

Nové možnosti rozvoje vzdělávání na Technické univerzitě v Liberci

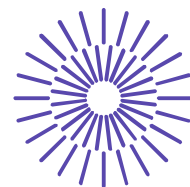
Specifický cíl A3: Tvorba nových profesně zaměřených studijních programů

NPO_TUL_MSMT-16598/2022



Téma 1: Příklad 2 – prosté rozdělení četností (číselná data)

Ing. Vladimíra Hovorková Valentová, Ph.D.



Zadání příkladu:

V distribučním centru jsou vedeny záznamy o počtu položek v odeslaných zásilkách. Roztřídte data (viz níže) o počtu položek v jednotlivých zásilkách do vhodné tabulky a vhodným způsobem graficky znázorněte.

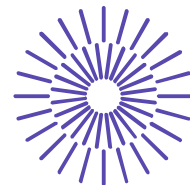
4	6	6	7	5
4	4	8	8	5
3	2	8	5	4
1	3	4	4	3
2	4	1	4	2
2	4	2	5	8
2	3	2	3	6
5	4	5	4	4

Řešení příkladu:

Sledovanou proměnnou je počet položek v zásilce. Tato proměnná je vyjádřena číselně a její hodnoty je možné hierarchicky seřadit a určit i rozdíly v jednotlivých hodnotách v kusech. Jde tedy o číselnou kardinální proměnnou. Hodnoty kardinální proměnné je možné třídit do tabulky prostého rozdělení četností nebo tabulky intervalového rozdělení četností. Záleží na tom, zda je daná proměnná spojitá nebo nespojitá, a také na tom, kolika variant nabývá. Vzhledem k tomu, že zde můžeme zjistit 8 variant a proměnná je nespojitá, pak vhodnou tabulkou je tabulka prostého rozdělení četností.

U kardinální proměnné má smysl uvádět v tabulce rozdělení četností absolutní, relativní i kumulativní četnosti výskytu jednotlivých variant.

Absolutní četnosti výskytu jednotlivých variant (n_i) doplníme tak, že spočítáme, kolikrát se na seznamu netříděných dat vyskytuje 1 položka v zásilce, potom kolikrát jsou 2 položky v zásilce, následně kolikrát odešly 3 položky v zásilce atd. V tabulce vidíme, že 1 položka byla v celkem dvou zásilkách. Dále pak 2 položky obsahovalo 7 zásilek, 3 položky byly zaznamenány v pěti zásilkách, 4 položky ve 12 zásilkách atd.

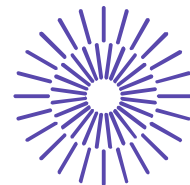


Počet položek v zásilce (x_i)	Počet zásilek (n_i)	Podíl zásilek (p_i)	Kumulativní četnosti absolutní	Kumulativní četnosti relativní
1	2	0,050	2	0,050
2	7	0,175	9	0,225
3	5	0,125	14	0,350
4	12	0,300	26	0,650
5	6	0,150	32	0,800
6	3	0,075	35	0,875
7	1	0,025	36	0,900
8	4	0,100	40	1,000
Celkem	40	1,000	x	x

Relativní četnosti (p_i) pak vypočítáme podle: $p_i = \frac{n_i}{n}$. Tj. $p_1 = \frac{n_1}{n} = \frac{2}{40} = 0,05$, dále $p_2 = \frac{n_2}{n} = \frac{7}{40} = 0,175$ atp. Relativní četnosti je možné interpretovat v procentuálním vyjádření, tj. 5,0 % zásilek obsahuje 1 položku, 17,5 % zásilek obsahuje 2 položky atd.

Kumulativní četnosti vznikají postupným načítáním četností absolutních, resp. relativních. Např. kumulativní četnost absolutní ve třetím řádku tabulky znamená, že 14 zásilek obsahuje maximálně 3 položky. A kumulativní četnost relativní ve čtvrtém řádku tabulky vyjadřuje, že 65 % zásilek obsahuje nejvýše 4 položky.

