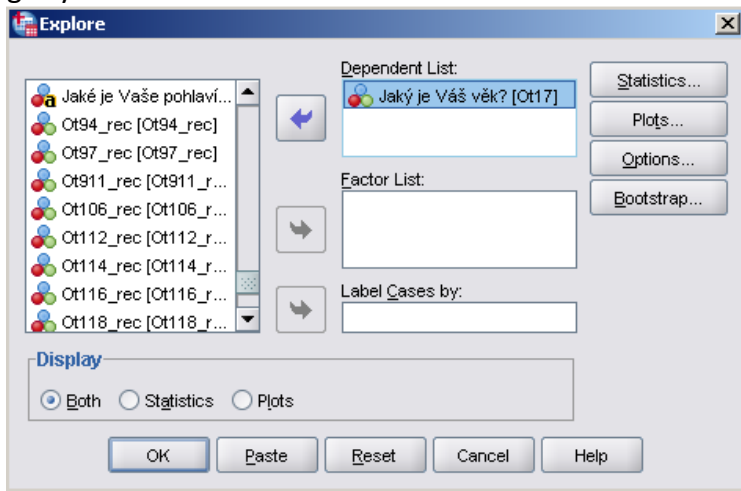


Ke zpracování kardinálních proměnných se hodí procedura EXPLORE

Lze ji použít pro jednorozměrnou i dvourozměrnou analýzu

Nevytváří tabulky hodnot proměnné (jako frequencies) ale zobrazuje souhrnné statistiky a grafy - Zadání: **ANALYZE -> DESCRIPTIVE STATISTICS -> EXPLORE**



V „Plots“ je možné nastavovat další grafy. Standardně je výstupem tzv. Box-plot a Stem and Leaf (ten je ale nepříliš přehledný). Lze také zvolit histogram (není součástí default nastavení)

Výstupy z EXPLORE:

Tabulka s hlavními statistikami:

Pro průměr máme vypočtenou také standardní chybu, která vyjadřuje spolehlivost dat. Za předpokladu prostého náhodného výběru bychom mohli říci, že průměr v základní populaci (z níž je náš soubor vzorkem) by ležel s 95% pravděpodobností v intervalu ± 2 standardní chyby průměru. Tedy zhruba mezi 22,99 a 23,29 roku. Všimněte si, že interval spolehlivosti je velmi úzký – je to tím, že pracujeme s velmi rozsáhlým souborem – při velikosti téměř 3000 respondentů je výběrová chyba poměrně malá.

Dále máme tzv. robustní průměr – bez 5 % odlehých případů. (například onen 100letý šprýmař zde vypadl a průměr se snížil)

Vedle tzv. měr centrální tendence (průměr, medián příp. modus) stojí u kardinálních proměnných vždy za povšimnutí míry variability. Rozptyl a směrodatná odchylka poukazují na to, jak moc jsou data rozházená kolem průměru. Malá hodnota = všichni v podobném věku, velká hodnota = velice rozmanité stáří studentů)

Descriptives

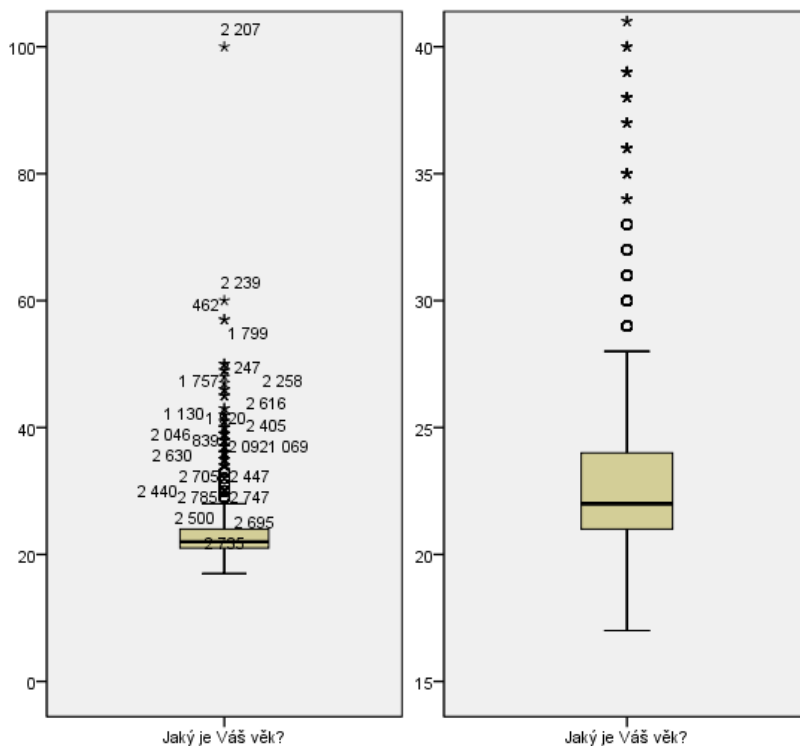
		Statistic	Std. Error	
Jaký je Váš věk?	Mean	23,14	,076	
	95% Confidence Interval for Mean	Lower Bound	22,99	
		Upper Bound	23,29	
	5% Trimmed Mean	22,68		
	Median	22,00		
	Variance	16,100		
	Std. Deviation	4,012		
	Minimum	17		
	Maximum	100		
	Range	83		
	Interquartile Range	3		
	Skewness	5,190	,047	
	Kurtosis	61,778	,093	

