

2.2 Tabulky

Tabulky jsou užitečné pro jasnou prezentaci a srovnání většího množství dat. Umožňují prezentovat data s přesností, které nelze dosáhnout prezentací grafem. Tabulky však v určitých případech zprostředkují i určitý grafický dojem.

Rozeznáváme demonstrační a referenční tabulky.

- **Demonstrační tabulky** mají ilustrovat určitý názor. Mohou se objevit v hlavní zprávě nebo na posteru (plakátu). Nemají obsahovat zbytečné podrobnosti. Například sumární tabulky jsou obvykle demonstrační tabulky. Rozeznáváme tabulky textové, statistické a prezentační.
- **Referenční tabulky** obsahují přesná a podrobná data a využívají se pro potenciálně potřebné výpočty. Často se objevují v přílohách prací nebo jako zvláštní datový soubor na internetu.

Mnoho aspektů dobré praxe je společných oběma těmito typům tabulek, ale existují také určité rozdíly. V návrhu tabulky vždy zohledňujeme její účel a také potenciálního čtenáře.

Obecná doporučení pro tvorbu tabulek

Nejdříve uvedeme heslovitě obecná doporučení a směr našeho úsilí při sestavování a popisu tabulek.

- Jasný název.
- Jasně označení řádek a sloupců.
- Pokud to je nutné, přidáváme poznámky.
- Indikujeme zdroj.
- Uvádíme informace o geografické oblasti, časový úsek, jednotky měření atd. Tyto údaje se mohou podle vhodnosti objevit v titulku nebo v názvech řádků a sloupců.
- Je jednodušší porovnávat data ve směru mezi sloupci než mezi řádky. Pokud tabulka porovnává data na dvou osách, pak důležitější osa určuje srovnání sešora dolů.
- Použijeme vhodné řazení sloupců a řádků. Časté řazení je chronologické, abecední, geografické nebo podle velikosti zobrazených dat (často u demonstračních tabulek) nebo podle jiných logických kritérií.
- Celkové údaje (součty nebo průměry, součty četností atd.) jsou dole nebo vpravo, ne nahoře nebo vlevo.
- Data udáváme v tisících či milionech, abychom se vyhnuli číslům s mnoha číslicemi.
- U dlouhých čísel používáme mezery (162 345, ne 162345).

- Mezery mezi řádky jsou větší po čtyřech až pěti řádcích větší tabulky.
- Je vhodné dostatečně zaokrouhlovat čísla a zároveň je vyrovnávat podle desetinné čárky.
- Vhodně používáme vodorovné čáry – čáry mezi řádky nejsou nutné, ale čáry pod názvy sloupců a na konci tabulky dodávají tabulce na přehlednosti.

Doporučení pro demonstrační tabulky

- Málo číselných údajů,
- Vhodně zaokrouhluje – tak, aby si údaje zachovaly rozlišující schopnost. Zásada je, že používáme dvě významná čísla. Lepší je tedy uvést 15 000 lidí než 15 234 lidí.
- Pokud to je vhodné, zmíníme, že součet zaokrouhlených čísel nemusí dát v součtu celkový platný součet.
- Textový souhrn. Klíčové informace je nutné popsat v textu zprávy.

Doporučení pro referenční tabulky

- Zpravidla mnoho čísel, včetně shrnujících údajů.
- Data zaokrouhluje velmi opatrně nebo vůbec ne.

V této kapitole se nebudeme zabývat uspořádáním referenčních tabulek. Tato problematika je důležitá především pro výzkumníky, kteří zakládají a připravují data pro ostatní zájemce o výsledky laboratorních pokusů a statistických šetření.

Příklady demonstračních tabulek

Uvedeme příklady textových tabulek, statistických tabulek pro prezentace výsledků analýzy a prezentaci dat.

Textové tabulky

Často potřebujeme tabulky, které obsahují textová data (viz tab. 2.1). To nastává obvykle tehdy, jestliže pracujeme s kvalitativními údaji. Proto má toto zobrazení význam v kvalitativním výzkumu. Ale při statistické analýze pracujeme i s kvalitativními proměnnými a potřebujeme je jako jiná data zobrazit pomocí tabulky, abychom usnadnili jejich porovnání a vyhodnocení. Takové tabulky se rovněž používají, jestliže chceme ukázat určité příklady, které bývají nějakým způsobem seskupeny. Také se používají, jestliže chceme ukázat kategorie určitých jednotek.

Statistické tabulky

Tyto tabulky mohou prezentovat popisné statistiky, výsledky statistické inferenční analýzy nebo obojí. Popisné statistiky obsahují v tabulce průměry, směrodatné odchylky, variační koeficienty, percentily atd. Tabulky výsledků inferenční statistické analýzy obsahují výsledky statistických testů a výsledky složitějších analýz, které jsou obvykle směsí popisných a inferenčních statistických údajů.