Vhodným grafem pro znázornění hodnot kardinální proměnné je spojnicový graf, označovaný jako polygon rozdělení četností. Postup třídění dat do tabulky prostého rozdělení četností a grafické znázornění si nyní představíme v programu SPSS. Data zadáme následujícím způsobem (zkráceno, aby obrázek nezabral celou stránku):



	Pocet_polozek
1	4
2	4
3	3
4	1
5	2
6	2
7	2
8	5
9	6
10	4
11	2
12	3
13	4
14	4
15	3
16	4
17	6
18	8
19	8
20	4
21	1
22	2
23	2

... a další data pod sebe do sloupce.

Potom zvolíme posloupnost procedur **Analyze – Descriptive Statistics – Frequencies** a vstupní panel vyplníme takto:

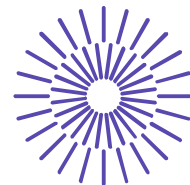


Financováno
Evropskou unií
NextGenerationEU

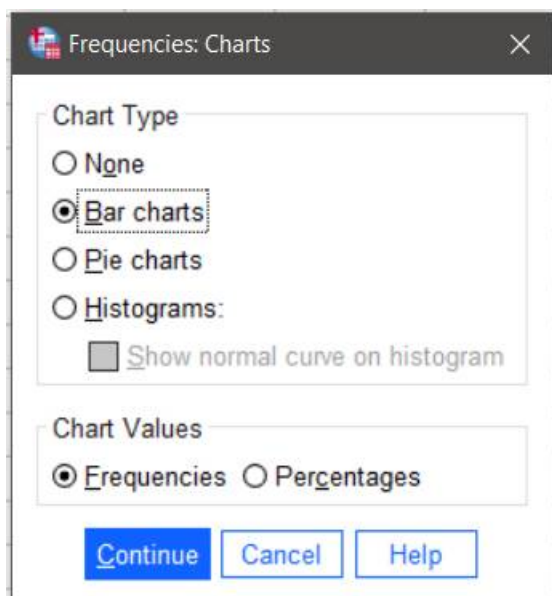


Národní
plán
obnovy

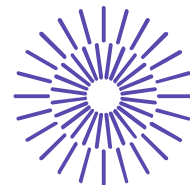
MS
MT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



V tlačítkové volbě vpravo zvolíme tlačítko *Charts ...* Tam zaškrtneme položku **Bar Chart** a **Frequencies**, chceme-li data znázornit pomocí sloupcového grafu s uvedením absolutních četností výskytu na ose y a následně jej upravit na polygon četností:

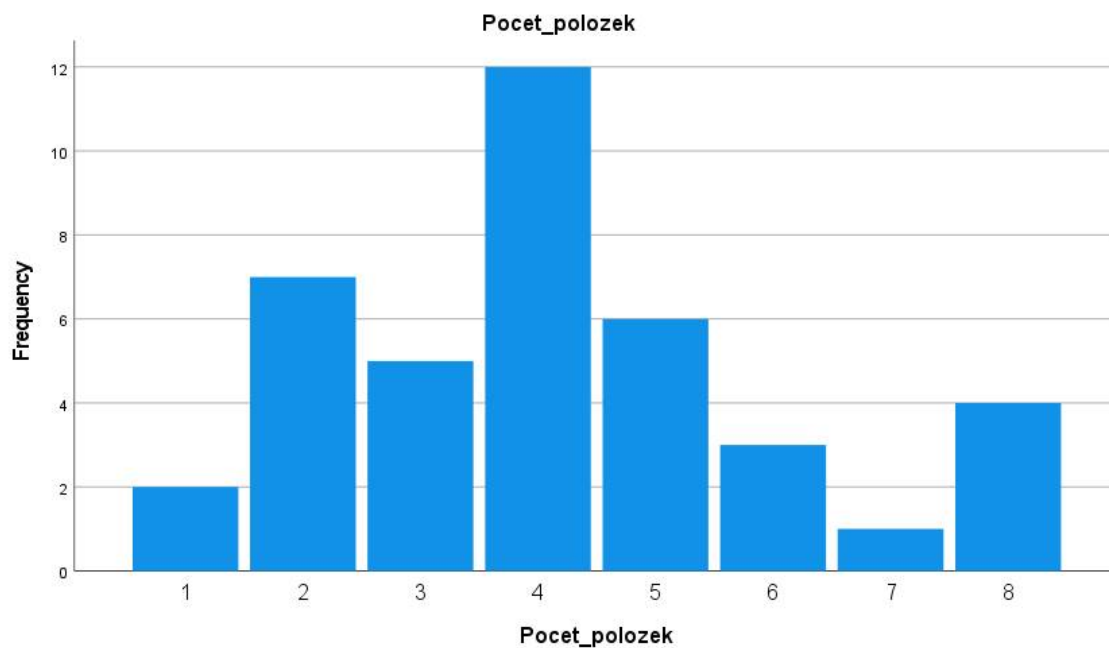


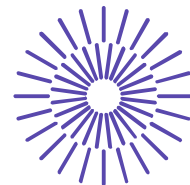
Po potvrzení volby klávesou Continue a následně OK dostaneme následující výstup:



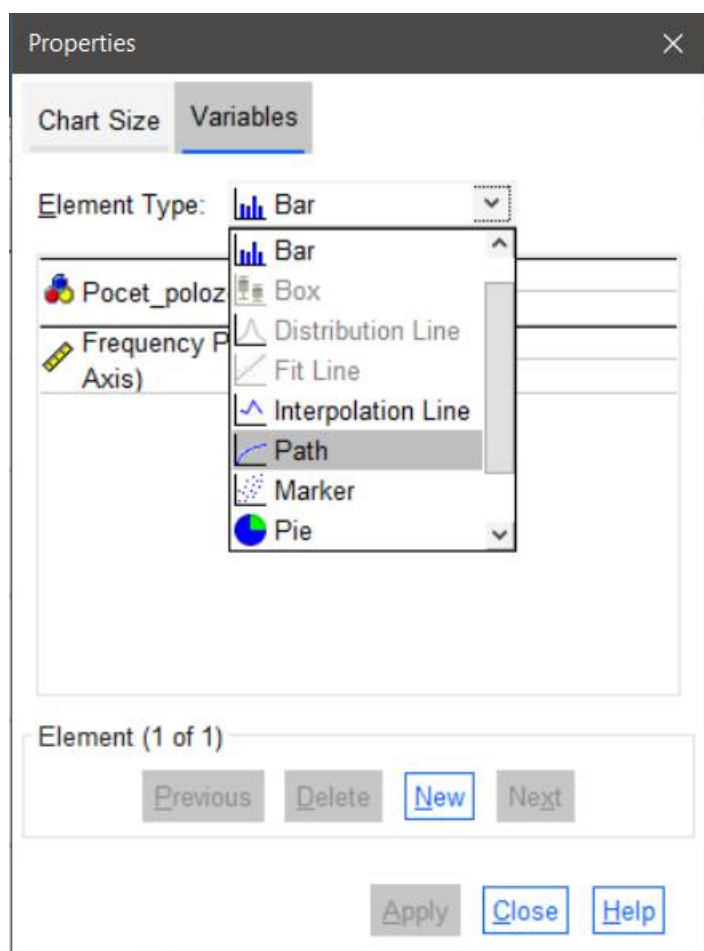
		Pocet_polozek			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	1	2	5,0	5,0	5,0
	2	7	17,5	17,5	22,5
	3	5	12,5	12,5	35,0
	4	12	30,0	30,0	65,0
	5	6	15,0	15,0	80,0
	6	3	7,5	7,5	87,5
	7	1	2,5	2,5	90,0
	8	4	10,0	10,0	100,0
Total		40	100,0	100,0	

Poznámka: Frequency = absolutní četnosti (n_i); Percent = relativní četnosti (p_i) vyjádřené v %; Valid Percent = relativní četnosti (p_i) vyjádřené v % vypočítané ze souboru bez chybějících dat; Cumulative Percent = kumulativní relativní četnosti v %. Varianty znaku jsou řazeny vzestupně.





V základním výstupu najdeme sloupcový graf, který je jistě také použitelný pro kardinální data, neboť i z něj je dobře poznat tvar rozdělení četností. Pokud bychom sloupcový graf chtěli upravit na polygon četností, pak musíme na graf 2x kliknout. Tím se aktivuje k úpravám. Pak zadáme *Edit – Properties* a na záložce *Variables* změním typ graf z Bar na Path, jak ukazuje následující obrázek:



Volbu potvrdíme tlačítkem Apply a následně Close. Sloupcový graf je transformován na polygon četností, jak vidíme na následujícím obrázku:

