

# Vizualizace multidimenzionálních dat

Jan Kupčák

12. července 2008

## Abstrakt

Velké množství dat bez vhodného způsobu prezentace ztrácí z hlediska uživatele svou informační hodnotu. Proto, aby uživatel mohl data správně interpretovat, pochopit vztahy mezi nimi a utvořit si představu o jejich možném budoucím vývoji, je nutné, aby údaje byly vhodně reprezentovány. Možných typů zobrazení již existuje mnoho. Tento dokument představuje a popisuje typy zobrazení, jenž lze použít pro vizualizaci multidimenzionálních dat charakteristických pro OLAP technologie.

## 1 Úvod

V dnešním moderním světě vzniká každý zlomek vteřiny velké množství informací. Totéž platí i obchodu – firmy nakupují a prodávají, mění se velikost jejich kapitálu, některé pobočky se ruší, jiné zase přesouvají nebo rozšiřují, zaměstnanci odcházejí a jsou přijímáni noví. O většině těchto a mnoha dalších operacích rozhodují manažeři, jejichž úkolem je rozvíjet firmu tak, aby vykazovala co nejvyšší zisk. Proto, aby mohli rozhodnout správně, však musí mít dostatečné množství relevantních a spolehlivých informací.

Protože většina firem ke své činnosti využívá informačních technologií, jsou všechny záznamy o těchto operacích ukládány v databázích. Databázová úložiště však nejsou vhodným zdrojem informací pro rozhodování, neboť mohou díky selhání lidského faktoru obsahovat neúplné nebo chybné informace. Nevýhodou je také fakt, že požadovaná informace může být rozmístěna ve více zdrojích a její skládání je při analýze náročné. Proto jsou data upravována a přenášena do datových skladů, kde jsou dlouhodobě uchovávána. Pomocí nástrojů OLAP a multidimenzionálního přístupu, který je pro OLAP charakteristický, je mnohem snazší k těmto informacím nejen přistupovat a zobrazovat je, ale také s nimi interaktivně pracovat, analyzovat je, hodnotit trendy, nebo třeba hledat ukryté zákonitosti v historických i aktuálních informacích, které z dat přímo nevyplývají.

Uvedené činnosti by však nebylo možné provádět bez sofistikovaných nástrojů. Díky toho, že datové sklady obsahují velké množství nejen současných, ale i historických dat, je potřeba, aby uživatel věděl, se kterými údaji potřebuje pracovat. Pokud by bylo uživateli vráceno větší množství údajů pouze v kontingenční tabulce, nebyl by zřejmě schopen udělat si jasný obrázek o jejich vztahu, rozložení a dalších parametrech, které chce sledovat. Data by opět ztrácela svou hodnotu.

Do hry vstupuje vizualizace dat, která převádí množství ve své podstatě surových údajů do snadněji pochopitelných grafických podob, jako jsou grafy, a další pokročilé typy zobrazení. Díky vizualizace značně stoupá schopnost interpretovatelnosti a použitelnosti nasbíraných obchodních dat. Data jsou zobrazována v podobě různých geometrických útvarů (body, čáry, obdélníky, . . .), jejichž atributy (velikost, barva, styl výplně, . . .) reprezentují skutečnou hodnotu. Čím názornější

je vizuální podoba dat, tím větší hodnotu mají nasbírané informace z procesu podnikání, čímž je lze snadněji pochopit a především následně i předpovídat jejich vývoj.

Hlavním cílem tohoto dokumentu je uvést velkou část možných zobrazení multidimenzionálních dat. Navíc je u každého typu reprezentace uvedeno, kolik dimenzí a faktů může být zobrazeno.

## 2 Multidimenzionální datový model

Multidimenzionální datový model, na kterém jsou technologie OLAP postaveny, představuje dva základní pojmy – dimenze a fakta (označována v anglickém textu jako *measures*).

Každá dimenze reprezentuje jistý úhel pohledu na data. Je tvořena jedním nebo více atributy, které danou informaci o pohledu popisují, přičemž může vytvářet logicky nebo organizačně uspořádanou hierarchii. Jestliže dimenze představuje pouze seznam nějakých záznamů, je automaticky doplněn nadřazený prvek, který je obsahuje. Příkladem běžné dimenze je geografická dimenze, ve které jsou uloženy například adresy poboček podniku. Hierarchie by pak mohla být členěna od nejvyšší úrovně na státy, regiony, města, až po konkrétní adresy. Mezi dimenzemi lze vždy najít časovou, která je obvykle dělena na roky, kvartály, měsíce, týdny, dny, atd. Uživatel má vždy možnost nadefinovat úroveň časové dimenze, které požaduje.

Samotné údaje, tedy konkrétní číselné hodnoty pocházející z činnosti podniku, se skrývají pod pojmem fakta. Každá hodnota se nachází v nějakém průsečíku nejnižších úrovní všech dimenzí dané kostky. Kromě těchto dat pocházejících z externích datových zdrojů se zde vyskytují navíc agregace, které jsou vypočítány až v datovém skladu. Funkce, kterou agregace bude představovat, je možno uvádět v definici příslušné dimenze. Může se jednat o součet, průměr, počet hodnot, maximální hodnotu a další. Tyto celkové hodnoty se nacházejí opět na průsečíku všech dimenzí, avšak alespoň jedna hodnota dimenze je z vyšší úrovně než nejnižší (nejdetailnější). Je zřejmé, že ne všechny průsečíky dimenzí budou obsahovat nějakou hodnotu, protože například při třech dimenzích místo, produkt a čas nelze očekávat, že každá pobočka prodá každý den všechny výrobky. Při velkém počtu dimenzí tak snadno vznikají řídké shluky dat.

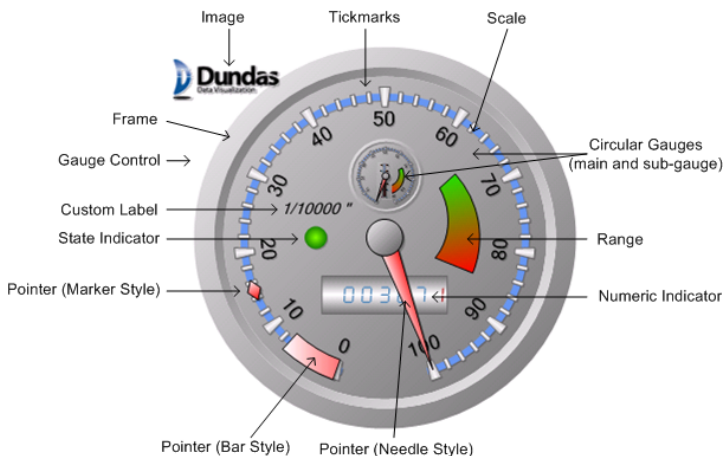
## 3 Reprezentace konkrétních hodnot

Konkrétní číselnou hodnotu lze reprezentovat dvěma způsoby: první možností je vypsát ji uživateli vhodným způsobem tak, aby byla „vizuálně atraktivní“. Většinou je tedy použit jiný typ a výška fontu než u okolního textu, mohou být použita barevná zvýraznění, okraje, atd. Příklad je vidět na obrázku 1a. K číslu může být dále připojena například šipka ukazující trend, minulá hodnota nebo informace o stavu aktuální hodnoty, tj. zda se hodnota nachází v povoleném rozsahu. Tento stav může být charakterizován barvou textu, nebo například vedle umístěným barevným polem – autoři se obvykle inspirovali semaforem, kde zelená barva znamená „hodnota je v rozsahu“, žlutá barva upozorňuje na uživatele na ostražitost a červená barva pak znamená hodnotu mimo povolený rozsah.

Hodnotu lze reprezentovat i čistě grafickou cestou, viz obrázky 1b, 1c. Nejčastěji se používají různé „metry“ a ukazatele (angl. *gauges*), které jsou obohaceny o stupnici s hodnotami. Skutečně zobrazovaná hodnota je pak tvořena sloupcem, nebo ručkou, jenž dosahují určité hodnoty na stupnici. Výhodou je možnost zobrazení více hodnot včetně jejich historie, rozsah očekávaných hodnot je názornější a celková reprezentace může působit příjemněji. Při jejich návrhu však nesmí dojít k přesycení uživatele informacemi a ztrátě přehlednosti. Příklad rozšíření ukazatele o další množství informací je na obrázku 3, zdroj [2].

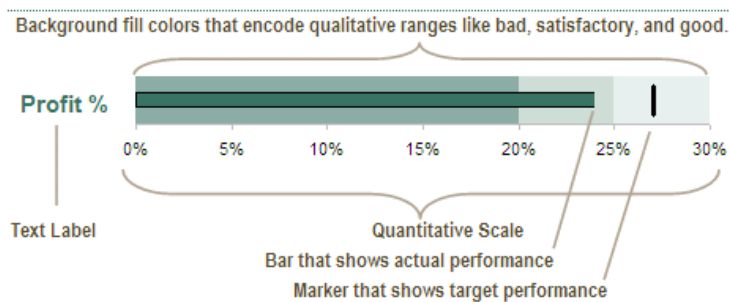


Obrázek 1: Možnosti reprezentace konkrétní hodnoty.



Obrázek 2: Kruhový ukazatel lze rozšířit o velké množství informací.

Ukazatel na obrázku 3 a další jemu podobné mají značnou nevýhodu, pokud mají být spolu s dalšími informacemi umístěny na displeji s malým rozlišením. V takovém případě zabírají zbytečně mnoho místa. Stephen Few upravil tyto ukazatele do podoby tzv. *bullet* grafu [4], viz obrázek ???. Lze vidět, že tento typ grafu ukazuje nejen aktuální hodnotu, ale i další, se kterou lze aktuální hodnotu snadno porovnat, a dále pak i barevně odlišené oblasti, které informují uživatele o stavu hodnoty. Navíc lze pod sebe umístit skupinu těchto grafů, a tak velice snadno srovnávat jejich hodnoty. To by u kruhových ukazatelů mohl být problém.



Obrázek 3: *Bullet graph*, jenž vytvořil Stephen Few. Obrázek převzat z [6].

## 4 Běžné typy grafů

Do této skupiny se řadí běžné grafy používané například v souhrnných zprávách (tzv. reportech). Jedná se o sloupcový, spojnicový, plošný, výsečový, paprskový a také X-Y graf.