Tab. 2.1 Příklad tabulky dat s kvalitativními údaji

Katalyzátor	Intenzita reakce	Teplota varu (ano/ne)
Organický		
A	žádná	ano
B	vysoká	ne
C	nízká	ne
Anorganický		
A	vysoká	ano
B	střední	ne
C	nízká	ano

Zdroj: modelová data

Tab. 2.2a Tabulka dat o sportování chlapců a dívek

	Zájem o sport		Řádkové součty
	Ano	Ne	
Chlapci	30	36	66
Dívky	11	63	74
Sloupcové součty	41	99	140

Zdroj: modelová data

Zvláštním případem jsou kontingenční tabulky, které obsahují popis závislosti kvalitativních proměnných pomocí četností, relativních četností, sloupcových a řádkových relativních četností.

Kontingenční tabulky

Popisujeme skupinu studentů a vytvoříme tabulku, která nás informuje o zájmu studentů o účast ve školním sportovním oddíle, přičemž nás zajímá vliv pohlaví. Zařazením všech 140 studentů podle zájmu o sport dostaneme četnostní dvojdímní tabulku 2.2a.

Z chlapců, resp. dívek mělo zájem o sportování ve školním oddíle 30, resp. 11 jedinců. V tabulce uvádíme také součty v jejich řádcích a sloupcích. Součty popisují, kolik je celkem chlapců a dívek a kolik studentů dohromady má zájem o sportování ve školním oddíle. Z této tabulky lze jednoduše odvodit tabulku procent, přičemž se můžeme rozhodnout, zda jsou pro nás důležitější procenta vzhledem k celkovým počtům sportujících a nespportujících, nebo vzhledem k celkovým počtům dívek a chlapců. Lze tak získat dvě tabulky (2.2b a 2.2c), které lépe objasňují poměry v datech. Obě tabulky jsme sestrojili z údajů primární tabulky. Každá z obou nových tabulek (2.2b a 2.2c) zdůrazňuje jiný aspekt vztahu proměnné pohlaví a proměnné zájem o sportování.

Tab. 2.2b Tabulka řádkových procent pro data z tabulky 2.2a

	Zájem o sport		Suma (řádková procenta)
	Ano	Ne	
Chlapci	45,5 %	54,5 %	100 %
Dívky	14,9 %	85,1 %	100 %
Suma	29,3 %	70,7 %	100 %

Zdroj: modelová data

Tab. 2.2c Tabulka sloupcových procent pro data z tabulky 2.2a

	Zájem o sport		Sloupcová procenta
	Ano	Ne	
Chlapci	73,2 %	36,4 %	47,1 %
Dívky	26,8 %	63,6 %	52,9 %
Suma (řádkové součty)	100 %	100 %	100 %

Zdroj: modelová data

Statistická tabulka 2.3 obsahuje statistické údaje o provedeném z-testu teoretické hodnoty relativní četnosti pomocí výběru o rozsahu 300. Počet úspěšných pokusů byl 100. Testujeme nulovou hypotézu 0,4 (stručný výklad k tomuto statistickému testu se nachází v kapitole 4). Zde si všímáme pouze toho, jak vypadá výstup z počítače.

Ve výzkumné zprávě upravujeme výstupy z počítače do zjednodušené podoby tak, aby obsahovaly pouze takové údaje, které jsou nutné pro pochopení výsledků (s překladem do jazyka zprávy).

Tab. 2.3 Výsledky statistického testu pro relativní četnost získané programem NCSS 2007 (přeloženo)

Rozsah vzorku (n)	Počet úspěchů (X)	Relativní četnost (P)	Hypotetická část (P0)	Spolehlivost alfa
300	100	0,333333	0,400000	0,050000
Sekce mezi spolehlivostí				
Výpočtová metoda	Dolní 95% mez spolehlivosti	Výběrový podíl (P)	Horní 95% mez spolehlivosti	
Aproximace (nekor.)	0,279990	0,333333	0,386677	
Aproximace (korigovaná)	0,278323	0,333333	0,388343	
Sekce testování hypotézy				
Alternativní hypotéza	Prav. úroveň	Rozhodnutí (5%)	Z-hodnota	
$P > P0$	0,021394	odmítní H_0	-2,2981	

Tab. 2.4 Měření tří vzorků ve dvou skupinách [cm]

a)

Vzorek 1	Vzorek 2	Vzorek 3
Ošetření 1		
40	44	65
29,7	33	61,5
33	30	36
23	36	67,4
33	65,6	59,6
29,7	25	54
Ošetření 2		
16	12	23
15	13	31
10	11	33
28	10,3	25
25	11	41

b)

