

Statistická analýza dat v psychologii I

Úvod do práce s jamovi

**Semináře 2-4**

---

# Popisné statistiky a z-skóry



# Obsah

---

- Četnosti
  - Míry centrální tendence a variability
  - Zobrazení rozložení
    - Normální rozložení
  - Z-skóry
  - (Robustní deskriptivy v modulu WALRUS)
-

# Četnosti

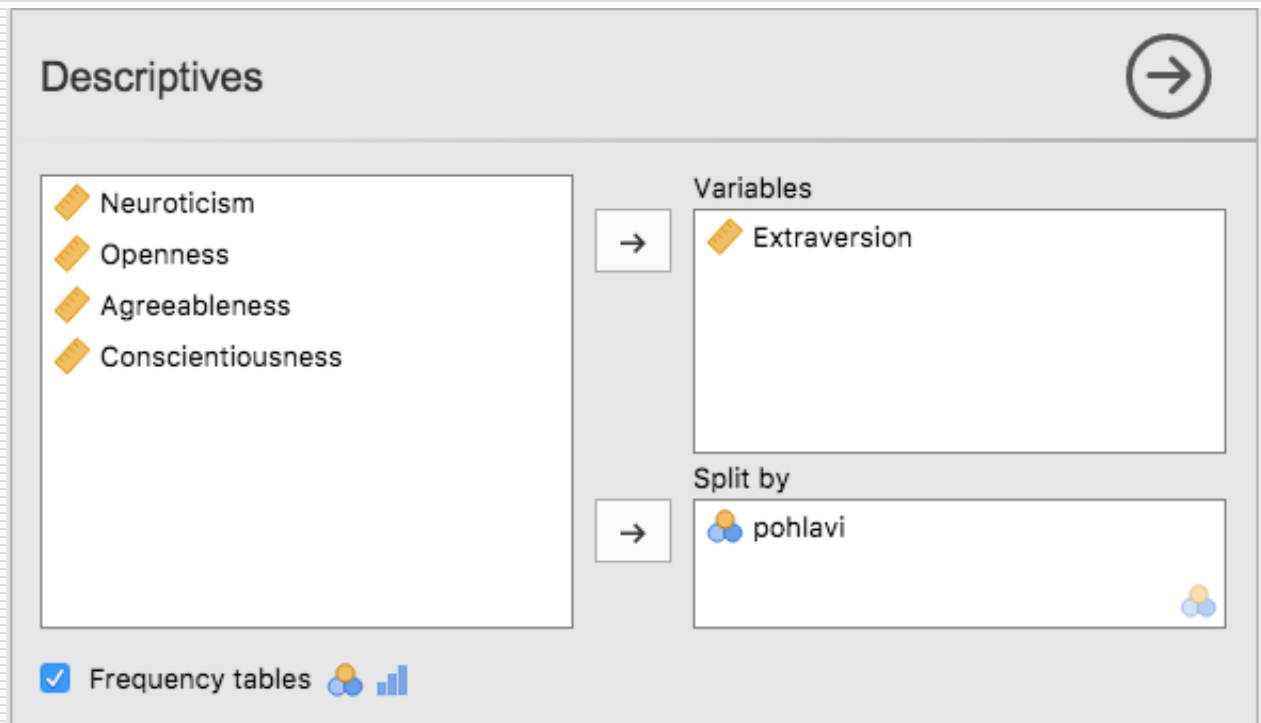
---

- Exploration – Descriptives
    - Četnosti lze zobrazit zaškrtnutím možnosti N ve Statistics
  
  - Př.: Po „měření“ extraverze nás zajímá, kolik bylo ve vzorku mužů a kolik žen.
-


# Četnosti

---

- Vybereme cílovou proměnnou (Extraverze) a rozdělíme pomocí pohlaví



The screenshot shows the 'Descriptives' dialog box in SPSS. The 'Variables' list contains 'Extraversion'. The 'Split by' list contains 'pohlavi'. The 'Frequency tables' checkbox is checked. The 'Descriptives' title bar has a right-pointing arrow icon.


Descriptives 

Variables

- Extraversion

Split by

- pohlavi

Frequency tables 

# Četnosti

---

- Ve vzorku je 202 mužů a 298 žen. Nenašel se nikdo, kdo by pohlaví nevyplnil.

