

Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Katedra informačních technologií

Student : Ondřej Černý
Vedoucí bakalářské práce : Ing. Helena Benáčanová, CSc.
Oponent bakalářské práce : Ing. Miroslav Lorenc

TÉMA BAKALÁŘSKÉ PRÁCE

**VYUŽITÍ MS EXCEL PRO ANALÝZU DAT Z OLAP
DATABÁZÍ**

ROK : 2010

Prohlášení

Prohlašuji, že jsem bakalářskou práci zpracoval samostatně a že jsem uvedl všechny použité prameny a literaturu, ze kterých jsem čerpal.

V Praze dne 29.6.2010

.....

podpis

Poděkování

Na tomto místě bych rád poděkoval paní Ing. Heleně Benáčanové, CSc. za cenné připomínky a odborné vedení.

Abstrakt

Tato bakalářská práce popisuje teoretické základy funkcí OLAP (OnLine Analytical Processing) databází a především se věnuje praktickému využití aplikace Microsoft Office Excel 2007 pro analýzu dat uložených v OLAP databázích. Cílem práce je popsat, jaké možnosti Excel pro práci s OLAP databázemi nabízí a jak při práci s těmito databázemi postupovat. Práce s OLAP databázemi je demonstrována praktickými ukázkami na vzorové databázi.

Tato práce může sloužit jako návod pro uživatele, kteří s problematikou OLAP databází ještě nejsou seznámeni a chtěli by Excel pro účel analýzy dat z OLAP databází použít.

Abstract

This bachelor thesis describes theoretical basis of OLAP (OnLine Analytical Processing) databases and focuses primarily on a practical use of Microsoft Office Excel 2007 for analysis of the data stored in OLAP databases. The aim of thesis is to describe, what possibilities Excel provides for the work with OLAP databases, and what is the working procedure of such work. Work of Excel with OLAP databases is demonstrated on sample database.

This thesis can be used as an user's guide for beginners with no experience in analysis of OLAP databases in Excel.

Obsah

1.	Úvod	6
2.	OLAP Databáze	8
2.1.	Fakta a dimenze	8
2.2.	MOLAP – binární OLAP databáze	8
2.3.	Relační databázový OLAP	10
2.4.	Hybridní OLAP (HOLAP)	12
3.	Práce s OLAP databází v MS Excel	13
3.1.	Použitá databáze	13
3.2.	Připojení k analytickému serveru	13
3.3.	Zpřístupnění OLAP krychle v režimu offline	17
3.4.	Práce s kontingenční tabulkou	20
3.4.1.	Zobrazení dat v kontingenční tabulce	22
3.4.2.	Filtrování dat v kontingenční tabulce	24
3.4.3.	Změna popisků v kontingenční tabulce	25
3.4.4.	Úprava vzhledu kontingenční tabulky	27
3.4.5.	Podmíněné formátování	30
3.5.	Práce s kontingenčním grafem	34
3.5.1.	Vytvoření kontingenčního grafu	34
3.5.2.	Zobrazení dat v kontingenčním grafu	35
3.5.3.	Formátování kontingenčního grafu	38
3.5.4.	Kopírování grafů a jejich použití v ostatních aplikacích	39
3.6.	Skupina funkcí CUBE pro práci s OLAP kostkou	40
3.7.	Omezení Excelu při práci s OLAP kostkou	46
4.	Závěr	47
5.	Použité zdroje	49
6.	Terminologický slovník	50

1. Úvod

Programových produktů, které slouží k analýze dat pomocí OLAP technologie, je velká řada. Díky funkcím datových krychlí a díky univerzálním možnostem aplikace můžeme pro analýzu dat z databáze OLAP použít i aplikaci Microsoft Office Excel 2007. Je to sice jeden z jednodušších nástrojů, ale je cenově dostupný, a mezi uživateli velmi rozšířený. Toto téma jsem si zvolil právě z těchto důvodů – protože je to produkt cenově dostupný a značně rozšířený, určitě se s ním ve svém budoucím zaměstnání setkám a budu moci znalosti, které jsem nabyl při psaní své bakalářské práce, uplatnit v praxi.

Cílem této práce je zjistit, jaké možnosti MS Excel 2007 bez žádných dalších rozšíření nabízí jako klient pro přístup k OLAP databázím a podrobně popsat, jakým způsobem je s nimi možné pracovat. Práce bude obsahovat popis funkcí, které Excel nabízí pro analýzu dat a jejich zobrazení a demonstrovat jejich použití na praktických ukázkách. Tato práce by měla sloužit uživatelům, kteří chtějí provádět analýzu dat pomocí aplikace MS Excel jako základní příručka, která je seznámí s funkcemi, které je v Excelu možné použít při práci s OLAP kostkou.

Při zpracování problému, který má řešit tato práce, budu postupně rozkrývat jednotlivé okruhy práce s OLAP databázemi, které je možné rozdělit do čtyř významnějších skupin: připojení k OLAP databázi, práce s kontingenčními tabulkami, práce s kontingenčními grafy a využití funkcí pro přístup k datům z OLAP databáze. Slovní komentáře budou doplněny názornými obrázky, a to jak ovládacích prvků MS Excel, použitých při práci s OLAP databázemi, tak samotných tabulek a grafů, které budou vytvořeny na datech ukázkové databáze. To umožní lepší pochopení postupu práce s OLAP databázemi.

Předpokládám, že po zpracování tématu budu moci zhodnotit výhody i nevýhody použití MS Excel 2007 pro analýzu dat z OLAP databáze – odhalené nedostatky či nevýhody budou zmíněny přímo v jednotlivých okruzích popisujících zpracování OLAP databází, na závěr pak budou ještě shrnuty v závěrečné kapitole praktické části.

Práce by neměla narazit na žádná omezení – aplikace Microsoft Office Excel 2007 je k dispozici, rovněž tak i databáze pro praktické ukázky je volně dostupná na internetu.

Struktura práce tedy bude následující:

- První část práce bude zaměřena na technologii OLAP obecně. V této části bych rád popsal, co to OLAP databáze vlastně jsou a jak fungují.
- V druhé části se pak chci zaměřit na praktické použití Excelu v kombinaci s OLAP kostkou a popsat práci s Excelem už od postupu připojení k databázi, přes provádění samotné analýzy dat, až po způsoby, jakými lze analýzy prezentovat. K praktickým ukázkám využiji vzorovou databázi FoodMart, která je volně dostupná na internetu.

2. OLAP Databáze

OLAP (OnLine Analytical Processing) je technologie uložení dat v databázi, která umožňuje uspořádat velké objemy dat tak, aby byla data přístupná a srozumitelná uživatelům zabývajícím se analýzou obchodních trendů a výsledků. Typické je využití systémů OLAP pro analýzu velkého množství údajů. Výsledkem analýzy jsou souhrny a reporty, které slouží manažerům jako podklady při jejich rozhodování, ať už v oblasti řízení firmy, řízení ekonomických a technologických procesů a podobně.[4]

OLAP technologie je založena na koncepci multidimenzionálních databází. Jejím hlavním principem je několikadimenzionální tabulka umožňující rychle a pružně měnit jednotlivé dimenze, a měnit tak pohledy uživatele na modelovanou ekonomickou realitu.[3]

OLAP databáze můžeme dělit do několika skupin podle technologií, které používají pro uložení dat. Nejčastěji používané jsou technologie MOLAP, ROLAP a HOLAP:

- Pro **MOLAP** (Multidimensional OLAP) je charakteristické uložení dat v multidimenzionálních – binárních OLAP kostkách. MOLAP databáze jsou vlastně vícerozměrná pole, kde každá dimenze (resp. každá úroveň dimenze) představuje jeden rozměr.
- **ROLAP** (Relational OLAP) řeší multidimenzionalitu uložením dat v relační databázi.
- **HOLAP** (Hybrid OLAP) je kombinací předchozích přístupů, kdy detailní data jsou uložena v relační databázi a agregovaná data jsou uložena v binárních OLAP kostkách.

2.1. Fakta a dimenze

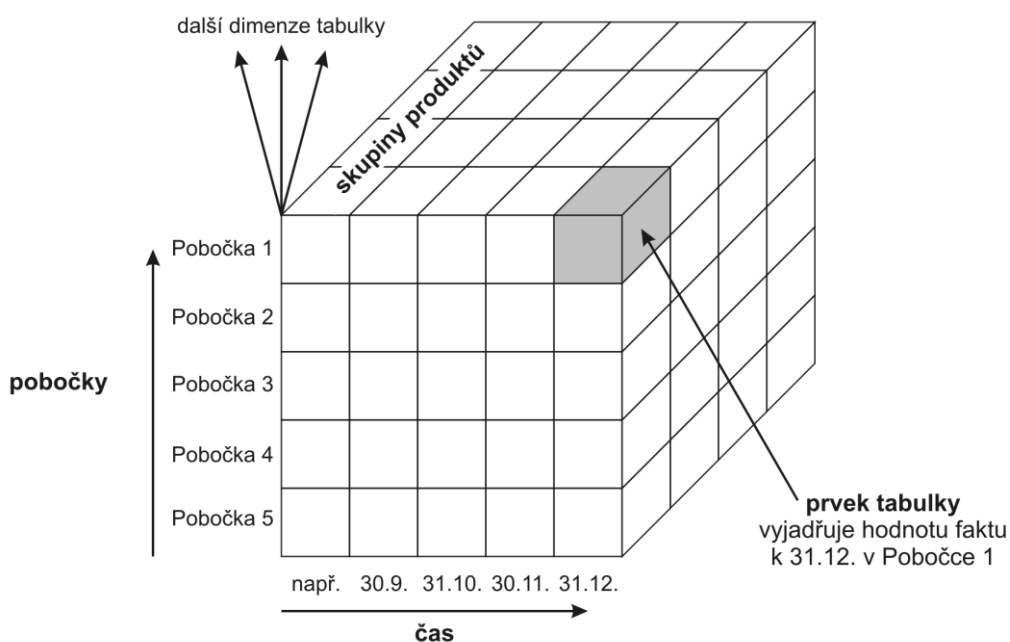
OLAP datová struktura se skládá z numerických faktických údajů odrážejících hodnotu sledovaných jevů. Dimenze reprezentují soubor tříd, podle nichž jsou fakta kategorizována.

2.2. MOLAP – binární OLAP databáze

Aby analytické systémy mohly poskytovat požadované analytické funkce a podporu rozhodování, je nutné, abychom se na jejich data mohli dívat z více hledisek současně. Mělo by tedy být možné vytvářet tzv. multidimenzionální pohledy. Navíc je nutné

procházet velké množství dat, vypočítávat agregace, rychle měnit pohledy na data a ukládat je do přehledných tabulek a grafů.

Multidimenzionální databáze jsou proto optimalizované pro uložení a interaktivní využívání multidimenzionálních dat. Výhodou multidimenzionality, resp. nasazení OLAP technologií, je rychlost zpracování a efektivní analýzy multidimenzionálních dat (tzv. drilling, slice and dice apod.). Základním principem je tak několikadimenzionální tabulka umožňující velmi rychle a pružně měnit jednotlivé dimenze, jak dokumentuje obrázek 2-1.



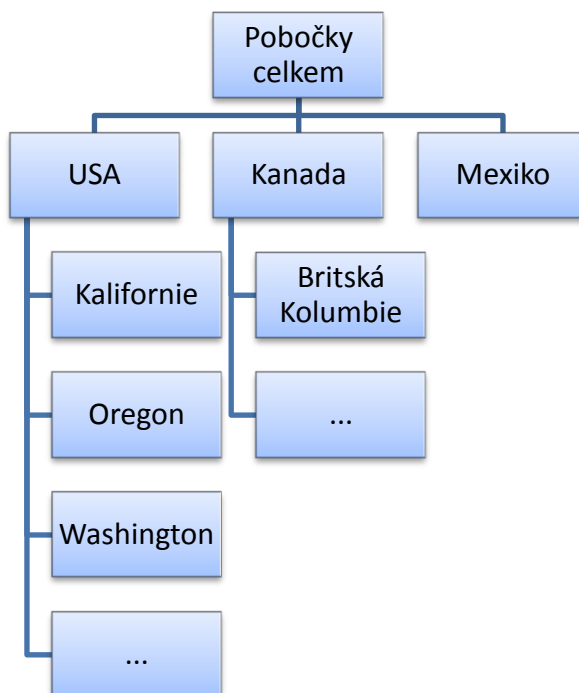
Obrázek 2-1 Princip multidimenzionální databáze¹

Každý prvek (buňka) datové kostky je definován konkrétními hodnotami jednotlivých dimenzí (třídami) a obsahuje údaj faktické hodnoty, například ekonomické proměnné, nebo předpisy (algoritmy) pro jejich transformace. Základní dimenzí datové kostky je zpravidla čas, ostatní dimenze se pro jednotlivé kostky definují podle potřeby, např. teritorium, organizační jednotka, zboží, zákazník, dodavatel, konkurent apod.

Prvky dimenzí jsou většinou uspořádány v hierarchické struktuře, tzn. že se rozdělují na skupiny prvků, podskupiny, až na jednotlivé prvky. Produkty Business Intelligence pak zajišťují automatické agregace hodnot (ekonomických proměnných), a to podle

¹ Jedná se o upravený obrázek ze zdroje [2].

definovaných hierarchických úrovní dimenzí. Jednoduchým příkladem takové struktury a odpovídajících agregací může být dimenze „pobočky“ – viz obrázek:



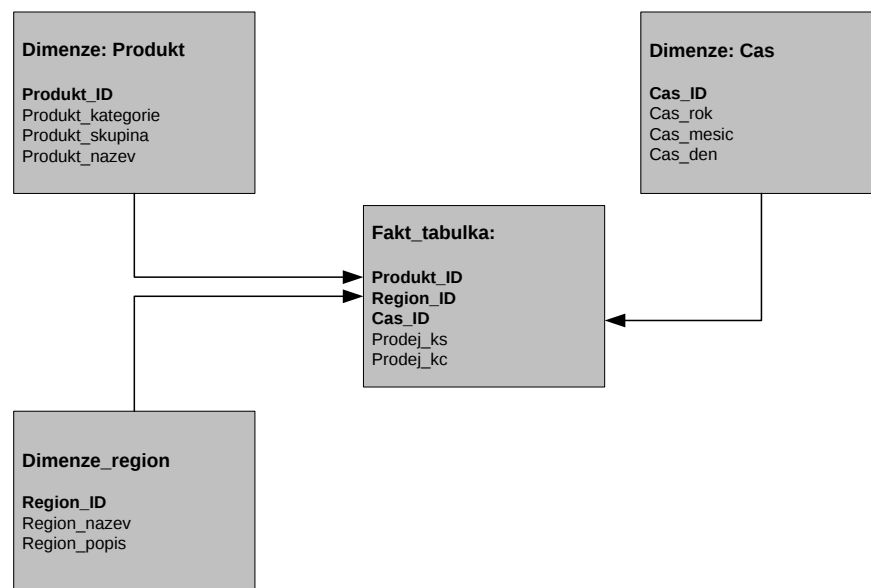
Obrázek 2-2 Struktura dimenze „pobočky“

Aplikace Business Intelligence budou automaticky sledovat např. hodnoty výroby, prodeje zboží, zásob, zisku apod., podle uvedené struktury kontinentů a zemí a zajišťovat agregace podle takto definované hierarchické struktury. Tato metoda je využívána pro urychlení odezvy systému na analytické požadavky. Pokud by totiž bylo nutné on-line vypočítávat součty při zobrazování tabulky či grafu ze statisíců či milionů hodnot, odezva systému by mohla být neúnosně dlouhá. Hierarchie uložení agregovaných dat pak uživateli umožňuje pružně se po požadovaných úrovních agregace pohybovat (na úrovni kontinentů, zemí, regionů atd.), aniž by bylo nutné vždy znovu požadované agregace počítat. Tento princip se označuje jako drill-down (pohyb – zpřístupnění dat na nižší úroveň agregace – detailnějších dat) nebo drill-up (v opačném směru).[2]

