

Pravidla pro návrh vizuálů

Zuzana Šedivá
Katedra informačních technologií
Vysoká škola ekonomická Praha
sediva@vse.cz

Abstrakt: článek se zabývá vizualizací dat v aplikacích podnikových informačních systémů. Způsob a nastavení vizualizace přímo ovlivňuje pochopení dat, která jsou zobrazována. Článek je zaměřen na základní pravidla pro vytvoření vizuálu tak, a by byl dobře srozumitelný, přehledný a aby uživatel získal očekávanou informační hodnotu. Text je zaměřen na základní typy grafů, jejich principy a časté chyby, které se v jejich používání vyskytují.

Klíčová slova: dashboard, graf, KPI, vizualizace dat, vizuál, piktogram.

Abstract: The article deals with data visualization in enterprise information system applications. The way and setting of the visualization directly affects the understanding of the data it displays. The article focuses on the basic rules for visual creation with the intention to create a comprehensible and clear visual and to give the user the expected information value. The text focuses on the basic types of graphs, their principles and common mistakes that occur in their use.

Key words: dashboard, chart, KPI, data visualization, visualization, icon.

Úvod

Vizualizace nebo také zkráceně označovaná jako vizuál je vizuální prezentací dat. Zahnuje různé formy jako je například vyjádření jedné hodnoty v podobě jednoho čísla (často používané pro zobrazení hodnoty ukazatele - KPI), dále data ve formě tabulek s barevným rozlišením oblastí vzhledem k velikosti zobrazovaných hodnot, různé typy grafů až po různé grafické prvky např. ve formě budíků a teploměrů. Grafické znázornění hodnot urychlí velmi často orientaci v datech, zvýrazní důležité limitní hodnoty či umožní zobrazit vztahy mezi daty. Návrh vizuálu není jen o návrhu typu grafu pro vybraná data, ale o jejich popisu či nastavení grafu. Při navrhování vizuálu je třeba se vyhnout některým chybám, které mohou vést buď jen k nepřehlednosti dat nebo i ke špatné interpretaci dat. Text tohoto článku je zaměřen na principy vybraných grafů, jejich užití a pravidel pro jejich nastavení. Článek tak volně navazuje na základní pravidla a zásady v grafickém návrhu dashboardů uvedené v (Šedivá, 2017).

Především u velkého množství dat je jejich vizualizace nezbytná, a to hlavně z důvodu jejich lepší přehlednosti a srozumitelnosti. Vizualizace dat se stává důležitým prostředkem pro rychlou práci s daty. Manažer si může volit nejen různé pohledy na data, ale i jejich granularitu. Je důležité, aby mu aplikace informačního systému umožnila určit úroveň detailu dat, tj. zda jsou pro něj v daný okamžik důležité souhrnné údaje nebo zda jsou zajímavá i detailní data na nižší úrovni hierarchie dat.

V procesu návrhu vizualizace je třeba zvažovat typ vizuálu, tj. zda budou data prezentována ve formě tabulky, ve formě grafu či v jiné grafické formě. Pokud je třeba zobrazit data v různých vzájemných vazbách (například zobrazit odchylky vůči určité hodnotě, limitu, apod.), potom je vhodnější zvolit formu grafické prezentace v grafu.

Pokud je třeba zobrazit i detaily, potom je vhodná prezentace formou tabulky dat s využitím podmíněného formátování pro zvýraznění důležitých informací k datům. Často se používají různé barevné výplně, ikonky, piktogramy či datové pruhy.

Data budou pro manažera uspokojivá z jeho pohledu informační potřeby jen tehdy, když jim bude rozumět a bude umět s nimi pracovat (Visage.co, 2014). I správná data mohou být chybně interpretována, pokud je zvolen chybný pohled na data nebo pokud jsou špatně popsána. Proto je forma vizualizace dat tak důležitá.

