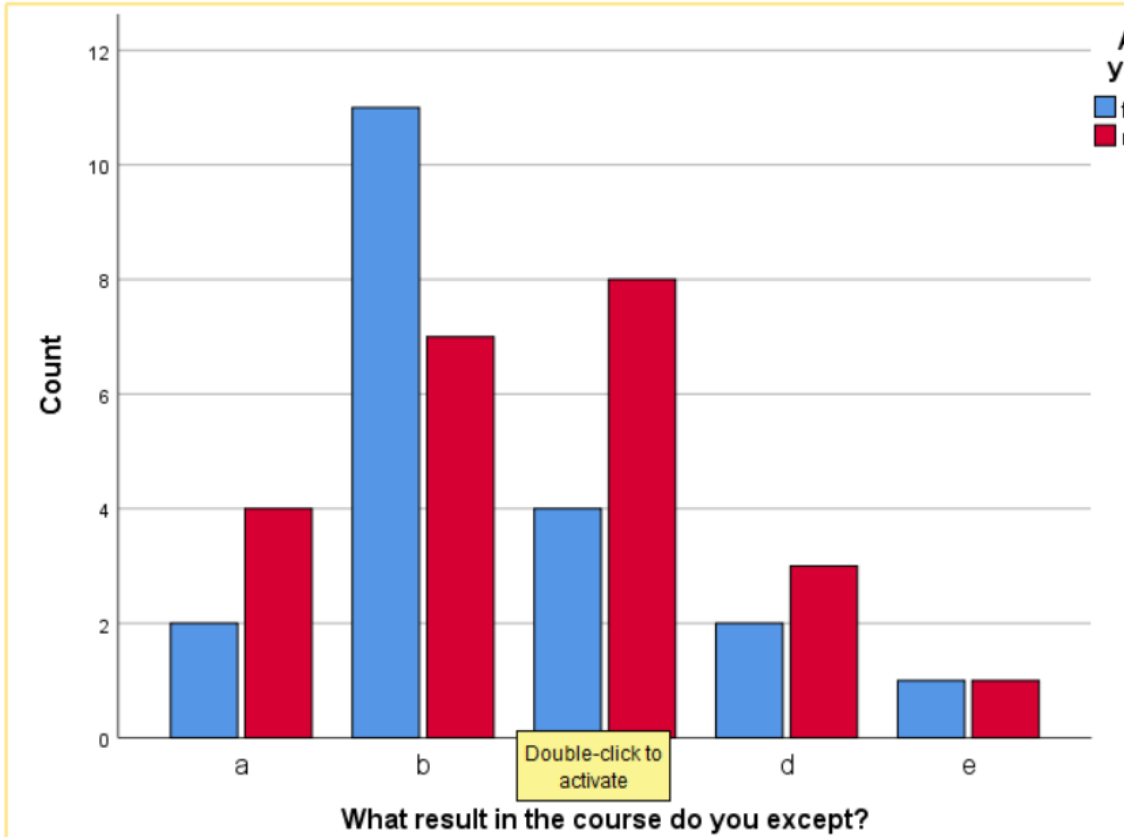
Nest variables (no empty columns)

Template

Use chart specifications from:

OK Paste Reset Cancel Help

# Sloupcový graf



Analyze **Graphs** Utilities Extensions Window Help

Chart Builder...  
Graphboard Template Chooser...

Width	Missing	Columns
2	None	12
4	None	12
4	None	12
1		
10	0	Which subjects... None
10	0	Which subjects... None
3	0	On a scale of 0... {0, bad}...
3	1	What result in t... {1, 0, a}...
1	0	How much stre... {1, very low
1	0	How high a leve... {1, very low
11	2	In your opinion,... None
18	0	Are you? {1, female}.

Legacy Dialogs

- Bar...
- 3-D Bar...
- Line...
- Area...
- Pie...
- High-Low...
- Boxplot...
- Error Bar...
- Population Pyramid...
- Scatter/Dot...
- Histogram...

Bar Charts

Simple

Clustered

Data in Chart Are

- Summaries for groups of cases
- Summaries of separate variables
- Values of individual cases

Define Cancel Help

Define Stacked Bar: Summaries for Groups of Cases

Bars Represent

- N of cases
- % of cases
- Cum. N
- Cum. %
- Other statistic (e.g., mean)

Variable:

Change Statistic...

Category Axis:

What result in the course do yo...

