



Ekonomická
fakulta
Faculty
of Economics

Jihočeská univerzita
v Českých Budějovicích
University of South Bohemia
in České Budějovice

Jihočeská univerzita v Českých Budějovicích
Ekonomická fakulta
Katedra aplikované matematiky a informatiky

Diplomová práce

Sbírka řešených příkladů pro výuku Business Intelligence pro studenty EF

Vypracovala: Bc. Miroslava Kozubková
Vedoucí práce: doc. Ing. Ladislav Beránek CSc.

České Budějovice 2016

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Miroslava KOZUBKOVÁ**
Osobní číslo: **E14733**
Studijní program: **N6209 Systémové inženýrství a informatika**
Studijní obor: **Ekonomická informatika**
Název tématu: **Sbírka řešených příkladů pro výuku Business Intelligence pro studenty EF**
Zadávající katedra: **Katedra aplikované matematiky a informatiky**

Z á s a d y p r o v y p r a c o v á n í :

Cílem práce je vytvořit sbírku řešených příkladů pro výuku předmětu Business Intelligence pro studenty EF v databázovém prostředí MS SQL server. V rámci práce budou navrženy jednotlivé úlohy odpovídající tématům probíraným v předmětu Business Intelligence. Bude se jednat o návrh datového modelu, ukázka rozdílu standardního normalizovaného datového modelu a datového modelu vhodného pro BI. Dále bude ukázána implementace tohoto modelu a dotazy nad tímto modelem, rovněž i možnosti reportování.

Metodický postup:

1. Studium odborné literatury, studium možností vývojového prostředí MS SQL.
2. Analýza jednotlivých úloh.
3. Návrh řešení, dokumentace k jednotlivým úlohám.
4. Zhodnocení, vypracování doporučení a závěrů.

Rozsah grafických prací: **dle potřeby**
Rozsah pracovní zprávy: **50 - 60 stran**
Forma zpracování diplomové práce: **tištěná**
Seznam odborné literatury:


1. **LABERGE, Robert.** *Datové sklady: agilní metody a business intelligence.* 1. vyd. Brno: Computer Press, 2012, 350 s. ISBN 978-80-251-3729-1.
2. **LACKO, L'uboslav.** *Business Intelligence v SQL Serveru 2008: reportovací, analytické a další datové služby.* Vyd. 1. Brno: Computer Press, 2009, 456 s. ISBN 978-80-251-2887-9.
3. **LARSON, Brian.** *Delivering business intelligence with Microsoft SQL server 2012.* 3rd ed. New York: McGraw-Hill, c2012, xviii, 813 p. ISBN 0071759387.
4. **POUR, Jan, Miloš MARYŠKA a Ota NOVOTNÝ.** *Business intelligence v podnikové praxi.* 1. vyd. Praha: Professional Publishing, 2012, 276 s. ISBN 978-80-7431-065-2.

Vedoucí diplomové práce: **doc. Ing. Ladislav Beránek, CSc.**
Katedra aplikované matematiky a informatiky

Datum zadání diplomové práce: **9. ledna 2015**
Termín odevzdání diplomové práce: **15. dubna 2016**


doc. Ing. Ladislav Rolínek, Ph.D.
děkan

JIHOČESKÁ UNIVERZITA
V ČESKÝCH BUDĚJOVICÍCH
EKONOMICKÁ FAKULTA
Studentská 13 (26)
370 05 České Budějovice


prof. RNDr. Pavel Tlustý, CSc.
vedoucí katedry

V Českých Budějovicích dne 30. března 2015

Prohlášení

Prohlašuji, že jsem diplomovou práci na téma Sbíрка řešených příkladů pro výuku Business Intelligence pro studenty EF vypracovala samostatně na základě vlastních zjištění s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že svoji diplomovou práci jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury. Prohlašuji, že v souladu s § 47 zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své diplomové práce, a to - v nezkrácené podobě/v úpravě vzniklé vypuštěním vyznačených částí archivovaných Ekonomickou fakultou - elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích dne 13. 4. 2016

Miroslava Kozubková

Poděkování

Chtěla bych poděkovat vedoucímu práce panu doc. Ing. Ladislavu Beránkovi CSc. za odborné vedení této bakalářské práce.

Obsah

1. Úvod.....	4
2. Cíl práce.....	5
3. Literární rešerše	6
3.1. Databázový systém	6
3.1.1. Relační databázový systém.....	7
3.1.2. Normalizace.....	7
3.1.3. Multidimenzionální databáze	8
3.2. Data.....	8
3.2.1. Metadata	9
3.2.2. Tabulka faktů a dimenze	9
3.2.3. Základní datové struktury	10
Sněhová vločka (Snowflake)	10
Hvězda (Star)	11
3.3. Business Intelligence	12
3.4. Komponenty Business Intelligence	13
3.4.1. OLTP (online transaction processing).....	14
3.4.2. ETL.....	14
3.4.3. EAI (Enterprise Application Integration).....	15
3.4.4. Datový sklad	15
3.4.5. Datový trh.....	16
3.4.6. OLAP.....	16
Rozdíly mezi analytickými systémy a OLTP	17
3.4.7. Dočasné úložiště dat (DSA)	17
3.4.8. Operativní úložiště dat (ODS).....	18
3.4.9. Reporting	18
3.4.10. Data mining.....	19

3.4.11.	Nástroje pro koncové uživatele.....	20
3.5.	Přístupy k řešení BI.....	21
3.5.1.	Přístup shora dolů	21
3.5.2.	Přístup zdola nahoru	22
3.6.	SW řešení pro Business Intelligence	23
4.	Praktická část	24
4.1.	Datové modely	24
4.1.1.	Normalizovaný relační datový model	24
4.1.2.	Rozdíly mezi datovým modelem hvězdy a sněhové vločky	28
4.2.	SQL Server Data Tools.....	30
4.2.1.	ETL.....	30
1.	Založení datového skladu a souboru se zdrojovými daty.....	30
2.	Transformace dat v SQL Server Data Tools	31
3.	Rozdělení.....	37
4.2.2.	Multidimenzionální kostka.....	38
4.2.3.	Reportování	42
Tvorba reportu bez průvodce	42	
Tvorba reportu s průvodcem.....	46	
Úprava reportu	48	
Přidání možnosti seřazení dat podle vybraného sloupce	48	
Podmíněné formátování.....	49	
4.3.	Email.....	51
4.3.1.	Odeslání emailu	52
4.3.2.	Seznam odeslaných emailů.....	52
5.	Závěr	54
	Summary.....	55
	Zdroje.....	56

Seznam obrázků.....	59
Seznam tabulek.....	60

1. Úvod

Informační systémy představují nedílnou součást většiny středních a velkých společností a prosazují se i v menších podnicích. Jejich role stále roste i v oblasti rozhodování. Nástroje Business Intelligence představují prostředky, které pomáhají hledat souvislosti v podnikových datech a predikovat budoucí vývoj.

Vzhledem ke stále vyššímu využívání informačních systémů a rozšiřujících technologií jako je např. právě Business Intelligence, je důležité, aby studenti Ekonomické fakulty byli s touto problematikou obeznámeni a aby byli schopni vypracovat alespoň základní úlohy.