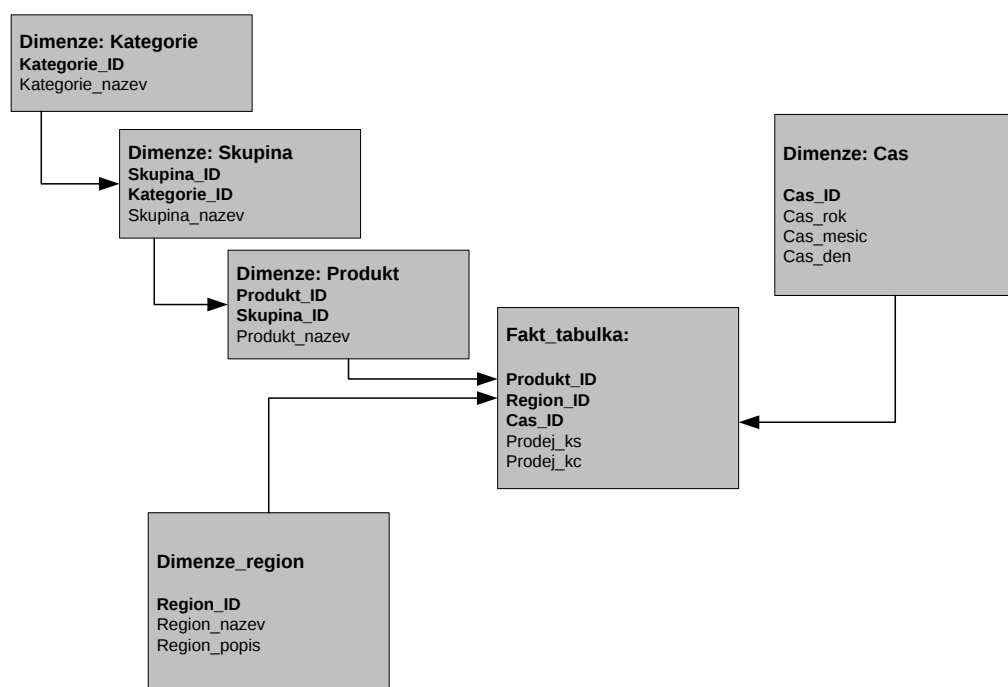
2.3. Relační databázový OLAP

Datové modely produkčních systémů jsou komplexní, obsahují mnoho tabulek a jejich vazeb. Takto organizovaná data jsou z hlediska jejich vytváření a aktualizace velmi efektivní, ale pro běžného uživatele se stávají velmi nepřehledná. Pro výše uvedený nedostatek se objevily snahy o zjednodušení takového uložení dat a jeho přizpůsobení pro

BI řešení. Vznikl tak relační dimenzionální model, kterému se také běžně říká Schéma hvězdy (STAR Scheme), resp. Schéma sněhové vločky (SNOWFLAKE Scheme).



Obrázek 2-3 Schéma hvězdy (STAR)



Obrázek 2-4 Schéma sněhové vločky (SNOWFLAKE)

V centru schématu je tzv. tabulka faktů, tedy tabulka sledovaných ekonomických a dalších ukazatelů identifikovaných klíčem složeným z klíčů tzv. dimenzionálních tabulek, v nichž jsou uloženy prvky jednotlivých dimenzí. Dimenzionální tabulky slouží pro uložení

textových informací o hodnotách uložených v tabulce faktů. Typicky si ji lze představit jako číselník. V některých případech je ovšem i řešení STAR schématu z řady důvodů nevýhodné. V těchto případech se proto dimenzionální tabulky upravují, resp. normalizují. To v tomto případě znamená, že se dimenzionální tabulka rozdělí podle hierarchických úrovní do více tabulek, aby se stejná data v tabulce neopakovala. Schéma, které takto vznikne, nazýváme SNOWFLAKE (schéma sněžné vločky).[2]

2.4. Hybridní OLAP (HOLAP)

Hybridní OLAP je kombinací úložišť MOLAP a ROLAP, přičemž jsou využívány výhody jednotlivých typů úložišť a do značné míry se eliminují nevýhody. Data jsou uložena v relačních databázích a vypočtené agregace (resp. jiné statistiky vypočtené z faktických údajů) se ukládají do multidimenzionálních struktur. V hybridním řešení relační databáze ukládá množství detailních dat a multidimenzionální model ukládá sumární data. [1]

3. Práce s OLAP databází v MS Excel

3.1. Použitá databáze

Pro zobrazení praktických ukázek v této práci je použita OLAP databáze FoodMart². Jedná se o ukázkovou databázi, která obsahuje data o hospodaření řetězce obchodů s potravinami.

Konkrétně se budeme připojovat ke kostce Sales and Employees, která obsahuje data o prodeích z roků 1997 a 1998. Kostka obsahuje 5 faktů:

- **Store Sales** (tržba)
- **Store Cost** (náklady)
- **Profit** (zisk, je vypočítán jako rozdíl tržby a nákladů)
- **Sales Count** (počet uskutečněných prodejů)
- **Unit Sales** (počet prodaných kusů)

Dále budeme používat tři dimenze:

Časová dimenze umožňuje zobrazovat data v závislosti na čase. Struktura úrovní dimenze je Rok → Čtvrtletí → Měsíc → Den.

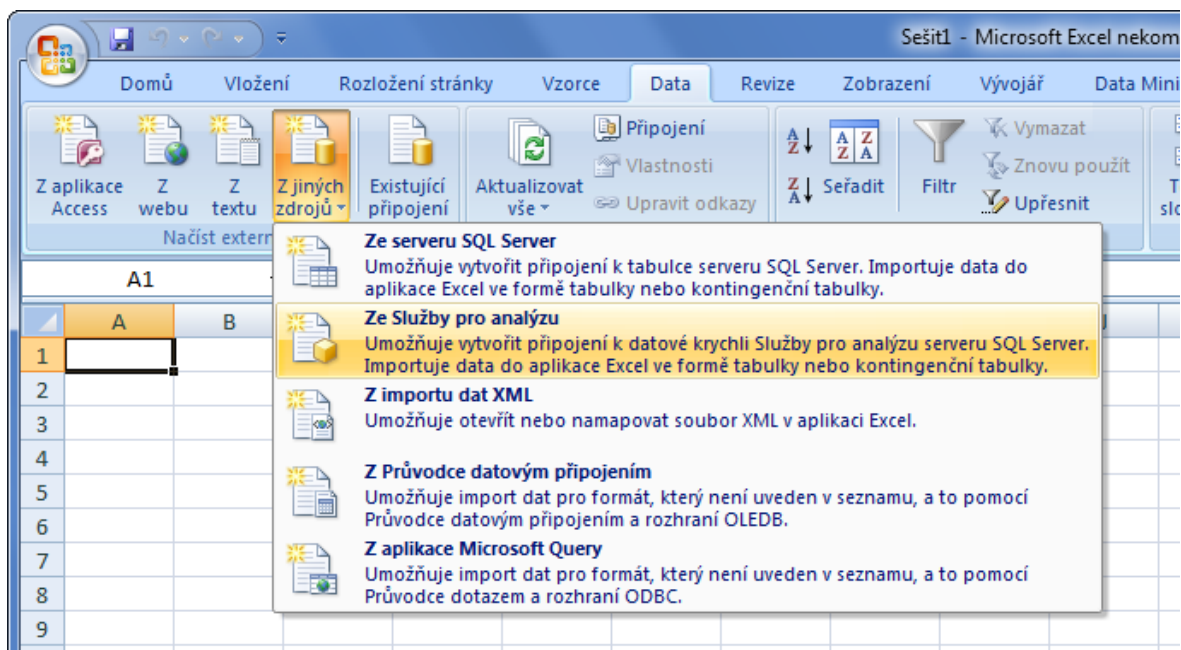
Dimenze Produktů – pomocí této dimenze je možné zobrazovat data v závislosti na kategoriích produktů, až po jednotlivé produkty. Databáze obsahuje přes 1500 různých produktů, které jsou zařazeny v této dimenzi.

Dimenze Stores – jedná se o dimenzi, podle které je možné zobrazovat data v závislosti na teritoriích, databáze obsahuje data pro 24 poboček, které jsou v USA, Mexiku a Kanadě.

3.2. Připojení k analytickému serveru

Pro připojení k analytickým službám SQL Serveru v aplikaci Excel 2007 zvolíme v hlavní nabídce programu položku **Data**. Po výměně panelu nástrojů klepneme na tlačítko **Z Jiných zdrojů** a ze zobrazené nabídky vybereme položku **Ze Služby pro analýzu**.

² Databáze FoodMart pro Microsoft SQL Server 2008 je ke stažení na adrese <http://www.informit.com/store/product.aspx?isbn=9780672330018>

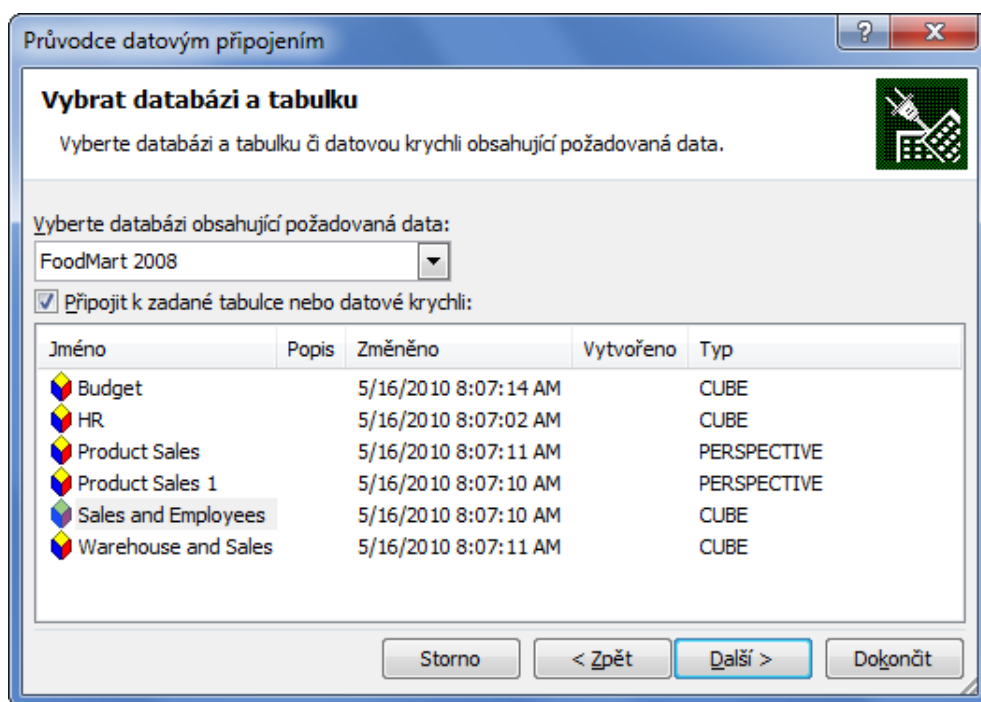


Obrázek 3-1 Možnosti připojení aplikace Excel k externímu zdroji dat

Tím se spustí průvodce datovým připojením. V úvodním kroku zadáme název serveru, uživatelské jméno a heslo. Kliknutím na tlačítko Další se připojíme k serveru.

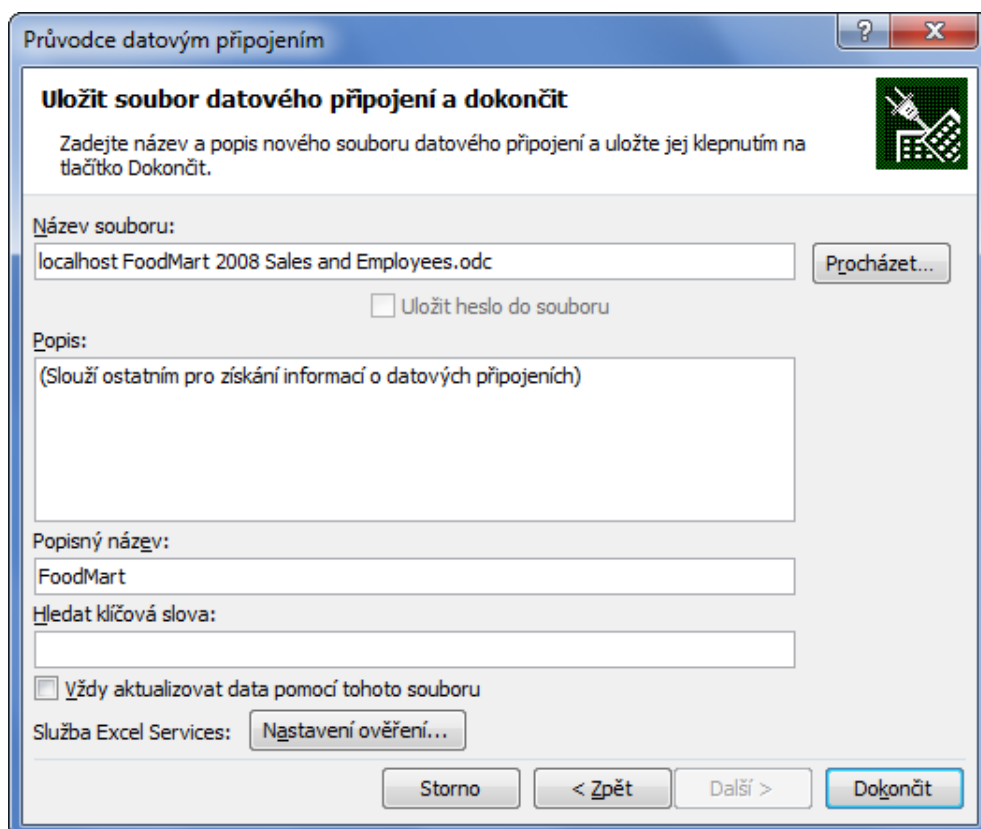
Po připojení následuje okno s výběrem databáze a OLAP kostky, ke které se chceme připojit. Na serveru se mohou nacházet dva typy datových kostek:

- CUBE představuje kompletní datovou kostku.
- PERSPECTIVE je zjednodušený pohled na kostku. Jedná se jen o logickou vrstvu využívající podmnožinu měřítek, dimenzí a hierarchií, takže žádná její fyzická realizace, například ve formě záznamů, nebo souborů, neexistuje.[1] Perspektiva se využívá hlavně v případech, kdy je kompletní kostka rozsáhlá a zobrazení všech měřítek a dimenzí by mohlo být pro koncového uživatele složité.



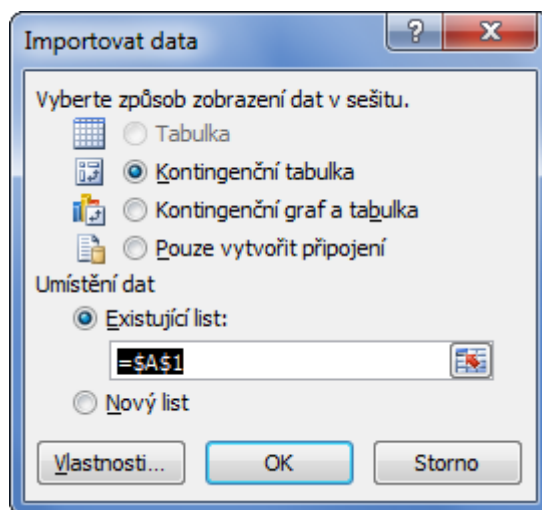
Obrázek 3-2 Dialog pro výběr analytické databáze a konkrétní kostky

Dalším krokem je nabídka uložení dat o připojení do souboru. Tento soubor je pak možné používat pro opětovné připojení k databázi.