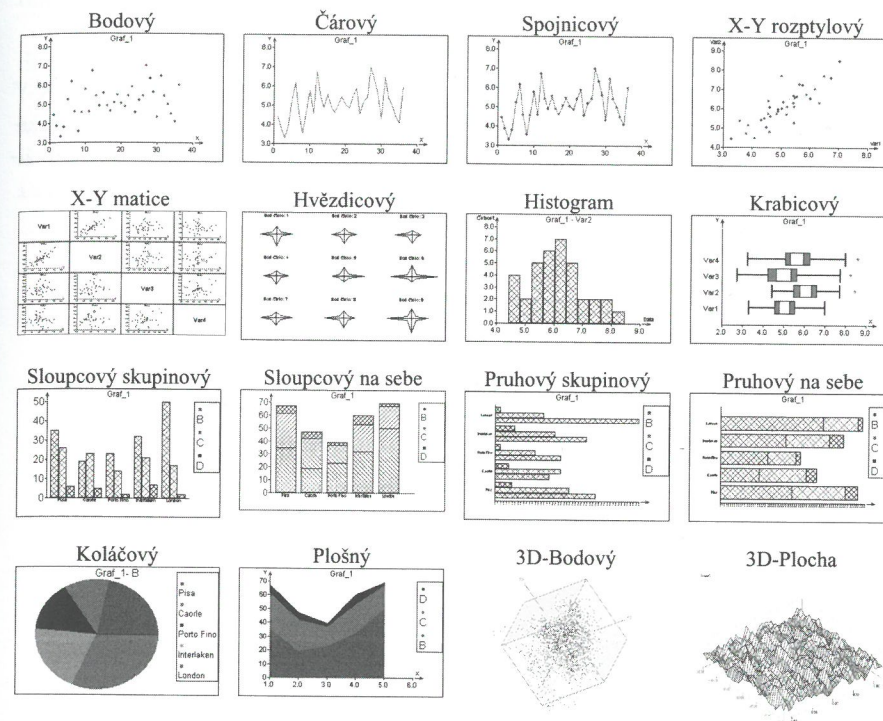
Vzorek 1	Vzorek 2	Vzorek 3
Ošetření 1		
40	44	65
29,7	33	61,5
33	30	36
23	36	67,4
33	65,6	59,6
29,7	25	54
Ošetření 2		
16	12	23
15	13	31
10	11	33
28	10,3	25
25	11	41

Zdroj: modelová data

Prezentační tabulky dat

Tyto tabulky zobrazují numerická data, někdy však mohou obsahovat kvalitativní i kvantitativní údaje. Používají se, pokud dat není velké množství a chceme upozornit na zajímavé konfigurace a vztahy pomocí primárních dat.

Ukazujeme dva příklady numerických tabulek s primárními daty, první příklad obsahuje zbytečné prvky prezentace (tab. 2.4a). Druhý je příklad správně upravené tabulky s numerickými údaji, která pomáhá při interpretaci rozdílů dvou skupin dat (tab. 2.4b). Takové tabulky mohou být paralelně doplněny grafickou prezentací dat.



Obr. 2.2 Příklady vybraných typů statistických grafů v dokumentaci statistického programu české firmy Trilobyte

Zdroj: Meloun, Militký, 2002

2.3 Statistické grafy

Statistické grafy mohou zaujmout a efektivně prezentovat naše data. Poskytují vizuální vjem rozdělení dat a jejich trendů a mohou dodat zajímavosti textu a tabulkám. Grafy slouží pro demonstrační cíle, a ne jako zdroj číselného materiálu. V moderní statistice se uplatňují i ve fázi analýzy dat. Poznamenejme, že s problematikou grafického

znázornění dat se setkáme ještě v kapitole 2.4, speciální grafy se také používají v mnohorozměrné analýze dat.

Známe mnoho typů grafů, některé často používané statistické grafy jsou schematicky ukázány na obrázku 2.2. Statistický software umožňuje snadno nakreslit všechny z nich barevně a v třídimenzionální (prostorové) podobě.

Základní statistické grafy

Sloupcové (sloupkové), koláčové, čárové (spojnicové), krabičkové (krabicové) grafy, X-Y bodové grafy, histogramy a polygony četnosti patří mezi nejčastěji používané statistické grafy. Popíšeme je podrobněji. Jejich cílem je zobrazit rozdělení dat, případně je použijeme pro znázornění závislosti na nějaké kategoriální proměnné nebo času.

Tab. 2.5 Modelová data o výdajích v podniku

Oddělení	Rozpočet (v tisících Kč)
Marketing	170
Výroba	1239
Řízení	250
Administrace	52

Zdroj: modelová data

Modelová data znázorněná v tabulce 2.5 zobrazíme pomocí sloupkového a koláčového grafu.



Obr. 2.3 Sloupcový graf modelových dat pro data z tabulky 2.5

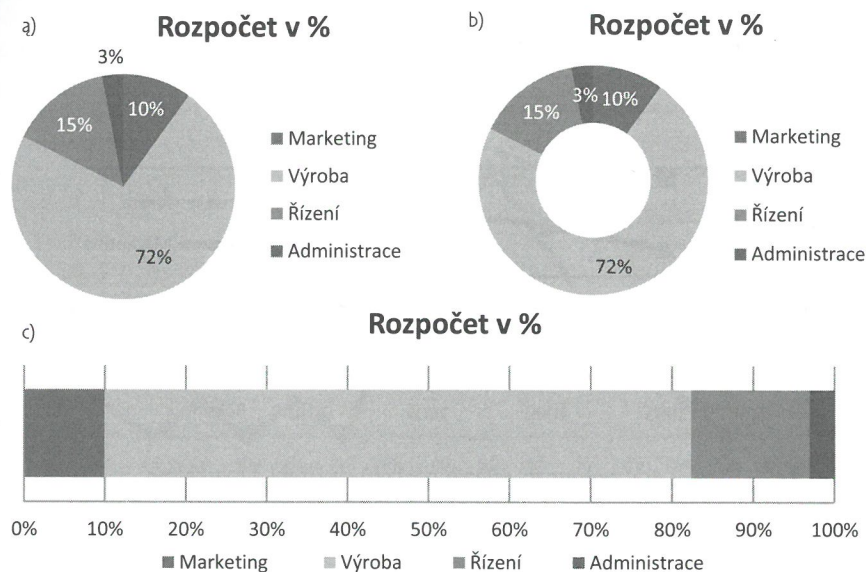
Zdroj: modelová data

Sloupcový (sloupkový) graf znamená znázornění množiny dat způsobem, že délka příslušného sloupku je funkcí znázorněného údaje. Vedle sebe lze zobrazovat několik množin dat (viz obr. 2.2: sloupkový graf).