2. Cíl práce

Cílem této diplomové práce je vytvořit sbírku řešených příkladů pro výuku předmětu Business Intelligence pro studenty Ekonomické fakulty Jihočeské Univerzity. Obsahem teoretické části bude popis jednotlivých komponent Business Intelligence. V praktické části budou zpracovány následující úlohy: bude navržen vhodný datový model, budou představeny a popsány základní úkony, které nabízí prostředí MS SQL Serveru a nástroje MS Server Data Tools, který je na fakultě využíván k výuce. Bude se jednat o úlohy ETL, základní práce s multidimenzionální kostkou a vytváření a úpravu reportů.

3. Literární řešerše

3.1. Databázový systém

Databázový systém je tvořen databází a systémem řízení báze dat (SŘBD). Databáze je množina vzájemně propojených dat. K těmto datům lze přistupovat s pomocí softwarového vybavení, které představuje SŘBD. (Gála, L., Pour, J. a Šedivá, Z. (2009))

„Databáze je kolekce strukturovaných informací. Databáze jsou specificky navrženy pro správu velkého objemu informací, přičemž data uchovávají uspořádaným a strukturovaným způsobem, který uživatelům usnadňuje jejich správu a získávání.“ (Agarwal V. V. a Huddleston, 2009, p. 45)

„Základními prvky databáze jsou data a program pro práci s nimi. Datový obsah tvoří množina jednotně strukturovaných dat uložených v paměti počítače nebo na záznamovém médiu, jež jsou navzájem v určitém vztahu a tvoří určitý celek z hlediska obsažených informací; data jsou přístupná výhradně pomocí speciálního programového vybavení – systému řízení báze dat. Na rozdíl od jiných souborů (např. texty, obrázky) uživatel nepracuje s databází jako s celkem, ale vybírá z ní data prostřednictvím dotazů.“ (Kučerová, 2004, p. 4)

V historii byly využívány rozličné typy databázových systémů. Již před příchodem informačních technologií byly využívány kartotéky a děrné štítky pro zavedení určitého systému do ukládání dat.

Mezi nejvýznamnější typy databázových systémů patří:

- hierarchické,
- síťové,
- relační,
- objektově orientované,
- objektově-relační.

Nejedná se o kompletní výčet všech existujících databázových systémů. I v současné době jsou vyvíjeny nové technologie, zejména kvůli nárokům na zpracování velkých dat

(Big Data) a zvyšujícímu se počtu uživatelů, kteří potřebují přístup k datům. Jako podnikové databáze se začínají využívat, např. NoSQL databáze či cloud databáze. (Pokorný, 2012)

3.1.1. Relační databázový systém

Relační databázový systém představuje systém, který ukládá data v tabulkách, které spolu souvisejí. Tento systém je založen na relačním modelu. V relačním modelu jsou uchovávány kromě dat i vztahy.

Agarwal a Huddleston (2009) uvádějí následující výhody použití tohoto systému:

- zamezení redundance dat,
- zamezení nekonzistence dat,
- integrita dat,
- atomičnost dat,
- zabránění aktualizace dat více uživateli ve stejném okamžiku,
- bezpečnost dat,
- zabezpečení zpracování transakcí,
- zotavení dat po selhání transakce,
- správa datového úložiště.

Relační databáze mají však i své nevýhody (Lacko L., 2003):

- *„absence komplexních analytických nástrojů,*
- *potenciální omezení objemu údajů, ke kterým je možné v rozumném čase přistoupit.“*

3.1.2. Normalizace

„Normalizace je proces uspořádání dat v databázi. Patří sem vytváření tabulek a vztahů mezi nimi podle pravidel navržených za účelem ochrany dat a za účelem vytvoření pružnější databáze odstraněním redundance a nekonzistentní závislosti.“ (Základy normalizace databáze)

Existuje několik normálních forem. V praxi se většinou používá první, druhá a třetí.

- „První normální forma (1NF) říká, že všechny atributy jsou atomické.
- Druhá normální forma (2NF) – tabulka splňuje 2NF, právě když splňuje 1NF a navíc každý atribut, který není primárním klíčem, je na primárním klíči úplně závislý.
- Třetí normální forma (3NF) – Relační tabulky splňují třetí normální formu (3NF), jestliže splňují 2NF a žádný atribut, který není primárním klíčem, není tranzitivně závislý na žádném klíči.“ (Skřivan J., 2000)

Normalizace má i své nevýhody. Dochází ke zvýšení počtu vytvořených tabulek a tím i ke zvýšení počtu jejich spojení. Tímto může dojít ke zpomalení vykonávání dotazů nad danými tabulkami. Aby se předešlo přílišné složitosti dotazů, bývají některé tabulky úmyslně denormalizovány. (Agarwal V. V. a Huddleston, 2009)

3.1.3. Multidimenzionální databáze

Multidimenzionální databáze jsou úložištěm sumarizovaných a agregovaných dat. Tyto databáze odstraňují dva výše uvedené nedostatky relačních databází:

- dovolují provádět komplexní analýzy,
- umožňují zpracovávat v kratším čase velké objemy dat.

Další výhodou multidimenzionálních databází je umožnění predikcí. Za nevýhodu lze označit zvýšené požadavky na kapacitu úložiště dat. Pro analytické databáze se používá i označení OLAP (Online Analytical Processing).

(Lacko L., 2003)

3.2. Data

Data tvoří důležitou část podnikových aktiv a pro nástroje business intelligence jsou klíčová. Aby však byla využívána efektivně a pomáhala jejich uživatelům v rozhodovacím procesu, je důležité dbát o jejich kvalitu. Prvků ovlivňujících kvalitu dat je vícero:

- Terminologie – je potřeba, aby všechny komponenty měly jednotné názvy a aby byly stejně chápány všemi odděleními společnosti.
- Struktura dat – Důležitým prvkem ovlivňujícím kvalitu dat a závislejícím na správně definované terminologii je struktura dat. Na začátku je potřeba se seznámit s cílem projektu business intelligence. Pro dosažení tohoto cíle bývá nutné přistoupit k určení dílčích cílů, které budou postupně navrženy a implementovány.

(Laberge, 2012)

3.2.1. Metadata

Metadata představují data o datech. Pro každý atribut datového skladu označují zdroj dat a jejich význam. Údaje uchovávané v metadatach by měly být neustále aktualizovány, aby byla zachycena jakákoliv modifikace ve struktuře datového skladu. Tyto údaje by měly být uživatelům datového skladu přímo přístupné. (Vercellis, 2009)

Humphries (2001) uvádí jako příklady dat uložených v metadatach:

- název atributu, pod kterým je uložen, a název, který uvidí uživatel,
- zdroj dat,
- informace, zda se jedná o tabulku faktů či dimenzi,
- apod.

3.2.2. Tabulka faktů a dimenze

„Tabulka faktů je centrální tabulka ve hvězdicovém schématu, která obsahuje podnikové metriky kvalifikované podle dimenzí. Datový trh s hvězdicovým schématem obvykle obsahuje jednu centrální tabulku faktů a několik dimenzí.“ (Laberge, 2012, p. 155)

„Dimenze je tabulka, která kvalifikuje metriky v tabulce faktů. ... Tabulky dimenzí obsahují sloupce ve smyslu fyzické databáze.“ (Laberge, 2012, p. 163) Rozlišují se tři typy dimenzí:

- Typ I – historie se neuchovává,
- Typ II – uchovává se úplná historie,

- Typ III – uchovává se aktuální a první nebo předchozí hodnota.