Obrázek 3-3 Dialog pro uložení souboru připojení

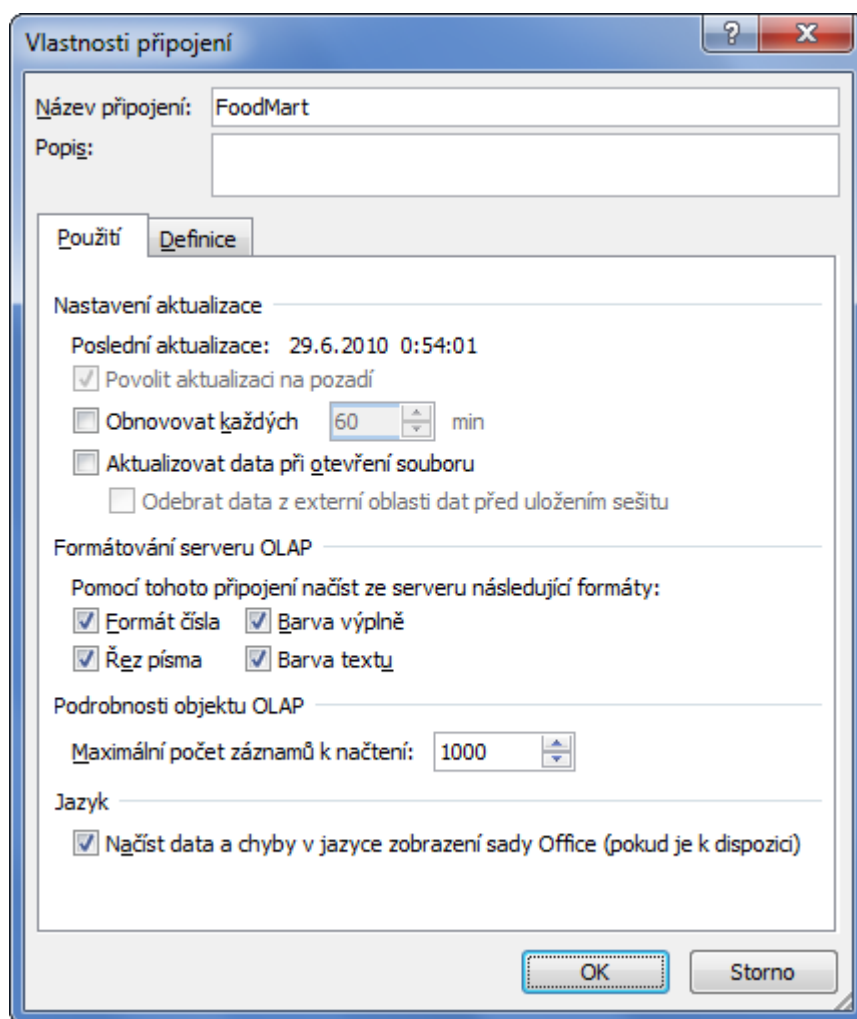
Poslední dialogové okno pak slouží pro výběr způsobu zobrazení dat z analytické databáze. Údaje z OLAP kostek můžeme zobrazit buď ve formě kontingenční tabulky, nebo kontingenčního grafu (mezi uvedenými volbami můžeme později libovolně přepínat).



Obrázek 3-4 Výběr způsobu zobrazení dat

Kliknutím na tlačítko **Vlastnosti** pak můžeme editovat nastavení připojení k analytickému serveru (toto nastavení je možné měnit i později v záložce **Data**). Na kartě **Použití** lze nastavit způsob použití informací o připojení používaných v kontingenční tabulce.

- Možnosti **Nastavení aktualizace** určují, jakým způsobem se budou aktualizovat data z OLAP databáze, v této nabídce je zajímavá možnost Odebrat data z externí oblasti dat před uložením sešitu, která zajistí, že při uložení sešitu se neukládají data z OLAP databáze, ale jen definice dotazu a při dalším otevření se data znovu načtou z databáze.
- Možnosti **Formátování serveru OLAP** umožňují zvolit, jestli se mají z OLAP serveru načítat informace o formátování dat.
- Podrobnosti objektu OLAP umožňují zadat maximální počet záznamů, který se může načíst z databáze.
- Poslední možnost pak určuje, jestli se ze serveru mají načítat data přeložená do jazyka, který používá Excel (v případě, že na OLAP serveru existují)



Obrázek 3-5 Vlastnosti připojení

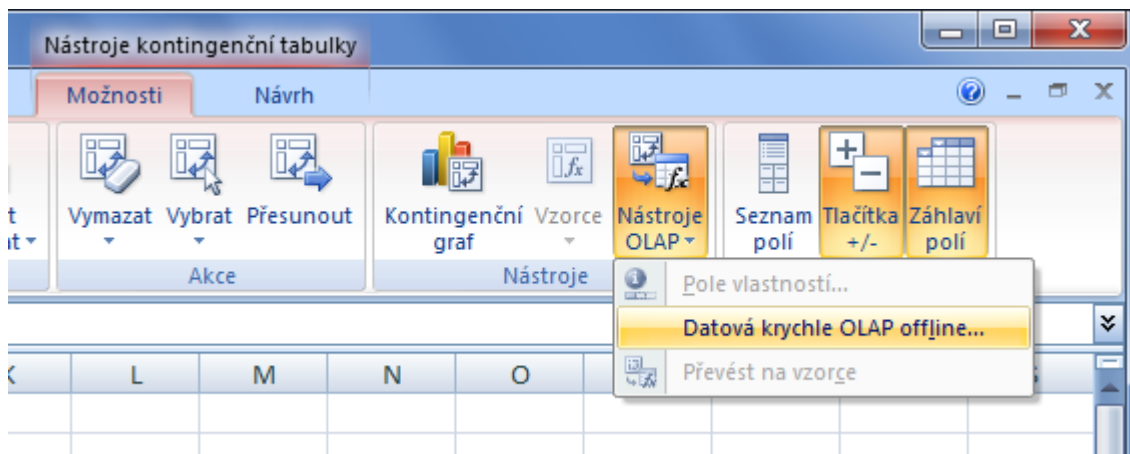
Na kartě **Definice** je pak možné měnit vlastnosti připojení.

3.3. Zpřístupnění OLAP krychle v režimu offline

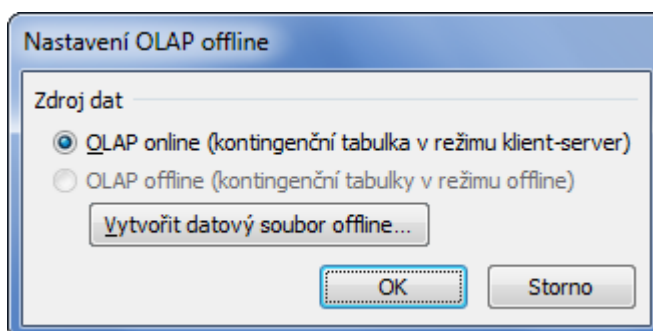
OLAP kostky uložené na pevném disku počítače umožňují pracovat v režimu offline, tedy i po odpojení od OLAP serveru. Kostka se uloží do souboru s příponou .cub a bude obsahovat všechny údaje potřebné k vytvoření sestav kontingenčních tabulek nebo grafů.

Při používání kostek v režimu offline je potřeba si uvědomit, že v uloženém souboru nejsou data nijak zabezpečená, a proto je potřeba dbát na to, že soubor může obsahovat citlivá data, která by neměla být přístupná nepovolaným osobám.

K uložení OLAP kostky slouží průvodce vytvoření souboru datové krychle, který spustíme v Excelu z karty **Možnosti** z nabídky **Nástroje OLAP** (tato položka je v nabídce dostupná jen tehdy, pokud je vybrána sestava s kontingenční tabulkou).

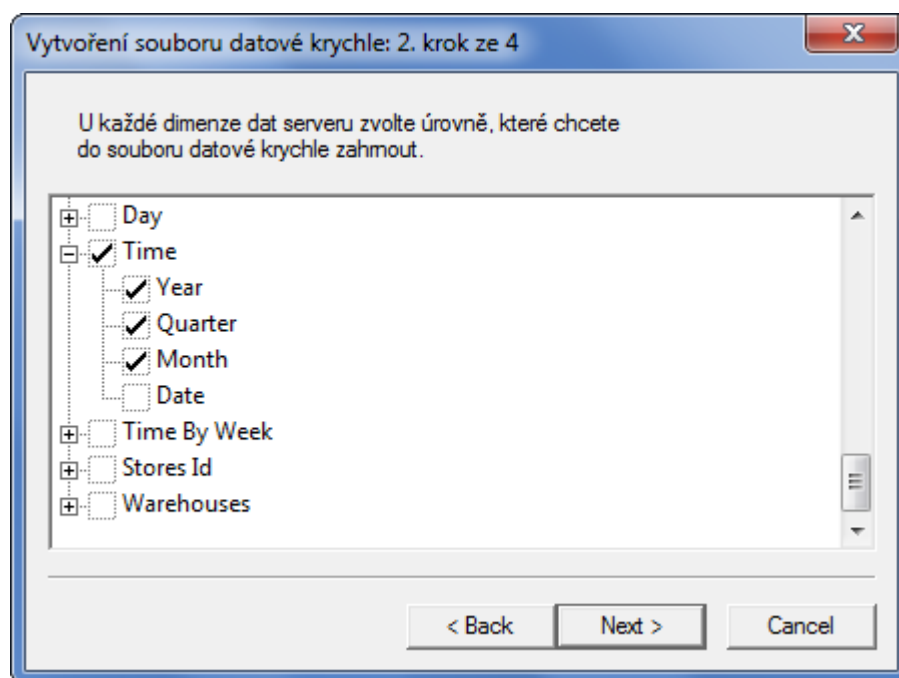


Obrázek 3-6 Uložení datové krychle v offline režimu



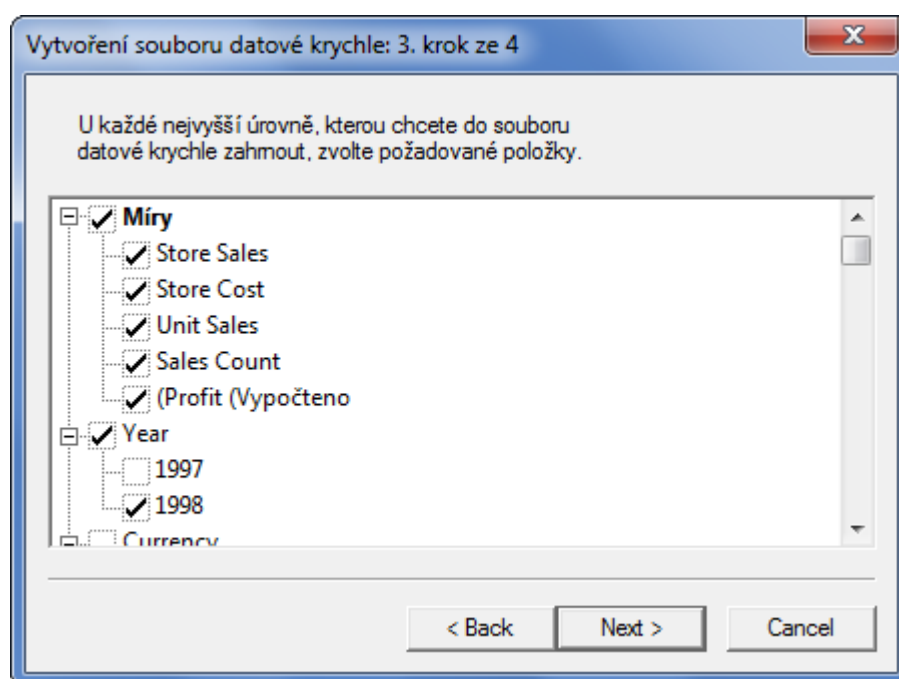
Obrázek 3-7 Nastavení OLAP offline

Kliknutím na tlačítko Vytvořit datový soubor offline spustíme průvodce vytvořením souboru datové krychle. Prvním krokem je výběr dimenzí, které chceme zahrnout do souboru kostky. Pomocí tlačítek + můžeme rozvinout jednotlivé dimenze na požadované hierarchické úrovně. V případě, že je potřeba zmenšit velikost uloženého souboru, můžeme vynechat nižší úrovně. Střední úrovně v rámci dimenze nelze vynechat.



Obrázek 3-8 Výběr dimenzí pro uložení

Dalším krokem je výběr měřítek, která budou uložena v souboru kostky. Měřítko se v kontingenční tabulce dají použít jako datová pole. Měřítko musí být vybráno alespoň jedno. Potom je potřeba vybrat položky jednotlivých dimenzí, které mají být zahrnuty v souboru kostky (pro každou dimenzi, kterou jsme vybrali v předchozím kroku, je potřeba vybrat alespoň jednu položku).



Obrázek 3-9 Výběr měřítek pro uložení

Posledním krokem je výběr umístění a názvu souboru, do kterého chceme kostku uložit.

Stejným způsobem, jakým jsme kostku vytvořili, lze kdykoliv aktualizovat data v souboru kostky na nejnovější, popřípadě přidávat další dimenze a měřítka do již vytvořené kostky.

3.4. Práce s kontingenční tabulkou

Kontingenční tabulka představuje interaktivní způsob rychlého shrnutí velkých objemů dat. Pomocí kontingenční tabulky lze do maximálních podrobností analyzovat numerická data a získat odpovědi na nepředpokládané otázky, které byste mohli k datům mít.

Kontingenční tabulka je primárně vytvořena k těmto účelům:

- dotazování na velké množství dat uživatelsky přívětivými způsoby;
- vytváření mezisoučtů a souhrnů číselných dat, sumarizace dat dle kategorií a podkategorií a vytváření vlastních výpočtů a vzorců;
- rozbalení a sbalení úrovní dat za účelem zaměření se na výsledky a přechod ze souhrnných dat k podrobnostem odpovídajícím dané oblasti zájmu;
- přesunutí řádků do sloupců nebo sloupců do řádků za účelem zobrazení různých souhrnů zdrojových dat;
- filtrování, řazení, seskupování a podmíněné formátování nejužitečnějších a nejzajímavějších podmnožin dat za účelem zaměření se na požadované informace[5].

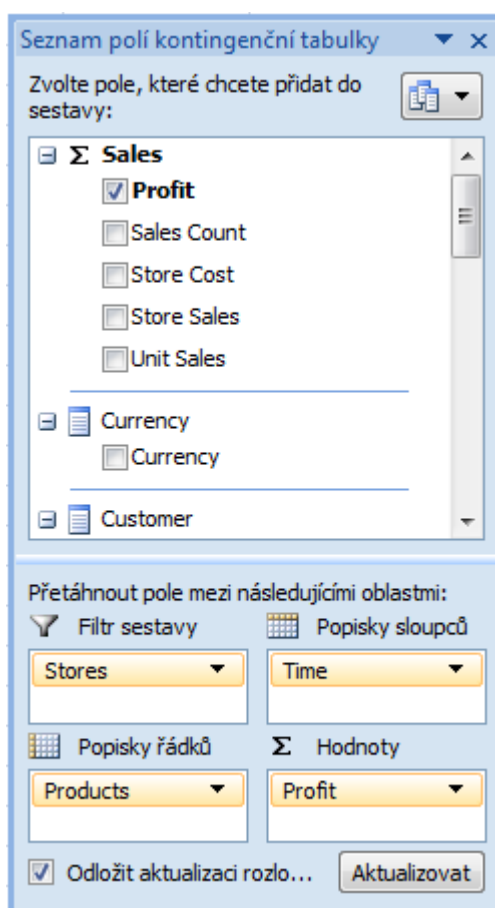
Při zobrazení dat z OLAP databází umožňuje kontingenční tabulka zobrazovat souhrny dat z tabulky faktů v souvislosti s dimenzemi. V kontingenční tabulce lze také jednoduše procházet jednotlivé úrovně dimenzí (tento princip - zpřístupnění detailnějších dat - se označuje jako **drill-down**[3]).

Základním ovládacím prvkem kontingenční tabulky je okno **Seznam polí kontingenční tabulky**. Toto okno má 5 částí. V horní části jsou zobrazena data a dimenze, které můžeme v kontingenční tabulce zobrazit. V dolní části jsou pak 4 pole, do kterých lze přetahovat položky z horní části okna:

- **Filtr sestavy** – do této části se dají přidat dimenze, podle kterých potom lze data v kontingenční tabulce filtrovat.

- **Popisky sloupců** – toto pole určuje, podle jakých dimenzí bude rozdělena hlavička tabulky
- **Popisky řádků** – určuje dimenze, podle kterých budou tříděna data v řádcích
- **Hodnoty** – do tohoto pole lze přidat hodnoty (fakta), která budou zobrazena v buňkách kontingenční tabulky

Pokud nechceme, aby se kontingenční tabulka aktualizovala po každé úpravě polí (což může být dost zdlouhavé kvůli pomalému načítání dat z analytického serveru), můžeme zaškrtnout položku Odložit aktualizaci rozložení, a tabulku aktualizovat ručně pomocí tlačítka **Aktualizovat** až když si vybereme, které pole chceme zobrazit.