Pro zjednodušení následující klasifikace zavedme matematické označení. Multidimenzionální kostka, se kterou uživatel pracuje, obsahuje množinu dimenzí  $D = \{D_1, D_2, \dots, D_u\}$  a množinu faktů  $M = \{M_1, M_2, \dots, M_v\}$ . Všechny níže uvedené dimenze i fakty budou vždy implicitně spadat do těchto dvou množin. Dále nechť:

- $D_d$  – představuje libovolnou dimenzi, přičemž  $d$  je přirozené číslo v rozsahu  $1, \dots, u$ . Jestliže  $d = time$ , pak jde právě o časovou dimenzi.
- $D_d^f = \{[D_d^f]_1, [D_d^f]_2, \dots, [D_d^f]_n\}$  – označuje množinu položek dimenze  $D_d$  odpovídajících výběrovému filtru  $f$ . Bez omezení obecnosti můžeme předpokládat, že si libovolná položka nese nejen své jméno, ale také kompletní informaci o svých předcích v hierarchii dimenze, vč. názvu dimenze, do které spadá. Například tedy  $[D_{time}^f]_2 = [Time].[2006].[Q1].[2]$  označuje položku únor, jenž spadá do prvního kvartálu roku 2006 v časové dimenzi. Výběrový filtr  $f$ , který by tuto položku vrátil, by mohl být například  $\{[ALL].[2006].[Q1].members\}$ . Ve většině níže uvedených případů je filtr reprezentován nedefinovanou proměnnou  $(\alpha, \beta, \gamma, \dots)$  – jedná se o filtr, jenž vrací přesně ty položky, jenž si uživatel přeje zobrazit.
- $M_m$  – představuje určitý fakt (údaj), kde  $m = 1, \dots, v$ .
- $M_m \left( [D_1^{f_1}]_{i_1}, [D_2^{f_2}]_{i_2}, \dots, [D_n^{f_n}]_{i_n} \right)$  – reprezentuje hodnotu průsečíku prvků uvedených dimenzí a zbývajících dimenzí, které nejsou uvedeny v tomto seznamu parametrů. Předpokládejme, že prvky těchto neuvedených dimenzí jsou vyjmenovány jinde, pro daný graf jsou považovány za statické, a nemají tudíž vliv na zobrazení.

Poznámka – grafy, které zobrazují celky rozdělené na určité části, tj. používají procenta, nejsou schopné korektně pracovat se zápornými hodnotami faktů. Na tuto skutečnost již dále není upozorňováno, v těchto případech se předpokládají pouze kladné hodnoty.

### 4.1 Sloupcový graf

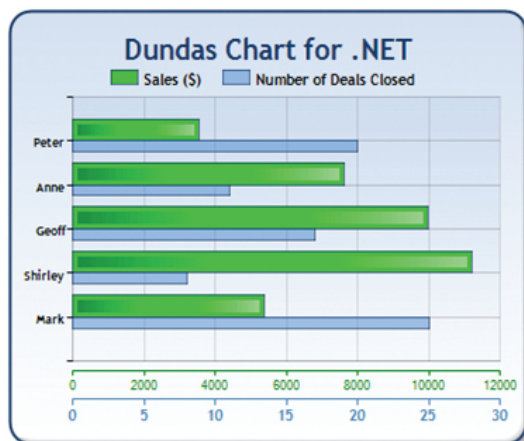
Sloupcový graf (*bar chart*) je složen ze sloupců obvykle obdélníkového tvaru, jejichž délka odpovídá hodnotám, které představují. To následně umožňuje kvantitativní porovnání těchto údajů. Existují různé modifikace – sloupce lze umístit vedle sebe, či nad sebe, mohou být také realizovány jako trojúhelníky, válce, případně jim přidělit jiné grafické symboly a graf obecně může být zobrazen ve dvou nebo tří rozměrném prostoru.

Z pohledu multidimenzionálního přístupu se v grafu může nacházet:

- 1 dimenze  $D_1$  a 1 údaj  $M_1$  – graf obsahuje sloupce  $b_1, b_2, \dots, b_m$ , kde  $m = |D_1^\alpha|$ , jenž odpovídají prvkům množiny  $D_1^\alpha$ . Délka jednotlivých sloupců odpovídá hodnotě zvoleného faktu  $M_1$ , tedy  $delka(b_i) = M_1([D_1^\alpha]_i)$  pro všechna  $i = 1, \dots, m$ . Je zřejmé, že se jedná o relativní délku, skutečná délka sloupce, například v pixelech, bude nějakým způsobem získávána z této této hodnoty a také z dalších parametrů aktuálního zobrazení grafu. Příklad tohoto typu grafu lze najít na obrázku 4, jenž byl převzat z [1].

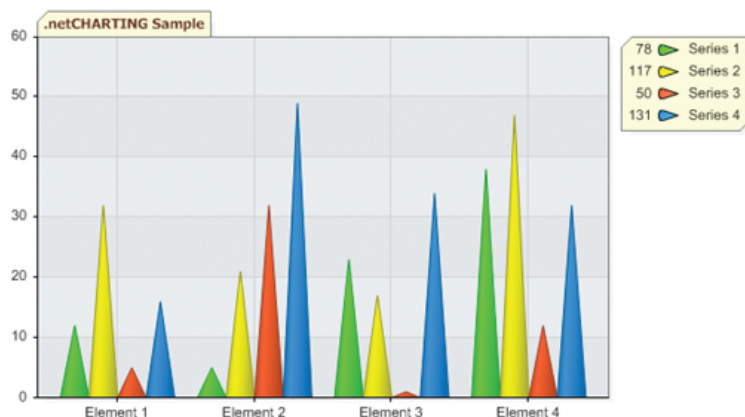


Obrázek 4: Sloupcový graf zobrazující jednu dimenzi a jeden fakt.



Obrázek 5: Sloupcový graf ukazující jednu dimenzi a více faktů.

- 1 dimenze  $D_1$  a  $r$  údajů  $M_1, \dots, M_r$  – graf obsahuje pro každý prvek z  $D_1^\alpha$  skupinu sloupců, každý z nich odpovídá jinému faktů. Graf tedy obsahuje sloupce  $b_{1,1}, b_{1,2}, \dots, b_{1,r}, b_{2,1}, \dots, b_{n,r}$ , přičemž  $n = |D_1^\alpha|$ . V tomto případě může být použito i více hodnotových os – každá by příslušela jiné sledované veličině. Relativní délka sloupce je dána vztahem  $delka(b_{i,j}) = M_j([D_1^\alpha]_i)$  pro všechna  $i = 1, \dots, n$  a  $j = 1, \dots, r$ . Viz obrázek 5, zdroj [3].
- 2 dimenze  $D_1, D_2$  a 1 veličina  $M_1$  – graf se skládá z  $n = |D_2^\beta|$  skupin sloupců, v každé této skupině se nachází  $m = |D_1^\alpha|$  sloupců odpovídajících prvkům z  $D_1^\alpha$ . Každému prvku z  $D_1^\alpha$  pak přísluší jiný typ výplně sloupce, čímž dojde ke snadnému odlišení. Celkově má graf sloupce  $b_{1,1}, b_{2,1}, \dots, b_{m,1}, b_{1,2}, \dots, b_{m,n}$ .  $delka(b_{i,j}) = M_1\left([D_1^\alpha]_i, [D_2^\beta]_j\right)$  pro všechna  $i = 1, \dots, m$  a  $j = 1, \dots, n$ . Příklad grafu je na obrázku 6.
- 2 dimenze  $D_1, D_2$  a 1 veličina  $M_1$  se zvýrazněním součtu – v grafu se nachází sloupce  $b_1, \dots, b_n$ , kde  $n = |D_2^\beta|$ , přičemž ty jsou dále rozděleny na  $m = |D_1^\alpha|$  částí. Jestliže jsou všechny získané hodnoty z průsečíku dimenzí kladné, pak je situace jednoduchá – jednotlivé



Obrázek 6: Sloupcový graf s dvěma dimenzemi a jedním faktem.

sekce daného sloupce umístěné nad sebou od počáteční osy postupně reprezentují prvky z  $D_1^x$ . Sloupce lze snadno porovnávat, protože prvky z  $D_1^x$  mají v celkovém sloupci stále stejnou pozici a jsou tak „vedle sebe“. Viz obrázek 7.

Situace se komplikuje, jestliže se mezi získanými hodnotami vyskytuje nějaká záporná. V takovém případě je zřejmě nejjednodušší řešení začít od počáteční osy, postupně pocházet prvky z  $D_1^x$  a získávat hodnoty faktu. Je-li hodnota kladná, pak je sloupec reprezentující danou hodnotu položen na nejvýše položený sloupec v kladné části sloupců. Je-li naopak získaná hodnota záporná, je sloupec umístěn pod nejnižše položený sloupec v záporné části sloupců. Celkový sloupec tedy neroste jen jedním směrem jako v předchozím případě, ale roste oběma směry od počáteční osy v závislosti na znaménku získané hodnoty.

Celková délka  $j$ -tého sloupce je dle očekávání:

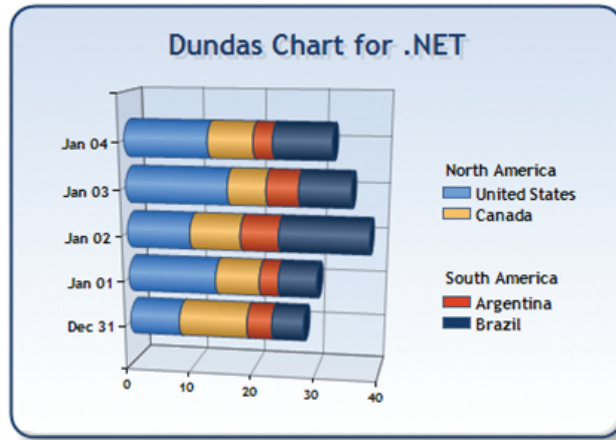
$$celkova\_delka(b_j) = delka\_sloupcu^+(b_j) - delka\_sloupcu^-(b_j) = \sum_i M_1 \left( [D_1^\alpha]_i, [D_2^\beta]_j \right) - \sum_j M_1 \left( [D_1^\alpha]_i, [D_2^\beta]_j \right),$$

kde  $i$  označuje indexy, kde je hodnota faktu kladná a  $j$  nabývá indexů, v nichž je získaná hodnota průsečíku záporná.

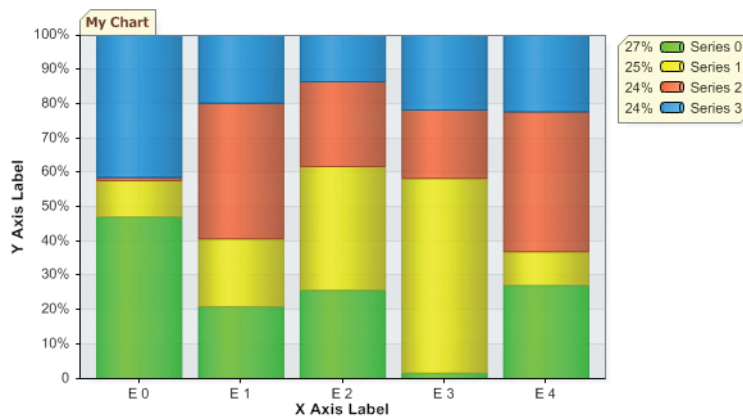
- 2 dimenze  $D_1, D_2$  a 1 veličina  $M_1$  se zobrazením procentuálního podílu – v předchozím případě byly celkové délky sloupců jiné, byl-li jejich součet různý. V tomto případě jsou všechny sloupce stejně dlouhé, součet hodnot v daném sloupci je chápán jako, tedy 100%. Označme  $sum_j = \sum_{i=1}^m M_1 \left( [D_1^\alpha]_i, [D_2^\beta]_j \right)$ . Pak určitá část  $p_i$  ve sloupci  $b_j$  zabírá  $\frac{M_1 \left( [D_1^\alpha]_i, [D_2^\beta]_j \right)}{sum_j} \%$  místa. Příklad grafu je uveden na obrázku 8.