Je třeba definovat, jaká data jsou prezentována a jakou mají v prezentaci roli. Ve vizualizaci se vyskytují kvantitativní číselné hodnoty, které můžeme měřit, používat je pro výpočty a vytvářet z nich souhrny. Jako příklad můžeme uvést počet odpracovaných dnů zaměstnance, počet vyrobených kusů, počet objednávek atd. Jinou roli ve vizualizaci mají hodnoty kvalitativní nominální. Kvalitativní hodnoty představují abstraktní vlastnosti, které nesou informaci o subjektu či objektu. Např. u produktu jsou kvalitativní hodnotami jeho identifikační číslo, váha, název produktu apod. Dalšími typy dat jsou hodnoty kategorické (ordinální). Tyto hodnoty mohou být číselné i textové a lze u nich zpravidla určit pořadí nebo rozsah hodnot v určitém rozmezí. Mohou jimi být například typ produktu, stupeň hořlavosti, mzdová tarifní třída apod. Datové typy uvedené v předchozím textu jsou jedním z určujících hledisek pro volbu vizualizálu.

Abychom mohli určit, který typ grafu je nevhodnější pro prezentaci dat, je třeba si uvědomit, s jakým cílem mají být data prezentována, neboť lze tatáž data prezentovat z různých úhlů pohledu a zdůraznit tím jinou informaci. Proto je důležité předem stanovit, pro koho je vizualizace připravována, z jakého hlediska na ni bude manažer pohlížet a jakou informaci z vizualizace má získat. Určujícím hlediskem výběru grafu by měl být vztah mezi daty, podle kterého chceme data prezentovat. Podle (visage.co, 2017) může být v grafu prezentován jeden i více z níže uvedených vztahů mezi daty:

- **nominální porovnání** graficky znázorňuje kvantitativní hodnoty podle různých kategorií, např. počet obyvatel v krajích v ČR, počet prodaných výrobků v jednotlivých státech apod.,
- **rozdíl (odchylka)** znázorňuje, jak se daná hodnota odchyluje od průměru nebo od očekávané hodnoty, např. skutečné náklady vs. rozpočet,
- **časová řada** prezentuje změny kvantitativních hodnot v časovém období, např. počet zakázek v jednotlivých měsících, výše cen v daném období atd.
- **distribuce** zobrazuje rozložení dat kolem dané centrální hodnoty, např. sledujeme vyšší mzdy v určité věkové kategorii, výšku hráčů v basketbalovém týmu atd.,
- **korelace** představuje vztah dvou a více hodnot, např. počet objednávek v závislosti na denní době prodeje, výše příjmu vzhledem k počtu členů v rodině,
- **vztah části vůči celku** porovnává podskupinu (podmnožinu) dat vůči celku, např. podíl fixních nákladů na celkových nákladech v rozpočtu projektu,
- **pořadí** vzájemně porovnává 2 a více hodnot v pořadí podle sledované veličiny, např. žebříček firem podle počtu zákazníků na trhu, žebříček akcií podle výše ceny na trhu atd.

Výše uvedené vztahy mezi daty jsou rozhodující pro správné zvolení typu grafu a jeho nastavení. V následujícím textu budou popsány často používané typy grafů, u který si přiblížíme časté chyby při jejich používání a zmíníme i základní pravidla pro jejich správné nastavení.

Sloupcový graf

Ačkoliv je sloupcový graf (*Column Chart*) velmi často používaným typem grafu, vyskytují se při jeho užití v prezentaci často chyby v nastavení či jeho grafické úpravě. Představuje způsob grafického zobrazení pro porovnávání kvantitativních a kvalitativních hodnot. Slouží pro nominální porovnání kvantitativních hodnot pomocí sloupců, přičemž využívá výšky sloupce pro znázornění hodnoty. Rozdíly v délce sloupce jsou pro manažera dobře rozpoznatelné, odchytky výšky sloupců mezi sebou lze rychle odečíst a uživatel tak získá rychle přehled při porovnávání hodnot. Hodnoty mohou být znázorněny v jedné či ve více datových řadách a v jedné či více kategoriích.