Obrázek 3-10 Seznam polí kontingenční tabulky

Takto se budou v naší ukázkové kontingenční tabulce zobrazovat souhrny hodnot zisku (profit) v závislosti na časové dimenzi a dimenzi produktů (products). Sestavu budeme filtrovat podle dimenze pobočky (stores).

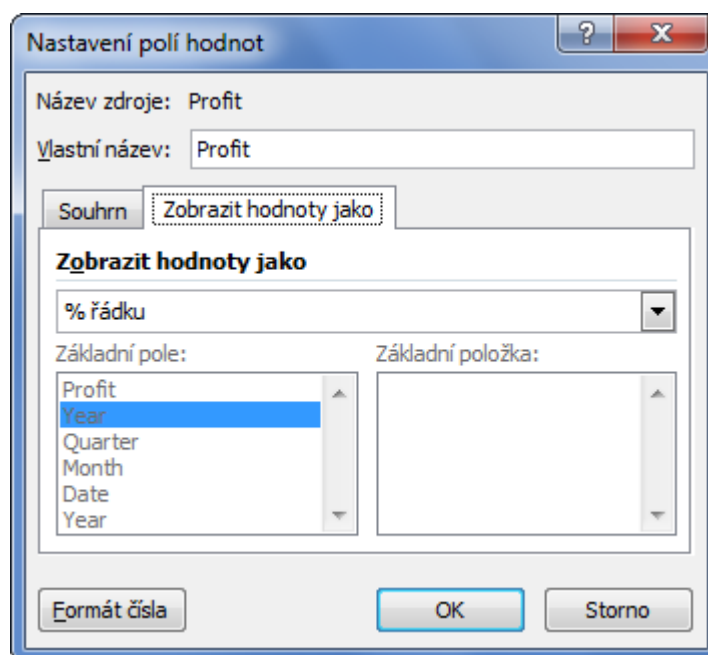
3.4.1. Zobrazení dat v kontingenční tabulce

Práce s kontingenční tabulkou je celkem intuitivní, pomocí tlačítek + a – můžeme zobrazovat různé úrovně dimenzí, v buňkách tabulky se zobrazují agregovaná data v závislosti na dimenzích v popiscích sloupců a řádků. V naší tabulce jsme zobrazili podrobněji data za 1. čtvrtletí roku 1997, a to na úroveň měsíců. Dokud se u libovolné dimenze zobrazuje tlačítko +, lze u ní stejným způsobem zobrazit podrobnější data až do nejnižší úrovně (v případě časové dimenze můžeme zobrazit data až pro jednotlivé dny).

Stores	All			
Profit	Popisky sloupců			
Popisky řádků	☐ Drink ☐ Food ☐ Non-Consumable			
☐ 1997	29358,9754	245764,8665	64487,0545	339610,8964
☐ Q1	6964,2966	60814,4714	16097,3415	83876,1095
☐ January	2261,6937	19827,935	5271,5687	27361,1974
☐ February	2191,2887	19289,4004	4978,4321	26459,1212
☐ March	2511,3142	21697,136	5847,3407	30055,7909
☐ Q2	7186,1098	57323,3736	15192,5618	79702,0452
☐ Q3	7203,3445	61262,5473	15901,1288	84367,0206
☐ Q4	8005,2245	66364,4742	17296,0224	91665,7211
☐ 1998	56243,491	466142,1581	124196,092	646581,7411
Celkový součet	85602,4664	711907,0246	188683,1465	986192,6375

Obrázek 3-11 Kontingenční tabulka

Pro lepší přehlednost lze změnit typ zobrazení hodnot v tabulce. Například pokud bychom chtěli mezi sebou porovnávat hodnoty pro jednotlivé skupiny výrobků, bylo by jednodušší srovnávat relativní hodnoty v procentech místo konkrétních údajů o zisku: kliknutím na položku hodnoty, kterou chceme změnit (v našem případě zisk (profit)), v okně **Seznam polí kontingenční tabulky** vyvoláme kontextové menu, ve kterém pak zvolíme položku **Nastavení polí hodnot**. Tím vyvoláme okno pro nastavení pole hodnot.



Obrázek 3-12 Nastavení polí hodnot

V tomto okně pak můžeme v listu **Zobrazit hodnoty jako** zvolit, jakým stylem se bude hodnota zobrazovat (způsob zobrazení v listu **Souhrn** nelze měnit, protože souhrny jsou předem vypočítané v OLAP kostce). Při práci s OLAP kostkou můžeme použít možnosti % řádku, % sloupce, % celku.

Stores	All				
Profit	Popisky sloupců				
Popisky řádků	+ Drink	+ Food	+ Non-Consumable	Celkový součet	
1997	8,64%	72,37%	18,99%	100,00%	
Q1	8,30%	72,51%	19,19%	100,00%	
January	8,27%	72,47%	19,27%	100,00%	
February	8,28%	72,90%	18,82%	100,00%	
March	8,36%	72,19%	19,45%	100,00%	
Q2	9,02%	71,92%	19,06%	100,00%	
Q3	8,54%	72,61%	18,85%	100,00%	
Q4	8,73%	72,40%	18,87%	100,00%	
1998	8,70%	72,09%	19,21%	100,00%	
Celkový součet	8,68%	72,19%	19,13%	100,00%	

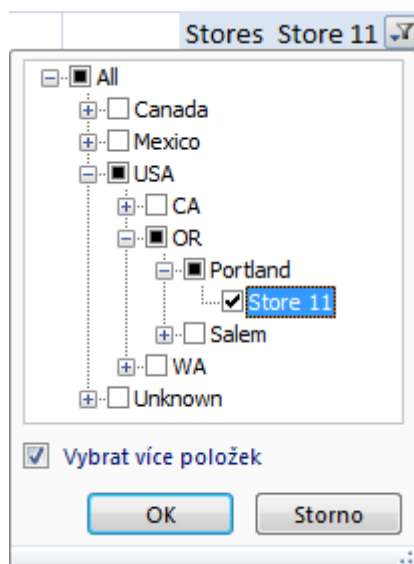
Obrázek 3-13 Zobrazení hodnot jako % z řádku

Takto zobrazené hodnoty mají určitě větší vypovídací hodnotu než na tabulce na obrázku 3-11, pokud chceme jednotlivé hodnoty vzájemně porovnávat. V jedné kontingenční tabulce bohužel nelze zobrazit jednu hodnotu dvakrát různým způsobem, takže nemůžeme zobrazit hodnotu v procentech a zároveň hodnotu v reálných číslech.

3.4.2. Filtrování dat v kontingenční tabulce

Data v kontingenční tabulce lze filtrovat nejen podle parametru zvoleného jako filtr sestavy, ale i podle popisků řádků a sloupců. Kliknutím na tlačítko  u popisků řádků, popisků sloupců nebo u filtru sestavy vyvoláme menu pro volbu dat k filtrování.

V případě filtrování podle položky v poli Filtr sestavy je možné pouze zvolit, jaká data se mají v tabulce zobrazovat. V našem příkladu jsme vybrali pro filtrování dimenzi pobočky (stores), pomocí které můžeme data v tabulce filtrovat podle regionů, ve kterých se nachází jednotlivé pobočky, až po konkrétní pobočku. Takto můžeme omezit použitá data například pouze na pobočku č. 11 (Store 11).



Obrázek 3-14 Filtr sestavy

V případě filtrování popisku řádků nebo popisku sloupců aplikace nabízí více možností:

- Řazení
- Filtry popisků
- Filtry hodnot
- Výběr úrovní dimenze, které mají být v tabulce zobrazeny

Řazení je možné provádět buď podle názvů úrovní dimenzí, nebo podle hodnot ekonomických údajů, které jsou v tabulce zobrazeny.

Filtry popisků umožňují v podstatě vyhledávat v popisích dimenze, v případě, že jsou popisky textové, můžeme použít operace s textovými řetězci. V případě, že jsou popisky

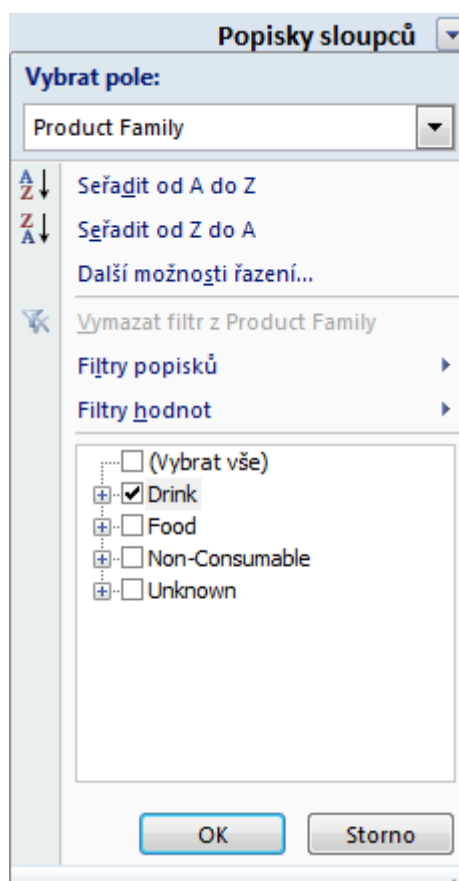
číselné, nabízí filtr popisků filtrování podle základních matematických operací (rovná se, nerovná se, je větší nebo menší než zadané číslo apod.).

V případě, že filtrujeme časovou dimenzi, nabízí Excel místo filtrů popisků filtry kalendářních dat. Ty umožňují omezit zobrazení dat na určité období, a to i v závislosti na aktuálním datu (např. zobrazit data za poslední týden).

Filtry hodnot slouží k filtrování tabulky podle hodnot faktů, které tabulka obsahuje.

Výběr úrovní dimenze pak opět slouží k filtrování hodnot v tabulce podle popisků jednotlivých úrovní dimenzí.

Pokud chceme v naší tabulce zobrazit pouze skupinu produktů nápoje (Drink), bude volba filtru vypadat následovně:



Obrázek 3-15 Filtr a řazení popisků sloupců

3.4.3. Změna popisků v kontingenční tabulce

Při načítání dat ze serveru do kontingenční tabulky obsahují popisky hodnoty, uložené na serveru. Všechny popisky, které kontingenční tabulka obsahuje, je však možné měnit –

pokud nám nevyhovuje název, jednoduše jej můžeme přepsat. Je to výhodné zejména v případech, kdy připravujeme obsahově stejnou kontingenční tabulku v různých jazykových mutacích. Je třeba mít na paměti, že tyto změny jsou provedeny jen v sešitu Excelu, v případě, že tabulku smažeme, změny, které jsme v popiscích provedli, nebudou zachovány.

Další zajímavou možností je vytváření vlastních skupin. Kliknutím na tlačítko Výběr skupiny je možné sloučit vybrané skupiny v popiscích sloupců nebo řádků. Tím je vlastně možné vytvořit vlastní úroveň v hierarchii dimenzí. Například můžeme přidat v časové dimenzi další úroveň tím, že sloučíme data za dvě čtvrtletí. Tak vznikne nová úroveň, kterou si pak můžeme pojmenovat.

V níže uvedené tabulce jsme už změnili popisky všech dimenzí z původních anglických na české a vytvořili nové úrovně **1. a 2. Pololetí** pro rok 1997 sloučením skupin příslušných čtvrtletí.

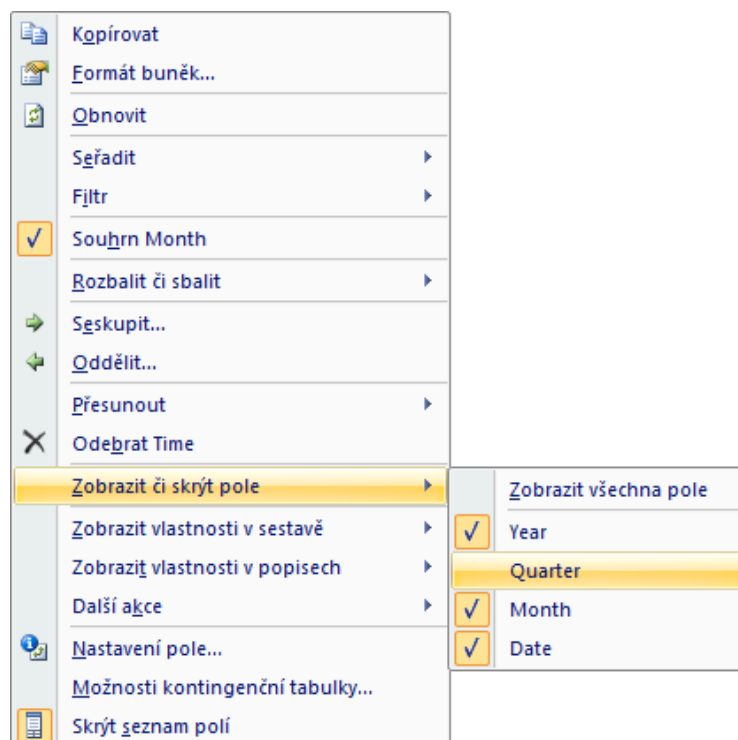
Je třeba dodat, že nově vytvořená úroveň se vztahuje pouze na skupiny, které jsme sloučili, takže pro rok 1998 bychom ji museli vytvořit znovu. Sloučení do skupin také nemusí být symetrické, například bychom mohli vytvořit novou úroveň, ve které by v první skupině byly sloučeny údaje za 1. až 3. čtvrtletí a ve druhé skupině by bylo pouze 4. čtvrtletí.

Zisk	Popisky sloupců			
Popisky řádků	Nápoje	Potraviny	Nespotřební zboží	Celkový součet
1997	29358,9754	245764,8665	64487,0545	339610,8964
1. Pololetí	14150,4064	118137,845	31289,9033	163578,1547
Q1	6964,2966	60814,4714	16097,3415	83876,1095
Leden	2261,6937	19827,935	5271,5687	27361,1974
Únor	2191,2887	19289,4004	4978,4321	26459,1212
Březen	2511,3142	21697,136	5847,3407	30055,7909
Q2	7186,1098	57323,3736	15192,5618	79702,0452
2. Pololetí	15208,569	127627,0215	33197,1512	176032,7417
1998	56243,491	466142,1581	124196,092	646581,7411
Celkový součet	85602,4664	711907,0246	188683,1465	986192,6375

Obrázek 3-16 Vlastní rozdělení do skupin

Při vytváření vlastních skupin je třeba počítat s tím, že agregovaná data pro takto vytvořené nové úrovně nejsou předem vypočítaná v OLAP databázi a dopočítává je až Excel. Proto se může každá aktualizace dat v tabulce značně prodloužit.

Podobně se také dají v tabulce úrovně skrýt. Skrytí tabulky je možné pomocí kontextového menu, které se zobrazí po kliknutí pravým tlačítkem na popisky sloupců nebo řádků a vybráním úrovně v menu **Zobrazit či skrýt pole**.

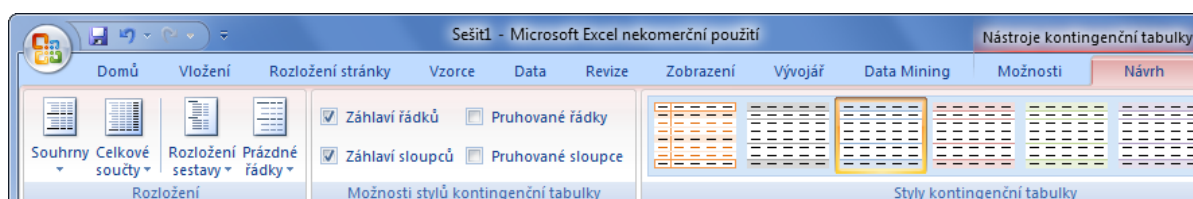


Obrázek 3-17 Kontextové menu popisků kontingenční tabulky

Po skrytí se úroveň v kontingenční tabulce nebude zobrazovat, například pokud v naší tabulce skryjeme úroveň Quarter (čtvrtletí), budou se po „rozkliknutí“ roku zobrazovat rovnou jednotlivé měsíce.

3.4.4. Úprava vzhledu kontingenční tabulky

Excel umožňuje tabulky velmi jednoduše upravovat, i co se týče vzhledu. K základním úpravám slouží nabídka **Návrh**.



Obrázek 3-18 Nabídka Návrh

Ke změně barevného zobrazení tabulky slouží pravá část nabídky, kde si můžeme ve skupině **Styly kontingenční tabulky** vybrat barevné provedení a ve skupině **Možnosti stylů kontingenční tabulky** označit, které sloupce a které řádky chceme mít vybarveny.

Skupina tlačítek **Rozložení** pak slouží k významnějším úpravám vzhledu tabulky, které jsou popsány níže:

- Tlačítko **Souhrny** slouží k určení, jakým stylem se budou zobrazovat souhrny jen pro aktuálně zobrazenou úroveň, nebo i pro vyšší úrovně. Souhrny lze buď úplně vypnout, nebo si zvolit, jestli se budou zobrazovat na prvním nebo na posledním řádku dané úrovně:

Zisk	Popisky sloupců	
Popisky řádků	Nápoje	Potraviny
1997		
1. Pololetí		
Q1	6964,2966	60814,4714
Q2	7186,1098	57323,3736
2. Pololetí	15208,569	127627,0215
Celkový součet	29358,9754	245764,8665

Obrázek 3-19 Tabulka s vypnutými souhrny

Zisk	Popisky sloupců	
Popisky řádků	Nápoje	Potraviny
1997		
1. Pololetí		
Q1	6964,2966	60814,4714
Q2	7186,1098	57323,3736
Celkem z 1. Pololetí	14150,4064	118137,845
2. Pololetí	15208,569	127627,0215
Celkem z 1997	29358,9754	245764,8665
Celkový součet	29358,9754	245764,8665

Obrázek 3-20 Tabulka se souhrny v dolní části skupiny

Zisk	Popisky sloupců	
Popisky řádků	Nápoje	Potraviny
1997	29358,9754	245764,8665
1. Pololetí	14150,4064	118137,845
Q1	6964,2966	60814,4714
Q2	7186,1098	57323,3736
2. Pololetí	15208,569	127627,0215
Celkový součet	29358,9754	245764,8665

Obrázek 3-21 Tabulka se souhrny v horní části skupiny

- Pomocí tlačítka **Celkové součty** je možné vypínat a zapínat zobrazení celkových součtů dat na posledním řádku a v posledním sloupci tabulky.
- Dalším tlačítkem je **Rozložení sestavy**, které ovlivňuje, jakým způsobem se budou zobrazovat popisky řádků kontingenční tabulky při zobrazování nižších úrovní. Výslednou tabulku je možné zobrazit v kompaktní formě, ve formě osnovy anebo ve formě tabulky. Jednotlivé způsoby zobrazení jsou ukázány na obrázcích 3-18, 3-19 a 3-20:

Zisk	Popisky sloupců		
Popisky řádků	Nápoje	Potraviny	Celkový součet
1997	29358,9754	245764,8665	275123,8419
1. Pololetí	14150,4064	118137,845	132288,2514
Q1	6964,2966	60814,4714	67778,768
Q2	7186,1098	57323,3736	64509,4834
2. Pololetí	15208,569	127627,0215	142835,5905
Celkový součet	29358,9754	245764,8665	275123,8419

Obrázek 3-22 Kompaktní forma

Zisk	Product Family		
Year	Year1	Quarter	Celkový součet
1997			29358,9754 245764,8665 275123,8419
	1. Pololetí		14150,4064 118137,845 132288,2514
		Q1	6964,2966 60814,4714 67778,768
		Q2	7186,1098 57323,3736 64509,4834
	2. Pololetí		15208,569 127627,0215 142835,5905
Celkový součet			29358,9754 245764,8665 275123,8419

Obrázek 3-23 Kontingenční tabulka ve formě osnovy

Zisk	Product Family				
Year	Year1	Quarter	+ Nápoje	+ Potraviny	Celkový součet
1997	1. Pololetí	Q1	6964,2966	60814,4714	67778,768
		Q2	7186,1098	57323,3736	64509,4834
	Celkem z 1. Pololetí		14150,4064	118137,845	132288,2514
	2. Pololetí		15208,569	127627,0215	142835,5905
Celkem z 1997			29358,9754	245764,8665	275123,8419
Celkový součet			29358,9754	245764,8665	275123,8419

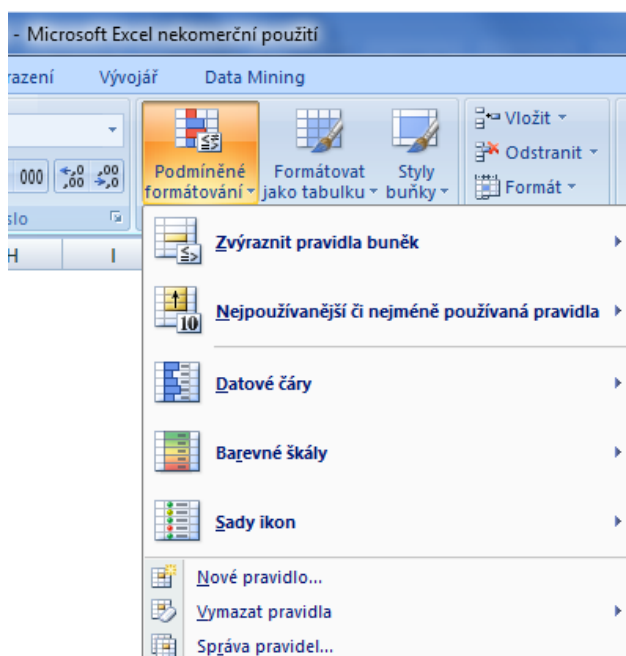
Obrázek 3-24 Kontingenční tabulka ve formě tabulky

- Tlačítko **Prázdné řádky** umožňuje přidat za každou úroveň skupiny prázdný řádek a zvýšit tím přehlednost tabulky.

3.4.5. Podmíněné formátování

Pomocí podmíněného formátování lze snadno upozornit na zajímavé buňky nebo oblasti buněk, zvýraznit neobvyklé hodnoty a vizualizovat data užitím datových čar, barevné škály a sad ikon. Při použití podmíněného formátování se změní vzhled oblasti buňky v závislosti na podmínce (nebo kritériu). Vyhovuje-li oblast buňky podmínce, zformátuje se dle podmínky. Nevyhovuje-li oblast buňky podmínce, nezformátuje se dle podmínky.[5]

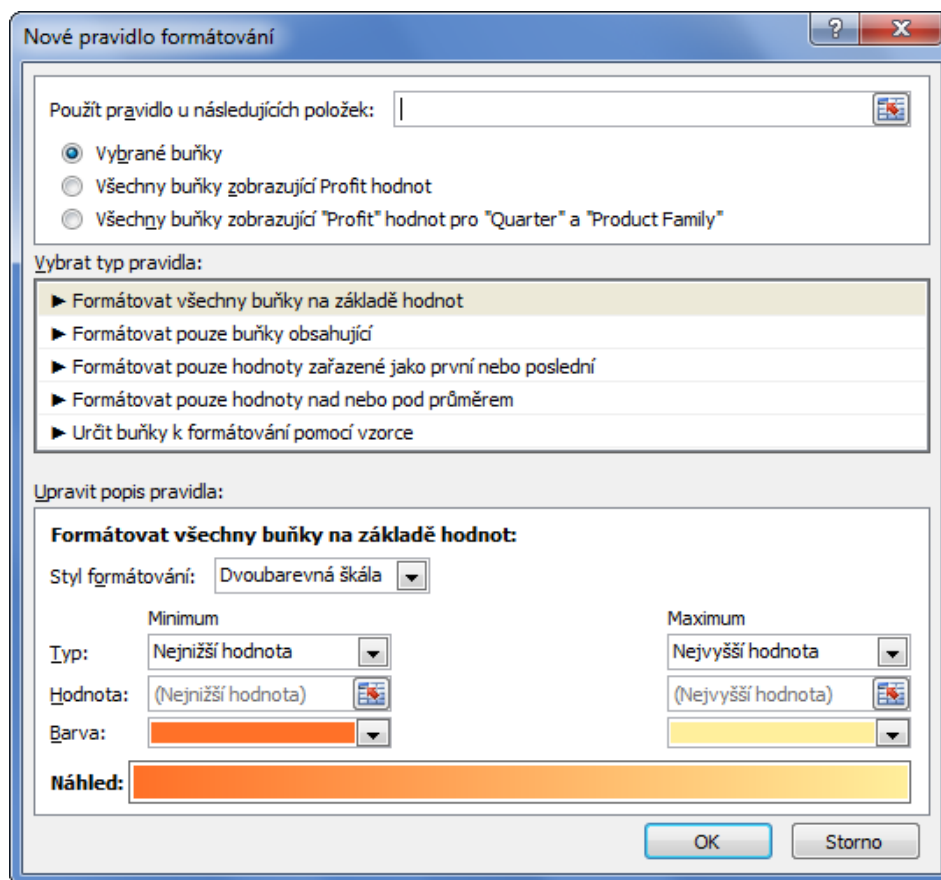
Tlačítko **Podmíněné formátování** lze najít v záložce **Domů**.



Obrázek 3-25 Nabídka Podmíněné formátování

V nabídce **Podmíněné formátování** najdeme rychlé odkazy pro vytvoření nejčastěji používaných způsobů zobrazení. Kliknutím na možnost **Nové pravidlo** pak zobrazíme okno

pro vytvoření nové podmínky, ve kterém můžeme nastavit více podrobností pro podmíněné formátování:



Obrázek 3-26 Přidání nového pravidla formátování

Ve vrchní části okna se nachází možnosti pro zvolení oblasti, kterou chceme formátovat. Výchozí způsob definování rozsahu polí v oblasti hodnot je výběrem buněk. Pomocí přepínače **Použít pravidlo u následujících položek** lze změnit způsob definování rozsahu odpovídajícím polem nebo hodnotou pole. Uvedené tři způsoby poskytují větší pružnost při řešení našich potřeb:

- **Definování rozsahu výběrem buněk** - tento způsob použijeme, když chceme, aby se podmíněné formátování zobrazovalo pouze v buňkách, které jsme označili.
- **Definování rozsahu hodnotou pole** - tento způsob použijeme, když chceme, aby se formátování zobrazovalo ve všech buňkách, které obsahují zvolený typ hodnoty (faktu) ve všech úrovních kontingenční tabulky (včetně celkových součtů v tabulce).

- **Definování rozsahu odpovídajícím polem** - tento způsob použijeme, když chceme, aby se podmíněné formátování zobrazovalo pouze v buňkách pro danou hodnotu, a konkrétní úrovni dimenzí v řádku i ve sloupci tabulky.

Dále musíme vybrat typ pravidla, které chceme použít. Excel nabízí 5 základních možností formátování, které se pak dají dále upřesnit v sekci **Upravit popis pravidla**.

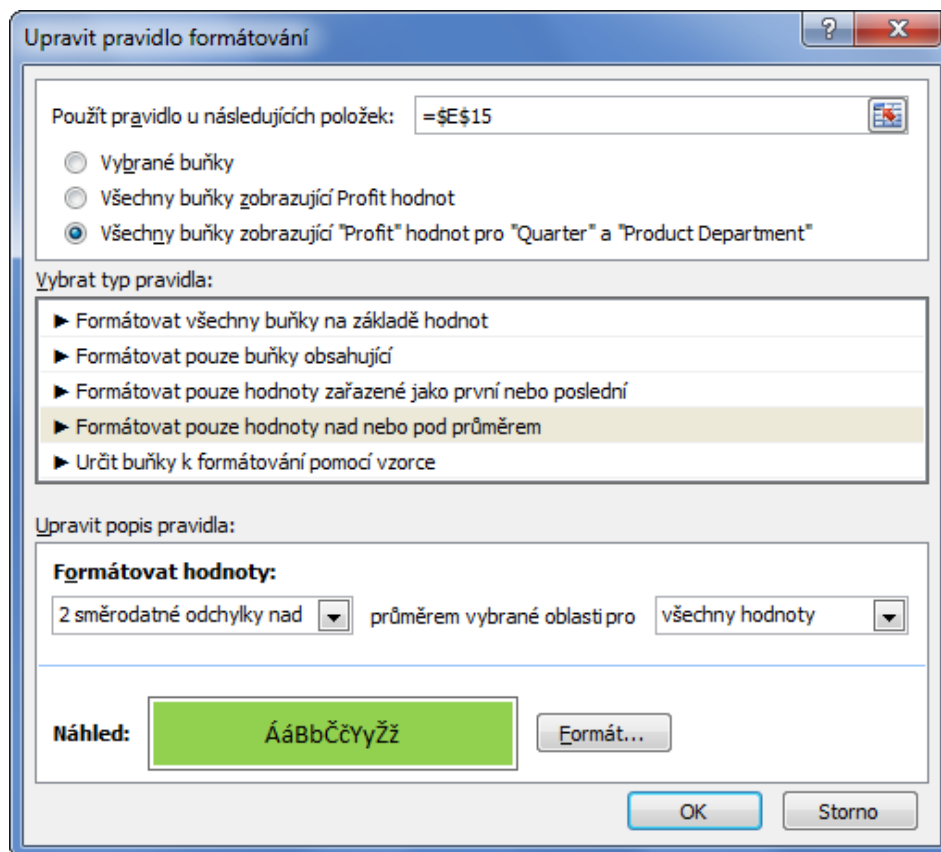
- Prvním typem pravidla je možnost formátovat všechny buňky na základě hodnot - takto lze barevně odlišit hodnoty ve formátovaných buňkách, a to s pomocí dvoubarevné nebo tříbarevné škály, datových pruhů, anebo pomocí sad ikon.

Počet prodejů	Popisky sloupců				
Popisky řádků	Nápoje	Potraviny	Nespotřební zboží	Celkový součet	
1997	7978	62445	16414	86837	
Q1	1959	15539	4090	21588	
Q2	1901	14594	3873	20368	
Q3	1959	15449	4045	21453	
Q4	2159	16863	4406	23428	
Celkový součet	7978	62445	16414	86837	

Obrázek 3-27 Různé typy zobrazení podmíněného formátování všech buněk na základě hodnot

- Druhým typem pravidla pro formátování je možnost formátovat pouze buňky, které vyhovují určitému pravidlu. Jako podmínku můžeme zvolit, zda hodnota v buňce vyhovuje danému výrazu (menší než, větší než, rovná se, spadá do daného intervalu apod.), dále je možné označit buňky, které jsou prázdné, nebo naopak ty, které nejsou prázdné. Také je možné označit buňky, které obsahují nebo neobsahují chyby.
- Třetím typem pravidla je možnost zvolit formátování pro vybraný počet (nebo procento) nejvyšších nebo nejvyšších hodnot ve formátované oblasti.
- Čtvrtý typ pravidla pro podmíněné formátování umožňuje označit hodnoty vzhledem k průměru v celé formátované oblasti. Kromě jednoduché možnosti zvýraznit hodnoty, které jsou nad nebo pod průměrem, také můžeme pomocí tohoto pravidla označit hodnoty podle toho, jak vysoká je jejich směrodatná odchylka v nad nebo pod průměrem (směrodatná odchylka v podstatě určuje, jak je daná hodnota vzdálená od průměru).

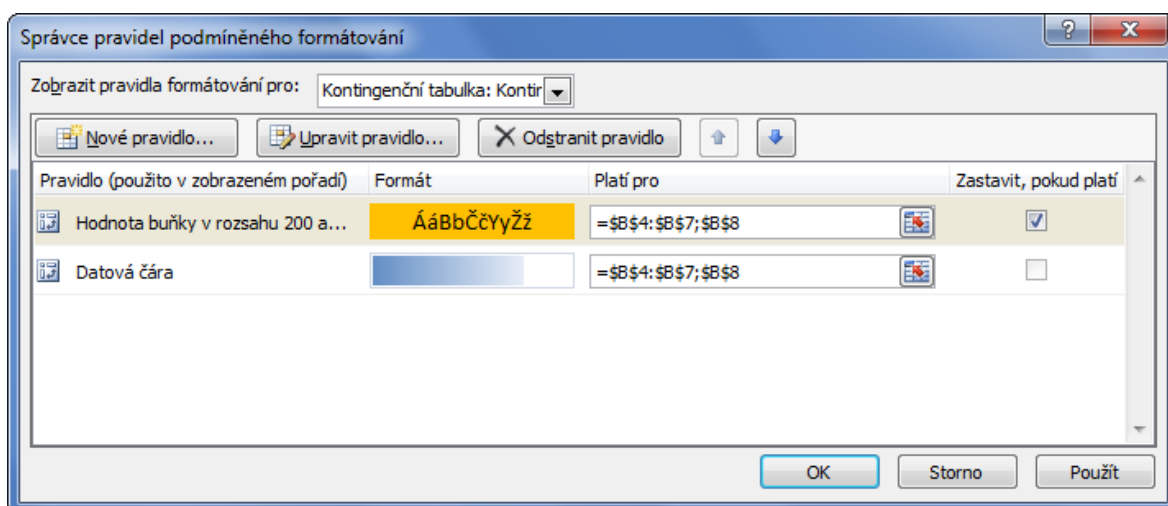
Pomocí tohoto typu formátování bychom například mohli zobrazit, z jakých skupin produktů plyne výrazně větší zisk ve srovnání s ostatními skupinami tak, že bychom použili podmíněné formátování pro hodnoty, které mají 2 směrodatné odchylky nad průměrem:



Obrázek 3-28 Podmíněné formátování ve vztahu k průměru

- Poslední typ pravidla pak umožňuje Podmíněné formátování podle vlastního vzorce.

Pravidel pro podmíněné formátování může být v jedné tabulce větší množství, pro zobrazení a jednoduchou správu pravidel podmíněného formátování slouží položka **Správa pravidel** v menu pro podmíněné formátování. V tomto okně je pak možné všechna pravidla upravovat, mazat a přidávat nová. Pomocí správy pravidel je také možné nastavit pořadí pravidel, což v kombinaci s políčkem **Zastavit, pokud platí**, umožňuje zabránit tomu, aby na jedné buňce bylo použito více pravidel podmíněného formátování – pravidla formátování se ověřují od shora dolů, pokud buňka vyhovuje více pravidlům, bude formátována víckrát. Buňku, která vyhovuje pravidlu, u kterého je zaškrtnuto políčko **Zastavit, pokud platí**, budou následující pravidla formátování vynechávat.



Obrázek 3-29 Správce pravidel podmíněného formátování

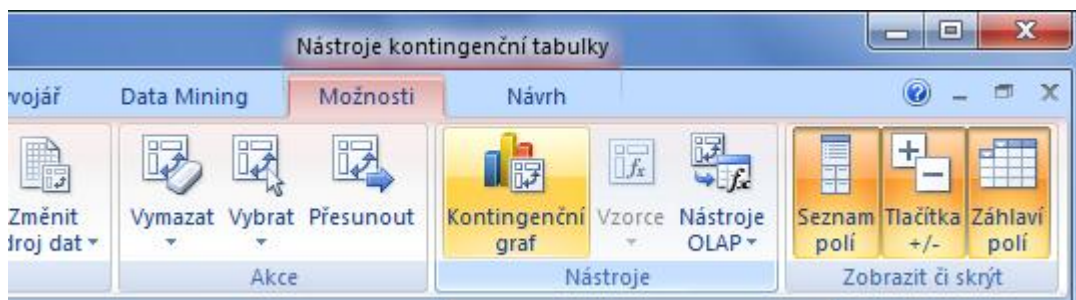
3.5. Práce s kontingenčním grafem

Pomocí kontingenčního grafu můžeme snadno graficky znázornit data z kontingenční tabulky. Kontingenční graf nabízí téměř stejnou funkcionalitu jako kontingenční tabulka. Mezi kontingenčním grafem a kontingenční tabulkou je oboustranná vzájemná vazba, takže změny rozložení sestavy v kontingenčním grafu se projeví i v kontingenční tabulce a naopak.

Schopnost vizuálně zobrazit a analyzovat trendy vývoje z dat je velkou výhodou kontingenčních grafů. Někteří uživatelé umí data vyhodnotit lépe, jestliže je mohou analyzovat v grafické podobě, nikoli jako prostou číselnou řadu.

3.5.1. Vytvoření kontingenčního grafu

Při práci s OLAP databází můžeme kontingenční graf vytvořit buď už při připojení k databázi, nebo kdykoliv později kliknutím na tlačítko **Kontingenční graf** v kartě **Možnosti** (tato položka je v nabídce dostupná jen tehdy, pokud je vybrána sestava s kontingenční tabulkou – s tou také bude kontingenční graf propojen).



Obrázek 3-30 Nabídka Nástroje kontingenční tabulky

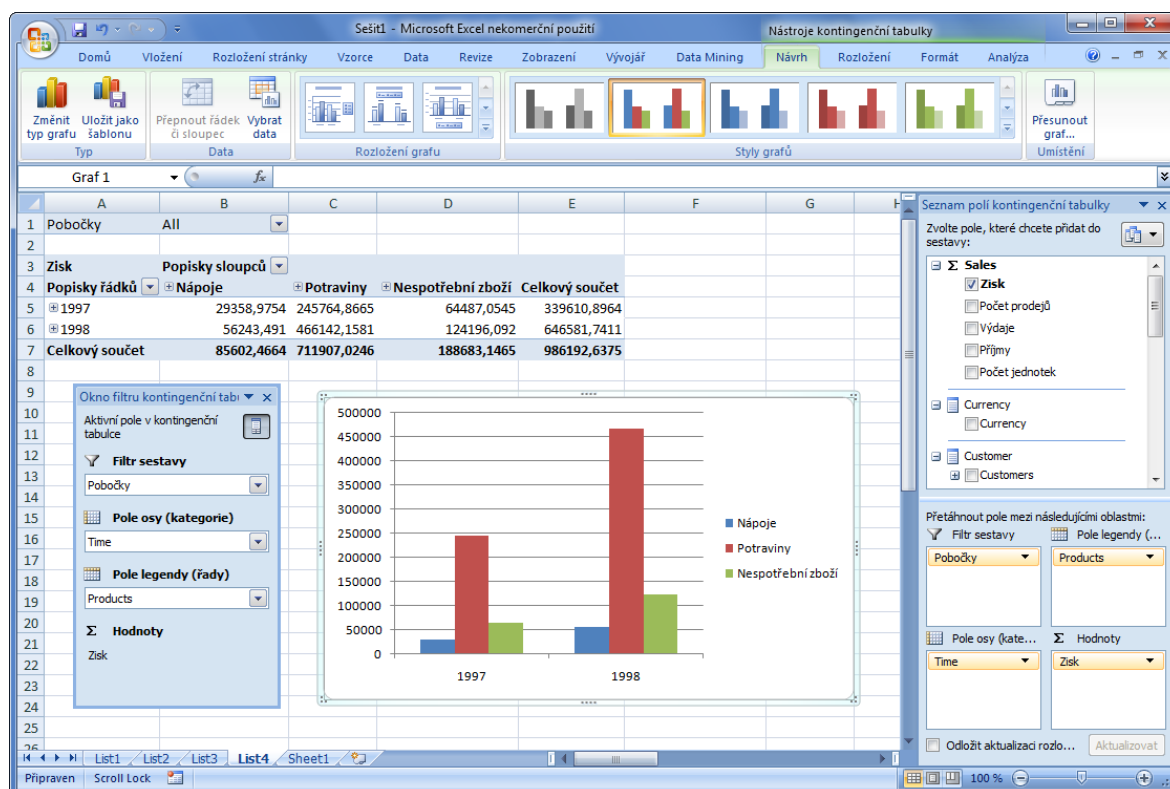
Pro zobrazení kontingenčního grafu máme na výběr z několika typů grafů a několika možností formátování. Asi nejčastěji používanými typy jsou grafy sloupcové, výšečové, spojnicové nebo pruhové.

Pro příklady zobrazení grafů použijeme stejnou strukturu tabulky, jako jsme používali v kapitolách zabývajících se prací s kontingenční tabulkou.

3.5.2. Zobrazení dat v kontingenčním grafu

Pro určení, jaké dimenze a hodnoty budou v grafu zobrazeny, se dá použít okno **Seznam polí kontingenční tabulky**, které ovládá jak kontingenční tabulku, tak i kontingenční graf. Při použití s kontingenčním grafem se pole, která jsou v kontingenční tabulce zobrazena jako popisky sloupců, zobrazují jako popisky legendy a pole, která jsou v tabulce použita jako popisky řádků, se zobrazují jako popisky jedné osy grafu. Na druhé ose grafu jsou pak zobrazeny hodnoty.

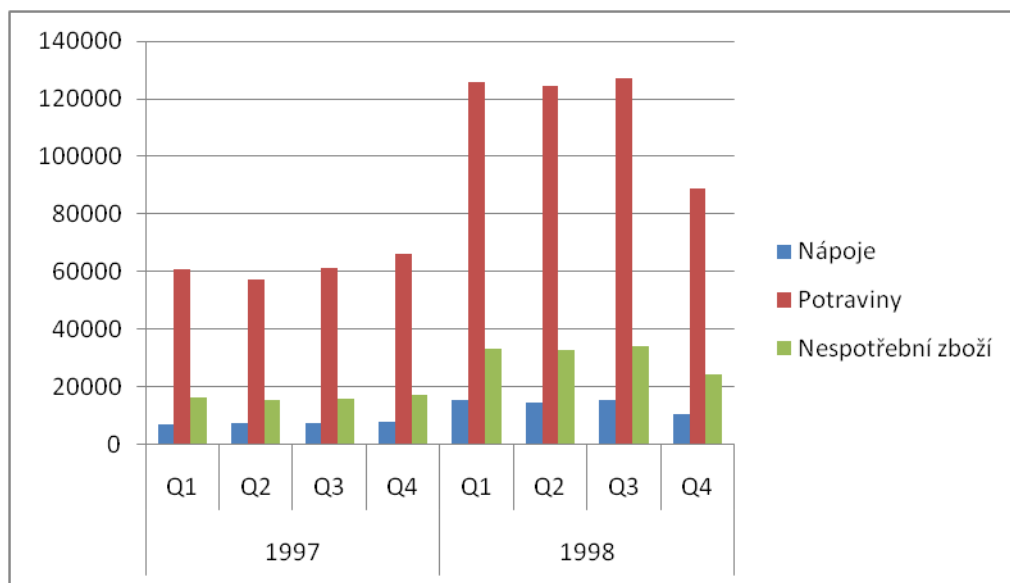
V našem grafu zobrazíme hodnoty zisku rozdělené podle jednotlivých let a skupin produktů (podle úrovně, které jsou viditelné v kontingenční tabulce, se kterou je graf spojen):



Obrázek 3-31 Kontingenční graf

Zobrazování dalších úrovní dimenzí je pak možné několika způsoby, zobrazení můžeme ovlivnit buď pomocí kontingenční tabulky, nebo jednoduše dvojklikem na legendu zobrazit nižší úroveň dimenze a dvojklikem na popisec osy (nebo přímo na sloupce grafu) zobrazit nižší úroveň dimenze, která je použita jako pole osy.

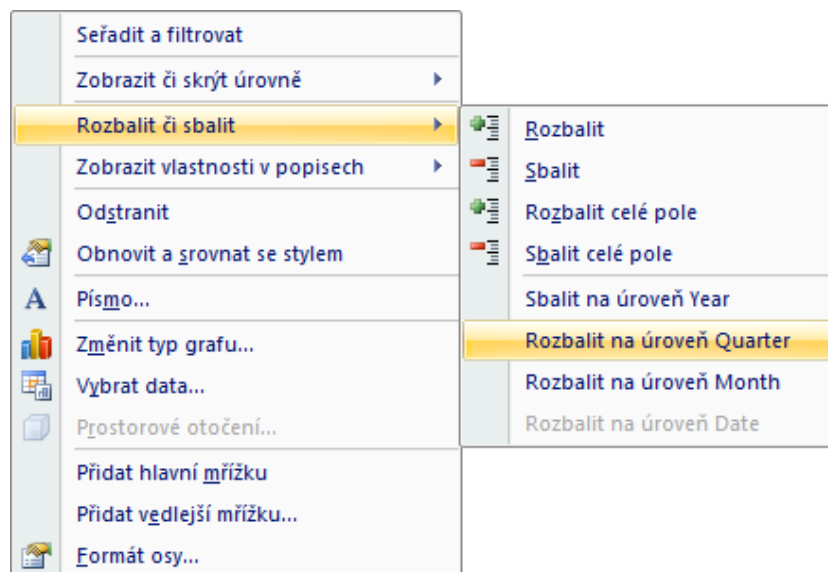
V našem grafu se po dvojkliku na popisec osy (nápis 1997 nebo 1998) zobrazí graf pro jednotlivá čtvrtletí obou let, který bude vypadat následovně:



Obrázek 3-32 Podrobnější graf

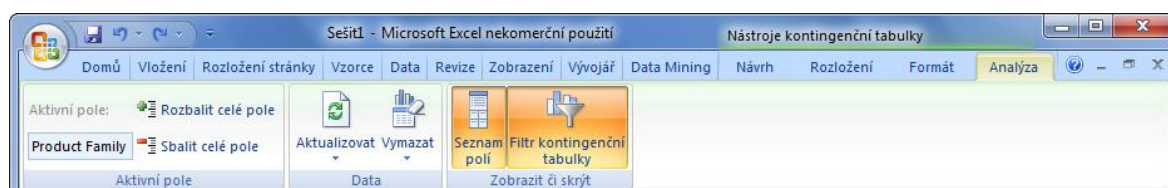
Při ovládání grafu pomocí kontingenční tabulky lze docílit i toho, že zobrazení v grafu bude nerovnoměrné. Například v našem grafu bychom mohli zobrazit 4 sloupce pro jednotlivá čtvrtletí roku 1997 a vedle nich jeden sloupec pro celý rok 1998. Takové zobrazení by ale nemělo smysl, protože graf zobrazuje souhrny hodnot, protože nelze dost dobře vizuálně porovnávat hodnoty, které jsou zobrazeny v různých časových jednotkách (čtvrtletí oproti roku). Ovládání přímo na kontingenčním grafu se v tomto ohledu chová logicky a při přechodu na nižší úroveň zobrazí všechny položky ve stejném měřítku.

Další možnosti pro ovládání zobrazení různých úrovní dimenzí jsou dostupné z kontextového menu, které vyvoláme kliknutím pravým tlačítkem myši na popisec legendy, popř. osy. V této nabídce je také možné přímo vybrat úroveň, kterou chceme zobrazit.



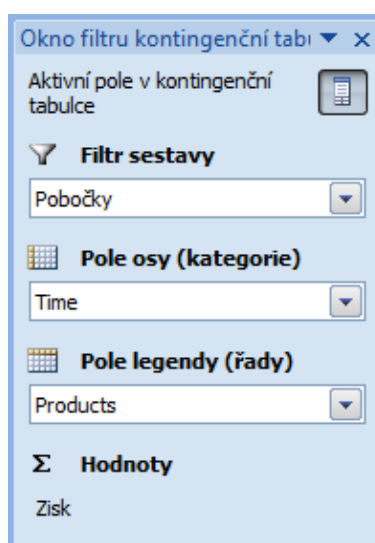
Obrázek 3-33 Kontextové menu

K pohybu mezi úrovněmi dimenzí také slouží nabídka Analýza, pomocí které můžeme sbalit nebo rozbalit skupiny, které máme zrovna označené. V této nabídce je také možnost aktualizovat data v grafu a znovu načíst data ze serveru.



Obrázek 3-34 Nabídka Analýza

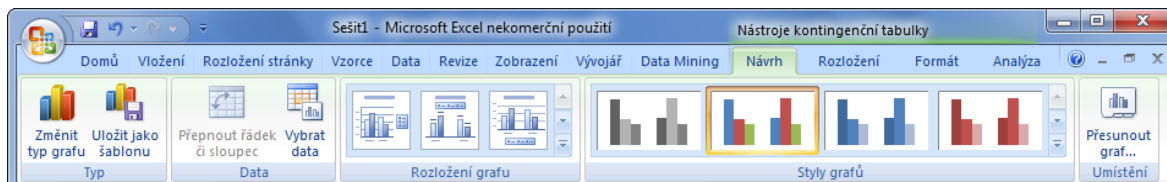
K filtrování dat zobrazených v kontingenčním grafu slouží **Okno filtru kontingenční tabulky**, pomocí kterého se dá zvolit, jaká data se mají v kontingenčním grafu zobrazit.