## 4.2 Spojnicový graf

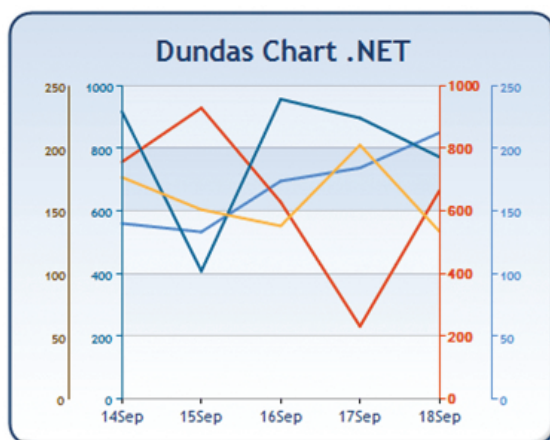
Spojnicový graf (*line chart*) umožňuje sledovat vztah dvou proměnných, přičemž jedna je závislá a druhá nezávislá. Ukazuje, jakým způsobem se mění závislá proměnná na základě změny stavu nezávislé proměnné pomocí sekvence propojených bodů reprezentujících hodnoty sledované veličiny. V grafech s více křivkami jsou pro přehlednost jednotlivé spojnice bodů obvykle odlišeny od ostatních jinou barvou, jiným typem (plná, čárkovaná, atd.), případně jsou také různě graficky



Obrázek 7: Sloupcový graf zobrazující dvě dimenze a jeden fakt s vyjádřením součtu.



Obrázek 8: Sloupcový graf reprezentující dvě dimenze a jeden fakt se zobrazením procentuálního podílu.



Obrázek 9: Spojnicový graf zobrazující časovou dimenzi a vybrané údaje.

ztvárněné body.

Tyto typy grafů jsou obvykle používány v případech, kdy je cílem zkoumání vývoje sledovaných veličin v čase. Uživatel si pak může udělat snadněji představu o aktuálních trendech, identifikovat vzory, a následně pak předpovídat jejich budoucí vývoj.

Spojnicový graf dokáže přehledně zobrazit následující případy:

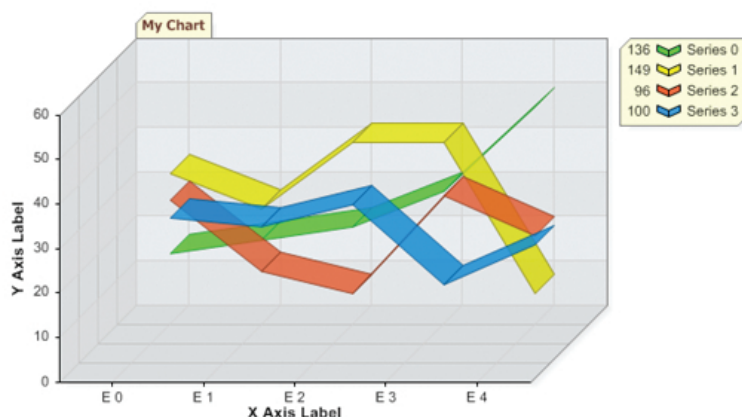
- 1 dimenze  $D_1$  a  $r$  faktů  $M_1, \dots, M_r$  – jak již bylo řečeno, nejčastěji se zde vyskytuje časová dimenze, tedy  $D_1 = D_{time}$ . Její vybrané prvky  $D_1^\alpha$  tvoří body na první ose grafu, v případě  $r > 1$  se jedná obvykle o osu x. Graf má  $r$  křivek, pozice jednotlivých bodů v případě  $D_1^\alpha$  zobrazených na ose x je  $[x_i, y_j = M_j([D_1^\alpha]_i)]$  pro  $i = 1, \dots, |D_1^\alpha|$ , kde  $j$  udává index sledovaného faktu. Situaci ukazuje příklad na obrázku 9.
- 2 dimenze  $D_1, D_2$  a 1 veličina  $M_1$  – stejně jako v předchozím případě, i zde bývá velice často  $D_1 = D_{time}$ . Na první ose jsou zobrazeny prvky  $D_1^\alpha$ , na druhé ose pak hodnoty. Graf obsahuje  $n = |D_2^\beta|$  křivek, každému prvku z  $D_2^\beta$  tedy přináleží právě jedna. Jestliže jsou prvky z  $D_1^\alpha$  umístěny na ose x, pak jednotlivé body jsou umístěny na pozicích  $[x_i, y_j = M_1([D_2^\beta]_i)]$ , pro  $j = 1, \dots, n$ , kde  $i$  udává index prvku z  $D_1^\alpha$ . Příklad je uveden na obrázku 10.

### 4.3 Plošný graf

Plošný graf (*area chart*) rozšiřuje vizuální vlastnosti spojnicového grafu. Křivky ze spojnicového grafu zde tvoří hranice ploch, které jsou jistým způsobem odlišeny od ostatních. Tím je možné zobrazovanou informaci zvýraznit.

Aby bylo možné zobrazit plochy grafu, musí být stanoveny jejich hranice. Pro tento účel definujeme množinu sestávající se vzájemně různých uspořádaných dvojic  $(a, b)$ , kde  $a, b$  představují rovnice dvou křivek ohraničujících příslušnou plochu. Tvar jejich rovnice je mimo jiné také závislý na tom, kde jsou zobrazeny prvky z  $D_1^\alpha$  – jestliže jsou umístěny na ose x, resp. y, pak  $a$  i  $b$  označuje horizontální, resp. vertikální okraj oblasti. Zbývající dvě hranice všech oblastí tvoří přímky procházející obvykle prvním a posledním prvkem z  $D_1^\alpha$ , jenž jsou kolmé na osu s těmito prvky.



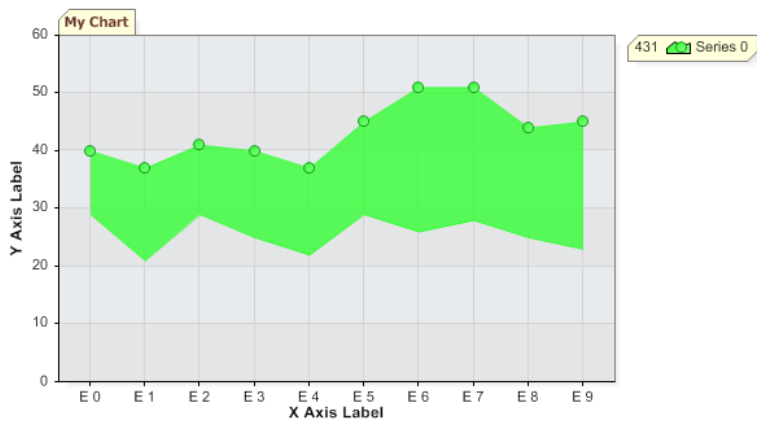


Obrázek 10: Spojnicový graf ukazující dvě dimenze a jeden fakt.

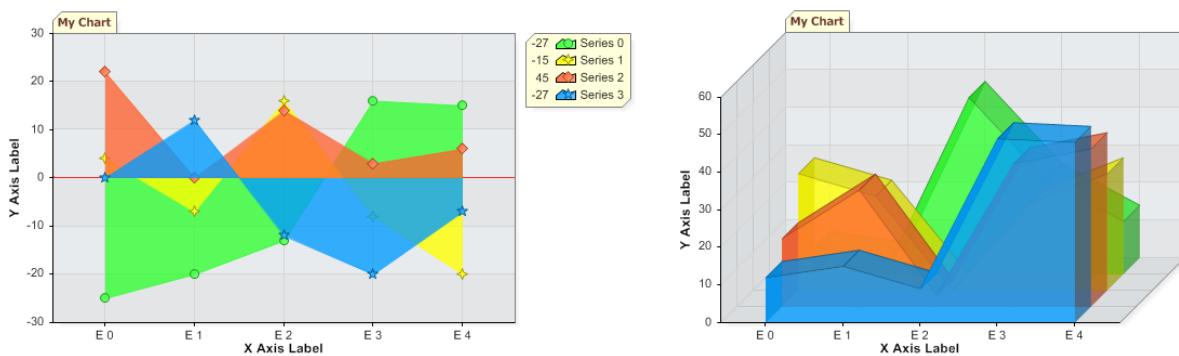
Rovnice křivky realizující hranici oblasti může být dána dvěma způsoby. Může jít o propojené body, jejichž poloha je závislá na aktuálním typu plošného grafu – jedná se v podstatě o křivky ze spojnicového grafu. Ve druhém případě je hranice dána rovnicí určité konstanty  $c$  – jestliže jsou prvky z  $D_1^\alpha$  umístěny na ose  $x$ , pak má rovnice tvar  $y = c$ , v opačném případě je to  $x = c$ . Dále již budeme uvažovat jen případ, kdy prvky z  $D_1^\alpha$  jsou stejnoměrně rozmístěny na diskretní ose  $x$ , hodnoty sledovaných faktů jsou pak na spojitě ose  $y$ .

Jak již bylo uvedeno výše, dvojice v množině definující oblasti musí být vzájemně různé. Navíc musí obsahovat všechny fakta, či prvky zvolené dimenze, které si uživatel přeje zobrazit. Oblastem, jenž jsou určeny jednotlivými dvojicemi, jsou pak přiřazeny určité vizuální reprezentace dle aktuálních potřeb. Je však nutné při tom pamatovat na to, aby případné vzájemně překrývající se oblasti byly nějakým způsobem jasně rozlišitelné.