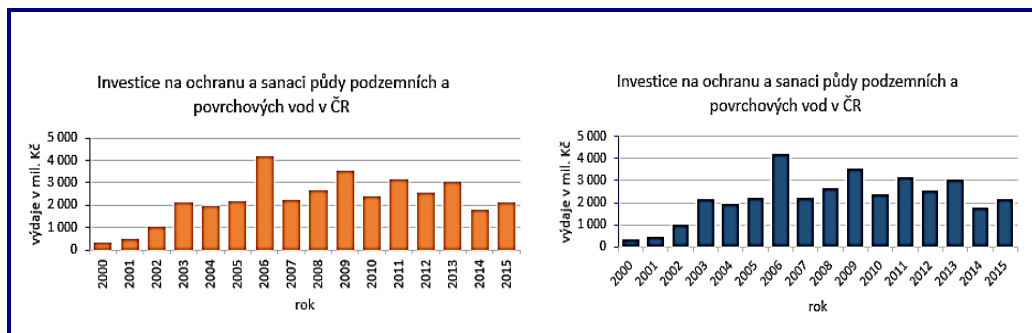
Obdobou sloupcového grafu je graf pruhový. Jedná se o horizontální variantu sloupcového grafu. Pruhový graf je vhodný, pokud zobrazujeme více položek, které by při omezené šířce na ose x byly špatně čitelné.

Sloupcový graf slouží velmi často také pro vyjádření vztahu hodnoty k celku, a to formou skládaného (nastavovaného) grafu. Hodnoty jednotlivých kategorií jsou ukládány do jednoho sloupce, a tak lze získat dobře přehled o podílu jednotlivých kategorií na celkovém výsledku. Tato forma sloupcového grafu se hodí pouze tehdy, když hodnoty vynášené za jednotlivé datové řady dávají v souhrnu celek, tedy 100%. Skládaný sloupcový graf potom lze zobrazit buď v nominálních hodnotách nebo v procentním vyjádření. V druhém případě mají sloupce stejnou výšku.

Pravidla pro návrh a nastavení sloupcového grafu jsou následující:

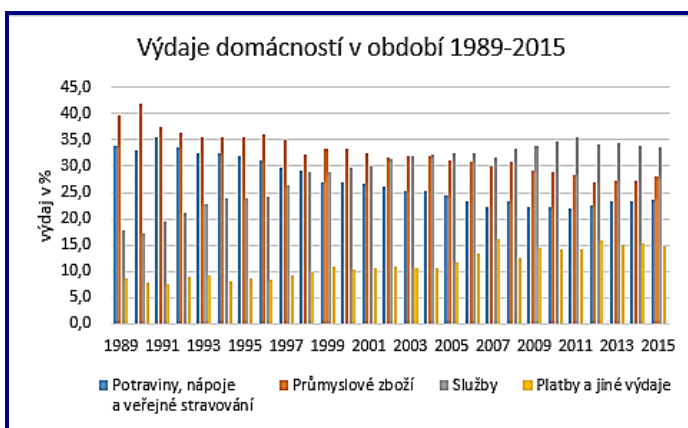
Nastavení šířky sloupců - z hlediska grafické úpravy grafu, je nutné zvolit takovou šířku sloupců a mezer mezi nimi tak, aby byl rozdíl hodnot co nejlépe přehledný a čitelný. Proto se z důvodu rychlé čitelnosti grafu používá pravidlo, podle kterého mezera mezi sloupci odpovídá přibližně polovině šířky sloupce. (visage.co, 2017) Oko manažera potom nemusí přeskakovat „velké“ vzdálenosti mezi sloupci a odchytky výšky sloupců jsou rychle měřitelné.

Popisky os x a y používané u převážné většiny grafů by měly být vždy dobře čitelné a měly by vhodně doplňovat prezentované hodnoty. Toto však může být problematické, když je na ose x v grafu zobrazeno mnoho hodnot, tudíž i popisků. Správně by měly být popisky v horizontálním směru pro snazší čtení, ale to vždy není možné (Singh, 2014). Obrázek 1 ukazuje problém nastavení směru popisků u hodnot na ose x. Horizontální ani vertikální směr zdola nahoru u popisků vlevo na obrázku není snadný pro čtení. V tomto případě lze použít kompromisní řešení – nastavit úhel cca 45° pro snadné čtení popisků osy x (řešení vpravo na obrázku).