Descriptives		
	pohlavi	Extraversion
N	muž	202
	žena	298
Missing	muž	0
	žena	0

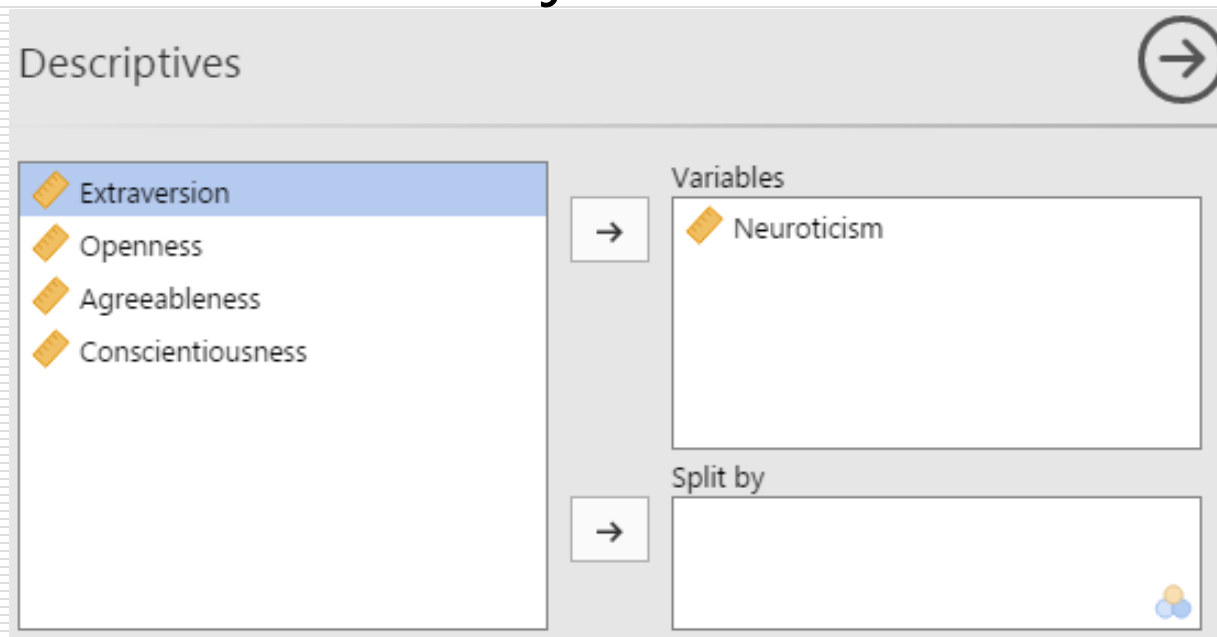
---

# Míry centrální tendence a variability

---

## □ Exploration – Descriptives

- Vybereme proměnnou, jejíž centrální tendence a variabilita nás zajímá



# Míry centrální tendence a variability

---

- Ve Statistics zadáme konkrétní míry centrální tendence a variability, které nás zajímají

The screenshot shows the 'Statistics' dialog box in SPSS. The 'Statistics' dropdown menu is open, showing the following options:

- Sample Size**
  - N
  - Missing
- Percentile Values**
  - Quartiles
  - Cut points for  equal groups
- Dispersion**
  - Std. deviation
  - Minimum
  - Variance
  - Maximum
  - Range
  - S. E. Mean
- Central Tendency**
  - Mean
  - Median
  - Mode
  - Sum
- Distribution**
  - Skewness
  - Kurtosis

# Míry centrální tendence a variability

---

- Svůj zrak upřeme vpravo, kde se nám zobrazil výstup, a obdivujeme ladné křivky APA formátu

## Descriptives

### Descriptives

	Neuroticism
Mean	2.83
Median	2.80
Standard deviation	0.453
Minimum	1.65
Maximum	4.38
25th percentile	2.52
50th percentile	2.80
75th percentile	3.10

