

BPM_STAE: Zadání 3. cvičení, 2. kapitola

1. Výzkumný ústav provedl dotazování obyvatel Japonska ohledně jejich oblíbené přísady na pizzu. Následující tabulka ukazuje odpovědi náhodně vybraného vzorku 36 lidí. Při výpočtu využijte datový soubor (Excel k leafletu 03, Příklad 1).

V	PI	B	PI	V	PO	S	PI	V	S	V	S	PI	S	V	V	V	PI
S	S	V	PI	C	V	V	V	C	V	S	PO	V	PI	S	PI	PO	PI

Pozn: vepřové (PI), hovězí (B), drůbeží (PO), mořské plody (S), zelenina a ovoce (V), sýr (C).

- Rozhodněte, zda je proměnná kvantitativní či kvalitativní. Následně sestavte odpovídající tabulku četností.
 - Vypočítejte relativní četnosti a procentuální zastoupení pro každou kategorii.
 - Určete, jaké procento respondentů uvedlo zeleninu a ovoce, drůbeží maso nebo sýr.
 - Vytvořte Paretův graf, který zobrazí rozložení relativních četností jednotlivých kategorií.
 - Na základě výsledků, co můžete říci o preferencích obyvatel Japonska ohledně přísad na pizzu? Diskutujte, zda výsledky naznačují nějaké kulturní nebo regionální vlivy na jejich výběr přísad.
2. Pracujte s daty (Excel k leafletu 03, Příklad 2), která uvádějí počet televizních přijímačů vlastněných 40 náhodně vybranými domácnostmi v Brně.
- Rozhodněte, zda je proměnná kvantitativní či kvalitativní. Následně sestavte odpovídající tabulku četností.
 - Vypočítejte relativní četnosti a procentní zastoupení pro každou třídu.
 - Vytvořte sloupcový graf zobrazující rozdělení četností.
 - Určete, jaké procento domácností vlastní dva nebo více televizorů.
 - Co může rozložení počtu televizorů v domácnostech naznačovat o životním stylu nebo socioekonomické situaci těchto domácností? Diskutujte, jaké faktory by mohly ovlivnit vlastnictví dvou nebo více televizorů.
3. V roce 2013 byl mezi zaměstnanci společnosti Financial Finesse Inc. proveden průzkum zaměřený na úroveň jejich finančního stresu. Výsledky průzkumu jsou uvedeny v následující tabulce. Při výpočtu využijte datový soubor (Excel k leafletu 03, Příklad 3).

Úroveň finančního stresu	Procento odpovědí
Žádný finanční stres	14
Nějaký finanční stres	63
Vysoký finanční stres	18
Nepřekonatelný finanční stres	5

- Vytvořte koláčový graf znázorňující procentuální zastoupení jednotlivých úrovní finančního stresu.
 - Vytvořte Paretův graf zobrazující procentuální zastoupení úrovní finančního stresu.
 - Jaký je nejčastěji uváděný stupeň finančního stresu mezi zaměstnanci? Na základě grafů vysvětlete, co tento výsledek může znamenat pro zaměstnavatele.
4. Čerpací stanice zaznamenala množství zakoupeného benzínu (v litrech) na denních účtenkách. Tabulka níže uvádí rozložení četností litrů benzínu zakoupených během jednoho dne. Pracujte s daty (Excel k leafletu 03, Příklad 4).

Litrů benzínu	Počet zákazníků
Od 0 do méně než 4	31
Od 4 do méně než 8	78
Od 8 do méně než 12	49
Od 12 do méně než 16	81
Od 16 do méně než 20	117
Od 20 do méně než 24	13