Obrázek 1 Čitelnost popisků os v grafu

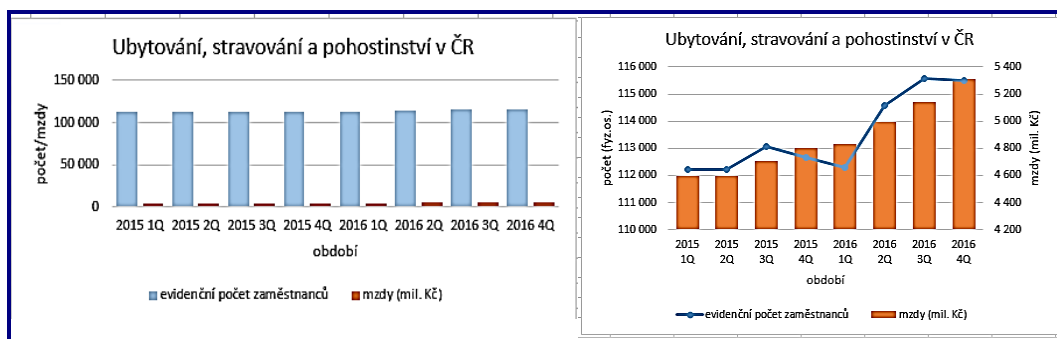
- Setřídění hodnot - pro lepší čitelnost a rychlou orientaci v grafu je rozumné, data setřídít buď podle hodnot popisků nebo vzestupně podle hodnot. Graf se opět stává přehlednější pro porovnání rozdílů hodnot navzájem. Především u pruhového grafu je nutné, aby byly hodnoty setříděné vzestupně, v horní části grafu se tak zobrazí nejvyšší hodnoty, v dolní části grafu jsou nejmenší.
- Velký počet datových řad – je velmi nevhodné použít sloupcový graf pro velký počet datových řad. Výhodou sloupcového grafu je, že je velmi dobře srozumitelný, ale to platí jen pro omezené množství položek na ose x. U velkého rozsahu kategorií, nejsou jednotlivé sloupce dobře rozeznatelné a manažer není často schopen rychle rozeznat jednotlivé kategorie i rozdíly mezi hodnotami. Obrázek 2 tento nedostatek demonstruje. Jsou zde vynášeny 4 kategorie pro velké množství datových řad, které odpovídají rokům. Sloupce jsou příliš úzké, velmi těžko rozeznatelné, téměř se slévají do sebe, popisky sloupců vůbec nelze zobrazit díky nedostatku místa na ose x.



Obrázek 2 Nevhodné použití sloupcového grafu

- Velký rozdíl mezi měřítkem rozsah hodnot na ose y u více datových řad. Velmi častou příčinou způsobující špatnou čitelnost grafu je nevhodné nastavení.

Pokud ve sloupcovém grafu vynášíme dvě datové řady s velmi rozdílným měřítkem, např. počet zaměstnanců řádově ve statisících a celkový objem mezd za řádově v mil. Kč, je sloupcový graf velmi nepřehledný. Sloupce pro mzdy jsou v grafu nepatrné vzhledem k počtu obyvatel. Vlevo na Obrázku 3 je zobrazeno nevhodné nastavení sloupcového grafu s jednou osou y, ve kterém jsou sloupce mezd téměř nečitelné. Je třeba tento problém vyřešit buď zvolením dvou rozdílných os y s různými měřítkem, nebo použijeme kombinovaný graf, kde pro každou řadu zvolíme odlišný typ grafu, např. sloupcový pro mzdy a spojnicový pro počet zaměstnanců. Alternativou řešení je proto použití kombinovaného grafu s hlavní a vedlejší osou y (Obrázek 3 vpravo).