Define Stacks by:

Are you? [Q10\_gender]

Panel by

Rows:

Nest variables (no empty rows)

Columns:

Nest variables (no empty columns)

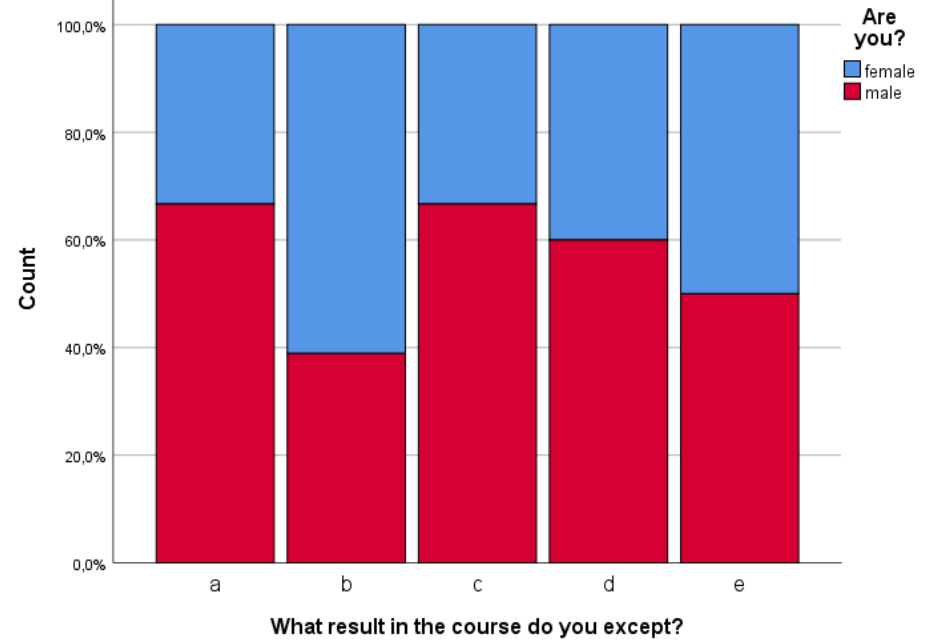
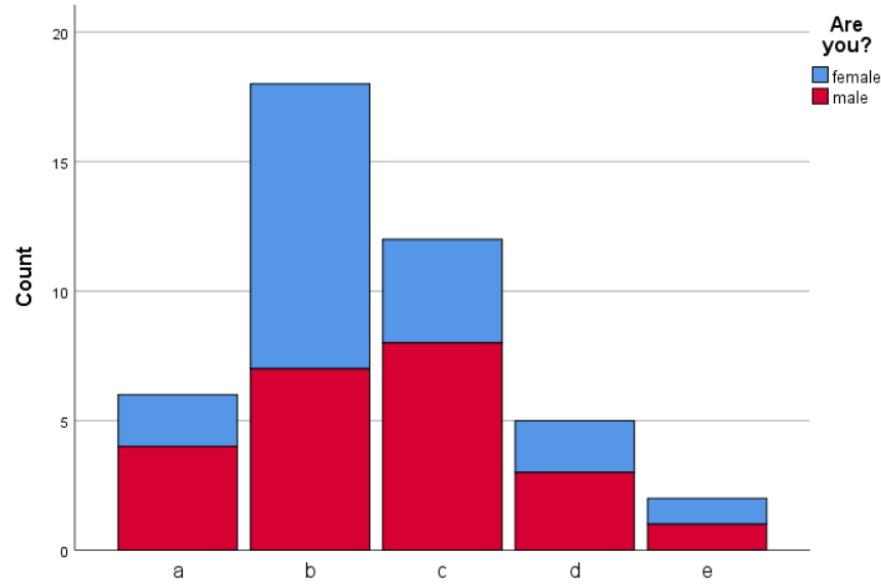
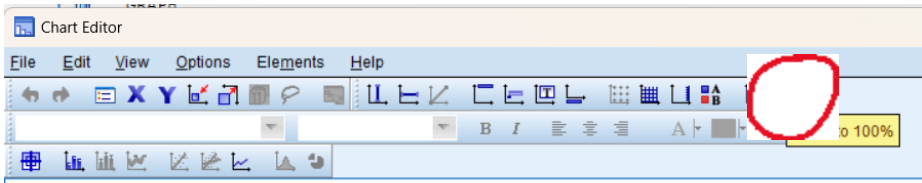
Template