---



# Zobrazení rozložení

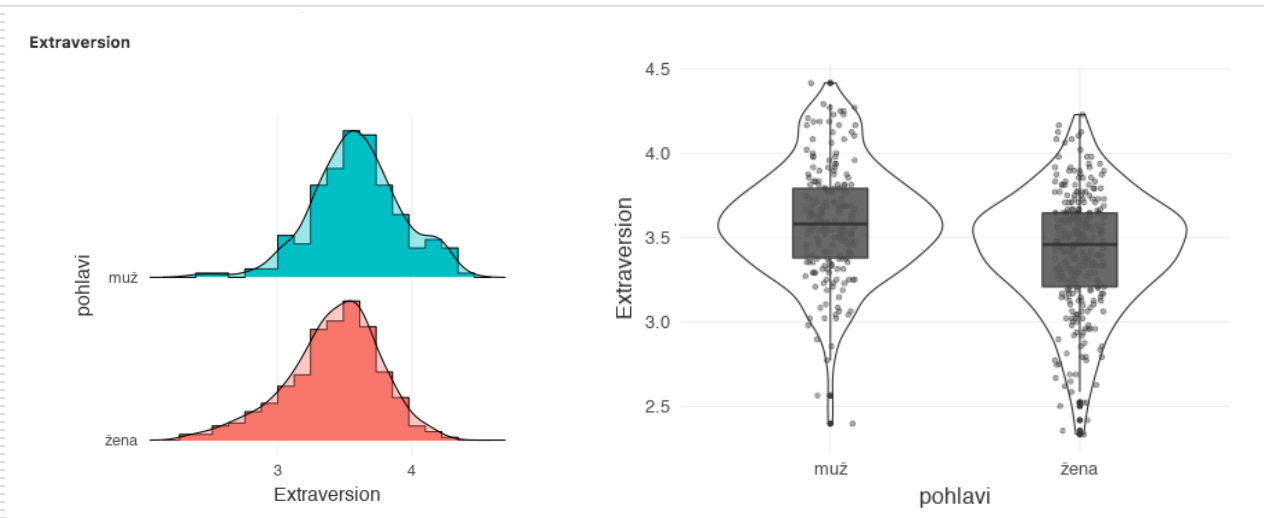
---

## Exploration – Descriptives

- Rozložení lze zobrazit v sekci Plots, která nabízí:
    - Histogramy spolu s možností na nich zobrazit hustotu dat (density)
    - Boxploty, opět s možností zobrazit hustotu dat skrze violin plot a s možností do grafu přidat jednotlivé případy jako body
    - Sloupcové grafy
-

# Zobrazení rozložení

- Př. Zajímá nás rozložení extraverze u mužů a u žen.
- Vybereme proměnnou, rozdělíme podle pohlaví a zvolíme některou z možností



# Normální rozložení

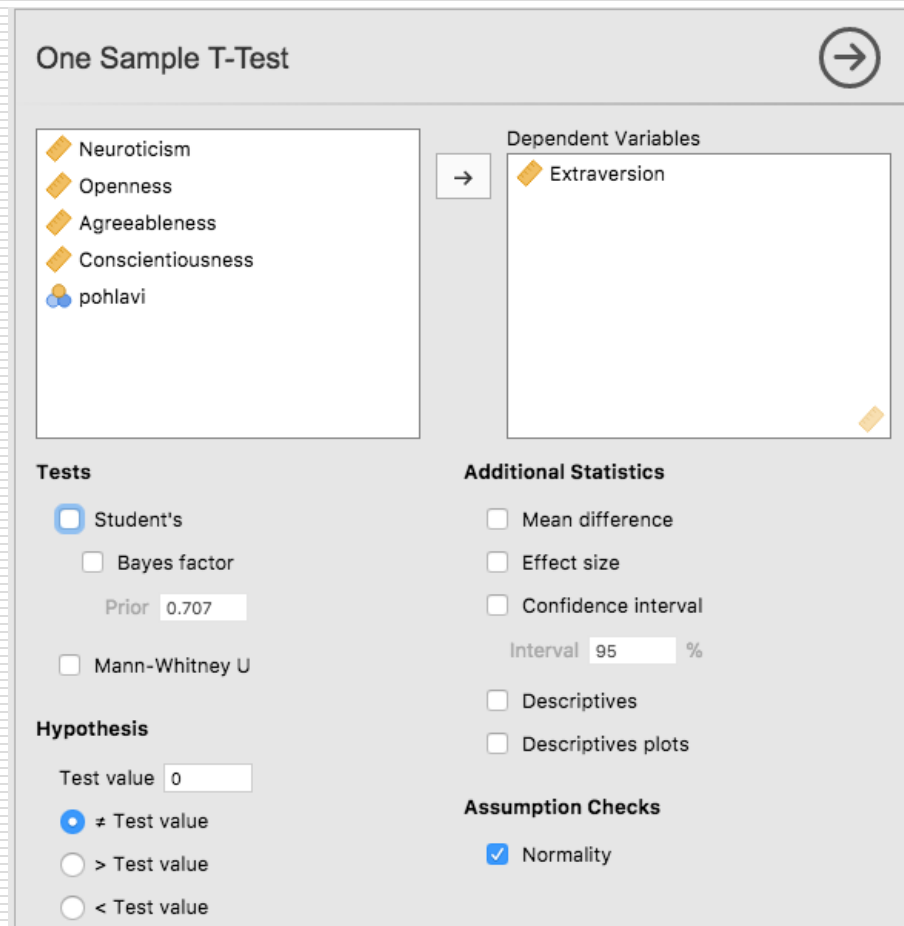
---

- jamovi nabízí několik způsobů ověření normality rozložení:
  - Vizually pomocí histogramů
  - Pomocí hodnot zešikmení (*skewness*) a špičatosti (*kurtosis*) v panelu Statistics u Descriptives