Koláčový graf znázorňuje hodnoty množiny údajů, které odpovídají velikosti části z celku, pomocí kruhových výsečí. Všechny výseče tvoří celý kruh. Často používané modifikace tohoto grafu jsou prstencový graf a 100% skládaný sloupcový graf.

Obrázek 2.5 představuje čárový graf vývoje nákladů v tisících Kč na řízení a marketing z tabulky 2.6 modelových dat.

V případě histogramu na osu X vynášíme hranice intervalů a na osu Y četnosti v jednotlivých intervalech. Histogram se podobá sloupkovému grafu s tím rozdílem, že jednotlivé sloupky k sobě přiléhají, tím je zdůrazněno, že jde o znázornění kvantitativních, a ne kategoriálních proměnných (viz obr. 2.6).



Obr. 2.4 Koláčový (a), prstencový (b) a 100% skládaný sloupcový graf (c) modelových dat

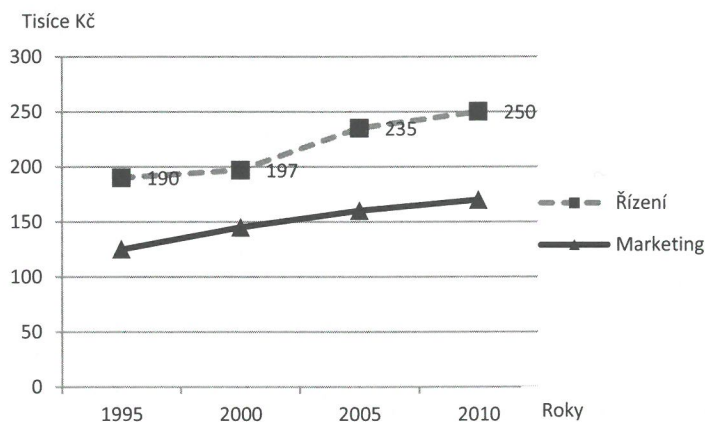
Zdroj: modelová data

Tab. 2.6 Modelová data o vývoji nákladů v podniku v letech 1995–2010

Roky	Řízení	Marketing
1995	190	125
2000	197	145
2005	235	160
2010	250	170

Zdroj: modelová data

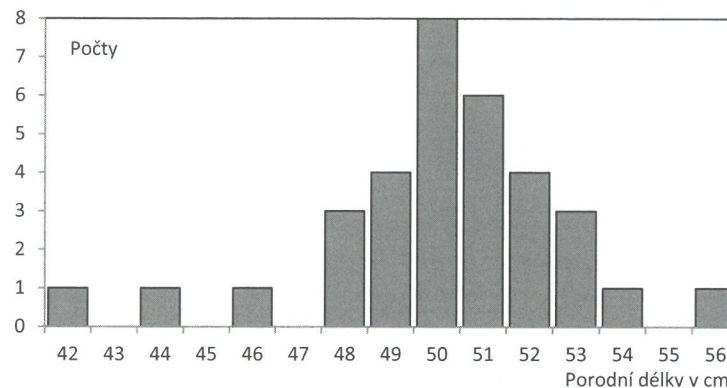
Polygon četností získáme obdobným způsobem jako histogram, pouze místo obdélníků spojujeme úsečkami četností vynesené pro sloupky histogramu ve středu jednotlivých intervalů.



Obr. 2.5 Čárový graf porovnání vývoje nákladů na řízení a marketing

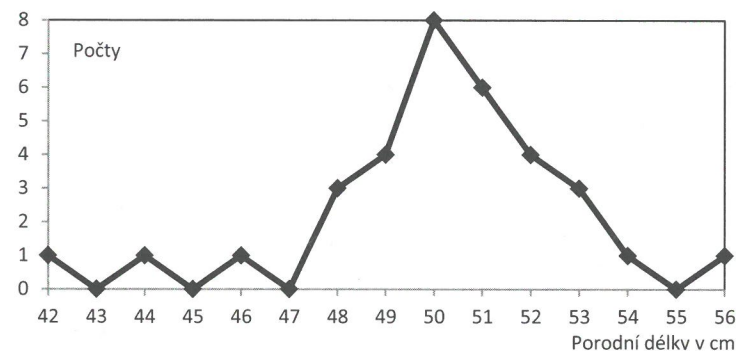
Zdroj: modelová data

Histogramy a polygony četností se používají pro znázornění rozložení nebo distribuci numerických údajů. Tvar grafu určuje charakter rozložení dat. Mluvíme o symetrickém rozložení, pokud histogram je symetrický, nebo o zešíkmeném rozložení, jestliže histogram nebo polygon četností či procent vykazují zešíkmení. Tyto grafy ale mohou mít i dva nebo více vrcholů (maxim), pak mluvíme o bimodálním nebo vícemodálním rozložení. Porovnáním těchto grafů sestavených pro různé skupiny získáme přehled o rozdílech průměrů nebo rozptýlenosti ve skupinách. Podrobnější zkoumání těchto grafů nám umožňuje identifikovat vychýlené hodnoty (outliers), které evidentně nepatří do skupiny. Takové údaje někdy vylučujeme z dalšího zpracování.