(Laberge, 2012)

Tabulky dimenzí většinou odpovídají primárním entitám uloženým v datovém skladu a ve většině případů jsou odvozeny z tabulek v OLTP systémech, např. zákazníci, prodej, čas atd. Jednotlivé tabulky bývají strukturovány podle předem stanovené hierarchie. (Vercellis, 2009)

3.2.3. Základní datové struktury

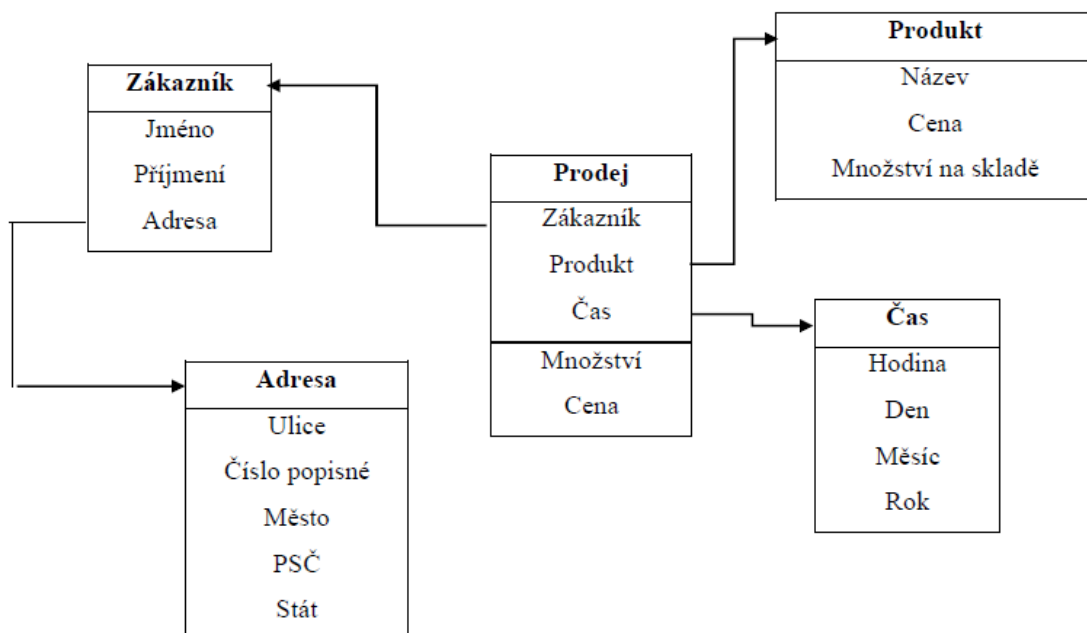
Pro zobrazení vztahů mezi jednotlivými tabulkami se používá zejména model hvězdy a sněhové vločky. Existují ale i další modely, např. galaxie.

Sněhová vločka (Snowflake)

V centru schématu je tabulka faktů, která je spojena s tabulkami dimenzí. Tabulky jsou normalizované, takže dochází k větvení, neboť na tabulky dimenzí jsou napojeny další tabulky. V důsledku vysokého množství operací spojování (join) je zpracování dotazů pomalejší.

(Gála, L., Pour, J. a Šedivá, Z. (2009))

Obrázek 1 Datový model Sněhová vločka



Zdroj: autorka dle Vercellis (2009), p. 56

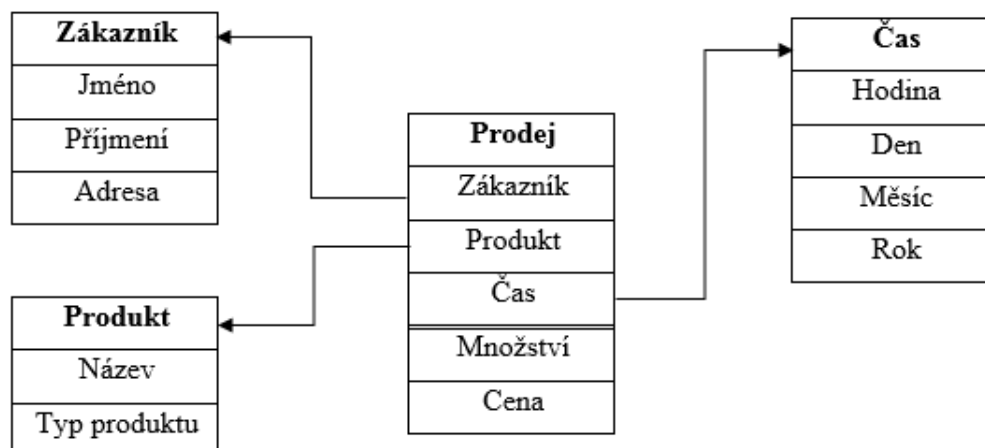
Obrázek 1 znázorňuje zjednodušené schéma sněhové vločky. Uprostřed je umístěna tabulka faktů, která je spojena s tabulkami dimenzí. Poslední dva řádky v tabulce faktů, které jsou odděleny dvojitou čarou a nejsou spojeny s žádnou tabulkou dimenzí, jsou metriky, které by sloužily k výpočtu celkových prodejů. V modelu došlo k větvení, neboť byla adresa zákazníka přesunuta do vlastní tabulky. Tato nová tabulka nezávisí na tabulce faktů, ale na dimenzi Zákazník.

Hvězda (Star)

Tento datový model je speciálním tvarem modelu sněhové vločky. Ve středu hvězdy je tabulka faktů, se kterou jsou spojeny tabulky dimenzí. Výhodou tohoto modelu je rychlejší zpracování dotazů, neboť zde nejsou využívány operace spojování. Data v tomto modelu nemusí být normalizovaná.

(Gála, L., Pour, J. a Šedivá, Z. (2009))

Obrázek 2 Datový model Hvězda



Zdroj: autorka dle Vercellis (2009), p. 56

Obrázek 2 znázorňuje zjednodušené schéma hvězdy. Uprostřed je umístěna tabulka faktů, která je spojena s tabulkami dimenzí. Poslední dva řádky v tabulce faktů, které jsou odděleny dvojitou čarou a nejsou spojeny s žádnou tabulkou dimenzí, jsou metriky, které by sloužily k výpočtu celkových prodejů. Všechny tabulky dimenzí závisí pouze na tabulce faktů Prodej.

3.3. Business Intelligence

Pojem Business Intelligence můžeme definovat následovně:

"Business Intelligence (BI) představuje komplex přístupů a aplikací IS/ICT, které téměř výlučně podporují analytické a plánovací činnosti podniků a organizací a jsou postaveny na principu multidimenzionality, kterým zde rozumíme možnost pohlížet na realitu z několika možných úhlů." (Novotný, 2005, p. 17)

Vercellis (2009) definuje BI jako sadu matematických modelů a analytických metodik, které využívají data pro generování informací a znalostí použitelných pro komplexní rozhodovací procesy.