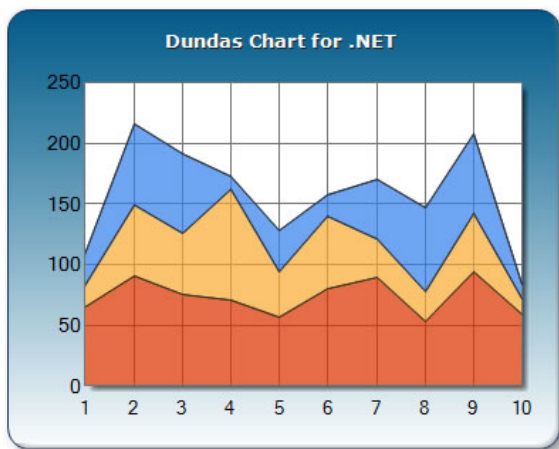
- 1 dimenze  $D_1$  a  $r$  veličin  $M_1, \dots, M_r$  se stejným rozsahem hodnot – v tomto typu zobrazení nelze počet dvojic množiny definující oblasti obecně stanovit, protože vychází z aktuálních požadavků a charakteru dat.  $a$  i  $b$  mohou být konstanty i rovnice ve tvaru  $y = M_i(x)$ , kde  $i = 1, \dots, r$ . Na obrázku 11 je uveden příklad s jednou dimenzí a dvěma fakty, oblasti jsou definovány množinou  $\{(M_1(x), M_2(x))\}$ .
- 2 dimenze  $D_1, D_2$  a 1 fakt  $M_1$  – graf obsahuje  $n = |D_2^\beta|$  oblastí, které jsou určeny množinou  $\{(0, M_1(x, [D_2^\beta]_1)), (0, M_1(x, [D_2^\beta]_2)), \dots, (0, M_1(x, [D_2^\beta]_n))\}$ . Každá oblast tedy přínáleží jednomu prvku z  $D_2^\beta$  a její hranice je dána přímkou kolmou na osu  $x$  procházející nulou na ose  $y$  a spojnicemi bodů, jejichž poloha je určena hodnotou průsečíku aktuálních prvků dimenzí. Na obrázku 12 jsou zobrazeny příklady dvou grafů tohoto typu – první je ve dvourozměrném a druhý v třírozměrném prostoru. Je vidět, že 2D graf může být náročnější na interpretaci, pokud se překrývá více ploch a přitom nejsou zvoleny vhodné výplně. Tuto nevýhodu lze částečně eliminovat zobrazením v 3D, avšak pak musí systém uživateli nabízet možnost rotování grafu kolem všech tří os. Druhou možností v obou případech je využít selektivního zvýrazňování jednotlivých vrstev, kdy většina vrstev má nastavenou vyšší hodnotu průhlednosti a jen vybrané se zobrazují jasněji.
- 2 dimenze  $D_1, D_2$  a 1 fakt  $M_1$  se zobrazením celkového součtu – má-li množina  $D_2^\beta$   $n$



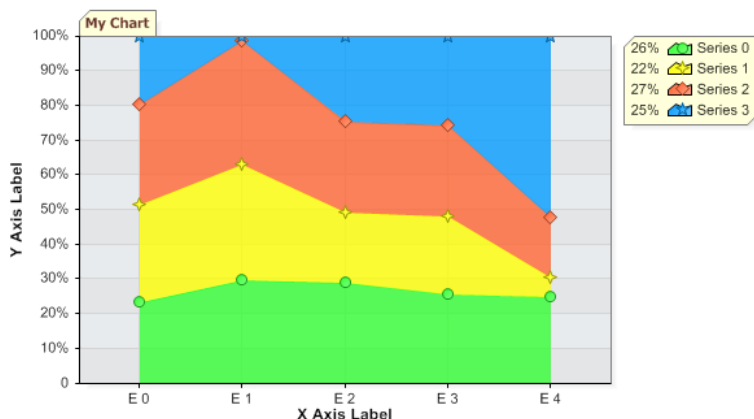
Obrázek 11: Plošný graf s jednou dimenzí a dvěma veličinami se stejným rozsahem tvořících oblast.



Obrázek 12: Srovnání dvou- a třírozměrného plošného grafu zobrazujících dvě dimenze a jeden fakt.



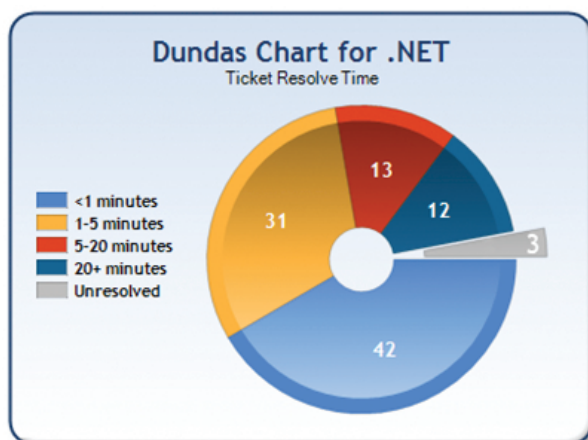
Obrázek 13: Plošný graf se dvěma dimenzemi a jedním součtovým faktem.



Obrázek 14: Plošný graf reprezentující dvě dimenze a jednu veličinu se zobrazením podílu.

prvků, pak  $i$ -tá oblast od výchozí x-ové osy příslušející  $i$ -tému prvku je reprezentována dvojicí  $(0, f_i(x))$  pro  $i = 1$  a  $(f_{i-1}(x), f_{i-1}(x) + f_i(x))$  pro  $i = 2, \dots, n$ , kde  $f_i(x) = M_1(x, [D_2^\beta]_i)$ .

- Uvedený vztah platí pouze v případě, že všechny průsečíky mají nezápornou hodnotu. V opačném případě je řešení popsáno u stejného typu zobrazení ve sloupcovém grafu. Příklad plošného grafu s kladnými hodnotami a součty je uveden na obrázku 13.
- 2 dimenze  $D_1, D_2$  a 1 fakt  $M_1$  s dělením celku na příslušné části – vizuální podoba grafu je podobná sloupcovému grafu s rozdělením celku, jen body příslušející jistému prvku  $[D_2^\beta]_i$  tvoří jednu hranici plochy. Obecný vzorec uvedený v předchozím typu zobrazení platí i zde, rozdíl je v  $f_i(x)$ . Zde platí  $f_i(x) = M_1\left(x, \frac{[D_2^\beta]_i}{\text{sum}(x)}\right) \cdot 100\%$ , kde  $\text{sum}(x) = \sum_{j=1}^n M_1\left(x, [D_2^\beta]_j\right)$ . Příklad tohoto typu lze najít na obrázku 14.



Obrázek 15: Výšečový graf zobrazující jednu dimenzi a jeden fakt.

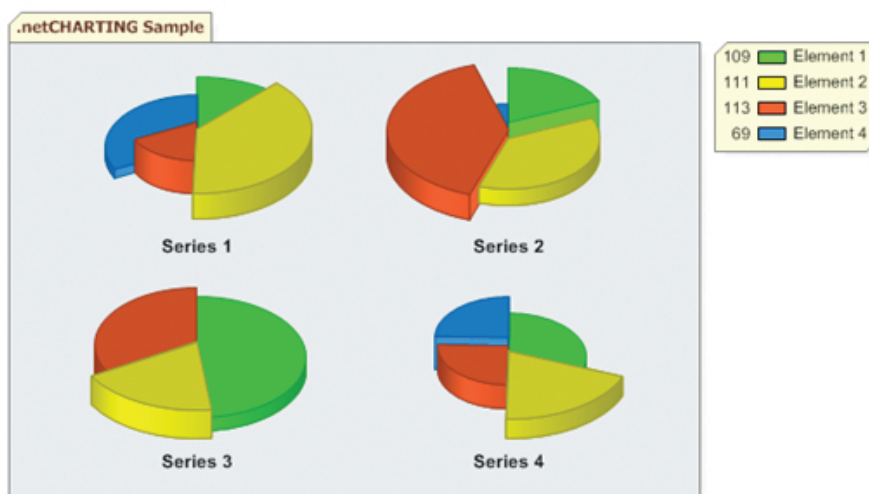
#### 4.4 Výšečový graf

Výšečový, nebo také koláčový graf (*pie chart*), znázorňuje dělení celé zkoumané množiny v podobě celého kruhu na jednotlivé části. Ty jsou reprezentovány kruhovými výsečemi, přičemž jejich úhel je přímo úměrný velikosti části vzhledem k celku.

Nejčastěji se tento graf používá právě pro jasné znázornění rozdělení určitého celku. V [10] však autor varuje, že v některých případech je použití tohoto typu grafu nevhodné. Porovnávání přibližně stejně velkých částí a odhadování skutečné hodnoty z grafu může být problematické díky různým úhlům výseků, a schopnost korektního úsudku se tak snižuje. Z tohoto vyplývá, že jednoduchý koláčový graf bez přesných hodnot je vhodný pouze pro orientaci, pro lepší odhad je doporučováno obohatit graf o přesné hodnoty, použít méně výseků, nebo zvolit jiný typ grafu (např. sloupcový).

Tento typ grafu může být použit k zobrazení:

- 1 dimenze  $D_1$  a 1 fakt  $M_1$  – nejjednodušší typ koláčového grafu, kde je celek rozdělen na  $m = |D_1^\alpha|$  částí, kruh je tedy složen z  $m$  výsečí. Jestliže hodnota  $sum = \sum_{i=1}^m M_1([D_1^\alpha]_i)$  představuje celek, je-li  $f_i = \frac{[D_1^\alpha]_i}{sum}$  a prvek  $[D_1^\alpha]_i$  tedy zabírá  $f_i \cdot 100\%$ , pak úhel výseče příslušející danému prvku je  $2\pi f$  rad. Příklad je na obrázku 15.
- 1 dimenze  $D_1$  a až 3 fakty  $M_1, M_2, M_3$  – představuje rozšíření předchozího typu zobrazení o další dvě veličiny. Jak již bylo řečeno, hodnotě prvního faktu  $M_1$  odpovídá úhel výseče. Druhý fakt, jehož hodnoty musí být vždy kladné, je zohledněn poloměrem daného výseku, tedy  $r_i = M_2([D_1^\alpha]_i)$ . Výška  $h_i$  dané výseče koresponduje s hodnotou třetího faktu, takže  $h_i = M_3([D_1^\alpha]_i)$ . V obou případech  $i$  označuje index prvku z  $D_1^\alpha$  neboli identifikuje příslušný výsek. Je zřejmé, že díky použití třetího rozměru, tedy výšky, je potřeba zobrazovat daný koláčový graf ve třírozměrném prostoru. Příklad lze najít na obrázku 16.
- 2 dimenze  $D_1, D_2$ , 1 údaj  $M_1$  a další možný agregovaný údaj  $M_2$  pro dimenzi  $D_2$  – v tomto případě je celý kruh rozdělen na  $n = |D_2^\beta|$  na sebe navazujících mezikruží. Do každé této části je promítnuto  $m = |D_1^\alpha|$  kruhových výsečí, jejichž střed je vždy ve středu koláčového grafu. Úhel těchto výsečí koresponduje s hodnotou faktu  $M_1$ , viz vzorec v předchozím typu.



Obrázek 16: Skupina trojrozměrných výšečových grafů reprezentujících jednu dimenzi a tři veličiny.

Pozn.: pro snadnější porovnávání všech  $n$  výšečí reprezentující první prvek z  $D_1^\alpha$  obvykle začíná na 0 rad.

Jestliže jsou šířky všech mezikruží stejné, pak graf zobrazuje pouze jeden fakt  $M_1$ . Do grafu však lze přidat ještě další informaci, jenž bude reprezentována šířkami mezikruží. Jestliže poloměr celého kruhu  $R$  je chápán jako celek, který odpovídá hodnotě 100%, pak šířka  $i$ -tého mezikruží je  $\frac{[D_2^\beta]_i}{sum} \cdot 100\%$ , kde  $sum = \sum_{i=1}^n [D_2^\beta]_i$ . Je však potřeba uvědomit si, že tento údaj je závislý pouze na prvcích  $D_2^\beta$ . Příklad tohoto typu grafu je zobrazen na obrázku 17.