Obr. 2.6 Porodní délky živě narozených děvčat (cm) v ústavu A – histogram

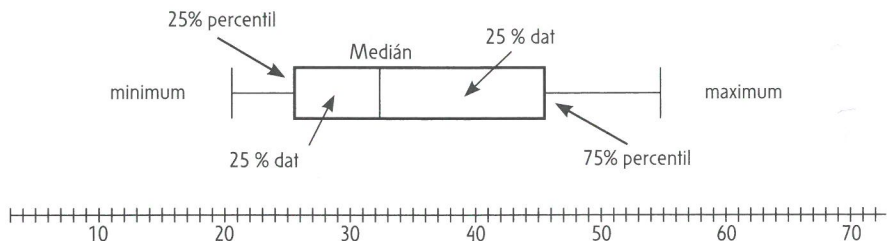
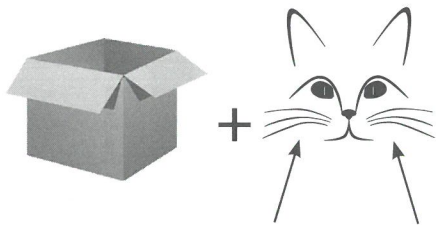
Zdroj: modelová data



Obr. 2.7 Polygon četností pro data z obrázku 2.6

Zdroj: modelová data

Na obrázku 2.8 je dále popsán krabičkový (krabicový) graf s vousy (podrobněji viz Hendl, 2012). Slouží pro grafické znázornění rozložení souboru dat a je vhodný pro porovnání několika souborů dat. Indikuje také vychýlené hodnoty. Krabička je vymezena dolním a horním kvartilem dat. Tento graf nás upozorňuje na nesymetrická rozložení dat tím, že medián neleží v prostředku krabičky.



Obr. 2.8 Krabičkový graf s vousy: piktogram a zobrazení jednoho souboru dat

Pro vysvětlení popíšeme stručně použité označení v grafu (příslušnými pojmy se zabýváme také v podkapitole o číselných charakteristikách 2.5): Medián je hodnota, pod kterou se nachází 50 % dat. Nad 75% percentilem (horní kvartil), resp. pod 25% percentilem (dolní kvartil) se nachází 25 % dat. To znamená, že krabička graficky znázorňuje 50 % dat.

Často se pro srovnání znázorňuje více souborů dat krabičkami v jednom obrázku. Tradičně se k tomuto účelu používá graf, v kterém sloupky zobrazují průměry proměnných. Sloupky jsou doplněny úsečkami pro zobrazení velikosti směrodatných odchylek nebo směrodatných chyb.

Příklad: Výpočty pro krabičkový graf s vousy

Nasbírali jsme data 144, 240, 153, 629, 540, 300. Minimum = 144, maximum = 629, medián = $(240+300)/2 = 270$, dolní kvartil = 153, horní kvartil = 540. Pomocí krabičkového grafu s vousy má význam zobrazovat data, jejichž počet je větší než 15.

Proces tvorby grafu

Špatná volba a návrh grafu mohou zkreslit informaci, kterou chceme zprostředkovat, naopak vhodný graf vede ke správnému a rychlému rozhodnutí. Uvedeme doporučení pro tvorbu grafu pro výzkumné zprávy. Proces návrhu grafu má čtyři kroky:

1. Určíme, co chceme říci.
2. Vybereme typ grafu.
3. Uspořádáme data.
4. Připravíme a formátujeme graf.

1. Určíme, co chceme říci

Jedná se o první a velmi důležitý krok. Tážeme se, jaký účel má náš graf. Pokud známe důvod, je snadnější zvolit správný typ grafu. Jestliže volíme vhodnou prezentaci pro naše představy o datech, rozlišujeme šest důvodů:

- porovnání;
- ukázání rozložení dat;
- objasnění částí celku;
- dokumentace trendu;
- odhalení odchylek;
- porozumění vztahu.

2. Vybereme typ grafu

Vysvětlíme jednotlivé účely grafického znázornění a zároveň identifikujeme, které základní typy grafů můžeme v každém případě využít. K demonstraci našeho postupu se zaměříme na grafy sloupkové, čárové, XY bodové a koláčové a krabičkové. Na tomto místě poznamenejme, že tabulky čísel často zprostředkují i určitý grafický dojem o charakteru dat.

Porovnání

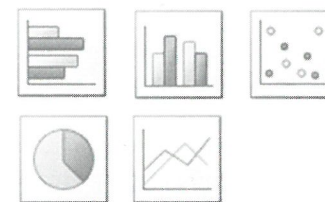
Co to znamená? Porovnáváme jednu množinu dat s jinou.

Příklady:

- Srovnání prodeje výrobku A a B v pěti oblastech.
- Výkon několika studentů v daném testu a porovnání různých skupin studentů.