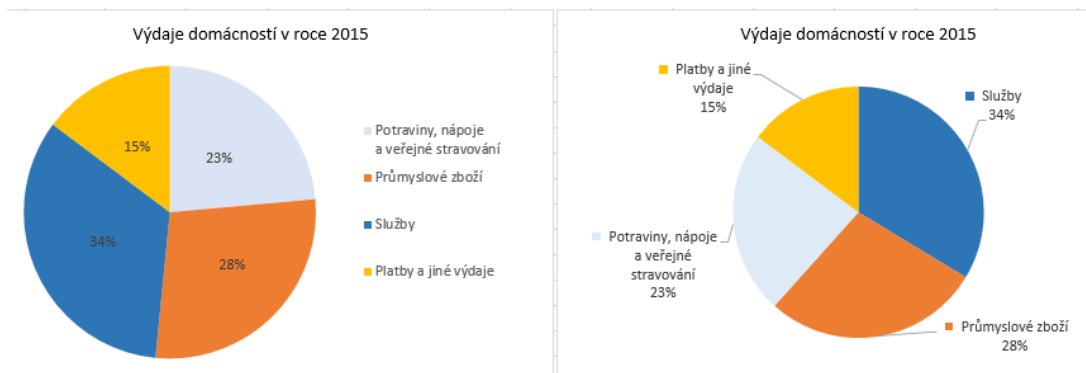
Obrázek 3 Řešení rozdílného měřítka na ose y

- Popisky hodnot musí být vždy dobře čitelné, jinak ztrácejí svůj význam. Proto se popisky hodnot umísťují v horizontálním směru (odpovídá to způsobu čtení zleva doprava), v kontrastní barvě a velikost písma popisek vzhledem k barvě pozadí, popisky umístíme tam, kde jsou nejlépe čitelné – mimo sloupec. Pokud je v grafu málo místa pro popisky hodnot, potom je lépe popisky hodnot vůbec nepoužít.

Výsečový graf

Výsečový nebo také koláčový graf (*Pie Chart*) je vhodný pro zobrazení vztahu částí k celku pouze pro jednu datovou řadu. Proto je vhodný pouze pro případ, kdy vynášené hodnoty dávají dohromady celek. Graf nemá žádné osy a hodnoty jsou zobrazeny jako výseče v kruhu, který tak tvoří celek, tj. 100 %. Úhel výseče představuje podíl části z celku.

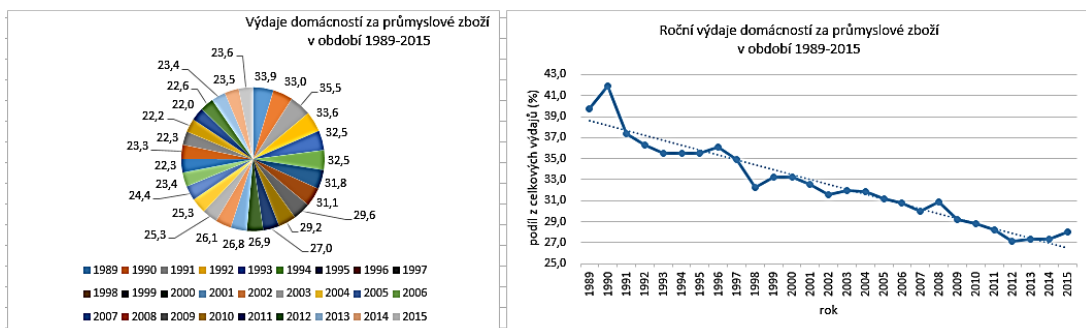
Zobrazování hodnot formou výsečí je na rozdíl od formy sloupců méně přehledné na první pohled. Především tehdy, když jsou výseče buď velmi malé nebo přibližně stejné. Manažer často těžko odečítá rozdíly mezi výsečemi. Proto zde může dojít snadno k chybné interpretaci hodnot. Z pohledu práce s daty je podle odborníků v oblasti vizualizace dat (Few, 2012) kruh nepřesný pro vnímání velikosti dat podle tvaru výsečí. Lidským okem zpravidla dobře rozpoznáme výseče po čtvrtinách kruhu, tedy po 90°, 180° a 270°. Ostatní stupně úhlů výsečí a rozdíly mezi nimi již nejsou snadno měřitelné a hodnoty nelze přesně vizuálně jednoduše určit (často se přeceňují tupé úhly a naopak podceňují se ostré úhly).