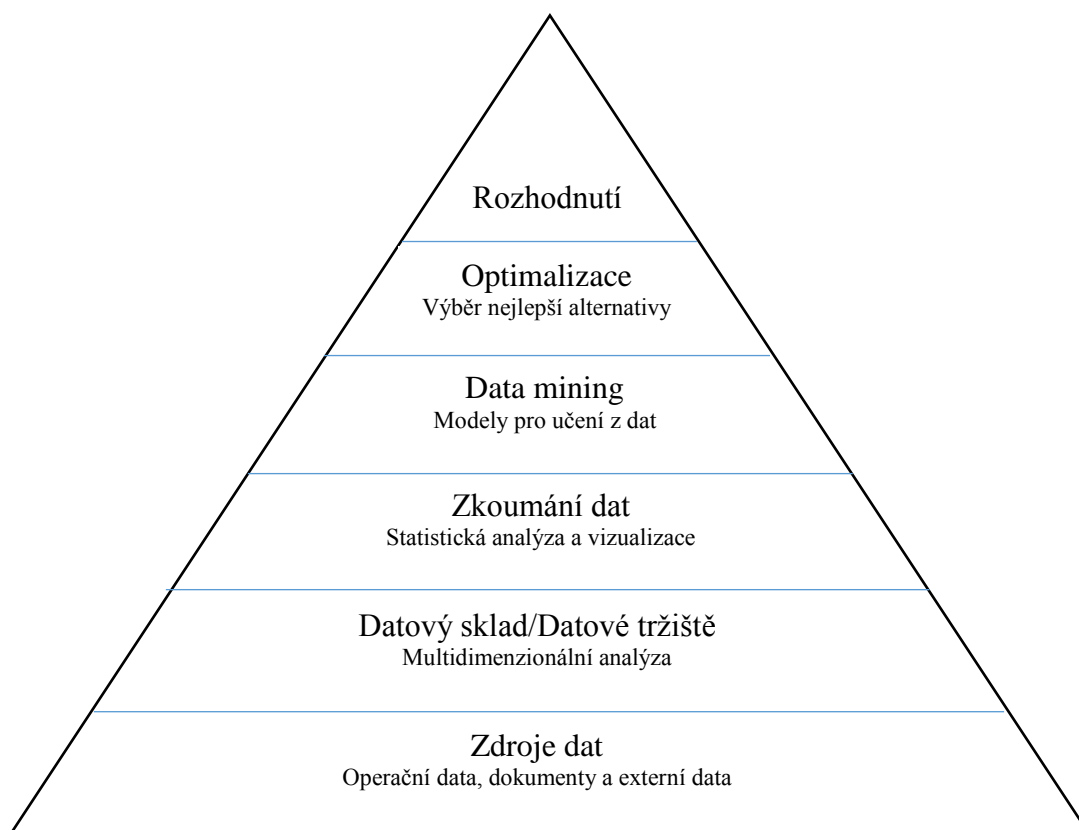
Jedná se o nástroje, které mají manažerům poskytovat odpovídající informace, které poslouží jako podklady pro důležitá rozhodnutí ovlivňující současné i budoucí fungování společnosti a umožní firmě si zachovat či zvýšit svou konkurenceschopnost.

Tyto nástroje by měly dodávat přehledné výstupy a umožňovat změnu kritérií, podle kterých jsou výstupy generovány. Kromě manažerů mohou tyto informace posloužit i jednotlivým oddělením podniku, např. oddělení analytickému či oddělení zákaznické podpory. (Novotný, 2005)

Laberge (2012, p. 29) uvádí jako oblasti využití business intelligence následující:

- *„Měření výkonu nebo stanovení základní úrovně,*
- *Analýza trendů a predikce,*
- *Sdružené seskupování neboli analýza tržního koše či segmentace,*
- *Řízení výkonu,*
- *Asociativní analýza neboli dolování dat,*
- *Analýza předmětných oblastí.“*

Obrázek 3 Části systému Business Intelligence



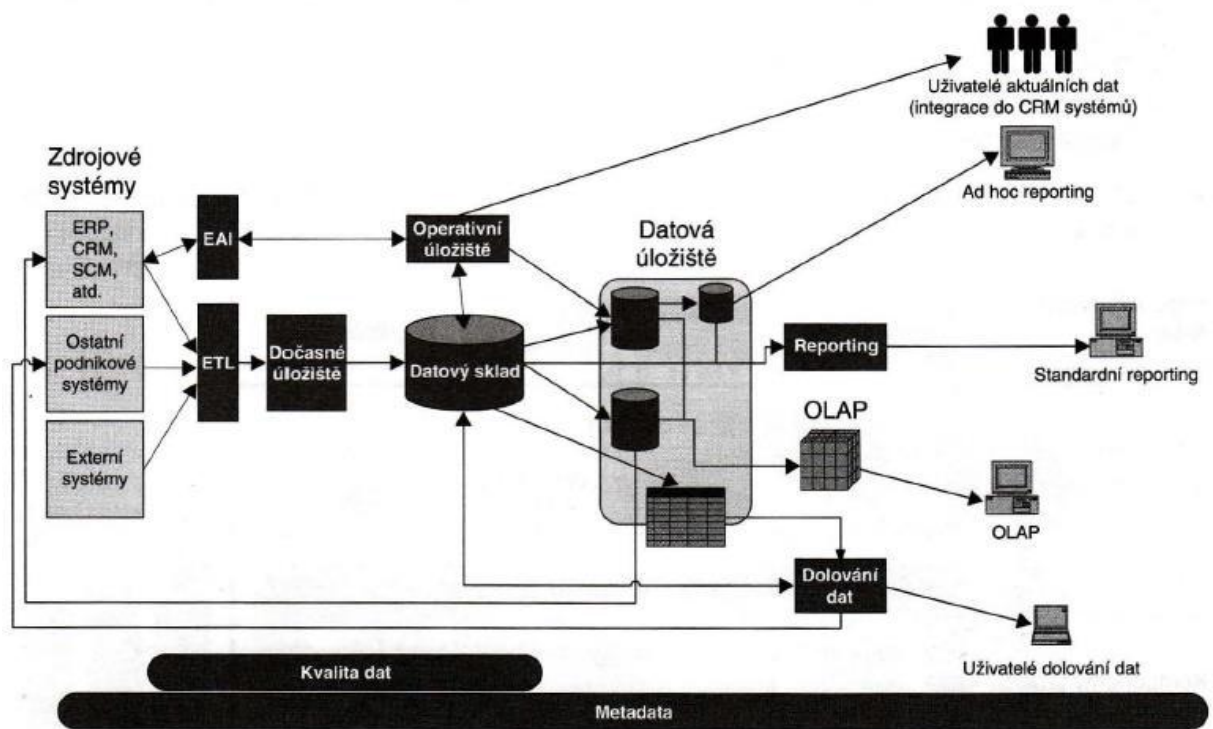
Zdroj: Vercellis (2009)

3.4. Komponenty Business Intelligence

Řešení Business Intelligence se skládá z několika různých vrstev (Novotný, 2005):

- zdrojové systémy – přestože nejsou součástí Business Intelligence, představují zdroj dat, která jsou následně zpracovávána jednotlivými komponentami BI,
- komponenty Business Intelligence:
 - komponenty datové transformace: ETL a EAI systémy,
 - databázové komponenty: datový sklad, datová tržiště, operativní úložiště, dočasné úložiště dat,
 - analytické komponenty: reporting, OLAP, Data Mining,
 - nástroje pro koncové uživatele.