Obrázek 3-35 Okno filtru kontingenční tabulky

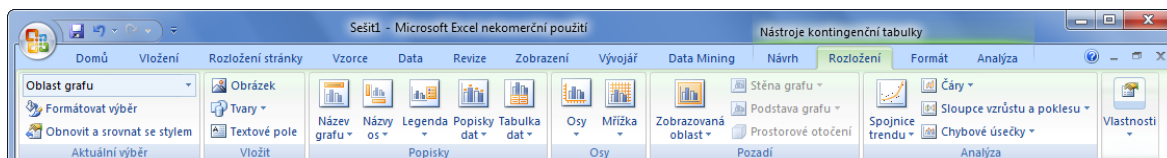
3.5.3. Formátování kontingenčního grafu

K úpravě formátování slouží další tři nabídky (které se zobrazí, pokud máme označenou kontingenční tabulku).



Obrázek 3-36 Nabídka Návrh

Nabídka **Návrh** slouží k rychlým změnám grafu podle předdefinovaných šablon, v této nabídce můžeme změnit typ grafu, měnit barevnou kombinaci grafu a vybrat jeden z předem definovaných způsobů rozložení grafu (to určuje, jak budou v grafu rozvrženy popisky legendy a os apod.). V této nabídce se také nachází důležité tlačítko **Přesunout graf**, pomocí kterého můžeme přesunout graf do jiného listu, popřípadě přesunout graf do nového grafického listu (v tomto listu je zobrazen pouze graf, a to na celé ploše listu).

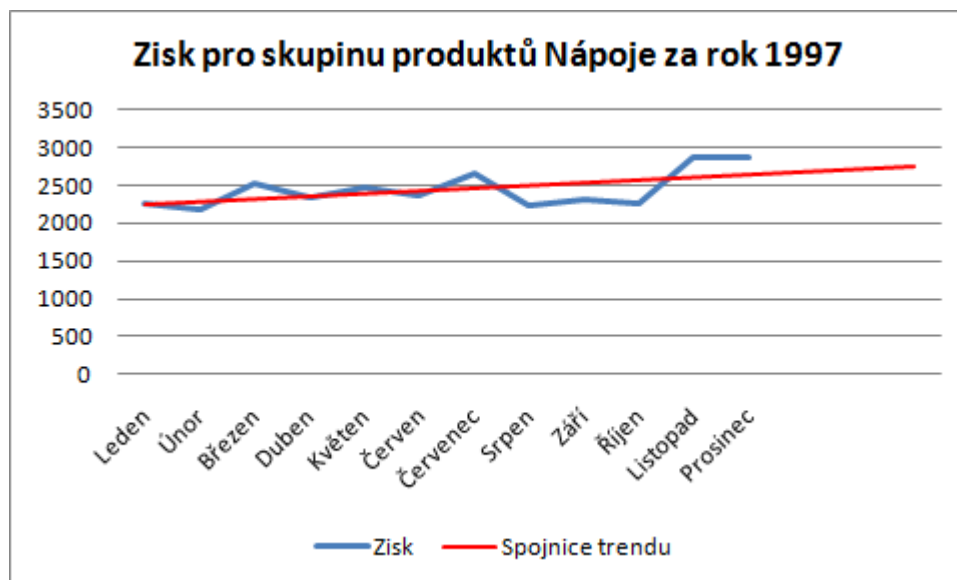


Obrázek 3-37 Nabídka Rozložení

V nabídce **Rozložení** pak můžeme dále upravovat jednotlivé části grafu – měnit název grafu, názvy os, vzhled legendy atd. Velice užitečné jsou možnosti **Popisky dat** a **Tabulka dat**, které zobrazí u každého bodu v grafu číselnou velikost hodnoty v daném bodě (v případě možnosti **Popisky dat**), nebo zobrazí pod grafem tabulku s číselnými hodnotami pro každý bod grafu.

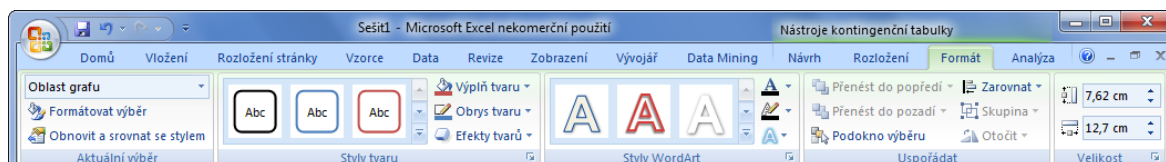
Další zajímavou možností v této nabídce jsou **Spojnice trendu**. Spojnice trendu je grafické znázornění trendů v datové řadě, například vzhůru stoupající křivka představující zvýšený prodej v průběhu měsíců. Spojnice trendu se používají při studiu problémů předpovědi, které se také nazývá regresní analýza[5]. Nabídka umožňuje vytvořit několik typů spojníc a vytvořit spojnice s předpovědí na určitý počet období. V této nabídce je také možné v grafu zobrazit klouzavý průměr za určitý počet období.

Jako příklad můžeme vytvořit spojnicový graf, který zobrazuje zisk pro skupinu produktů Nápoje za rok 1997 a přidat do něj spojnici trendu, která vypovídá o nárůstu prodeje nápojů a zobrazuje směr trendu, kterým se hodnota ubírá.



Obrázek 3-38 Graf s lineární spojnici trendu s předpovědí na další 3 období

Poslední nabídkou pro formátování kontingenčního grafu je nabídka **Formát**:



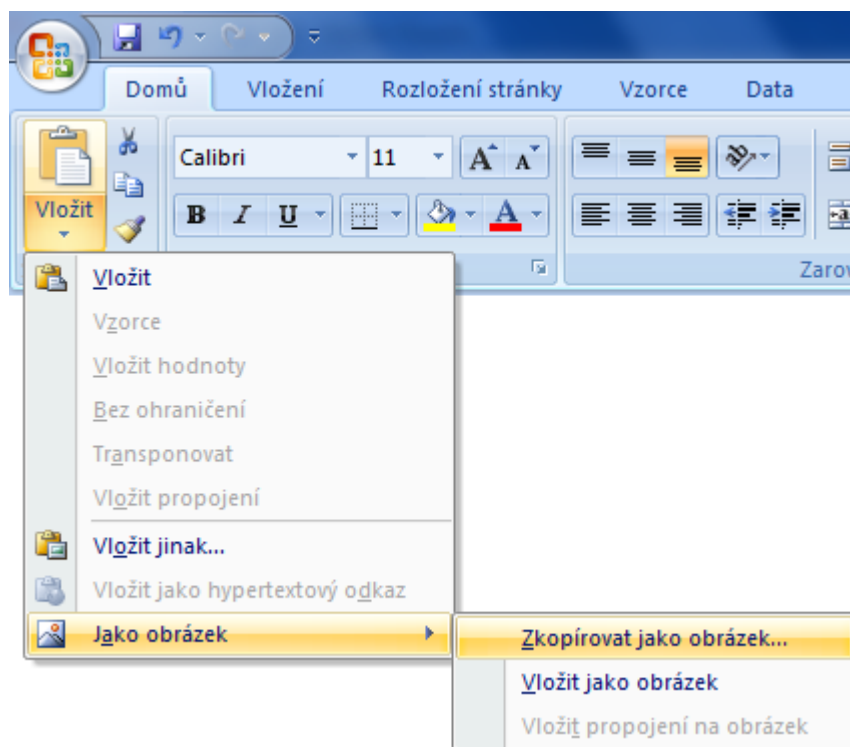
Obrázek 3-39 Nabídka Formát

Tato nabídka obsahuje nástroje pro grafické formátování jakéhokoliv prvku v grafu (je možné formátovat všechny popisky linky v grafu, např. i mřížku na pozadí grafu atp.).

3.5.4. Kopírování grafů a jejich použití v ostatních aplikacích

V případě, že chceme vytvořený graf použít v jiných aplikacích (např. do přehledové zprávy, vystavit na webových stránkách apod.), můžeme ho jednoduše zkopírovat, a to buď pravým kliknutím na graf a vybráním položky **Kopírovat**, nebo využitím menu **Vložit**, které je na kartě **Domů**. Pokud takto zkopírovaný graf vkládáme do aplikace, která umožňuje práci s vektorovými obrázky, bude zkopírován jako vektorový (v případě použití s dalšími aplikacemi z balíku Microsoft Office je dokonce možné zachovat propojení s podkladovou kontingenční tabulkou a změny v tabulce se budou dále v grafu projevovat).

V případě, že chceme graf kopírovat pouze jako obrázek, můžeme použít v menu Vložit položku **Jako obrázek**.

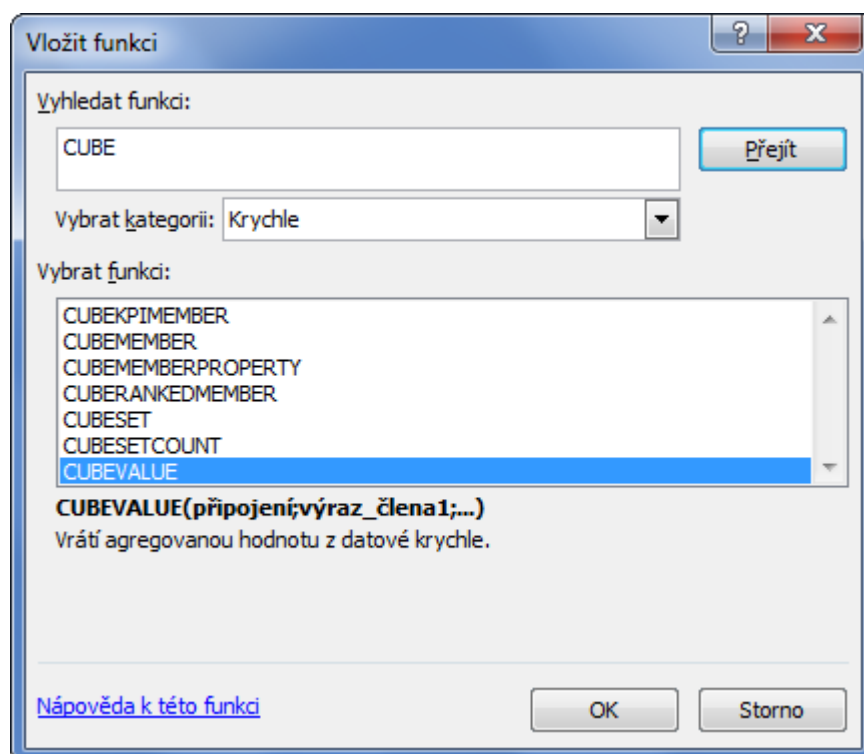


Obrázek 3-40 Menu Vložit

Toto menu je také ve všech ostatních kancelářských aplikacích Microsoft Office, a tak v nich lze při vkládání vybrat, v jakém formátu se graf vloží.

3.6. Skupina funkcí CUBE pro práci s OLAP kostkou

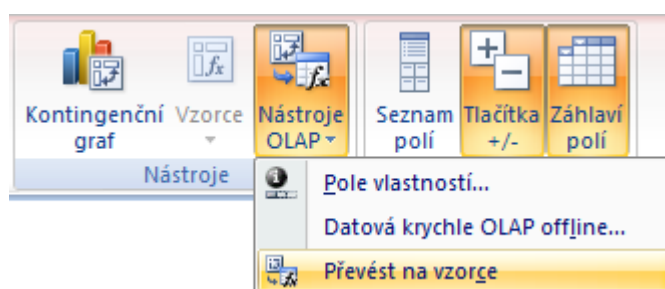
Excel pro práci s OLAP databázemi nabízí kromě kontingenční tabulky a kontingenčního grafu také skupinu funkcí CUBE, které slouží pro práci s daty z OLAP kostky. Možnosti zobrazení dat pomocí těchto funkcí jsou v podstatě stejné, jako nabízí kontingenční tabulka.



Obrázek 3-41 Skupina funkcí pro práci s OLAP krychlí

Alternativně lze také použít funkci ZÍSKATKONTDATA, která slouží k získání dat z kontingenční tabulky.

Kontingenční tabulku je také možné převést na vzorce pomocí menu **Nástroje OLAP** na kartě **Možnosti**.



Obrázek 3-42 Převedení kontingenční tabulky na vzorce

Převědeme-li buňky na vzorce, mají tyto vzorce přístup ke stejným datům jako kontingenční tabulka a mohou být aktualizovány, aby byly zobrazeny aktuální výsledky. Nebudeme ovšem již mít přístup, s výjimkou filtru sestavy, k interaktivním funkcím kontingenční tabulky, například k filtrování, řazení nebo k úrovním rozbalení a sbalení.[5]

V případě, že je v kontingenční tabulce použit filtr sestavy, je po převedení na vzorce možné tento filtr v sešitu ponechat. Data zobrazovaná pomocí CUBE funkcí budou pak filtrována podle tohoto filtru.

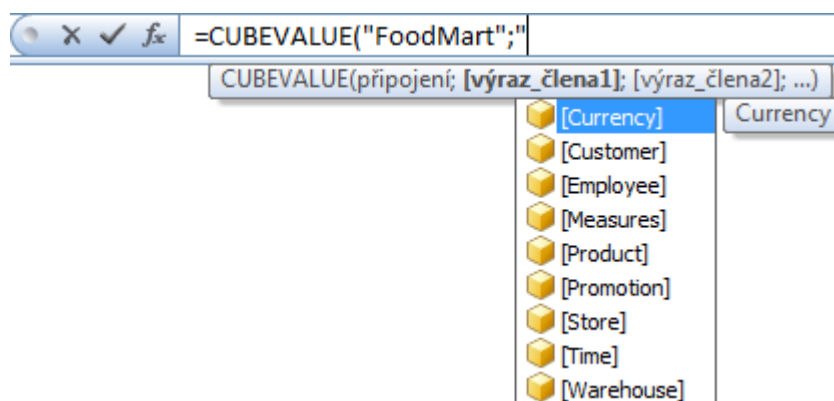
Stores	All			
Profit	Popisky sloupců			
Popisky řádků	Drink	Food	Non-Consumable	Celkový součet
1997	29358,9754	245764,8665	64487,0545	339610,8964
Q1	6964,2966	60814,4714	16097,3415	83876,1095
January	2261,6937	19827,935	5271,5687	27361,1974
February	2191,2887	19289,4004	4978,4321	26459,1212
March	2511,3142	21697,136	5847,3407	30055,7909
Q2	7186,1098	57323,3736	15192,5618	79702,0452
Q3	7203,3445	61262,5473	15901,1288	84367,0206
Q4	8005,2245	66364,4742	17296,0224	91665,7211
1998	56243,491	466142,1581	124196,092	646581,7411
Celkový součet	85602,4664	711907,0246	188683,1465	986192,6375

Obrázek 3-43 Tabulka převedená na vzorce

Využití těchto funkcí nabízí všechny možnosti, které máme při práci s buňkou. Mezi hlavní výhody použití funkcí CUBE patří:

- Větší možnosti při změně uspořádání dat v listu – buňky lze libovolně přesouvat a zobrazená data budou vždy aktuální
- Data z OLAP kostky lze použít v kombinaci s dalšími funkcemi, které Excel nabízí
- Pomocí těchto funkcí je možné zobrazovat data z více různých OLAP kostek v jednom listu a libovolně je seskupovat

Velkou pomocí při používání CUBE funkcí je automatické doplňování vzorců v Excelu. Jelikož k navigaci v dimenzích a faktech je v CUBE funkcích potřeba použít jména dimenzí a faktů tak, jak jsou definována v OLAP kostce, bylo by jejich zapisování ručně velmi složité. Automatické doplňování vzorce po vyplnění pole připojení nabízí list pro výběr dimenzí (pole připojení obsahují všechny funkce pro práci s OLAP krychlí, je to textový řetězec názvu připojení ke krychli).