- (a) Kolik zákazníků bylo během tohoto dne obsluženo na čerpací stanici?
- (b) Najděte středy jednotlivých tříd. Mají všechny třídy stejnou šířku? Pokud ano, jaká je tato šířka? Pokud ne, určete různé šířky tříd.
- (c) Sestavte tabulku relativních četností a procentního zastoupení pro každou třídu.
- (d) Jaké procento zákazníků zakoupilo 12 litrů benzínu nebo více?
- (e) Vysvětlete, proč nelze přesně určit počet zákazníků, kteří zakoupili 10 litrů benzínu nebo méně.
- (f) Vytvořte tabulky kumulativních četností, kumulativních relativních četností a kumulativního procentního zastoupení.
- (g) Na základě rozložení množství zakoupeného benzínu, co můžete usoudit o nákupním chování zákazníků této čerpací stanice? Diskutujte, jaké faktory by mohly ovlivnit, že většina zákazníků nakoupila více než 12 litrů benzínu.
5. Pracujte s daty (Excel k leafletu 03, Příklad 5), která uvádějí dobu jednosměrného dojíždění (v minutách) z domova do práce pro náhodný vzorek 50 pracovníků.
- (a) Sestavte tabulku četností pro následující třídy: 0–9, 10–19, 20–29, 30–39 a 40–49 minut.
- (b) Spočítejte relativní četnosti a procentní zastoupení pro každou z těchto tříd.
- (c) Vytvořte histogram znázorňující procentní zastoupení jednotlivých tříd dle části b.
- (d) Určete, jaké procento pracovníků z tohoto vzorku dojíždí 30 minut nebo déle.
- (e) Na základě tabulky z části a. sestavte tabulky kumulativní četnosti, kumulativní relativní četnosti a kumulativního procentního zastoupení.
- (f) Na základě výsledků, jaký závěr byste mohli vyvodit o typické době dojíždění pro pracovníky v tomto vzorku? Diskutujte, zda je doba dojíždění pro většinu pracovníků přijatelná, nebo zda by mohla být považována za dlouhou.
6. Následující tabulka ukazuje rozdělení počtu pokut za parkování, které obdrželi studenti na univerzitním kampusu během minulého týdne. Celkem bylo dotázáno 200 studentů.

Počet lístků	0	1	2	3	4
Počet studentů	59	44	37	32	28

Pracujte s daty (Excel k leafletu 03, Příklad 6) a sestrojte dva sloupcové grafy pro tato data. První graf vytvořte bez úpravy osy četností (začíná od nuly). Druhý graf vytvořte s upravenou (ořezanou) osou četností, kde osa začíná od hodnoty 25. Stručně porovnejte oba sloupcové grafy. Zaměřte se na rozdíly ve vnímání dat v závislosti na způsobu zobrazení.

7. Následující tabulka uvádí věkové rozložení řidičů, kteří způsobili autonehody během jednoho týdne ve městě.

Věk (let)	Počet případů	Věk (let)	Počet případů
18 až méně než 20	7	40 až méně než 50	15
20 až méně než 25	12	50 až méně než 60	16
25 až méně než 30	18	60 a více	35
30 až méně než 40	14		

- (a) Vytvořte histogram relativních četností na základě této tabulky, pracujte s Excel k leafletu 03, Příklad 7.
- (b) Jaké závěry můžete vyvodit o věkové skupině řidičů, kteří způsobili nejvíce autonehod? Diskutujte možné důvody, proč právě tato věková skupina dominuje v počtu nehod.
- (c) Popište, jakým způsobem může být tento histogram zavádějící.
- (d) Navrhněte, jak byste mohli změnit rozdělení četností, aby výsledný histogram poskytoval jasnější obraz.

8. Pracujte s daty (Excel k leafletu 03, Příklad 8), která uvádějí uvádějí časy (v minutách), které potřebovalo 50 studentů k dokončení zkoušky ze statistiky. Zkouška měla maximální dobu trvání 75 minut.

- (a) Vytvořte bodový graf (dotplot) pro uvedená data.
- (b) Na základě bodového grafu identifikujte, zda existují nějaké odlehlé hodnoty (outliers), které se výrazně liší od ostatních výsledků.
- (c) Jaký závěr můžete vyvodit o rozložení časů potřebných k dokončení zkoušky? Zvažte, zda většina studentů dokončila zkoušku rychleji, nebo zda potřebovali téměř maximální dobu.

9. Následující data představují daně zaplacené v roce 2014 vzorkem 30 rodin (zaokrouhлено na tisíce).

11	17	35	3	15	9	21	13	5	19
5	12	8	16	10	8	12	6	14	18

- (a) Sestavte Stem and Leaf diagram pro tato data. Uspořádejte listy pro každý stonek ve vzestupném pořadí.
- (b) Vytvořte zdvojený Stem and Leaf diagram. Rozdělte každý stonek na dvě části. První část by měla obsahovat listy 0 až 4, a druhá část by měla obsahovat listy 5 až 9.
- (c) Co nám tato vizualizace může říct o rozložení dat?