Grafy, které je možné v této situaci použít:

1. Sloupkové grafy s různou orientací.
2. Dvojrzměrné bodové grafy.
3. Koláčové grafy.
4. Čárové grafy.



Ukážeme rozložení

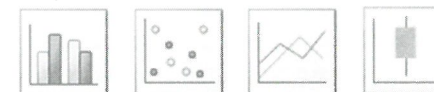
Co to znamená? Chceme ukázat rozložení množiny hodnot, abychom si udělali představu o vychýlených hodnotách, intervalu, kde se nalézá většina hodnot, atd.

Příklady:

- Rozložení čekacích dob v telefonním centru.
- Rozložení chyb v desetidenním období ověřování nového softwaru.

Grafy, které je možné v této situaci použít:

- Sloupkové grafy.
- Dvojrzměrné bodové grafy.
- Čárové grafy.
- Krabičkové grafy.



Objasnění částí celku

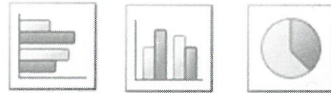
Co to znamená? Chceme ukázat, jaký podíl z celku tvoří různé části.

Příklady:

- Ukazujeme relativní podíl jednotlivých produktů na celkovém prodeji.
- Typy prohlížečů u návštěvníků našich internetových stránek.

Grafy, které ukazují části celku:

- Sloupkové grafy s různou orientací.
- Koláčové grafy.



Trend v čase

Co to znamená? Chceme porozumět trendu v čase u zvolených proměnných.

Příklady:

- Vývoj zisku podniku za posledních 365 dnů.
- Ceny produktu za posledních 15 let.

Grafy, které je možné použít k ukázání trendu v čase:

- Sloupkové grafy.
- Čárové grafy.



Odchylky

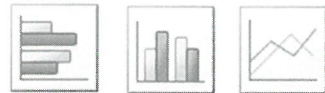
Co to znamená? Chceme dokumentovat, které hodnoty se odchylují od normy.

Příklady:

- Chyby zjištěné pomocí kontroly kvality.
- Objemy prodeje na různých prodejních místech.

Grafy používané k prokázání odchylek:

- Sloupkové grafy s různou orientací.
- Čárové grafy.



Vztah

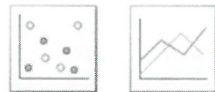
Co to znamená? Chceme ukázat vztah mezi dvěma nebo více proměnnými.

Příklady:





- Vztah mezi dvěma atributy jedinců.
- Vztah mezi vynaloženými náklady a ziskem.

Grafy, které ukazují vztahy:

- Dvozměrné bodové grafy.
- Čárové grafy.



Obrázek 2.9 obsahuje v přehledu informace usnadňující volbu grafu.

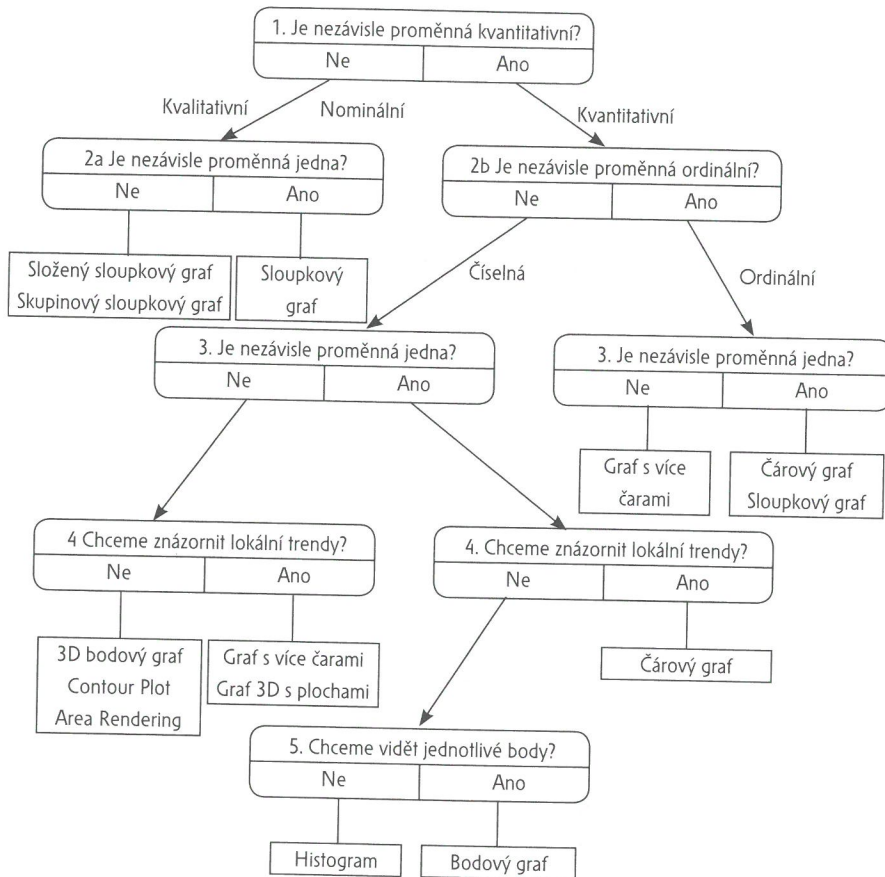
	Zobrazovací prostředky			
	Body	Čáry	Sloupky	Krabičky
Posuzovaná konfigurace				
Časová řada Hodnoty popisují, jak se mění věci v čase (ročně, měsíčně, atd.)	Ano (jako bodový graf, pokud nemáme hodnotu pro každý časový interval)	Ano (zobrazení obecných trendů a jejich srovnání)	Ano (pouze vertikální sloupky k zobrazení individuálních hodnot k podpoře jejich srovnání)	Ano (vertikální krabice k zobrazení, jak se mění rozdělení v čase)
Řazení Hodnoty jsou řazeny dle velikosti	Ano (jako bodový graf, zvláště pokud škála nezačíná v nule)	Ne	Ano	Ano (k zobrazení rozdělení seřazených skupin dle určitého ordinálního kritéria)
Část z celku Hodnoty reprezentují části (např. v procentech)	Ne	Ano (zobrazení, jak se mění části celku v čase)	Ano	Ne
Odchylky Rozdíly mezi dvěma množinami hodnot	Ano (jako bodový graf, zvláště pokud škála nezačíná v nule)	Ano (pokud také změny v čase)	Ano	Ne
Rozdělení Četnost hodnot v intervalech (např. počty lidí v věkových intervalech)	Ano (pomocí proužků v závislosti na velikosti individuální hodnoty)	Ano (jako polygon četností k zobrazení celkového tvaru rozdělení)	Ano	Ano (pokud srovnání několika rozdělení)
Korelace Vztahování dvou proměnných (např. váhy a výšky skupiny lidí)	Ano (jako XY bodový graf)	Ne	Ano (jako tabulka grafických symbolů o různé velikosti dle hodnoty proměnné, pokud čtenář nezná XY grafy)	Ne
Prostorové Hodnoty jsou zobrazeny na mapě podle polohy	Ano (jako kružnice o různé velikosti na mapě)	Ano (k zobrazení cest na mapě)	Ne	Ne
Srovnání hodnot podle kategorií Srovnání hodnot pro množiny objektů různého typu	Ano (jako bodový graf, zvláště pokud škála nezačíná v nule)	Ne	Ano	Ne