- 2 dimenze  $D_1, D_2$  a 2 fakty  $M_1, M_2$  se zobrazením procentuálního podílu – jako ve všech předchozích případech, i zde je fakt  $M_1$  reprezentován úhlem kruhové výšeče. V tomto typu grafu je však úhel závislý pouze na prvcích z  $D_1^\alpha$ . Každá výšeč je rozdělena na  $n = |D_2^\beta|$  mezikruží, které dohromady tvoří celek – poloměr kruhu je opět chápán jako 100%.

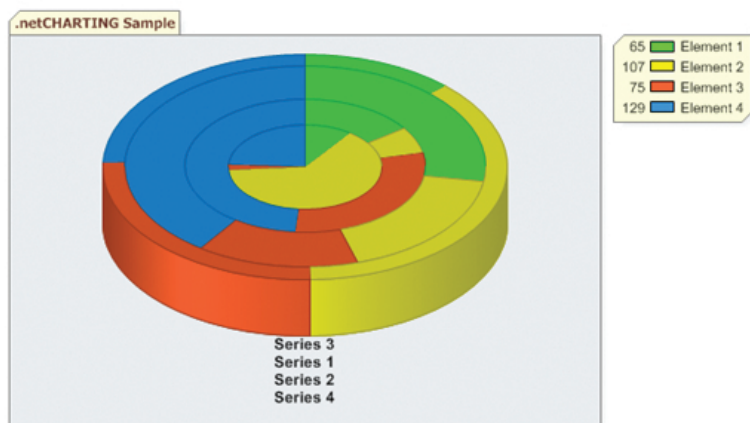
Označíme-li  $sum_i = \sum_{j=1}^n M_2 \left( [D_1^\alpha]_i, [D_2^\beta]_j \right)$  jako celkový součet hodnot daného výseku

s indexem  $i$ , pak šířka mezikruží  $j$  v tomto tvoří  $\frac{M_2 \left( [D_1^\alpha]_i, [D_2^\beta]_j \right)}{sum_i} \cdot 100\%$ . Je zřejmé, že lze také zobrazit skutečné hodnoty (místo procentuálních), a to tak, že povolíme různý poloměr jednotlivých výšečí – viz případ sloupcového grafu s  $D_1, D_2, M_1$  se zvýrazněním součtu. Příklad tohoto typu koláčového grafu s procenty je na obrázku 18.

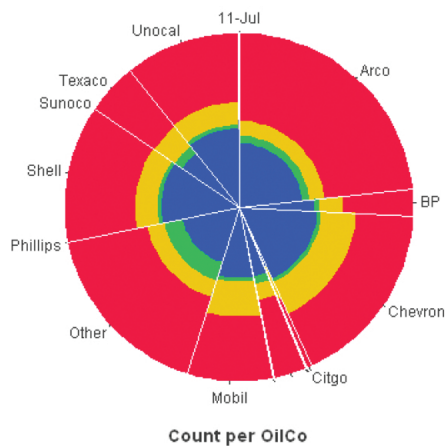
## 4.5 Paprskový graf

Paprskový graf (*radar, star chart*) je založen na centrálním bodu, z něhož vychází skupina os se stejnými hodnotami i měřítkem. Každá tato osa reprezentuje hodnotu nezávislé veličiny a na ni se vynáší hodnota sledované (závislé) veličiny. Osy obvykle bývají rozprostřeny přes celou plochu grafu rovnoměrně, tj. úhel mezi dvěma sousedícími osami je stejný.

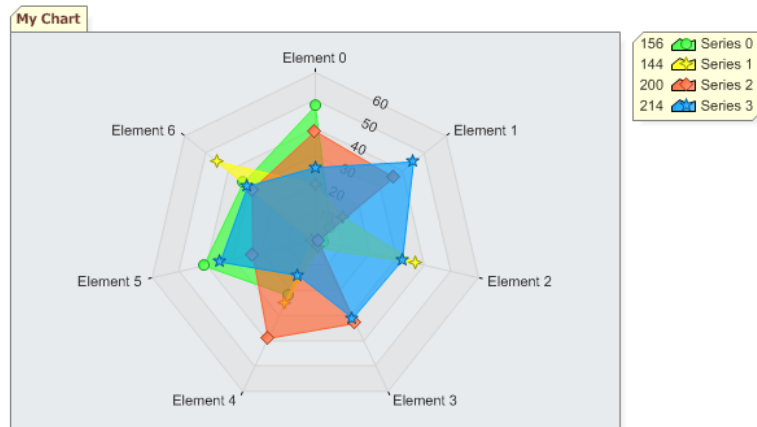
Graf dokáže zobrazit:



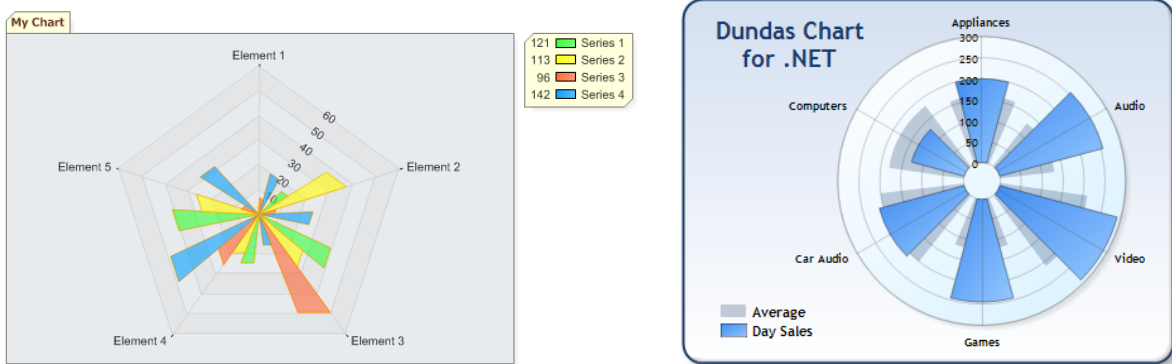
Obrázek 17: Výsečový graf zobrazující dvě dimenze a dva fakty, kde druhý údaj je agregován pro zvolenou dimenzi.



Obrázek 18: Výsečový graf ukazující dvě dimenze a dva fakty s důrazem na srovnání podílu.



Obrázek 19: Paprskový graf ukazující dvě dimenze a jeden fakt pomocí oblastí.

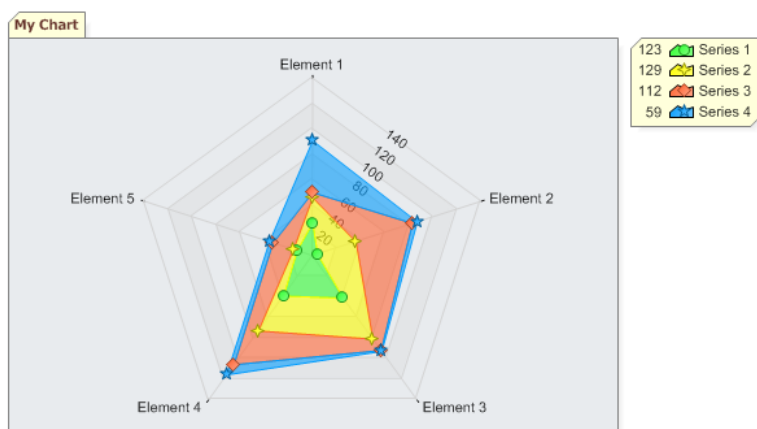


Obrázek 20: Paprskové graf ukazující dvě dimenze a jeden fakt pomocí sloupců.

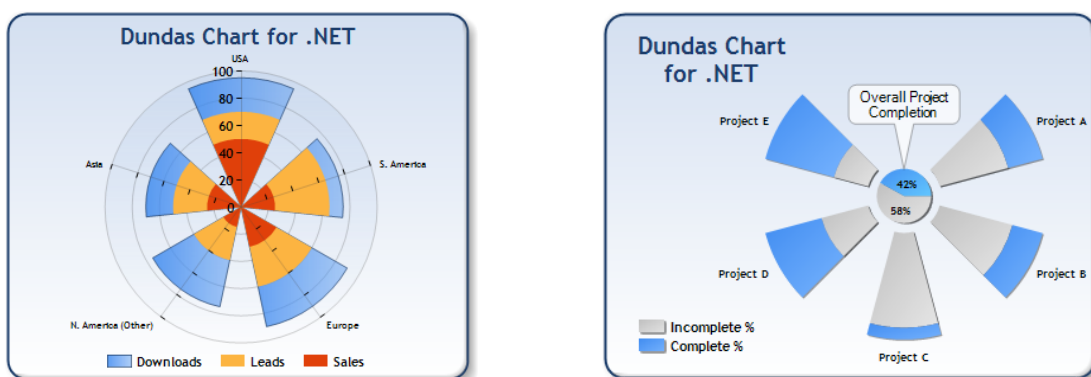
- Až 2 dimenze  $D_1, D_2$  a 1 údaj  $M_1$  – graf se skládá z  $|D_1^\alpha|$  os. Jestliže úkolem grafu je zobrazit hodnoty vztahující se pouze k jedné dimenzi  $D_1$ , pak jsou hodnoty vynášeny na tyto osy dle vztahu:  $hodnota_i = M_1([D_1^\alpha]_i)$ , kde  $i$  udává index prvku z  $D_1^\alpha$ .

V případě, že je potřeba zobrazit hodnoty závislé na prvcích obou dimenzí, jsou prvkům z  $D_1^\alpha$  opět přiřazeny jednotlivé osy. Hodnota sledované veličiny je dle očekávání  $hodnota_{i,j} = M_1\left([D_1^\alpha]_i, [D_2^\beta]_j\right)$ , kde  $i$ , resp.  $j$  je index prvku z  $D_1^\alpha$  (a také tedy index osy), resp.  $D_2^\beta$ . Nyní existují dvě možnosti reprezentace  $D_2^\beta$  – graf může obsahovat na každé ose  $n = |D_2^\beta|$  bodů, přičemž každý bod přináleží jinému prvku  $D_2^\beta$ . Příklad tohoto typu, v němž jsou body se stejným indexem  $j$  propojeny a vzniklé oblasti barevně odlišeny, je zobrazen na obrázku 19. Druhou možností je přiřadit každému prvku z  $D_2^\beta$  sloupec, jenž bude od ostatních vizuálně odlišen a jehož výška bude korespondovat s příslušnou hodnotou. Příklad je ukázán na obrázku 20.

- 2 dimenze  $D_1, D_2$  a 1 údaj  $M_1$  se součtem – jde o kombinaci výše uvedených způsobů zobrazení dvou dimenzí s jedním faktem v paprskovém grafu a sloupcového grafu zobrazujícího dvě dimenze a jeden fakt se součtem. Na obrázku 21 je uveden příklad grafu, ve kterém



Obrázek 21: Paprskový graf reprezentující dvě dimenze a jeden součtový fakt pomocí oblastí.



Obrázek 22: Paprskové grafy realizující dvě dimenze a jeden součtový fakt pomocí sloupců.

jsou místo jednotlivých částí sloupců použity body umístěné na osách. Další obrázek 22 již ukazuje příklad se sloupci. Lze také vidět, jakým způsobem je možné do paprskového grafu přidat další pohled na zobrazenou informaci – do středu grafu je možné umístit například koláčový graf.