Descriptives	
Extraversion	
Skewness	-0.388
Kurtosis	0.528

# Normální rozložení

- Pomocí statistických testů v rámci ověřování předpokladů analýz, které normalitu očekávají
- Zvolíme např. jednovýběrový  $t$ -test v nabídce  $t$ -tests a zaškrtneme *Normality*



The screenshot shows the 'One Sample T-Test' dialog box in SPSS. The 'Dependent Variables' list contains 'Extraversion'. The 'Tests' section has 'Student's' selected. The 'Hypothesis' section has '≠ Test value' selected. The 'Assumption Checks' section has 'Normality' checked.

One Sample T-Test

Dependent Variables

- Neuroticism
- Openness
- Agreeableness
- Conscientiousness
- pohlaví

Extraversion

Tests

- Student's
- Bayes factor
- Prior 0.707
- Mann-Whitney U

Hypothesis

Test value 0

- ≠ Test value
- > Test value
- < Test value

Additional Statistics

- Mean difference
- Effect size
- Confidence interval
- Interval 95 %
- Descriptives
- Descriptives plots

Assumption Checks

- Normality

# Normální rozložení

---

- Hodnota  $p$  menší než .05 se tradičně interpretuje jako ukazatel odchylky od normálního rozložení.
- Tento postup je však **nejméně spolehlivý**, samotný test pro ověření normality nikdy nestačí.

	W	p
Extraversion	0.986	<.001

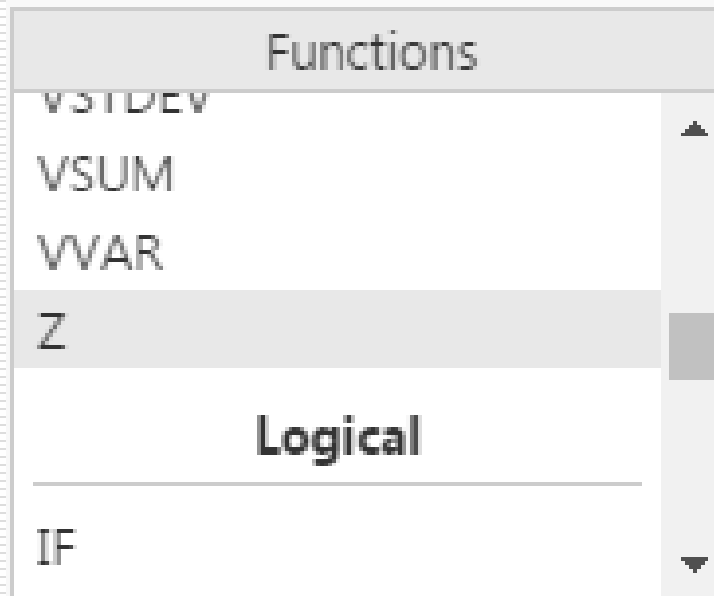
Note. A low p-value suggests a violation of the assumption of normality

---

# Z-skóry

---

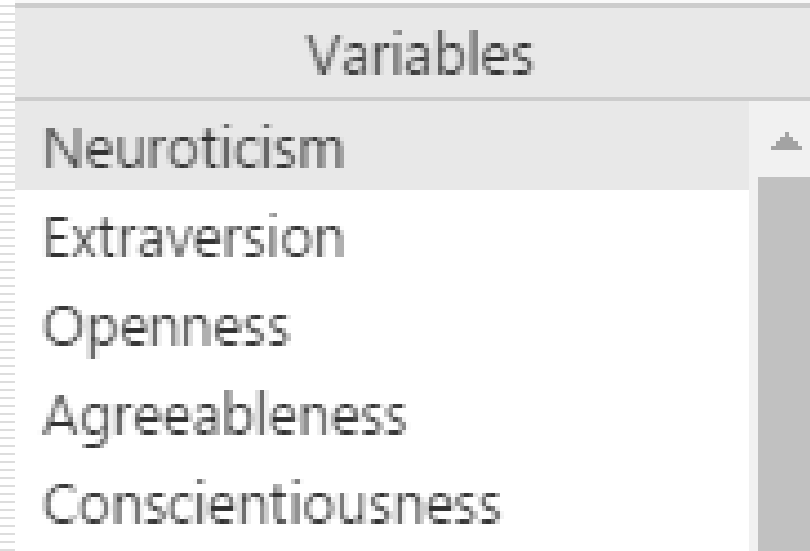
- Z-skóry vytvoříme pomocí New computed variable (viz předchozí prezentace)
  - Ve Functions vybereme dvojklikem funkci Z