Obrázek 4 Hlavní komponenty Business Intelligence a jejich vazby



Zdroj: (Novotný, 2005, p. 28)

3.4.1. OLTP (online transaction processing)

„OLTP systémy uchovávají záznamy o jednotlivých uskutečněných (typicky obchodních) transakcích a jsou obvykle realizovány pomocí dnes nejběžnější relační databázové technologie. Data uchovávaná v OLTP databázovém systému jsou (zpravidla periodicky) agregována (typicky sumarizována) a poté ukládána do datového skladu. Data se získávají z transakcí primárních informačních systémů. Jsou to veškeré pohyby dat, které tyto systémy zachytí do své relační databáze. Nejčastějším zdrojem to jsou ERP systémy, účetní a ekonomické informační systémy.“ (OldanyGroup – OLTP)

3.4.2. ETL

„Z obecného hlediska lze ETL popsat jako extrakci dat ze vstupního zdroje, transformaci těchto dat do příslušného formátu pro načítání a následné načítání dat do cílové databáze. ... Termín „transformace“ zahrnuje mnoho dílčích procesů a kroků, které závisí na požadavcích transformace. Může se jednat o některé z následujících

aspektů: čištění, slučování, třídění, definování jedinečných identifikátorů, zajištění pořizovacích časových razítek, zajištění datových razítek období platnosti, zpracování delta, vytváření dat, ověřování dat, zajištění referenční integrity, agregování, sumarizace a profilování dat.“

(Laberge, 2012, p. 242)

Během procesu transformace bývá potřeba vyřešit několik problémů, mezi něž patří (Lacko L., 2003):

- nejednoznačnost údajů – data označující stejné skutečnosti mohou být zapsána různým způsobem, př. pohlaví muž může být zapsáno: M, m, Muž, muž,
- duplicitní záznamy,
- chybějící hodnoty,
- různá terminologie,
- formát atributů – k potížím může dojít např. u čísel, která mohou být uložena jako číselný datový typ, ale i jako řetězec,
- chybějící časové údaje.

3.4.3. EAI (Enterprise Application Integration)

EAI představuje metody a nástroje využívané pro modernizaci, sjednocení a koordinaci aplikací používaných ve společnosti. Společnost již většinou využívá nějaké aplikace a databáze a zároveň chce začít používat i další aplikace či technologie. EAI umožňuje posoudit, jak existující prostředky zapadají do nového modelu a navrhnout jejich efektivní využití ve spojení s novými aplikacemi a daty. (EAI – Enterprise Application Integration)

3.4.4. Datový sklad

„Datový sklad je podnikově strukturovaný depozitář subjektivě orientovaných, integrovaných, časově proměnlivých, historických dat použitých na získávání informací a podporu rozhodování. V datovém skladu jsou uložena atomická a sumární data.“

(Lacko, L.,2003, p.48)

Data jsou do datového skladu získávána z provozních databází, jsou dále zpracována a následně uložena do datového skladu. (Lacko L., 2003)

3.4.5. Datový trh

Datový trh je část datového modelu, která je vytvořena pro konkrétní oblast dat. Modely datových trhů bývají znázorněny pomocí struktur sněhové vločky, hvězdy nebo jsou normalizovány do třetí normální formy. (Laberge, 2012)

3.4.6. OLAP

„OLAP je volně definovaný řád principů, které poskytují multidimenzionální rámec pro podporu rozhodování. ... Analýza OLAP slouží ke zpracování údajů uložených v datovém skladu do podoby pro koncové uživatele, tedy manažery a analytiku.“ (Lacko, 2003, p. 111)

Multidimenzionální databázové modely dělíme na:

- Multidimenzionální OLAP (MOLAP) – data jsou brána z operačních databází či z datového skladu, následně jsou získané údaje uloženy ve vlastních datových strukturách,
- Relační databázový OLAP (ROLAP) – data jsou uložena v relačním datovém skladu a zpracovaná jsou uživateli prezentována v podobě multidimenzionálního pohledu, relační databáze umožňují vyhnout se redundanci dat, ke které dochází u předcházejícího modelu,
- Hybridní OLAP (HOLAP) – tento model umožňuje využít přednosti obou předchozích modelů a omezit dopad jejich nevýhod, data jsou zde uchovávána v relačních databázích a do datových skladů či tržišť se ukládají až vypočítané agregace.
- Desktopová OLAP (DOLAP) – tato architektura umožňuje stažení vybraných dat a jejich zpracování bez nutnosti stálého připojení k serveru. Jedná se řešení výhodné především pro uživatele, kteří chtějí k analytickým nástrojům přistupovat přes smartphony či tablety.

(Lacko, 2003, Novotný, 2005)

Rozdíly mezi analytickými systémy a OLTP

Novotný (2005) uvádí následující rozdíly mezi analytickými systémy a OLTP:

- Základním cílem OLTP systémů je zejména pořizování dat. Tomu je podřízen i databázový model, který bývá ve třetí normální formě a obsahuje velké množství tabulek. Je zde také snaha vyhnout se redundanci dat.
- Analytické systémy jsou naopak zaměřeny na zejména na podporu dotazování. Databázový model bývá tedy denormalizován a obsahuje menší množství tabulek. Pro tento model je typická i redundance dat.
- OLTP systémy jsou plněny daty v reálném čase, takže dochází k jejich rovnoměrnému a stálému zatěžování. U analytických systémů je zatížení spíše nárázově, neboť zde data bývají aktualizována v určitých intervalech. Záleží také na náročnosti a četnosti dotazů prováděných nad těmito systémy.

Tabulka 1 Rozdíly mezi OLTP a OLAP systémy

OLTP	OLAP
dynamická data	statická data
současná data	současná a historická data
detailní data	agregovaná a konsolidovaná data
operační funkce	analytické funkce
opakující se dotazy	nepředvídatelné dotazy
nízká flexibilita	vysoká flexibilita
vysoký výkon, pár vteřin na provedení dotazu	pro komplexní dotazy může být výkon nižší
prioritou je vysoký výkon	prioritou je vysoká flexibilita
uživatelé jsou zaměstnanci	uživatelé jsou znalostní pracovníci

Zdroj: Vercellis (2009), upraveno autorkou

3.4.7. Dočasné úložiště dat (DSA)

Dočasné úložiště dat představuje úložiště a sadu procesů, které čistí, transformují, kombinují, zbavují duplikátů, uchovávají, archivují a připravují data pro použití v

datových skladech. Většinou se nejedná o jedno zařízení, ale bývá součástí zařízení několika. (Kimball, 1998)

Dočasné úložiště dat je dočasné úložiště, kam jsou nahrána data z produkčních databází před uložením do datového skladu. Jedná se o nepovinnou komponentu Business Intelligence. Její výhoda tkví především v tom, že lze předejít narušení funkčnosti produkčních systémů během čištění dat, neboť se tento krok odehrává až v DSA a ne v daném produkčním systému. Dočasné úložiště dat je výhodné také v případě, že data v produkčním systému jsou aktualizována častěji než v datovém skladu. Všechny změny jsou ukládány do DSA, takže nedojde k jejich ztracení. Tato komponenta může sloužit také jako úložiště dat pro budoucí využití, kdy tvůrci skladu zatím neví, jak určitá data využijí, ale jsou si jisti, že je budou potřebovat.

(Lans, 2012)

3.4.8. Operativní úložiště dat (ODS)

Stejně jako DSA je i operativní úložiště dat nepovinnou komponentou Business Intelligence. Novotný (2005, p. 30-31) uvádí dva typy definic ODS:

- 1) ODS je *„jednotné místo datové integrace aktuálních dat z primárních systémů. Jedná se o zdroj pro sledování konsolidovaných agregovaných dat s minimální dobou odezvy po zpracování (tedy sledování téměř v reálném čase). V mnoha případech takové ODS slouží jako centrální databáze základních číselníků (zákaznický, produktový) nebo pro podporu interaktivní komunikace se zákazníkem.*
- 2) ODS je *databáze navržená s cílem podporovat relativně jednoduché dotazy nad malým množstvím aktuálních analytických dat. ...ODS vzniká jako derivace již existujícího datového skladu.“*

3.4.9. Reporting

Reporting představuje proces získávání a poskytování informací nebo reportů uživatelům či aplikacím pomocí řešení Business Intelligence. Většinou se jedná o část software, která slouží pro podávání sumarizovaných a strukturovaných reportů k analýzám či operacím provedeným nad určitou sadou dat. Většinou se jedná o

automatizovaný proces, jehož výstup může mít podobu statistických dat, grafů nebo textu.