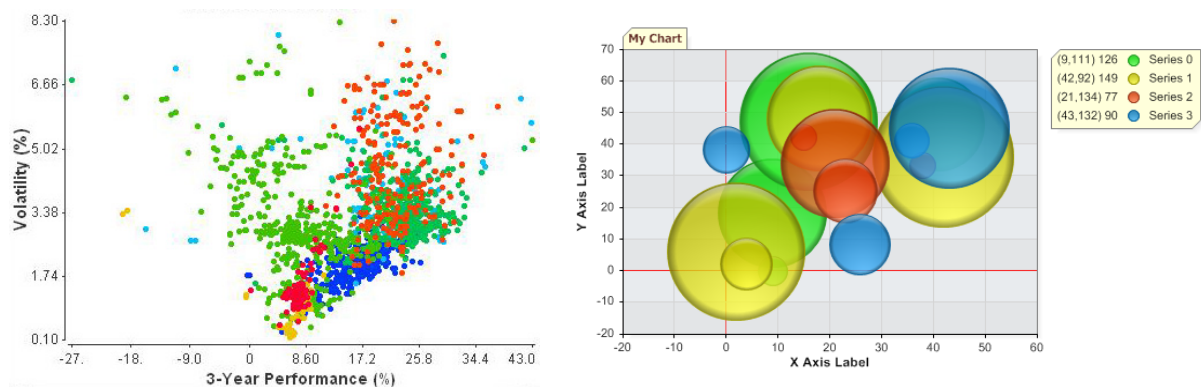
## 4.6 X-Y graf

X-Y graf, v některých případech označován také jako bodový graf nebo rozptylogram (angl. *scatterplot*) [11], je vhodný pro zkoumání možných vztahů mezi více proměnnými, které jsou spolu nějakým způsobem svázaný.

V X-Y grafech je možné zkoumat:

- $q$  dimenzí  $D_1, \dots, D_q$  a minimálně 2 fakty  $M_1, M_2$  – údaje v grafu jsou reprezentovány zcela jiným způsobem, než tomu bylo u předchozích typů. V naprosté většině výše uvedených grafů byly alespoň na jedné ose umístěny prvky dimenze. Zde obě osy představují zvolené fakty –  $M_1$  je umístěna na ose x,  $M_2$  na ose y. Body grafu jsou vykresleny na pozici  $[x_{i_1, i_2, \dots, i_q}, y_{i_1, i_2, \dots, i_q}] = \left[ M_1 \left( [D_1^{f_1}]_{i_1}, \dots, [D_q^{f_q}]_{i_q} \right), M_2 \left( [D_1^{f_1}]_{i_1}, \dots, [D_q^{f_q}]_{i_q} \right) \right]$  pro





Obrázek 23: X-Y graf s libovolným nenulovým počtem dimenzí dvěma fakty zobrazenými na osách (a) a případným třetím faktem jako velikostí kruhu (b). Body jsou obarveny dle prvků vybrané dimenze.

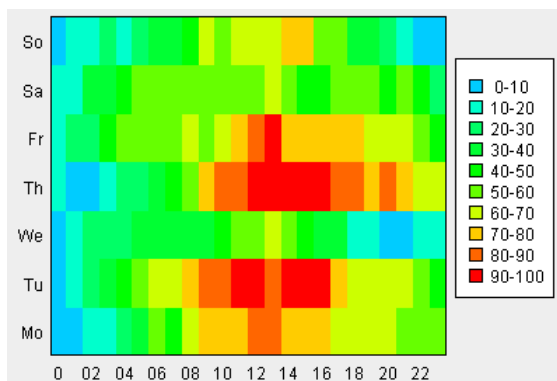
všechna  $i_k = 1, \dots, |D_k^{f_k}|$ , kde  $k = 1, \dots, q$ . Celkový počet bodů v grafu je tedy  $\prod_{k=1}^q |D_k^{f_k}|$ .

Jednotlivé body mohou být obarveny – obvykle je zvolena jedna dimenze, například vyberme  $D_1^{f_1}$ , jejímž prvkům jsou přiřazeny vzájemně různé barvy. Přísluší-li prvku  $[D_1^{f_1}]_k$  barva  $c_k$ , pak všechny body na pozicích  $[x_{k,i_2,\dots,i_q}, y_{k,i_2,\dots,i_q}]$  budou obarveny právě barvou  $c_k$ . Viz obrázek 23a.

Do grafu lze přidat ještě další sledované fakty.  $M_3$  může být reprezentován velikostí bodu (pak je označován také jako bublinkový graf), viz příklad na obrázku 23b,  $M_4$  pak třeba tvarem bodu (čtverec, kruh, trojúhelník, atd.). Pokud je graf promítán v trojrozměrném prostoru, lze hodnotu dalšího fakt reprezentovat výškou bodu. Důležitou roli v X-Y grafech s těmito volitelnými fakty hraje míra průhlednosti bodu a hloubka, ve které se bod nachází, protože při větším počtu bodů se podstatné informace mohou ztrácet.

- 2 dimenze  $D_1, D_2$  a alespoň 1 veličina  $M_1$  – tento typ sice není typický zástupce bodového grafu<sup>1</sup>, ale přesto jej sem zařadíme kvůli své vizuální povaze. Graf má prvky množin  $D_1^\alpha$ , resp.  $D_2^\beta$  umístěné na ose x, resp. y. Celkový rozsah hodnot dané veličiny je namapován na barevnou škálu, takže skutečné hodnotě  $hodnota_{x,y} = M_1(x,y)$  je přiřazena jistá barva z této škály.  $x$ , resp.  $y$  jsou proměnné nabývající hodnot z  $D_1^\alpha$ , resp.  $D_2^\beta$ . Příklad takto utvořeného grafu lze spatřit na obrázku 24. Graf může být obohacen o další informaci – místo barevných ploch by se v jednotlivých průsečících mohly vyskytovat obarvené tvary, či značky, které by odpovídaly hodnotám volitelné veličiny  $M_2$ .

<sup>1</sup>Lze jej totiž chápat také jako sloupcový graf v 3D pohledu s prvky z  $D_1^\alpha$  na ose x, prvky množiny  $D_2^\beta$  jsou na ose y a hodnoty sloupce rostou v ose z, přičemž výšce sloupce odpovídá jeho barva. Na tento graf se pak uživatel dívá kolmo nad rovinou x-y, tj. paralelně k ose z.



Obrázek 24: X-Y graf s dvěma dimenzemi a jednou sledovanou veličinou, jejíž hodnota je zakódována do barvy.

## 5 Další možnosti reprezentace dat

### 5.1 Hierarchická síťová mapa

Zajímavý způsob, jak zobrazit údaj o velikosti a další údaje prvků vybrané dimenze představuje hierarchická síťová mapa [7]. Každému prvku zvolené části dimenze odpovídá jistá část celku v podobě obdélníku. Velikost této dílčí plochy koresponduje s hodnotou faktu vztahující se k velikosti. Dále je možné dané ploše přiřadit grafické atributy, jenž budou reprezentovat hodnoty ostatních sledovaných údajů.

Lze si také všimnout, že v tomto typu grafu je možné při dělení celku na části uplatnit jistá pravidla. Například celek může představovat celý svět, ten je horizontálně rozdělen na části reprezentující jednotlivé kontinenty, ty mohou být dále děleny na státy. Pravidly lze tedy vyjádřit například hierarchii dimenze, což může být velice výhodné.

Příklad na obrázku 25 ukazuje jednotlivé státy a počet provozovaných IP adres v těchto zemích – čím více jich je v dané zemi využíváno, tím větší mu přísluší plocha obdélníku. Barva výplně pak odpovídá aktuální využívané přenosové kapacitě.

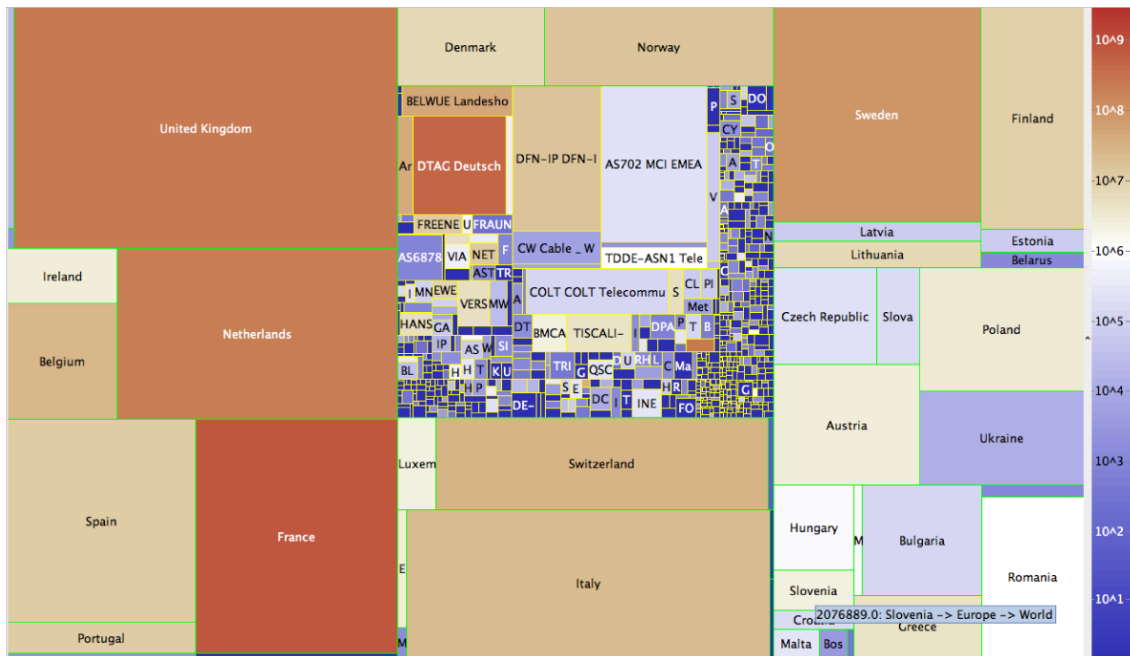
### 5.2 Histogram

Histogram rozděluje množinu hodnot do tříd a zobrazuje frekvenci, s jakou se každá třída vyskytuje v této množině. Ve spojitém grafu je obvykle na ose x umístěn rozsah hodnot, jenž sledovaná veličina nabývá (nebo jednotlivé třídy v případě diskrétního grafu), a na osu y je vynášen počet hodnot, jenž do daného intervalu spadá, tj. frekvence. Příklad histogramu je uveden na obrázku 26.