Use chart specifications from:

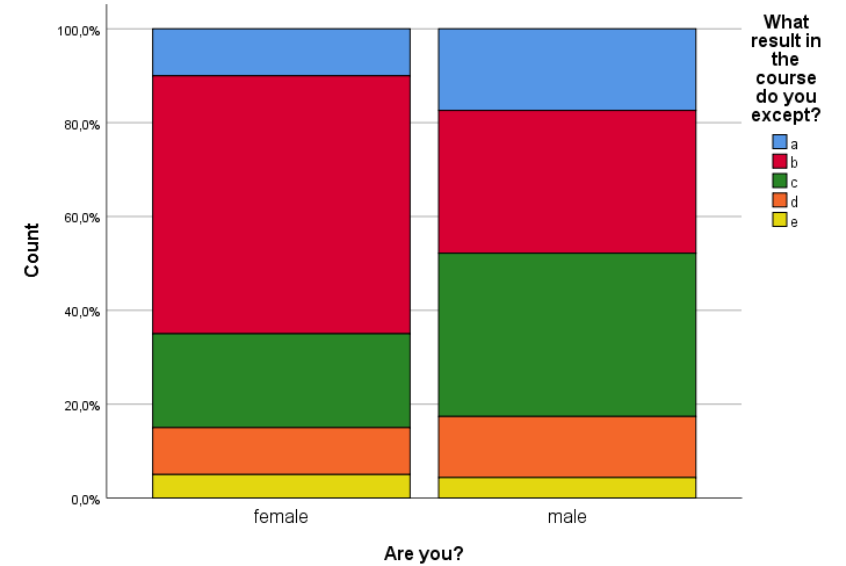
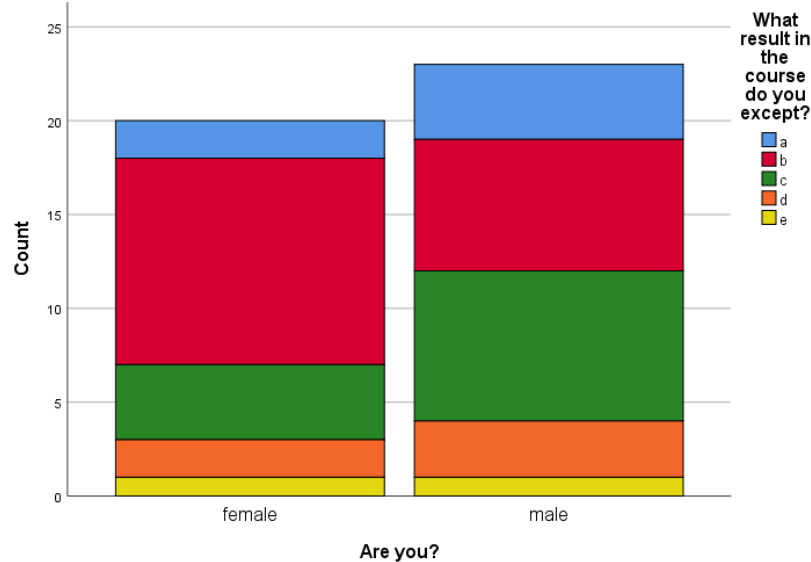
File...

OK Paste Reset Cancel Help

# Stacked bar chart



What result in the course do you expect?



Are you?



Utilities Extensions Window Help

Chart Builder...  
Graphboard Template Chooser...  
Weibull Plot...  
Compare Subgroups  
Dependent Variable Plots