(Business Intelligence Reporting)

„Prostřednictvím reportingu získá oprávněná osoba informace ve formě, která je mu graficky příjemná, funkcionálně dostatečná a se snadnou obslužností tak, aby nebyl uživatel zdržován a zahlcován úkony, které nejsou bezpodmínečně nutné.“

("Business intelligence a reporting")

3.4.10. Data mining

„Data mining je proces analýzy dat z různých perspektiv a jejich přeměna na užitečné informace. Z matematického a statistického hlediska jde o hledání korelací, tedy vzájemných vztahů nebo vzorů v údajích.“

(Lacko L., 2003, p. 313)

Pro data mining se většinou využívají normalizované datové struktury, obvykle se jedná o třetí normální formu. Během dolování dat dochází k zatížení základní databáze a snížení jejího výkonu, a jelikož se databáze používá i pro běžné dotazy na data, je potřeba vybrat vhodný datový model, který umožní danému řešení co největší efektivitu.

(Laberge, 2012)

Data mining využívá matematické a statistické techniky a modely, mezi které patří např.

(Novotný, 2005):

- Rozhodovací stromy – jedná se o prediktivní modely, ve kterých jsou data rozložena do podoby stromu. Strom obsahuje uzly, které určují další přerozdělení dat do větví a následně listů. Listy představují skupiny dat, které sdílejí určité vlastnosti.
- Neuronové sítě – snaží se napodobovat počínání lidského mozku, využívají se pro hledání vzorů v datech.
- Genetické algoritmy,
- Clustering a klasifikace.

3.4.11. Nástroje pro koncové uživatele

Mezi nástroje pro koncové uživatele patří zejména EIS (Executive Information Systems), které představují nástroje pro podporu manažerských procesů. EIS lze definovat jako počítačové informační systémy, které představují podporu komunikace, koordinace, plánování a dohledu nad výkonem společnosti pro manažery a řídicí pracovníky. Na EIS se běžně nahlíží jako na specializovanou formu DSS (Decision Support Systems – nástroje pro podporu rozhodování), která je navržena pro nejvyšší pracovníky.

(Herschel, 2012)

„EIS je ta část celkového IS firmy, která pracuje s vybranými nebo upravenými daty a která se těmito úpravami stává nositelem komplexních informací, charakterizujících příslušné procesy ve firmě. Primárně slouží k identifikaci a lokalizaci určitých jevů ve firmě, v dalším kroku pak k jejich podrobné analýze.

Typickou vlastností EIS je multidimenzionalita, která umožňuje rychle a jednoduše vytvářet nové pohledy na data, jejich řazení do nových souvislostí, vyhledávání zákonitostí (trendových charakteristik), indikaci odchylek klíčových ukazatelů od plánovaných hodnot, práci s historií a anticipaci budoucího vývoje.“

(„EIS - nezbytná součást“)

Mezi komponenty Business Intelligence mohou být zařazeny i další nástroje, např.:

- Systémy pro podporu rozhodování (DSS) - DSS jsou aplikace, které sbírají, organizují a analyzují obchodní data pro usnadnění rozhodování. Dobře navržené DSS kompilují data z mnoha zdrojů: nezpracovaná data, dokumenty, znalosti zaměstnanců a business modelů. DSS pomáhají společně identifikovat a řešit problémy.

(Decision Support System)

- Expertní systémy (ES – Expert Systems) - Expertní systém je program, který využívá umělou inteligenci. ES má rozsáhlé znalosti z určité oblasti a umí tyto znalosti využít k odpovědi na daný problém.

(WHAT IS EXPERT SYSTEM)

3.5. Přístupy k řešení BI

Mezi přístupy k řešení BI patří (Laberge, 2012):

- Shora dolů,
- Zdola nahoru,
- Hybridní, který kombinuje oba předchozí přístupy a využívá jejich výhod.

3.5.1. Přístup shora dolů

Autorem tohoto přístupu je Bill Inmon. „*Tento přístup spočívá v jednorázovém vybudování celkového řešení, tedy všech potřebných komponent pokrývajících současně všechny definované analytické potřeby zadavatelské organizace. Přístup se skládá z následujících kroků:*

- *Celková analýza a dokumentace všech relevantních uživatelských potřeb.*
- *Návrh a implementace celkového řešení. Týká se zejména vybudování konsolidovaného datového skladu, pokrývajícího výše zmapované potřeby uživatelů a tvorby základních závislých datových tržišť.*
- *V případě dalších uživatelských potřeb jsou tyto nároky pokryty tvorbou nových datových tržišť, za využití konsolidovaného datového skladu vybudovaného v rámci předchozího kroku.“*

(Novotný, 2005, p. 48)

Výhody:

- nižší riziko redundance dat,
- definice pojmů je stejná pro celou společnost.

Nevýhody:

- vyšší počáteční investice,
- vytvoření a implementace datového skladu trvá delší dobu, neboť je zde delší přípravná fáze.

3.5.2. Přístup zdola nahoru

Autorem tohoto přístupu je Ralph Kimball. „*Jeho princip spočívá v relativně nezávislém vytváření jednotlivých datových tržišť pro specifické útvary podniku (divize, oddělení, pobočky, závody). Každé takové tržiště je typicky kompletně života schopné, tzn., obsahuje veškeré vrstvy a komponenty, které umožní získat data z primárních systémů, zpracovat je, uložit v datovém tržišti, případně analyzovat pomocí OLAP aplikací či Data Mining komponent, a prezentovat je uživateli.*“ (Novotný, 2005, p. 45)

Výhody:

- nižší počáteční investice,
- datový trh lze rychleji implementovat, neboť jeho vytvoření zabere méně času.

Nevýhody:

- riziko redundance dat,
- každý podnikový útvar může mít jednotlivé pojmy definovány různými způsoby, což může vést ke komplikacím při budoucím propojení jednotlivých tržišť,
- každý útvar zajímá historie jiných dat a může je aktualizovat v různých intervalech, takže v budoucnu zde opět může dojít k nesrovnalostem.

(Dva způsoby budování datového skladu)

Nástroje BI by měly umožňovat:

- sledování vývoje podnikových ukazatelů,
- analýzu podnikových ukazatelů na základě vybraných dimenzí a jejich vývoje v čase,
- podporovat rozhodovací a řídicí procesy.

(BI v podnikové praxi)

3.6. SW řešení pro Business Intelligence

V současné době je na trhu řada řešení, která nabízejí široké spektrum vlastností a doplňků pro potřeby nástrojů Business Intelligence. Společnost Gartner uvádí jako nejvýznamnější společnosti následující: MicroStrategy, Microsoft, SAP, IBM, SAS, Oracle a další.

(Magic Quadrant for Business Intelligence and Analytics Platforms, 2015)

Ve své práci jsem se zaměřila na nástroje od společnosti Microsoft, které jsou využívány pro výuku na Ekonomické fakultě. Jedná se zejména o MS SQL Server. Společnost Microsoft nabízí také nástroj Power BI, který rozšiřuje možnosti programu MS Excel o funkce Business Intelligence.