Explore umí: Krabicový graf - je užitečný:

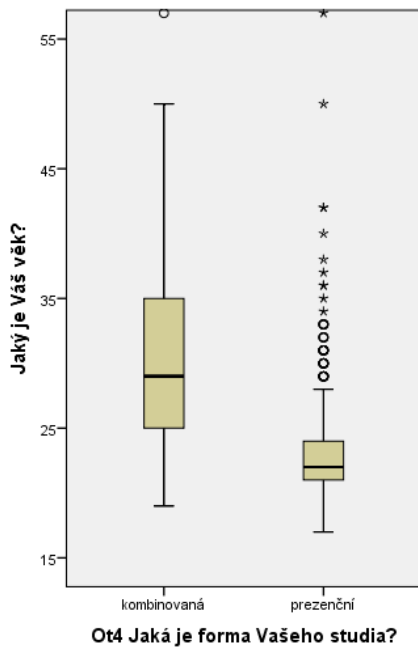
- když chceme rychle identifikovat odlehlé případy (čísla případů jsou přímo v grafu – pak by stačilo najít daný řádek v matici a zkontrolovat, zda respondent nevyplnil podivně i jiné otázky, nebo nejde o překlep)
- když chceme získat rychlý přehled o rozložení hodnot (ne jen průměr či medián, ale i to, zda je rozložení symetrické, špičaté, šikmé...)
- když chceme porovnat více rozložení navzájem (třeba věk podle typu studia) Tady už se jedná o dvourozměrnou analýzu a musíme použít kategorizovanou proměnnou vloženou do „factor list“

Vpravo je tentýž graf s rozumným měřítkem. Tlustá čára uprostřed znázorňuje medián, krabice je definována 25. a 75. percentilem. Uvnitř krabice leží 50 % případů a její výška je dána tzv. interkvartilovým rozpětím.

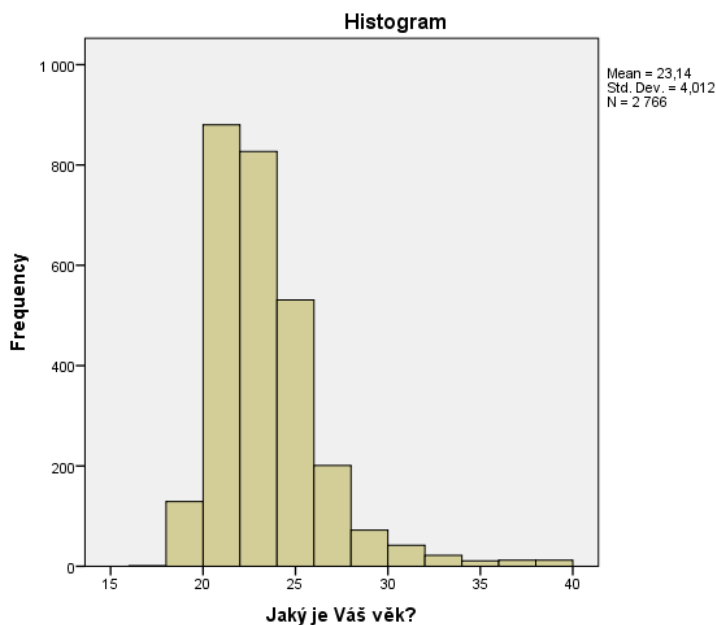
Blíže k tomu viz: <http://cs.wikipedia.org/wiki/Boxplot> nebo <http://www.eistat.cz/popis/boxplot/index.htm>



Krabicový graf použitý pro srovnání – vidíme nejen vzájemnou pozici mediánů, ale můžeme srovnat i základní charakteristiky rozložení