Obrázek 4 Výšečový graf a jeho rozdílné nastvení

Vizualizace na Obrázku 4 představuje rozdílné nastavení parametrů téhož výšečového grafu. Graf zobrazuje procentní podíly výdajů domácností v ČR členěné do 4 kategorií. Ukázka vlevo zobrazuje nastavení grafu s legendou a popisky hodnot uvnitř výšečí. Popisky hodnot ve výšečích musí být vždy dobře čitelné, což není dodrženo v levé ukázce Obrázku 4. Popisek hodnoty u výšeče Služby je hůř čitelný díky tmavé výplni výšeče. Také používání legendy v tomto případě není vhodné, protože manažer musí očima přeskakovat mezi grafem a legendou. Ukázka vpravo obsahuje popis kategorií přímo u výšečí a procentní hodnot, což je pro přehlednost dat vhodnější způsob prezentace. Data grafu byla seříděna sestupně, a proto první největší výšeč (Služby) začíná na stupni 0, následuje druhá největší výšeč (Průmyslové zboží), atd. Manažer tak získá ihned informaci o žebříčku ve výdajích domácností.

Příklad zcela nevhodného použití výšečového grafu představuje Obrázek 5 v ukázce vlevo, který zobrazuje procentní podíly výdaje domácností za průmyslové zboží z celkových ročních výdajů v letech 1989 až 2015. V souboru dat se vyskytuje příliš velké množství kategorií (roků). Graf je zcela nepřehledný a manažer nezíská žádnou informační hodnotu.



Obrázek 5 Nevhodné použití výšečového grafu a jeho alternativa

Druhou hrubou chybou použití tohoto grafu je jeho neadekvátní použití vůči datům. Data obsahují procentní podíly výdajů za průmyslové zboží z celkových ročních výdajů

v jednotlivých letech. Tyto hodnoty však netvoří celek 100%, a proto je výšečový graf zcela nevhodný.

Vzhledem k tomu, že se jedná o výdaje v časovém období, zvolíme buď spojnicový graf nebo bodový graf. V ukázce vpravo na Obrázku 5 byl zvolen spojnicový graf. Z vizualizace je jasná základní informace, že podíly výdajů za průmyslové zboží v jednotlivých letech klesají, čemuž odpovídá i trendová křivka. Všimněte si, že bylo změněno i nastavení měřítka osy y, která nezačíná od hodnoty nula, ale až od 25%. Toto nastavení osy y odpovídá faktu, že meziroční změny v procentních podílech jsou v řádech desetin až jednotek procent. Pokud by osa y začínala od nuly, potom by rozdíly byly v grafickém vyjádření nepatrné.

Pro lepší přehlednost a čitelnost výšečového grafu je vhodné dodržet následující pravidla (Few, 2012), (Singh, 2013):

- výšečový graf používat pro malý počet kategorií, optimálně do 5 kategorií, pro větší počet kategorií zvolit jinou alternativu vizuálu, např. skládaný sloupcový graf,
- popisky uvádět přímo u výsečí, legendu používat jen pokud je to nezbytně nutné,
- pro lepší přehlednost hodnot grafu, je účelné hodnoty seřadit sestupně od největší po nejmenší hodnotu, manažer tak rychle získá kromě informace o jednotlivých podílech i informaci o žebříčku (tedy kdo je nejlepší a kdo nejhorší).
- první výseč umístit od 0°, čímž se zajistí lepší přehlednost a čitelnost grafu. Člověk je zvyklý číst hodnoty shora dolů ve směru hodinových ručiček, proto jsou výseče v tomto směru přehlednější.

Variantou výšečového grafu je prstencový graf. Také prstencový graf zobrazuje podíly částí z celku, kdy jednotlivé podíly jsou představovány délkou oblouku. Pro umístění oblouků a jejich popis platí stejná pravidla jako u výsečí ve výšečovém grafu. I když lze v prstencovém grafu na rozdíl od výšečového grafu vynášet i více proměnných (datových řad), je jejich počet omezen. U více než 3 proměnných se prstencový graf stává nepřehledným.