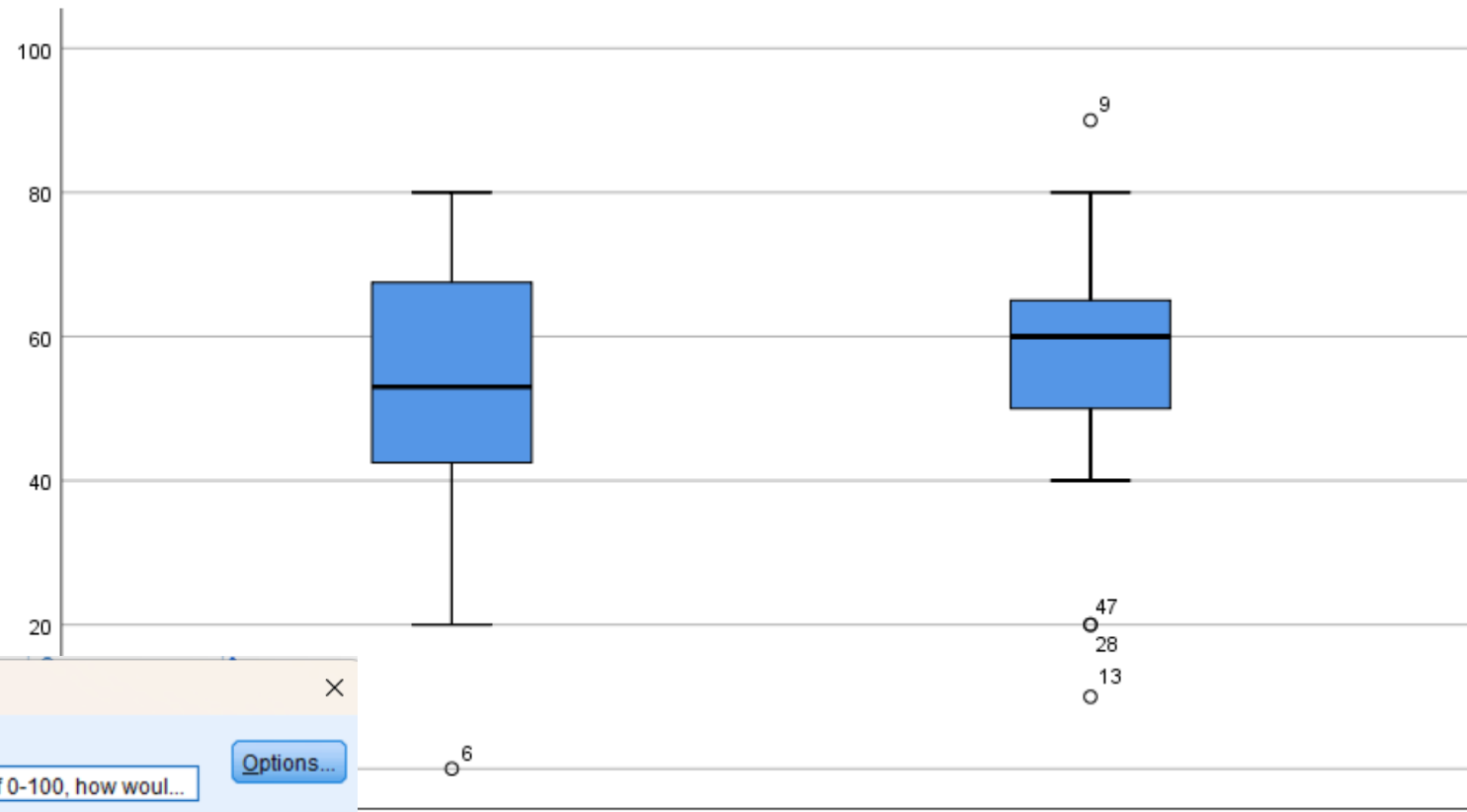
Missing	Columns
None	12
None	12
None	12

Bar...  
3-D Bar...  
Line...  
Area...  
Pie...  
High/Low...

0	Which subjects...	None
0	Which subjects...	None
0	On a scale of 0...	{0, bad}...
1	What result in t...	{1,0, a}...
0	How much stre...	{1, very low
0	How high a leve...	{1, very lo
2	In your opinion,...	None
0	Are you?	{1, fem

TS

scale of 0-100, how would you rate your mathematical ability?



female male

Are you?

Boxplot

Simple  
Clustered

Data in Chart Are

Summaries for groups of cases  
 Summaries of separate variables

Define Cancel Help

Define Simple Boxplot: Summaries for Groups of Cases

Variable: On a scale of 0-100, how woul...