Podle Few, S.: Show Me the Numbers, Analytics Press, 2014, modifikováno

Obr. 2.9 Doporučení pro volbu typu grafu

Pokud jsme proměnné klasifikovali na závislé a nezávislé proměnné, volíme vhodný graf podle toho, na jaké škále byly jednotlivé proměnné měřeny. Přihlížíme k tomu, zda proměnné jsou kvantitativní a nabývají mnoho hodnot, nebo nabývají pouze málo hodnot a jsou kvalitativní. V mnoha situacích pracujeme se závisle proměnnou kvantitativního

typu. V takovém případě nám může k volbě vhodného grafu pomoci rozhodovací strom zobrazený na obrázku 2.10. Přitom zohledňujeme, kolik je nezávislých proměnných.



Obr. 2.10 Rozhodovací schéma pro volbu typu grafu

Zdroj: NC State University 2004 (www.ncsu.edu/labwrite/res/gh/gh-graphtype.html), modifikováno

Jestliže pro volbu grafu existuje více důvodů, pak nám při rozhodování o typu grafu pomáhá jednoduchá úvaha o užitečnosti zobrazení. Ptáme se, zda uděláme dva grafy, nebo kombinaci grafů.

3. Uspořádání dat

Jestliže jsme vyjasnili účel grafu a zvolili jeho typ, zkoumáme data, která máme k dispozici. Často nejsou data ve vhodné podobě. Musíme je uspořádat, přepočítat nebo transformovat. Tabulkové procesory i statistické programy nám poskytují mnoho prostředků, jak to udělat. Teprve pak můžeme přistoupit ke kroku 4.

4. Připravíme graf a formátujeme graf

Pomocí statistických grafů vypravujeme o datech určitý příběh, hodnotíme alternativy, hledáme trendy nebo se snažíme přijít na to, co je běžné v průměru. Účel grafu je rozhodující pro jeho správné formátování.

Mnoho aspektů návrhu statistických grafů je společných pro všechny grafy. Proto je nejdříve uvedeme, pak se věnujeme speciálním doporučením pro sloupkové, čárové, koláčové grafy a X-Y bodové grafy.

Obecné zásady pro tvorbu grafů ve výzkumné zprávě

Data mají být především jasně odlišena a zdůrazněna před pozadím grafu. V grafu má být jasně popsáno:

- název a účel grafu;
- co znamenají jednotlivé osy, třídy, sloupky atd.;
- škály každé osy s vyznačením začátku.

Je třeba uvést zdroj dat. V grafu nemá být nadbytek zbytečných čar a grafických prvků.

Nejčastější chyby v grafech:

- některá z os není popsána;
- zbytečně se zdůrazňují trendy (např. vhodnou volbou, změnou škály a počátku);
- zavádějící jednotky na jednotlivých osách;
- nevyhází se z přesných dat.