Spojnicový graf

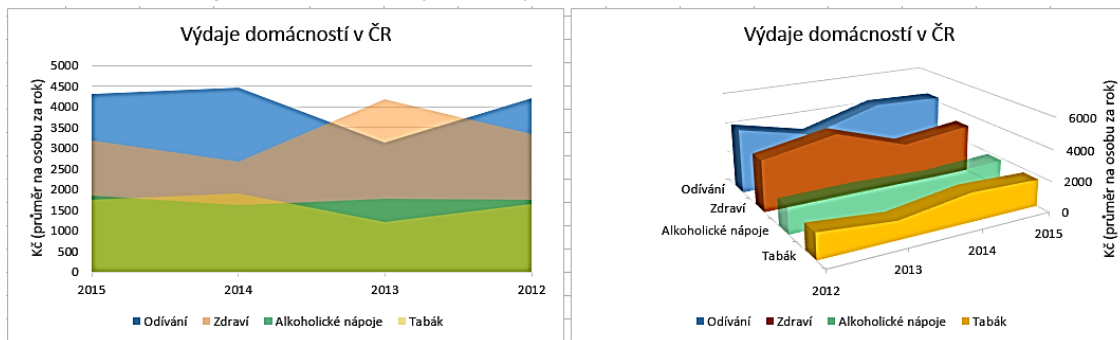
Také spojnicový graf (*Line Chart*) zobrazuje vztah kvantitativních a kvalitativních hodnot v jedné či více datových řadách. Často se využívá čas jako kvalitativní ordinální hodnota. Proto se využívá především pro prezentaci časové řady pro zachycení korelace hodnot (např. plánované vs. skutečné hodnoty v čase). Spojnicový graf se používá také pro kombinaci porovnání hodnot včetně odchylek.

Plošný graf

Tento typ grafu (*Area Chart*) je alternativou pro spojnicový graf a slouží pro zobrazení vývoje kvantitativních hodnot v časovém období s tím, že plocha grafu opticky ještě více zdůrazňuje objem dat. (Singh, 2016) Je důležité, aby data vynášená v zobrazované ploše byla spojitá, nikoliv diskrétní. Tento typ grafu umožňuje i verzi pro skládaný graf a 100 % skládaný plošný graf.

Obrázek 6 zobrazuje ukázkou plošného grafu, ve kterém jsou zobrazena data o ročních výdajích domácností v ČR v průměru na osobu pro vybrané kategorie výdajů. Na

obrázku vlevo je zobrazena 2D varianta grafu, ve které byla data seříděna sestupně podle roků vzhledem k faktu, že výdaje rostou. Tím bylo zajištěno, aby se plochy co nejméně překrývaly a byly tak dobře viditelné. Vpravo je zobrazena 3D varianta plošného grafu, která zobrazuje hodnoty ve vzestupném pořadí roků.



Obrázek 6: Plošný 2D a 3D graf

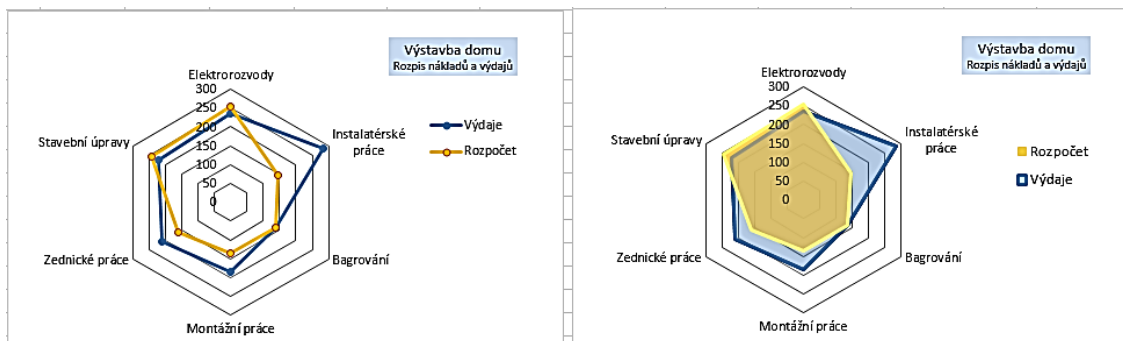
Pro vytváření plošného grafu je vhodné dodržet určitá pravidla (Singh, 2013) pro zajištění jeho čitelnosti:

- graf používat do maximálně 5 datových řad, jinak se stává nepřehledný,
- graf se nehodí pro data, kde se nevyskytují výraznější změny, plochy se tak z velké části překrývají a nejsou dobře viditelné,
- pro lepší přehlednost dat je nutné hodnoty seřadit vzestupně tak, aby v popředí byly plochy, které jsou menší než plochy v pozadí grafu,
- čitelnost lze v grafu dosáhnout použitím transparentní barvy pro výplně ploch.