Poskytované služby:

- *„Power Query – umožňuje snadné zjišťování a připojování k datům z veřejných a podnikových zdrojů dat. To zahrnuje možnosti pro snadnou transformaci a sloučení dat z několika zdrojů dat tak, aby se daly dále analyzovat v Excelu.*
- *PowerPivot – umožňuje sestavovat sofistikované datové modely z těchto dat v Excelu vytvářením relací, vlastních měřítek, hierarchií a klíčových ukazatelů výkonu.*
- *Power View – umožňuje vytvářet sestavy a analytická zobrazení prostřednictvím interaktivních tabulek a grafů, díky kterým můžete data zkoumat a vizuálně prezentovat v Excelu.*
- *Power Map – umožňuje zkoumat a procházet geografická data na 3D mapě v Excelu.“*

(Příručka Začínáme s Power BI, 2015)

4. Praktická část

4.1. Datové modely

Tato část obsahuje ukázkou, jak by mohl vypadat relační model procesu objednávky a k němu vytvořené schéma hvězdy a sněhové vločky. Obrázek 8 pak zobrazuje konkrétní datový model, který byl využit v mé práci.

4.1.1. Normalizovaný relační datový model

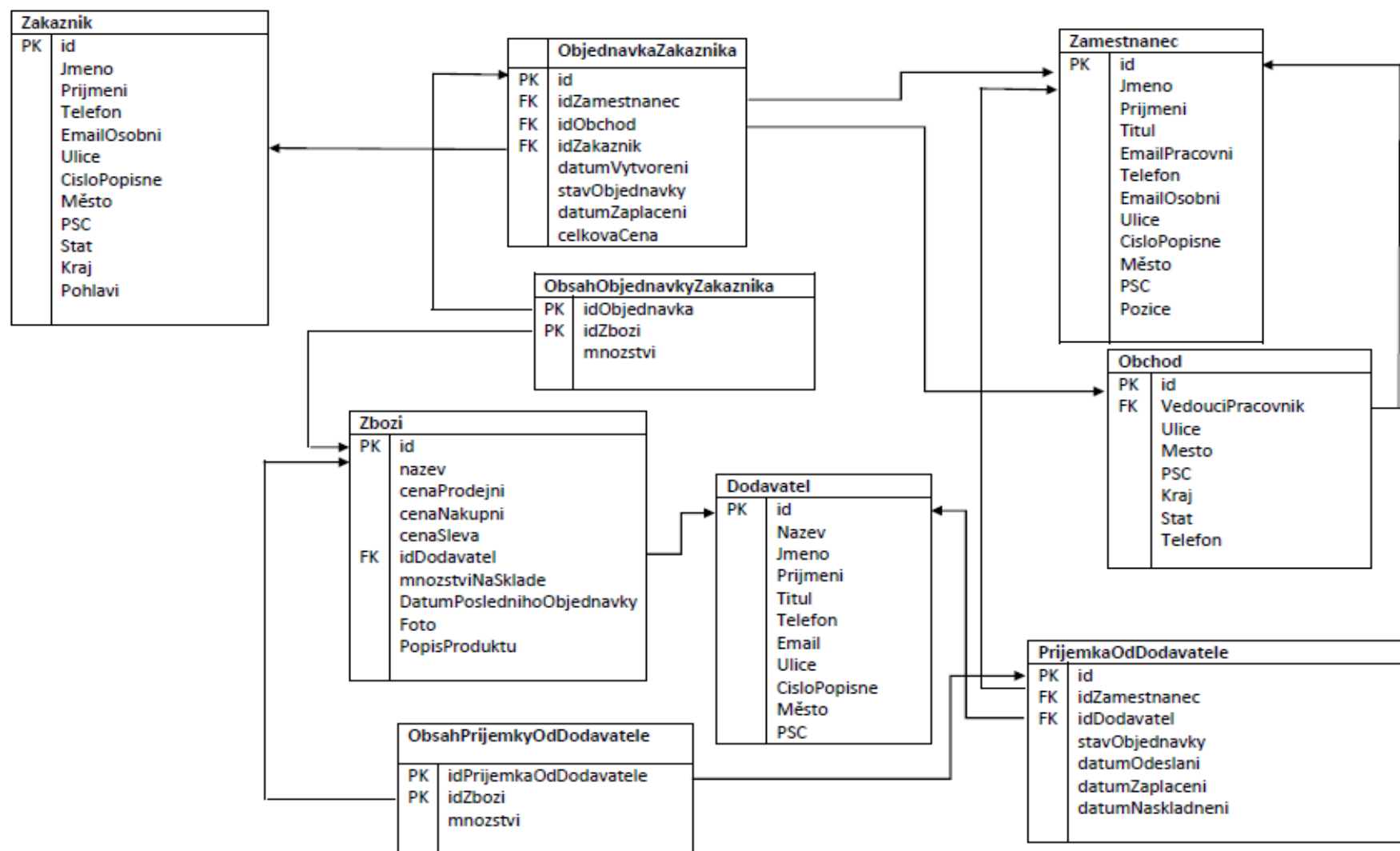
Na obrázku 5 je datový model ve 3 normální formě. Jedná se o model procesu objednávky a její provedení. Jedná se o zjednodušený vzor¹, který obsahuje následující entity:

- Zakaznik – tabulka, která obsahuje základní identifikační údaje o zákaznících,
- ObjednavkaZakaznika – tabulka, ve které jsou zachyceny jednotlivé zákaznické objednávky,
- ObsahObjednavkyZakaznika – vazební tabulka, jež obsahuje seznam položek jednotlivých nákupů,
- Zamestnanec – tabulka se základními identifikačními údaji o zaměstnancích,
- Obchod – tabulka se základními údaji o jednotlivých prodejnách,
- Zbozi – tabulka se základními údaji o zboží,
- Dodavatel – tabulka, která obsahuje základní identifikační údaje o dodavatelích,
- PrijemkaOdDodavatele – tabulka, ve které jsou zachyceny jednotlivé objednávky zboží od dodavatelů,
- ObsahPrijemkyOdDodavatele – vazební tabulka, jež obsahuje seznam položek jednotlivých příjemek.

Primární klíče jsou v modelu označeny PK, cizí klíče jsou označeny FK.