Při posuzování kvality zobrazení si před provedením interpretace klademe devět otázek:

- Jsou vlastní data jasně patrna a zdůrazněna?
- Je evidentní účel grafu a je popsán názvem?
- Je uveden zdroj dat v obrázku a ve vlastním textu?
- Pocházejí informace v obrázku ze spolehlivých zdrojů?
- Je vše jasně popsáno?
- Je patrné, kde začínají osy?
- Nejsou na osách provedena přerušení? Jsou přerušení jasně vyznačena?
- Mají osy konstantní škálu, nebo proměnlivou škálu?
- Neobsahuje obrázek matoucí prvky, které znehodnocují prezentaci?

Dále uvedeme aspekty, kterých si máme všimnout u třech základních typů grafů.

Sloupkové grafy

- Mohou být tvořeny horizontálními nebo vertikálními sloupky.
- Osy mají začínat v nule a mají být jasně označeny. Výjimku je nutné zdůvodnit a vyznačit.
- Sloupky mají mít stejnou šířku.
- Sloupky mají být vhodně seřazeny. Obvykle nejlepší je pořadí určené délkou sloupku. Zohledňujeme však také logiku řazení, např. čas.

Čárové grafy

- Osy mají začínat v nule a mají být jasně označeny. Výjimku je nutné zdůvodnit a vyznačit.
- Je povoleno více čar. Velký počet čar ale znesnadňuje přehled.
- Pokud v datech scházejí údaje, úsek v datech se má vynechat nebo se vykresluje jiným typem čáry.

Koláčové grafy

- Vyhýbáme se mnoha kategoriím, povoluje se maximálně pět kategorií. Málo zastoupené kategorie sdružujeme do kategorie „ostatní“.
- Jednotlivé kategorie znázorňujeme ve vhodném pořadí – často podle velikosti.
- První segment začíná na 12. hodině, pak postupujeme ve směru hodinových ručiček.
- Opatrně postupujeme u více komparativních koláčových grafů. Jejich srovnání je často obtížné. Lepší je pak použít sloupcové grafy.

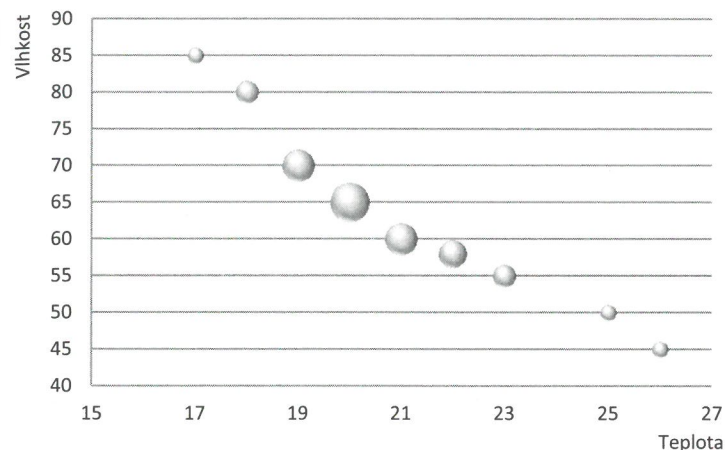
Bodové X-Y grafy

Bodové X-Y grafy se používají, pokud:

- prezentujeme větší počet naměřených v párech X a Y hodnot;
- uvažujeme dvě proměnné (často jedna je závisle a druhá nezávisle proměnná);
- lze zařadit i třetí proměnnou změnou typu a barvy bodů podle kategoriální proměnné nebo velikosti bodů podle metrické proměnné. Bublínový graf zobrazuje velikost bodů v závislosti na velikosti třetí proměnné (viz obr. 2.11).

2.4 Mapy

Mapy jsou výborným prostředkem pro prezentaci statistické informace v geografických souvislostech. Jsou vizuálně přitažlivé, zjednodušují rozpoznání vztahu dat k danému místu a pomáhají uživatelům identifikovat geografické trendy v datech lépe, než se to může podařit pomocí běžného statistického grafu nebo tabulky. Známým příkladem datových map jsou geografická zobrazení výskytu nemocí. Jiným příkladem takového zobrazení dat jsou klimatologické mapy nebo mapy o demografické situaci v různých oblastech.



Obr. 2.11 Závislost relativní vlhkosti na teplotě, velikost bodů odpovídá velikosti srážek v oblasti

Zdroj: modelová data

Rozeznáváme tři typy map.

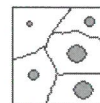
1. Barevné mapy a mapy s různou výplní

Jedná se o nejčastější typ mapy. Používá se k zobrazení procent a intenzit. Různé barvy se použijí pro jednotlivá rozmezí hodnot, takže uživatel jednoduše identifikuje vysoké, střední a nízké hodnoty.



2. Proporční symbolické mapy

Tyto mapy používají symboly s různou velikostí tak, že velké symboly jsou v oblasti s velkou hodnotou cílové proměnné a malé symboly prezentují nízké cílové hodnoty. Můžeme použít kolečka, sloupky nebo symboly předmětů, o které jde. Jsou vhodné pro četnostní hodnoty.



3. Bodové mapy

Jednotlivé výskyty událostí se vyznačují bodem v dané oblasti. Známá jsou tato zobrazení v souvislosti s propukajícími epidemiemi infekčních nemocí.

Uvedeme některé obecné zásady a poté doporučení týkající se jednotlivých hlavních typů map.