Obrázek 3-44 Automatické doplňování vzorce

Základní funkce pro práci s kostkou je **CUBEVALUE**, která slouží k zobrazení agregované hodnoty podle zadaných dimenzí a faktů. Například funkce `=CUBEVALUE("Foodmart"; "[Measures].[Profit]"; "[Time].[Time].[All].[1998].[Q1]")` vrátí agregovanou hodnotu zisk za 1. čtvrtletí roku 1998. V této funkci můžeme použít libovolné množství dimenzí, jednou podmínkou je, že musí být použit pouze jeden fakt a každá dimenze pouze jednou.

Funkce ze skupiny CUBE, se od ostatních funkcí Excelu liší v tom, že se každá funkce dá použít jako odkaz na dimenzi, která je definovaná v argumentech funkce. K zobrazení stejné hodnoty jako v předchozím vzorci bychom tak mohli dojít např. použitím funkce **CUBEMEMBER**, která je určena ke kontrole, jestli dané dimenze v kostce existují a k odkazování na ně. Funkce CUBEVALUE pak může pouze odkazovat na buňky, ve kterých jsou popsány jednotlivé dimenze:

	A
1	=CUBEMEMBER("FoodMart"; "[Measures].[Profit]"; "Zisk")
2	=CUBEMEMBER("FoodMart"; "[Time].[Time].[All].[1998].[Q1]"; "1. Čtvrtletí 1998")
3	
4	=CUBEVALUE("FoodMart"; A1; A2)

Obrázek 3-45 Použití funkce CUBEMEMBER

Funkce **ZÍSKATKONTDATA** může sloužit jako alternativa pro funkci CUBEVALUE. Tato funkce obecně slouží k načtení souhrnné hodnoty z kontingenční tabulky. Při použití s naší databází by například funkce, která zobrazuje celkovou hodnotu zisku za rok 1997, vypadala následovně:

`=ZÍSKATKONTDATA("[Measures].[Profit]"; A1; "[Time].[Time]"; "[Time].[Time].[Year].[1997]")`

První argument funkce označuje název hodnoty, kterou chceme zobrazit. Druhý argument pak odkazuje na kontingenční tabulku (může odkazovat na jednu buňku nebo oblast v tabulce). Po těchto dvou argumentech pak následuje libovolný počet dvojic (může jich být až 126), ve kterých vždy první položka označuje dimenzi, kterou chceme zobrazit, a druhá položka označuje konkrétní úroveň této dimenze, pro kterou chceme zobrazit data. Pokud bychom chtěli zobrazit například hodnotu zisku za rok 1997 v pobočkách v USA za skupinu produktů nápoje, obsahovala by funkce tři dvojice, ze kterých by každá dvojice označovala jednu dimenzi a její úroveň (pro časovou dimenzi, dimenzi pobočky a dimenzi produktů).

Použití této funkce při psaní argumentů ručně by bylo dosti složité, protože argumenty funkce, které popisují dimenze, je potřeba uvádět v jazyce MDX (a navíc při jejich zadávání nefunguje automatické doplňování jako u CUBE funkcí), nicméně se tato funkce dá použít, pokud si chceme rychle vytvořit odkaz na hodnotu z kontingenční tabulky.

Vzorec totiž můžeme rychle zadat vložením znaménka = do buňky, ve které chceme hodnotu zobrazit, a poté klikneme na hodnotu v kontingenční tabulce, na kterou chceme odkazovat. Do buňky se pak automaticky doplní funkce ZÍSKATKONTDATA s odkazem na tuto hodnotu.[5]

Použití dalších funkcí pro práci s kostkou si ukážeme na složitějším příkladu tabulky, která bude zobrazovat seznam prvních pěti poboček, které měly v roce 1998 největší zisk:

	A	B	C	D
1		Zisk		
2		1998		
3				
4		Obchody s největším ziskem	Město	Zisk v roce 1998
5	1	Store 8	Merida	47403,42
6	2	Store 12	Hidalgo	47276,41
7	3	Store 19	Vancouver	46666,79
8	4	Store 17	Tacoma	45114,40
9	5	Store 13	Salem	44931,36

Obrázek 3-46 Tabulka 5 neúspěšnějších poboček

Po zobrazení vzorců (pomocí klávesové zkratky CTRL + `) tabulka vypadá takto:

	A	B	C	D
1		=CUBEMEMBER("FoodMart";"[Measures].[Profit]")		
2		=CUBEMEMBER("FoodMart";"[Time].[Time].[All].[1998]")		
3				
4		=CUBESET("FoodMart";"[Store].[Store].[All].children"; "Obchody s největším ziskem";2;D4)	Město	=CUBEMEMBER("FoodMart";B1:B2; "Zisk v roce 1998")
5	1	=CUBERANKEDMEMBER("FoodMart";\$B\$4:\$A5)	=CUBEMEMBERPROPERTY("FoodMart";\$B5;"Store City")	=CUBEVALUE("FoodMart";\$B5;D\$4)
6	2	=CUBERANKEDMEMBER("FoodMart";\$B\$4:\$A6)	=CUBEMEMBERPROPERTY("FoodMart";\$B6;"Store City")	=CUBEVALUE("FoodMart";\$B6;D\$4)
7	3	=CUBERANKEDMEMBER("FoodMart";\$B\$4:\$A7)	=CUBEMEMBERPROPERTY("FoodMart";\$B7;"Store City")	=CUBEVALUE("FoodMart";\$B7;D\$4)
8	4	=CUBERANKEDMEMBER("FoodMart";\$B\$4:\$A8)	=CUBEMEMBERPROPERTY("FoodMart";\$B8;"Store City")	=CUBEVALUE("FoodMart";\$B8;D\$4)
9	5	=CUBERANKEDMEMBER("FoodMart";\$B\$4:\$A9)	=CUBEMEMBERPROPERTY("FoodMart";\$B9;"Store City")	=CUBEVALUE("FoodMart";\$B9;D\$4)

Obrázek 3-47 Tabulka se zobrazenými vzorci

K řazení položek v množině slouží funkce **CUBESET**, v našem případě jsme použili v buňce B4 CUBESET, který obsahuje množinu všech poboček seřazených sestupně podle zisku v roce 1998 (řazení je řešeno odkazem na funkci CUBEMEMBER v buňce D4).

Výraz, který se v těchto funkcích používá pro definování dimenze, je v jazyce MDX³, a proto je možné používat další možnosti tohoto dotazovacího jazyka. V této tabulce je použit výraz **[Store].[Store].[All].children**, kde výraz **children** znamená, že v množině budou zahrnuty všechny podúrovně této dimenze.

Pro získání určitého člena ze seřazené množiny pak slouží funkce **CUBERANKEDMEMBER**, pomocí které zobrazujeme jednotlivé pobočky ve sloupci B. K určení, jakou položku v pořadí chceme zobrazit, jsou využita čísla ve sloupci A.

Další funkcí, kterou jsme v tabulce použili je funkce **CUBEMEMBERPROPERTY**, která slouží k zobrazení vlastností člena krychle. V našem případě jsme pro každou pobočku v seznamu získali vlastnost Store City, tedy název města, ve kterém se pobočka nachází.

Pomocí funkcí pro práci s kostkou můžeme vytvářet přehledné reporty, které budou vždy zobrazovat aktuální hodnoty z databáze. Hlavní výhodou při použití těchto funkcí je možnost vlastního rozvržení buněk v listu a větší volnost ve způsobech zobrazování dat. Nevýhodou je však ztráta možnosti jednoduše a rychle měnit pohled na data (jako to umožňuje kontingenční tabulka). Tyto funkce se tedy hodí spíše pro vytváření statických pohledů na data (například listů, ve kterých se budou vždy zobrazovat souhrnná data o důležitých údajích z databáze) než k hledání souvislostí v datech.

³ MDX (MultiDimensional eXpressions) je dotazovací jazyk pro OLAP databáze (podobně, jako je SQL dotazovací jazyk pro relační databáze). Vysvětlení dalších funkcí tohoto jazyka je dostupné ve zdrojích [8] a [9]

3.7. Omezení Excelu při práci s OLAP kostkou

Hlavním nedostatkem Excelu jsou omezené možnosti, pomocí kterých můžeme zobrazit data v kontingenční tabulce při práci s OLAP kostkou. V podstatě jedinou možností je zobrazení dat jako procentuální podíl z řádku, sloupce nebo celé tabulky. V normální kontingenční tabulce (té, která má jako zdroj dat tabulku v sešitu Excelu) jsou tyto možnosti daleko větší, data lze zobrazit například jako průměrné hodnoty, lze zobrazit minimální a maximální hodnoty ze skupin, apod. To, že tyto možnosti při práci s OLAP kostkou není možné využít, je ale logické - OLAP databáze obvykle obsahuje daleko větší množství dat, a výpočet, který by Excel musel provést, by trval moc dlouho.

Dalším omezením je následující - pokud hodnotu v kontingenční tabulce zobrazíme jako % řádku, % sloupce anebo % celku, nemůžeme k těmto relativním hodnotám zároveň zobrazit i hodnoty absolutní – to je další možnost, která je dostupná při práci s normální kontingenční tabulkou. Uživatel tak ztrácí kontinuitu s tím, co vlastně relativní hodnota vyjádřená v % představuje. Úplnější obrázek o datech by dala určitě tabulka, která obsahuje jak absolutní, tak i relativní čísla.

Další možností, která při práci s OLAP kostkou chybí, je zobrazení podrobností – pokud v normální tabulce klikneme pravým tlačítkem na jakoukoliv hodnotu v kontingenční tabulce, můžeme zvolit položku **Zobrazit podrobnosti**, pomocí které se pak zobrazí přehledná tabulka, která obsahuje podrobnosti o datech, ze kterých je tato hodnota složena (například pokud bychom zobrazili podrobnosti pro hodnotu, která zobrazuje celkový součet za jeden měsíc, zobrazila by se tabulka s jednotlivými dny a podrobnostmi o nich).

4. Závěr

Tato bakalářská práce poskytuje jak základní teoretické informace o principech, na kterých fungují OLAP databáze, tak rovněž návod na praktické užití programu Microsoft Office Excel 2007 pro analýzu dat z těchto databází. V práci jsem se pokusil popsat všechny funkce, které by uživatel mohl při práci s OLAP databází využít.

Práce pomáhá vytvořit čtenáři ucelenou představu o možnostech práce s OLAP databázemi v aplikaci Excel. Přehledně s pomocí řady obrázků provede čtenáře jednotlivými okruhy analýzy dat z OLAP databáze už od procesu připojení k databázi až k publikaci výsledků. Díky názorným ukázkám se čtenář snadno orientuje v aplikaci a je schopen sám provádět základní operace s OLAP databázemi. Proto by tato práce mohla být užitečná pro čtenáře, kteří s problematikou OLAP databází ještě nejsou seznámeni a chtěli by Excel pro účel analýzy dat z těchto databází použít.

Vzhledem k tomu, že jsem dosud nepracoval s žádnou další aplikací, která slouží pro analýzu dat z OLAP databází, nemůžu popsat výhody a nevýhody použití Excelu v plné míře, protože nemám s čím srovnávat. Nicméně jsem se v závěrečné kapitole praktické části pokusil shrnout omezení, na která jsem narazil při srovnání práce s OLAP databází a normální kontingenční tabulkou.

Po napsání této práce vidím hlavní nedostatek Excelu v malém množství možností pro změnu zobrazení číselných dat (v podstatě jedinou možností kromě základního zobrazení hodnot ze serveru je vyjádření dat v podobně procentuálních podílů vztažených k součtovým hodnotám řádků, sloupců anebo celku).

Funkce CUBE, které jsou novinkou ve verzi 2007, pak jsou podle mě vhodné zejména k vytváření statických zpráv, což považuji spíše za doplňkovou funkci a je otázkou, zda by místo nich nebylo lepší využít raději reportovací služby, které nabízí snad každý analytický server, na kterém se dají provozovat OLAP databáze.

Na tuto práci by mohly navazovat další práce, které by se zabývaly rozšířeními, které pro práci s OLAP databázemi byly v Excelu vytvořeny (většinou se jedná o rozšíření, která jsou dodávána jako součást celého řešení Business Intelligence). Dalším námětem by pak

mohlo být porovnání funkcí, které nabízí Excel ve verzi 2007 s funkcemi, které přináší nová verze Excel 2010 (ten je v prodeji od 15. června 2010[10]).

5. Použité zdroje

- [1] LACKO, Luboslav. *Business Intelligence v SQL Serveru 2008*. Praha: Computer press, a.s., 2009. ISBN 978-80-251-2887-9.
- [2] GÁLA, Libor; POUR, Jan; TOMAN, Prokop. *Podniková informatika*. Praha: Grada publishing, a.s., 2006. ISBN 80-247-1278-4.
- [3] NOVOTNÝ, Ota; POUR, Jan; SLÁNSKÝ, David. *Business intelligence : jak využít bohatství ve vašich datech*. Praha: Grada Publishing, a.s., 2005. ISBN 80-247-1094-3.
- [4] LACKO, Luboslav. *Databáze: datové sklady, OLAP a dolování dat s příklady v Microsoft SQL Serveru a Oracle*. Brno: Computer Press, 2003. ISBN 80-7226-969-0.
- [5] MS Excel 2007, Help
- [6] ZAPAWA, Timothy. *Microsoft Excel: Získávání, analýza a prezentace dat*. Brno: Computer press, a.s., 2007. ISBN 978-80-251-1535-0.
- [7] GAINER, David. *Microsoft Excel 2010 : The official blog of the Microsoft Excel product team* [online]. 2006 [cit. 2010-06-21]. CUBE Functions 2: A Few Examples. Dostupné z WWW: <<http://blogs.msdn.com/b/excel/archive/2006/02/06/cube-functions-2-a-few-examples.aspx>>.
- [8] *MSDN Library* [online]. 2010 [cit. 2010-06-21]. MDX Overview. Dostupné z WWW: <<http://msdn.microsoft.com/en-us/library/aa216775%28SQL.80%29.aspx>>.
- [9] *Arena Retail* [online]. 2004 [cit. 2010-06-21]. MDX Tutorial. Dostupné z WWW: <http://www.anselmiconsulting.com/webtools/wtcontent/acdr/tutorials/mdxreporttutorial/mdxtutorial/mdx_tutorial.htm>.
- [10] *Microsoft News Center* [online]. 2010-06-15 [cit. 2010-06-25]. Microsoft Office 2010 Now Available for Consumers Worldwide. Dostupné z WWW: <<http://www.microsoft.com/Presspass/press/2010/jun10/06-152010OfficeLaunchPR.mspx>>.

6. Terminologický slovník

Agregovaná data	Vypočítaná a uložená data pro vyšší prvky hierarchie dimenzí. Značně urychlují přístup k těmto údajům.
Business Intelligence (BI)	Sada procesů, aplikací a technologií, jejichž cílem je účinně a účelně podporovat rozhodovací procesy ve firmě. Podporují analytické a plánovací činnosti podniku a organizací, a jsou postaveny na principu multidimenzionality.
Dimenze	Dimenze dávají kontext faktům. Obsahují většinou textové atributy popisující fakta.
Dotazovací jazyk	Umožňuje uživateli formulovat dotazy, či získávat informace z databáze pomocí požadavku podle definovaných standardů.
Fakt	Metrika (zpravidla číselná hodnota), která je v podniku sledována.
Fakt tabulka	Tabulka obsahující obchodní fakta a hodnoty, zpravidla číselné.
Kostka	Označení pro multidimenzionální databázi (nebo její princip). Kostka je realizována buď prostřednictvím multidimenzionální databáze, nebo zvláštní struktury relační databáze.
Multidimenzionální databáze	Databáze, kde jsou data uložena na principu vícerozměrné matice. Hodnoty jsou přístupné přímo pro danou kombinaci prvků dimenzí.
OLAP	OnLine Analytical Processing. Provádění analýzy dat v reálném čase pomocí technologií, umožňující konzistentní přístup k informacím datového skladu.
OLAP server	Technologie pro uložení a správu dat v multidimenzionálních databázích, pro jejich vytváření a plnění obsahující metody přístupu k datům.
Relační databáze	Typ databáze ukládající data v relacích. Každý vztah je složen ze

	záznamů a polí.
Report	Výstup ze software s výsledky hodnot ukazatelů. Může být vyjádřen tabulkou, grafem atd.
Reportovací služba	Software na straně serveru, který slouží pro zpřístupnění dat a výsledků analýz uživatelům ve vhodné formě, obsahu a rozsahu.