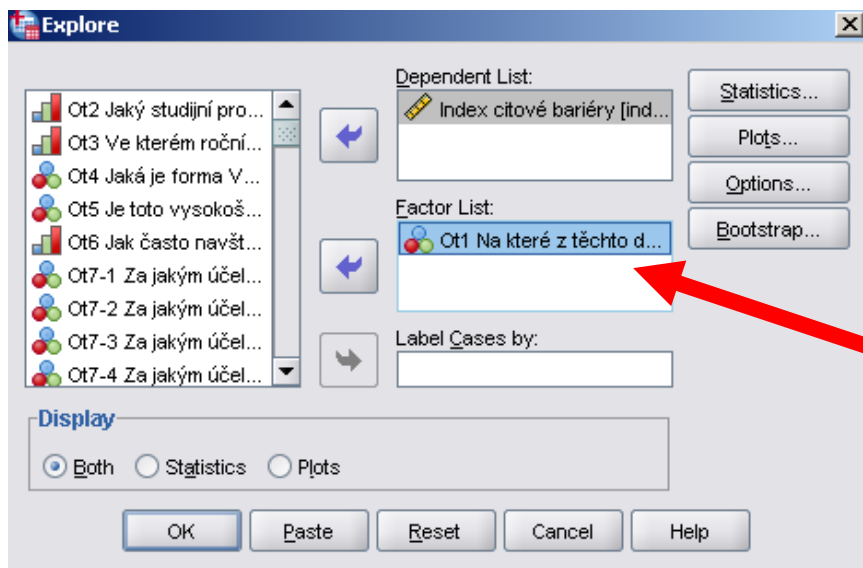


Explore umí: Histogram - je ještě přehlednější a zobrazuje celé rozložení proměnné (zde opět měřítko upraveno na 15 – 40 let). Histogram lze získat také v proceduře Frequencies. (od sloupcového grafu se liší tím, že má lineární osu x - nezkresluje tedy vzdálenosti).



Dále je procedura explore užitečná ještě pro zjišťování, zda je proměnná normálně rozložena. K tomu jsou určeny speciální tzv. kvantilové grafy (Q_Q Plot) a testy (kolmogorov-smirnov). Zadat lze v submenu „Plots“ -> „normality plots with tests“

EXPLORE V DVOUROZMĚRNÉ ANALÝZE

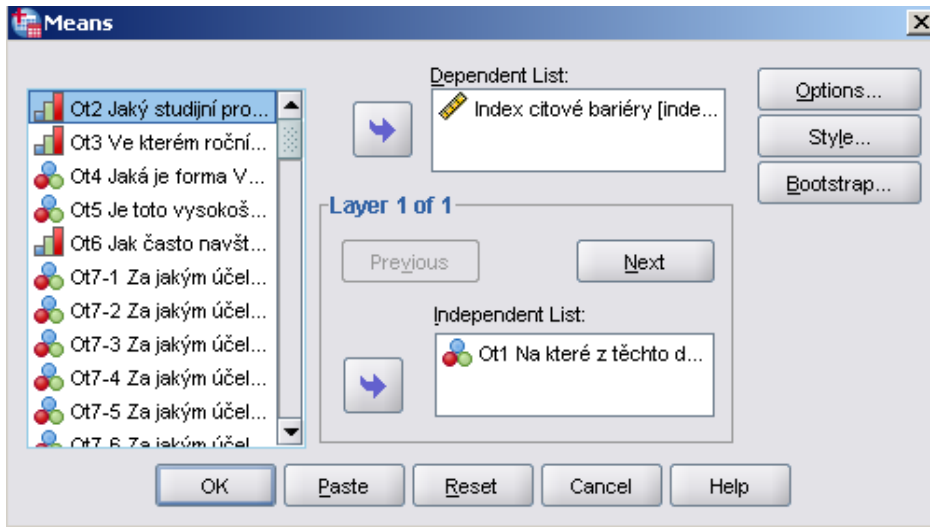


Kategorizovaná proměnná, podle které třídím výsledky

Výsledek: tytéž statistiky spočítané zvlášť pro studenty FF a PŘF

Descriptives				Statistic	Std. Error
Ot1 Na které z těchto dvou následujících fakult Masarykovy univerzity v současné době studujete?					
Index citové bariéry	Filozofická fakulta	Mean		48,2386	,18236
		95% Confidence Interval for Mean	Lower Bound	47,8809	
			Upper Bound	48,5963	
		5% Trimmed Mean		48,5986	
		Median		49,0000	
		Variance		55,334	
		Std. Deviation		7,43868	
		Minimum		16,00	
		Maximum		60,00	
		Range		44,00	
		Interquartile Range		10,00	
		Skewness		-,703	,060
		Kurtosis		,380	,120
	Přírodovědecká fakulta	Mean		47,9601	,26432
		95% Confidence Interval for Mean	Lower Bound	47,4413	
			Upper Bound	48,4789	
		5% Trimmed Mean		48,3275	
		Median		49,0000	
		Variance		57,779	
		Std. Deviation		7,60127	
		Minimum		22,00	
		Maximum		60,00	
		Range		38,00	
		Interquartile Range		11,00	
		Skewness		-,630	,085
		Kurtosis		-,034	,170

**Jednodušší varianta pro srovnání průměrů je procedura MEANS:
Analyze -> Compare means -> Means**



Zadání je téměř totožné, terminologie mírně jiná (místo „factor“ „independent“), vždy jde ale o třídění průměrů spojité (kardinální) proměnné podle kategorií kategorizované proměnné.

Výsledek obsahuje tři základní věci: průměry (informaci o centrální tendenci), počty případů (informaci o velikosti souboru) a směrodatnou odchylku (informaci o rozmanitosti v dané skupině).

Report

Index citové bariéry

Ot1 Na které z těchto dvou následujících fakult Masarykovy univerzity v současné době studujete?	Mean	N	Std. Deviation
Filozofická fakulta	48,2386	1664	7,43868
Přírodovědecká fakulta	47,9601	827	7,60127
Total	48,1461	2491	7,49267