Paprskový graf

Grafickou formou pro zachycení vzájemného vztahu několika kategorií jedné či dvou proměnných je paprskový graf (*Radar Chart*). Opět se jedná o vztah kvalitativní a kvantitativní proměnné. Graf je vhodnou alternativou ke sloupcovému skládanému grafu nebo pruhovému dvojitému grafu. Plochu grafu tvoří jakási pavučina, kterou tvoří ramena, na kterých jsou vynášeny hodnoty kvantitativních proměnných.

Paprsky odpovídají jednotlivým kategoriím datové řady. Lze vynášet více proměnných, ale pouze jen na jedné ose y, tudíž musí odpovídat jednomu měřítku. Tento typ grafu je velmi oblíbený například pro porovnání plánovaných a skutečných hodnot z nějakého celku, např. čerpání výdajů členěných podle položek vzhledem k plnění rozpočtu, podíly tržeb za vybrané druhy produktů, apod.



Obrázek 7: Paprskový graf

Příklad na Obrázku 7 představuje údaje o čerpání rozpočtu na výstavbu domu. Graf zobrazuje data šesti základních kategorií plánovaných nákladů a skutečných výdajů v projektu. Náklady jsou zobrazeny v tisících Kč. Pavučinu mohou tvořit spojnice (viz graf vlevo na obrázku) nebo vybarvené plochy (ukázka vpravo na Obrázku 7). V tomto příkladě byla použita poloprůhledná barva výplní plochy pavučiny. Díky vizualizaci dat v tomto typu grafu lze velmi rychle odhalit shodu či naopak odchylky porovnávaných hodnot – v tomto příkladu lze rychle zjistit, že kde byl rozpočet přečerpán.

Při rozhodování, zda zvolit paprskový graf, je třeba pamatovat na to, že při vynášení více než dvou proměnných se tento typ grafu stává méně přehledným.

Závěr

V současné době, kdy jsou manažeři doslova zahlceni velkým množstvím dat a kdy jsou současně nuceni se rychle rozhodovat, je vizualizace dat jedním z nástrojů, jak získat rychlé a přehledné informace. Z výše uvedeného textu vyplývá, že vizuál není jen grafickou alternativou „surových“ číselných či znakových dat, ale že může zásadně přinést novou informaci, která z dat vyplývá. Ve vizuálu se kombinuje více vztahů mezi daty najednou, např. korelace, pořadí i časová řada. Právě díky různým grafickým prvkům (tvar, barva, ikona, textura výplně atd.) lze tyto vztahy zachytit. Jak jsou data prezentována ovlivní i jejich přehlednost a srozumitelnost. To znamená, že velmi záleží i na nastavení parametrů vizuálu. U grafů se jedná o měřítka os a jejich popisky, popisky hodnot, volba barvy i typ písma.

Použité zdroje

SINGH, V., 2013: *14 Tips to present awesome charts*. E-book. (online). Jazzfactory.in. (07.10. 2017). Dostupné z: <http://www.24point0.com/ppt-shop/ppt-tips-present-awesome-charts>

Few, S., 2012: *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. 2nd edition. Burlingame, AnalyticsPress. ISBN 978-0-970-60197-1

Pour, J., Maryška, M., Novotný, O., 2012: *Bussiness Intelligence v podnikové praxi*. Praha, Professional publishing. ISBN 978-80-7431-065-2

Šedivá, Z., 2017: Vizualizace dat v návrhu dashboardu v oblasti BI. *Systémová integrace* 3/2017, str. 52 – 62. (07.12. 2017). ISSN1804-2716 (online). Dostupné z: <http://www.cssi.cz/cssi/vizualizace-dat-v-navrhu-dashboardu-v-oblasti-bi>

Visage.co, 2014: *Data Visualization 101: How to Design Charts and Graphs*. E-book. Visage.co. (online) (10.12.2017). Dostupné z: <https://visage.co/data-visualization-101-design-charts-graphs/>

JEL Classification: C80, Y10