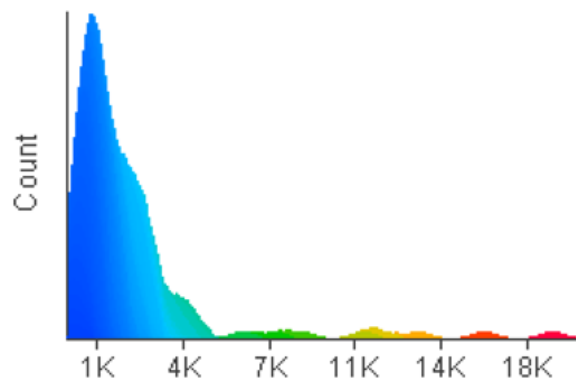
Histogram se využívá v situacích, kdy uživatele nezajímají konkrétní hodnoty, ale chce si udělat přehled o distribuci hodnot v celé množině. Díky něž lze tak snadno identifikovat rozsah odlehlých hodnot a jejich přibližný počet.

### 5.3 Parabox

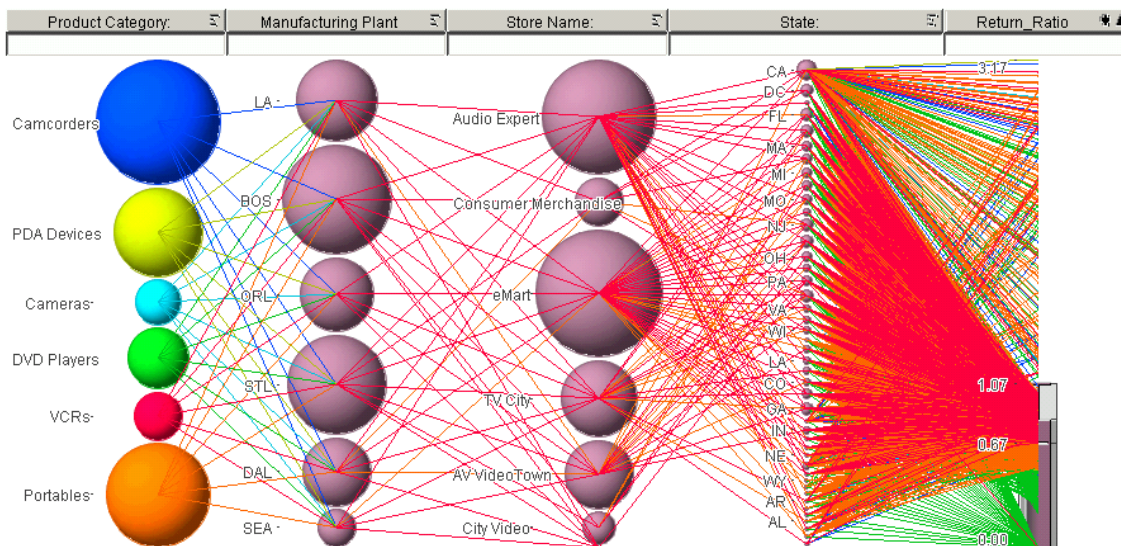
Parabox představuje spojení dvou typů reprezentace dat – krabicového grafu (*box plot*), jenž popisuje množinu hodnot pomocí pěti sumarizačních údajů [12], a grafu paralelních souřadnic (*parallel coordinate plot*) zobrazujícího vztahy mezi hodnotami více proměnných [5].



Obrázek 25: Hierarchická síťová mapa, v níž velikost plochy obdélníku a barva výplně odpovídá dvěma údajům prvků zvolené dimenze.



Obrázek 26: Příklad histogramu.



Obrázek 27: Parabox, v němž první až předposlední sloupec náleží různým dimenzím a poslední sloupec reprezentuje veličinu.

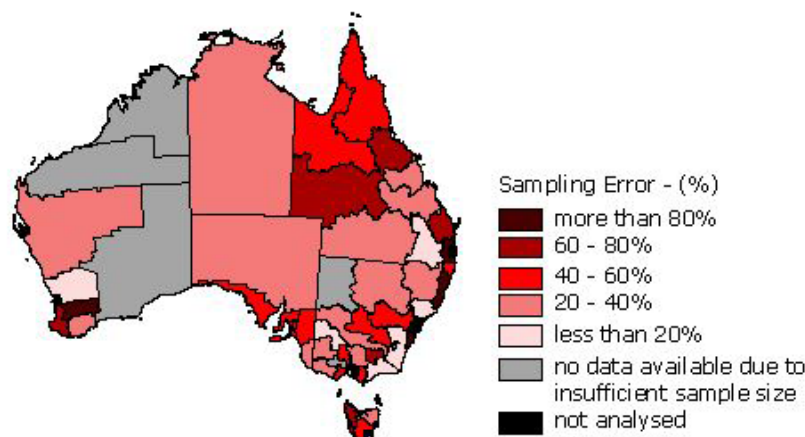
Tento typ grafu vychází z uživatelem daného seznamu dimenzí a jednoho vybraného faktu, přičemž každému z nich, tj. i dimenzím, i faktu, přísluší v grafu právě jeden sloupec. V prvním sloupci, který je chápán jako řídicí, jsou uživatelem zvolené prvky dimenze zobrazeny pomocí různě barevných kruhů s průměrem korespondujícím s počtem záznamů, kterých se týkají. Poslední sloupec je tvořen požadovaným faktem, jehož množina hodnot je zde reprezentována stejným způsobem jako v krabicovém grafu. Zbývající sloupce, tj. druhý až předposlední, obsahují ty prvky dimenzí, na jejichž průsečíku spolu s libovolným prvkem první dimenze je hodnota požadovaného faktu definována. Spojnicemi jsou v grafu reprezentovány jednotlivé průsečíky multidimenzionální kostky, vč. hodnoty průsečíku, která je zobrazena bodem v posledním sloupci. Příklad je ukázan na obrázku 27.

Cílem paraboxů je nabídnout uživateli přehled o rozložení hodnot sledované veličiny v závislosti na zvolených prvcích více dimenzí. Navíc umožňuje pokročilejší analýzu dat a odhalit tak složitější vztahy mezi dimenzemi a fakty.

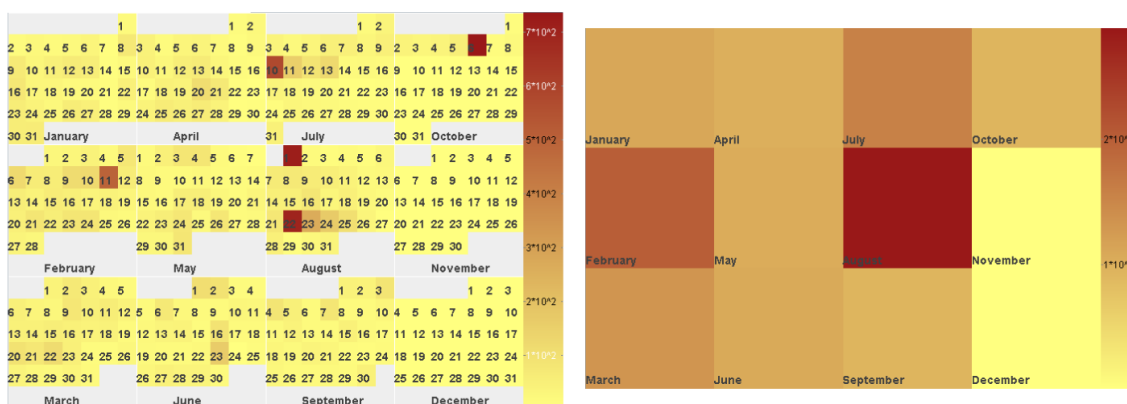
## 5.4 Zobrazení geografické dimenze

Ve většině podnikových multidimenzionálních krychlic se nachází minimálně tři typy dimenzí – produkty, místo a čas. Zde se zaměříme na prezentaci údajů závislých na dimenzi uchováající místo (například prodeje produktů), která je také známá pod názvem geografická dimenze.

Zřejmě nejpřirozenější cestou, jak prohlížet data závislá na poloze, je zobrazit uživateli mapu oblasti, do níž spadají všechny požadované prvky geografické dimenze (tj. pozice). Způsob zobrazení příslušných údajů pak závisí na typu geografických dat a úrovni mapy – kontinenty, státy a jiné správní celky se zobrazují na mapě jako plochy, města, poštovní směrovací čísla, nebo konkrétní adresy jsou obvykle zobrazeny jako body a pomocí čar mohou být označeny například řeky, ale také i trasy dopravních spojů, telekomunikační spoje a mnoho dalších. Těmto geometrickým reprezentacím se pak přiřazuje barva, tloušťka, výplň a další grafické atributy, které kódují skutečnou hodnotu sledovaného údaje. Příklad rozdělení Austrálie na jednotlivé části, které jsou vyplněny barvou dle hodnoty jistého faktu, je ukázan na obrázku 28. Poznamenejme, že v ma-



Obrázek 28: Reprezentace údajů závislých na poloze je nejpřirozenější na mapách.



Obrázek 29: Údaje závislé na čase lze přehledně zobrazit v kalendáři. Zdroj [8].

pách lze s výhodou využít i agregovaných údajů – například barva ploch jednotlivých států bude označovat agregovaný údaj, zatímco styl bodů reprezentující města bude představovat konkrétní údaje.

## 5.5 Zobrazení časové dimenze

Jak již bylo řečeno dříve, vývoj veličin v čase lze sledovat například pomocí sloupcových, spojnicových, ale i ostatních typů grafů. Další možností je pak zobrazit si požadovaný časový úsek s požadovanou granularitou (všechny dny v určitém roce, hodiny v týdnu,…) v kalendáři, viz obrázek 29. Do jednotlivých polí mohou být následně přidány hodnoty sledovaných faktů. Ty mohou být v číselné formě, lze je vyjádřit barvou pole, symbolem, nebo například jednodušším podgrafem.

## 5.6 Table Lens

Table Lens představuje techniku vertikální komprese údajů v tabulce, díky čehož lze na stejné velké ploše zobrazit podstatně větší množství dat. V přímé číselné, resp. textové podobě jsou zde zobrazeny pouze ty hodnoty, které leží ve vybraných řádcích, tj. uživatel projevil zájem o kon-

	Year	Quarter	Product	Channel	Region	Salesperson	Units	Revenue	Profits
126	1993	2	ForeCode Pro	Direct Sales	Southwest	Kevin Polen	1029	439898	171561
444	1993	4	ForeCode Pro	VAR	West	Tom Tuttle	302	122310	51371

Obrázek 30: Zobrazení velkého množství dat pomocí *Table Lens*. Obrázek převzat z [9].

krétního hodnoty. Zbývajícím řádkům je na výšku přiřazen obvykle jeden obrazový pixel, fakty jsou prezentovány jako grafické sloupce, jejichž délka koresponduje s jejich hodnotou. Jednotlivým prvkům dimenzí jsou přiřazeny různě barevné (ale stejně dlouhé) úsečky, přičemž i pozice jejich začátku na řádku udává příslušnost k prvku. Příklad je zobrazen na obrázku 30.

## 6 Úprava počtu dimenzí a faktů

V předchozích kapitolách byly u jednotlivých typů zobrazení uvedeny počty dimenzí a faktů, které v nich lze reprezentovat. Tyto údaje však nejsou pevné a neměnné. V některých případech je vhodné zvýšit počet dimenzí, či faktů ve vybraném typu zobrazení.