Category Axis: Are you? [Q10\_gender]

Label Cases by:

Panel by

Rows:

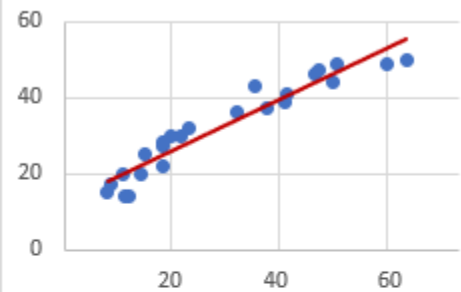
Columns:

Nest variables (no empty rows)  
Nest variables (no empty columns)

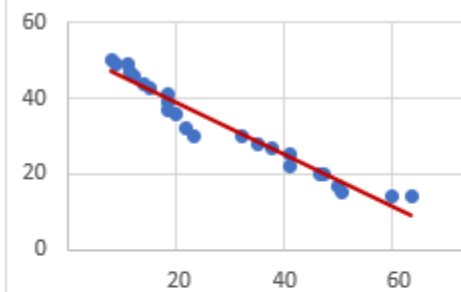
Scatter plot

- Vztah 2 kardinálních proměnných
- korelace
  
- Výška x váha
- Nezaměstnanost x kupní síla

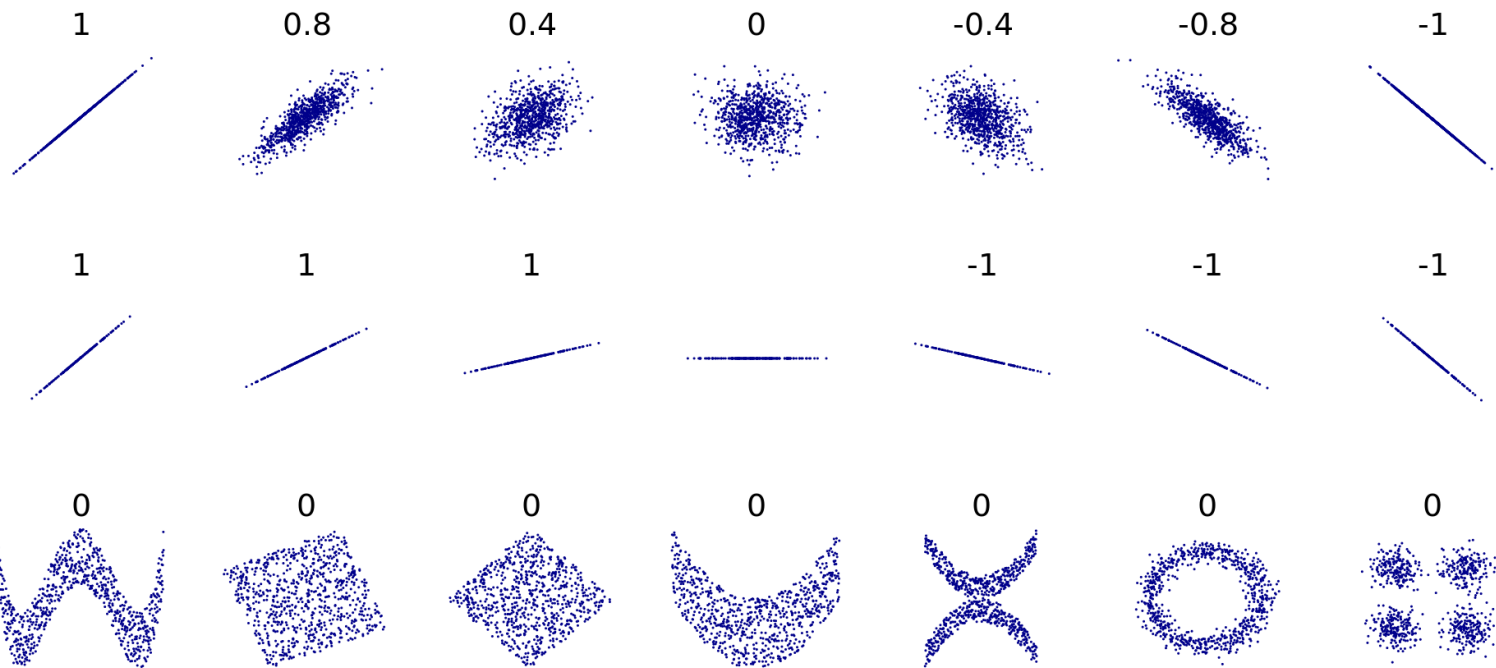
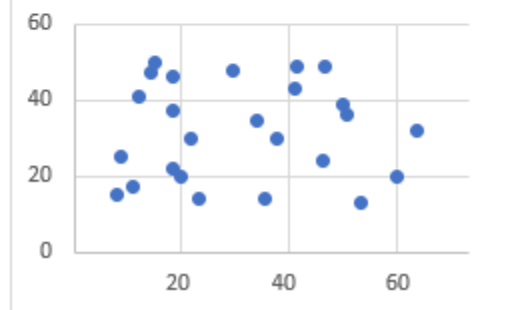
*Positive Correlation*