# Z-skóry

---

- Ve Variables vybereme proměnnou, jejíž Z-skóry chceme spočítat



- Vzorec tedy vypadá: `= Z(Neuroticism)`
-

# Z-skóry

---

- Odklikneme fajfku vpravo nahore



- V naší datové matici se zobrazí nová proměnná obsahující z-skóry

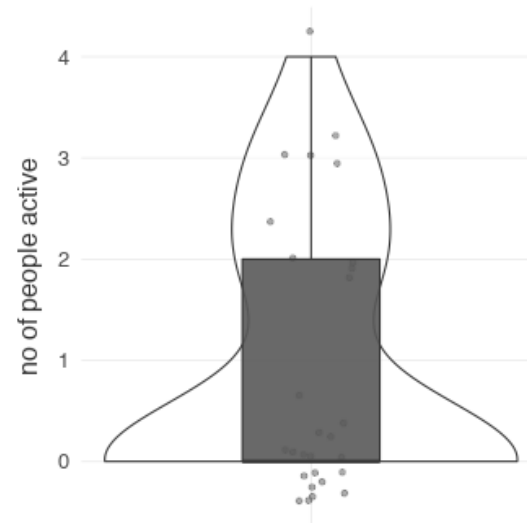
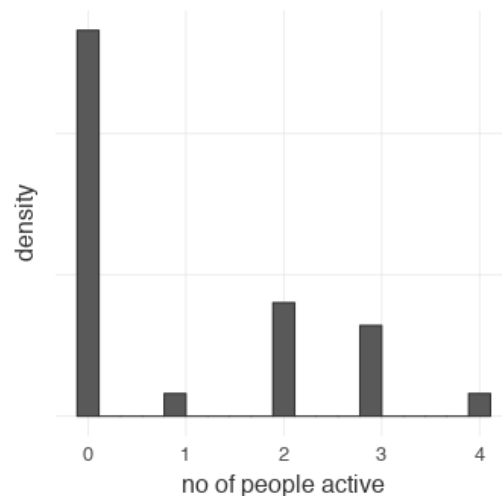
$f_x$ Z_Neuroti...
-0.767
-0.491
-0.031
0.153
0.429



# Robustní deskriptivy

---

- V případě velmi atypicky rozložených dat přestává být průměr jakožto ukazatel centrální tendence použitelný (zde  $M = 0,96$ )



# Robustní deskriptivy

---

- Pomoci může modul Walrus, který lze do programu přidat z jamovi library v menu Modules.
- Kromě deskriptiv nabízí také paletu robustních (= odolných vůči porušení předpokladů) alternativ k tradičním testům



## Walrus 1.0.2

Jonathon Love, Patrick Mair

A toolbox of common robust statistical tests, including robust descriptives, robust t-tests, and robust ANOVA. Walrus is based on the [WRS2 package](#) by Patrick Mair, which is in turn based on the scripts and work of Rand Wilcox. These analyses are described in depth in the book [Introduction to Robust Estimation & Hypothesis Testing](#). Walrus is also [available for R](#) from CRAN.

INSTALLED

# Robustní deskriptivy

## □ Walrus – Robust descriptives

- Vanilla mean poskytuje aritmetický průměr
- Trimmed mean je průměr počítaný na hodnotách, které by v proměnné zůstaly po ostřížení (*trim*) horních a dolních několika procent extrémů
- Tato procenta se specifikují v Trim proportion, používá se např. 25 % (0.25)

Robust Descriptives

		SE	
no of people active	Mean	0.964	0.249
	Trimmed mean	0.500	0.366

# Robustní deskriptivy

---

## □ Walrus dále nabízí:

- Medián
- Winsorized mean funguje podobně jako trimmed s tím rozdílem, že extrémní konce nahrazuje nejnižší a nejvyšší zbývající hodnotou

### Robust Descriptives

		SE	
no of people active	Winsorized mean	0.750	0.290
	Median	0.00	0.388

---