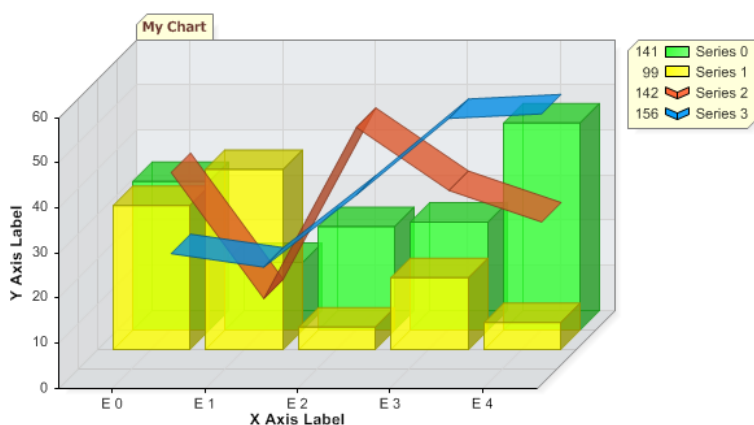
Zvýšení počtu sledovaných faktů u jednotlivých grafů lze buď přímo (pokud to graf umožňuje), například přidáním křivky, skupiny sloupců a nebo například upravením grafické reprezentace výplně (k barvě přidat například šrafování). Druhou možností je pak umístit dva stejné pohledy na data vedle sebe, přitom každý bude ukazovat jinou veličinu.

Zvýšení počtu dimenzí se provádí pomocí znásobení počtu zobrazení téhož grafu. Příklad řešení tohoto případu je vidět na obrázku 16 – jeden graf přísluší pouze prvkům jedné dimenze. Dodáním prvků jiné dimenze však dojde k přidání nezávislé proměnné, kterou je potřeba v grafech určitým způsobem zobrazit. Za každým grafem, nebo skupinou grafů, se pak skrývá konkrétní prvek nově přidané dimenze, původní dimenze jsou již uchovány v dílčích reprezentacích.

Ukazatele, grafy i další způsoby zobrazení umožňují vzájemné skládání a doplňování. Každé výšce koláčového grafu je možné přiřadit externí sloupcový graf, v mapách se někdy zobrazují místo bodů reprezentujících města například jednodušší koláčové grafy, do sloupcového grafu může být přidána skupina křivek ukazujících jiné fakty. Možností kombinací jsou opravdu široké, avšak nesmí přitom dojít ke ztrátě přehlednosti a přesycení uživatele informacemi. Díky této vlastnosti je snadné vytvořit zajímavý graf s požadovaným počtem dimenzí i faktů, viz příklad na obrázku 31. Uvedený příklad může zobrazovat od jedné dimenze tvořené prvky na ose x a čtyřmi fakty, až po tři dimenze (prvky první dimenze jsou opět na ose x, z druhé dimenze jsou vybrány dva prvky, kterým přísluší rozdělení na křivky/sloupce, a třetí dimenze je opět reprezentována dvěma prvky, kterým náleží zvolené dvojice barev) a jeden fakt. Vhodným popisem os a uvedením korektní legendy by mělo ihned vyplynout, jaké informace graf zobrazuje.

## 7 Použití v reportech

Všechny výše uvedené typy zobrazení mohou být součástí přehledových zpráv, neboli reportů. Jejich úkolem je poskytnout uživateli souhrnné informace o dané oblasti zájmu. Report je tedy složen ze skupiny grafů, které často prezentují stejné typy údajů pomocí různých grafů, čímž ukazují problematiku z různých úhlů. Navíc mohou být doplněny o další grafy doplňující celkovou



Obrázek 31: Vhodnou kombinací grafů lze snadno zobrazit potřebný počet dimenzí i faktů.

informaci, o tabulky s vybranými přesnými údaji, atd. U reportů se často klade také důraz na kvalitní vizuální zpracování.

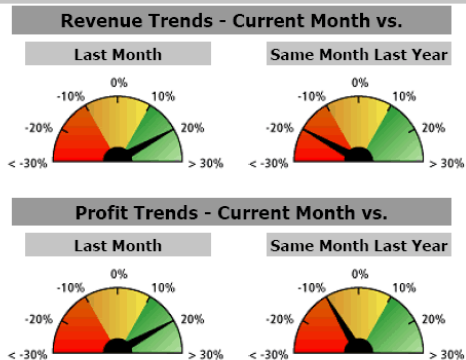
Obecně existují dva typy reportů. Za statické reporty lze považovat například vytisknuté reporty na papíru, nebo ty, jenž jsou zobrazeny na monitoru, avšak umožňují jen zvětšování/zmenšování grafů, nebo změnu barev. Druhou skupinu tvoří dynamické reporty, které uživateli nabízí bohaté funkce – samozřejmě je změna detailu grafů a dalších vizuálních vlastností, report nabízí funkci filtru, zobrazení detailních hodnot, umožňuje pohyb v hierarchii zvolené dimenze a další. Příklad reportu je na obrázku 32.

## 8 Závěr

Různé typy vizualizací dat umožňují, aby uživateli nebyla prezentována data pouze v prosté tabulkové podobě, avšak byla převáděna do intuitivních grafických reprezentací, čímž se značně zvyšuje schopnost uživatele předkládané informace hodnotit, srovnávat a analyzovat. Je však potřeba, aby pro konkrétní množinu dat byla vybrána správná technika zobrazení. Nevhodným výběrem by totiž mohlo dojít ke zmatení uživatele, v horším případě k chybné interpretaci. Tento dokument uvedl jistě ne zcela vyčerpávající seznam možných zobrazení se zaměřením na data, která jsou ukládána v multidimenzionální podobě. Jak lze ale vidět, možnosti reprezentace jsou skutečně bohaté.

## Monthly Operational Metrics December 2003

### CORPORATION OPERATIONAL METRICS

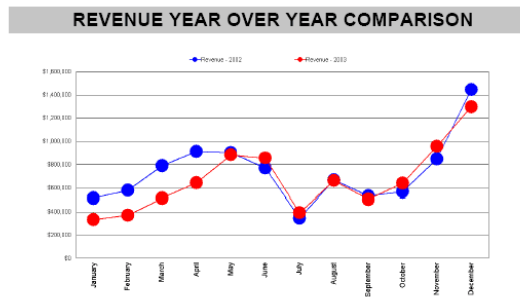
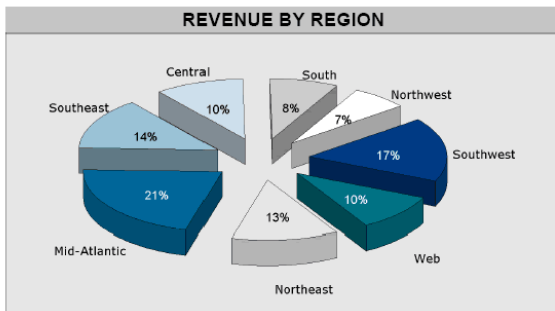


Metrics	This Month	Status	Trend
Revenue	\$1,296,667	▲	▲
Profit	\$312,376	▲	▲
Margin	24.09%	▶	▶
Units Sold	32,122	▲	▲
Order Count	21,420	▲	▲
Avg Revenue per Order	\$61	▲	▼
Customer Count	10,000	▼	▲
Avg Revenue per Customer	\$130	▼	▲

Abbreviations: LM = Last Month, LY = Last Year, TM = This Month

Notes on Metric Definitions:

- Status indicators are based on corporate monthly targets for each metric.
- Trend indicators are based on the % increase or decrease from last month (% Diff from LM).
  - ▼ indicates a percentage change of less than or equal to -5%.
  - ▶ indicates a percentage change between -5% and 5%.
  - ▲ indicates a percentage change greater than or equal to 5%.



Obrázek 32: Příklad obchodního reportu.



## Reference

- [1] CORPORATE WEB SOLUTIONS LTD. *ASP.NET Chart and Charting for .NET* [online]. [cit. 8. 7. 2008]. Dostupné z: <<http://www.dotnetcharting.com/>>.
- [2] DUNDAS DATA VISUALIZATION, INC. *Main Elements of Dundas Gauge* [online]. [cit. 9. 7. 2008]. Dostupné z: <<http://support.dundas.com/OnlineDocumentation/RSGauge/MainElementsofDundasGauge.html>>.
- [3] DUNDAS DATA VISUALIZATION, INC. *Advanced Data Visualization for Microsoft Technologies* [online]. [cit. 8. 7. 2008]. Dostupné z: <<http://www.dundas.com>>.
- [4] FEW, S. *Bullet Graph Design Specification* [online]. [cit. 9. 7. 2008]. Dostupné z: <[http://www.perceptualedge.com/articles/misc/Bullet\\_Graph\\_Design\\_Spec.pdf](http://www.perceptualedge.com/articles/misc/Bullet_Graph_Design_Spec.pdf)>.
- [5] FEW, S. *Multivariate Analysis Using Parallel Coordinates* [online]. [cit. 12. 7. 2008]. Dostupné z: <<http://www.b-eye-network.com/view/3355>>.
- [6] KYD, C. *An Excel Tutorial: How to Create Bullet Graphs To Replace Gauges in Excel* [online]. [cit. 9. 7. 2008]. Dostupné z: <<http://www.exceluser.com/explore/bullet.htm>>.
- [7] MANSMANN, F. *Interactive Exploration of Data Traffic with Hierarchical Network Maps and Radial Traffic Analyzer* [online]. [cit. 11. 7. 2008]. Dostupné z: <<http://infovis.uni-konstanz.de/index.php?region=research&reg2=hnmap>>.
- [8] MANSMANN, S. – SCHOLL, M. H. Exploring OLAP Aggregates with Hierarchical Visualization Techniques. In *SAC '07: Proceedings of the 2007 ACM symposium on Applied computing*, s. 1067–1073, New York, NY, USA, 2007. ACM. doi: <http://doi.acm.org/10.1145/1244002.1244235>. Dostupné z: <<http://www.inf.uni-konstanz.de/cgip/bib/files/MaSc07b.pdf>>. ISBN 1-59593-480-4.
- [9] RAO, R. *Understanding Information Collections with Maps and Visualizations* [online]. [cit. 11. 7. 2008]. Dostupné z: <<http://www.ramanarao.com/articles/2001-12-online-info/cviz.html>>.
- [10] REITER, B. *Warning against using piecharts* [online]. [cit. 2. 7. 2008]. Dostupné z: <<http://www.usf.uni-osnabrueck.de/~breiter/tools/piechart/warning.en.html>>.
- [11] TVRDÝ, L. *Metodika pro tvorbu grafů* [online]. [cit. 9. 7. 2008]. Dostupné z: <[http://www.spss.cz/files/ruzne/vsb/report\\_metodika\\_grafy.pdf](http://www.spss.cz/files/ruzne/vsb/report_metodika_grafy.pdf)>.
- [12] WIKIPEDIA. *Box plot* [online]. [cit. 12. 7. 2008]. Dostupné z: <[http://en.wikipedia.org/wiki/Box\\_plot](http://en.wikipedia.org/wiki/Box_plot)>.