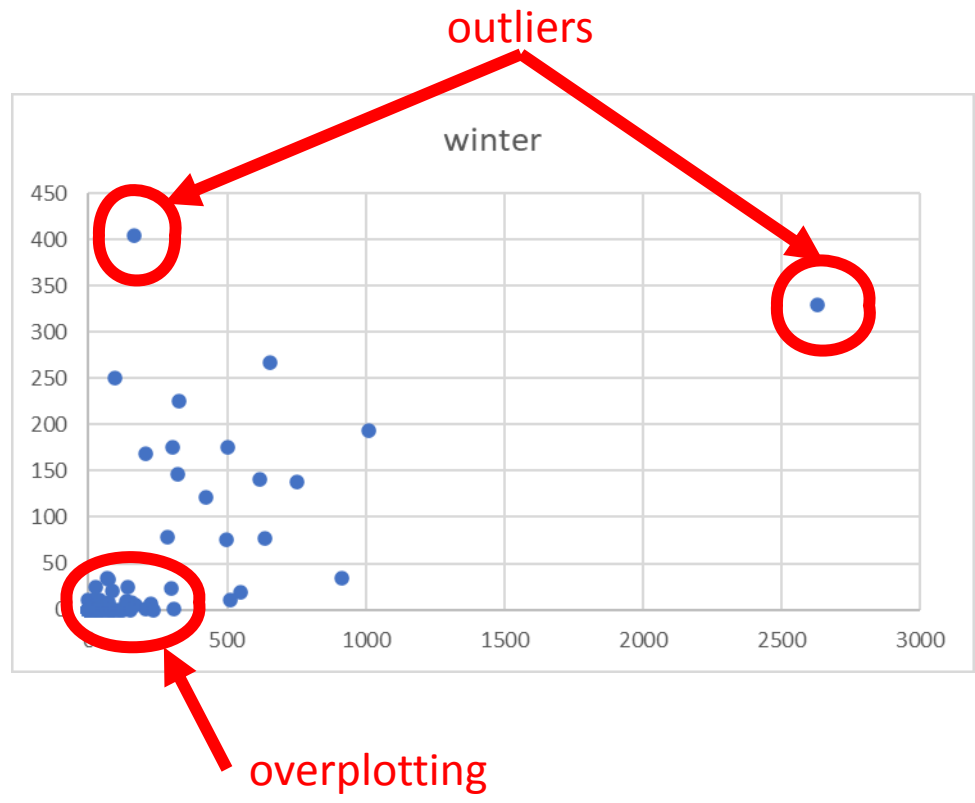


*Negative Correlation*



*No Correlation*





Graphs Utilities Extensions Window Help

Chart Builder...  
Graphboard Template Chooser...  
Weibull Plot...  
Compare Subgroups  
Regression Variable Plots

Legacy Dialogs

	Missing	Columns
None	12	
None	12	
None	12	

Bar...  
3-D Bar...  
Line...  
Area...  
Pie...  
High-Low...  
Boxplot...  
Error Bar...  
Population Pyramid...  
Scatter/Dot...  
Histogram...

0	Which subjects...	None
0	Which subjects...	None
0	On a scale of 0...	{0, bad}...
1	What result in t...	{1,0, a}...
0	How much stre...	{1, very low
0	How high a leve...	{1, very low
2	In your opinion,...	None
0	Are you?	{1, female}

Simple Scatterplot

id  
How far from Brno i...  
Calculated distance...  
Have you passed a...  
Which subjects did ...  
Which subjects did ...  
How much stress d...  
How high a level of ...  
In your opinion, wha...  
Are you? [Q10\_gend...

Y Axis:  
What result in the course do yo...

X Axis:  
On a scale of 0-100, how woul...

Set Markers by:

Label Cases by:

Panel by

Rows:

Nest variables (no empty rows)

Columns:

Nest variables (no empty columns)

Template

Use chart specifications from:  
File...

OK Paste Reset Cancel Help

Titles...  
Options...

Scatter/Dot

Simple Scatter  
Matrix Scatter  
Simple Dot  
Overlay Scatter  
3-D Scatter

Define Cancel Help