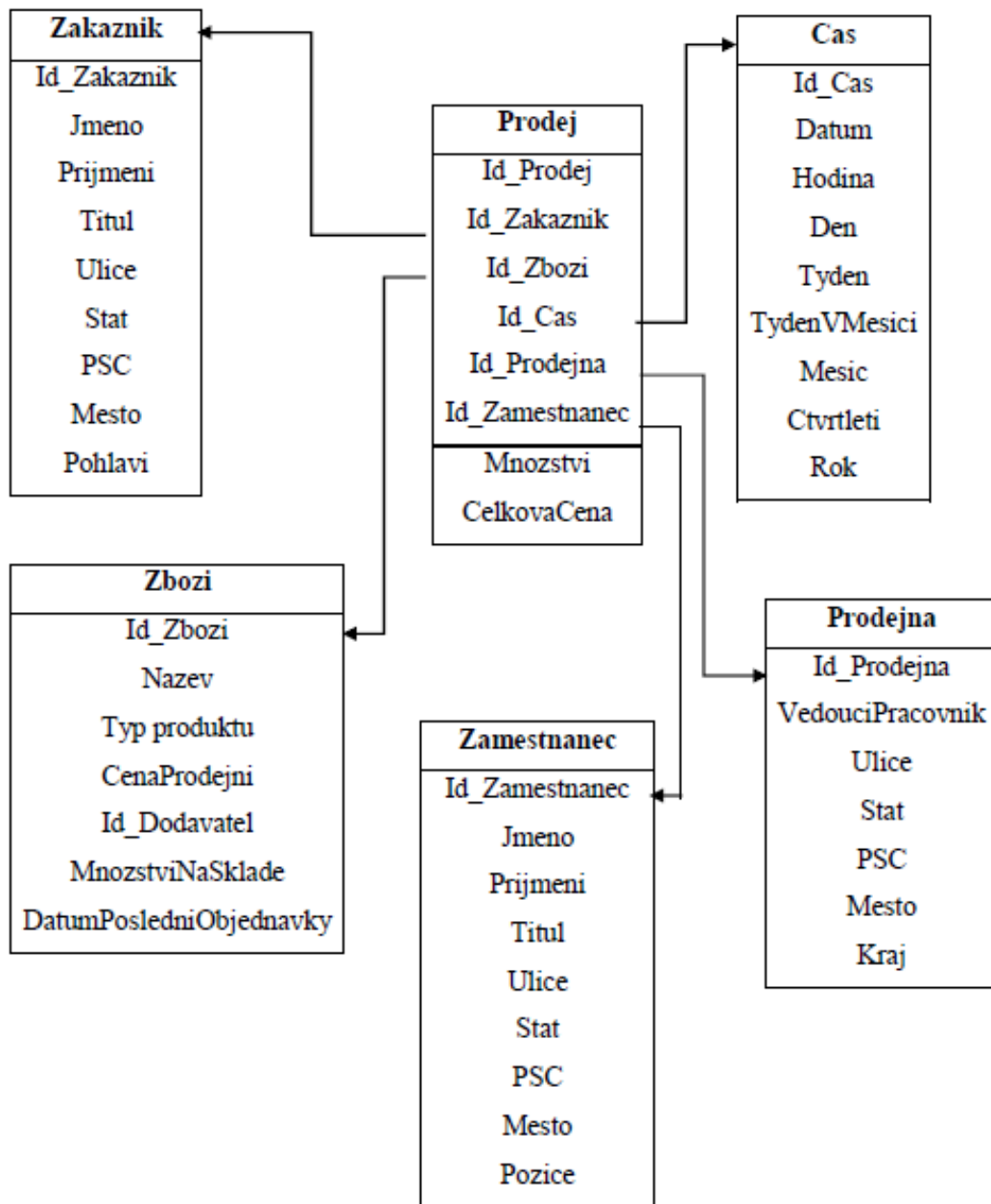
¹ Daný model předpokládá, že každý atribut nabývá pouze jedné hodnoty. Ve skutečnosti mohou mít zákazníci, např. více telefonních čísel.

Obrázek 5 Normalizovaný datový model



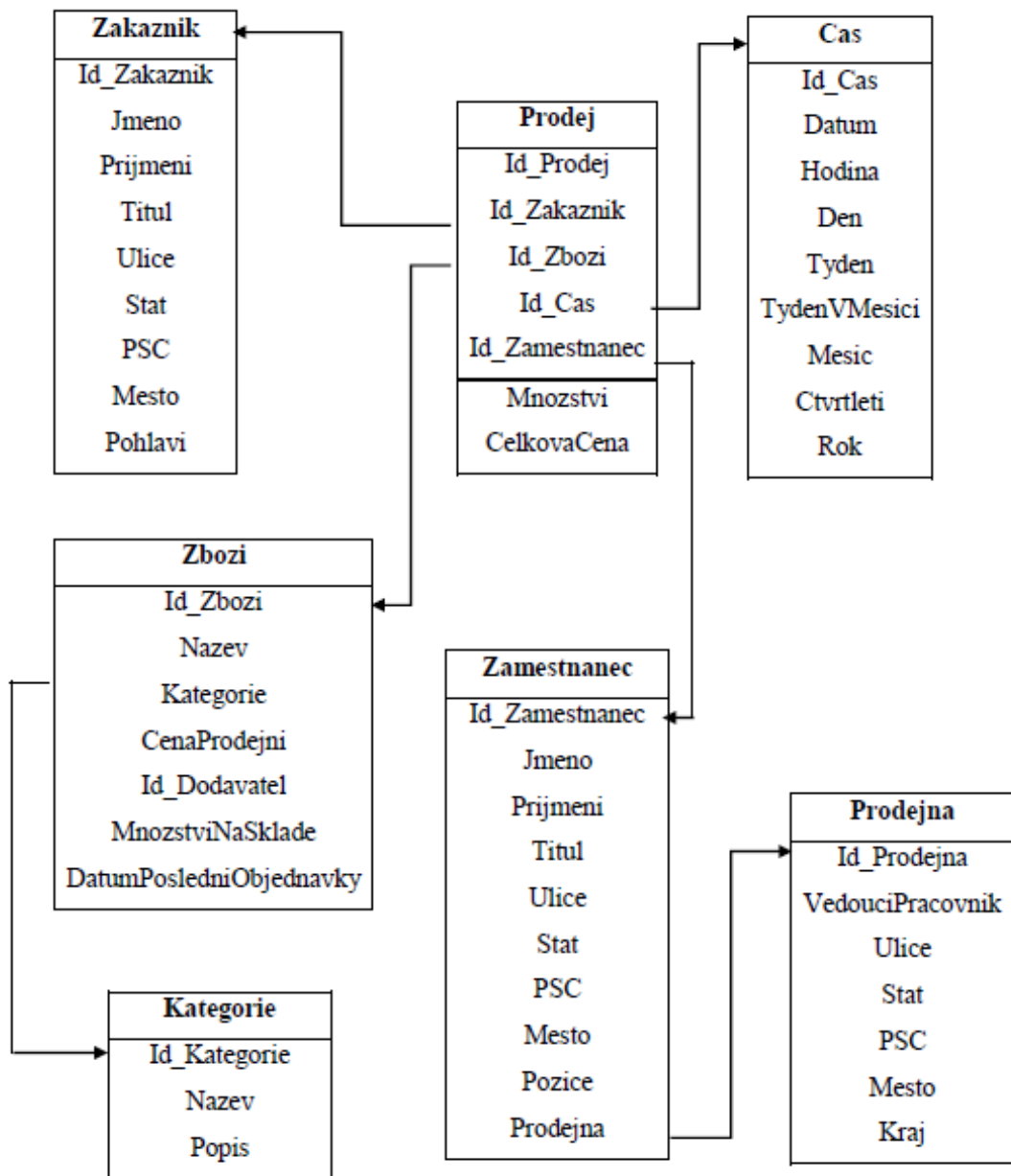
zdroj: autorka

Obrázek 6 Schéma Hvězda



zdroj: autorka dle Vercellis (2009)

Obrázek 7 Schéma Sněhová vločka



zdroj: autorka dle Vercellis (2009)

4.1.2. Rozdíly mezi datovým modelem hvězdy a sněhové vločky

Na obrázcích 6 a 7 jsou modely používané pro úlohy business intelligence. Obrázek 6 zobrazuje schéma hvězdy a obrázek 7 schéma sněhové vločky. Zatímco u hvězdy jsou s tabulkou faktů (tabulka Prodej), spojeny všechny dimenze, u sněhové vločky již došlo k větvení, takže s tabulkou faktů jsou spojeny pouze některé dimenze a zbývající dimenze jsou spojené s některými z nich, jedná se o dimenzi Kategorie, která je spojená s dimenzí Zboží a dimenzi Zaměstnanec, která je spojena s dimenzí Prodejna.

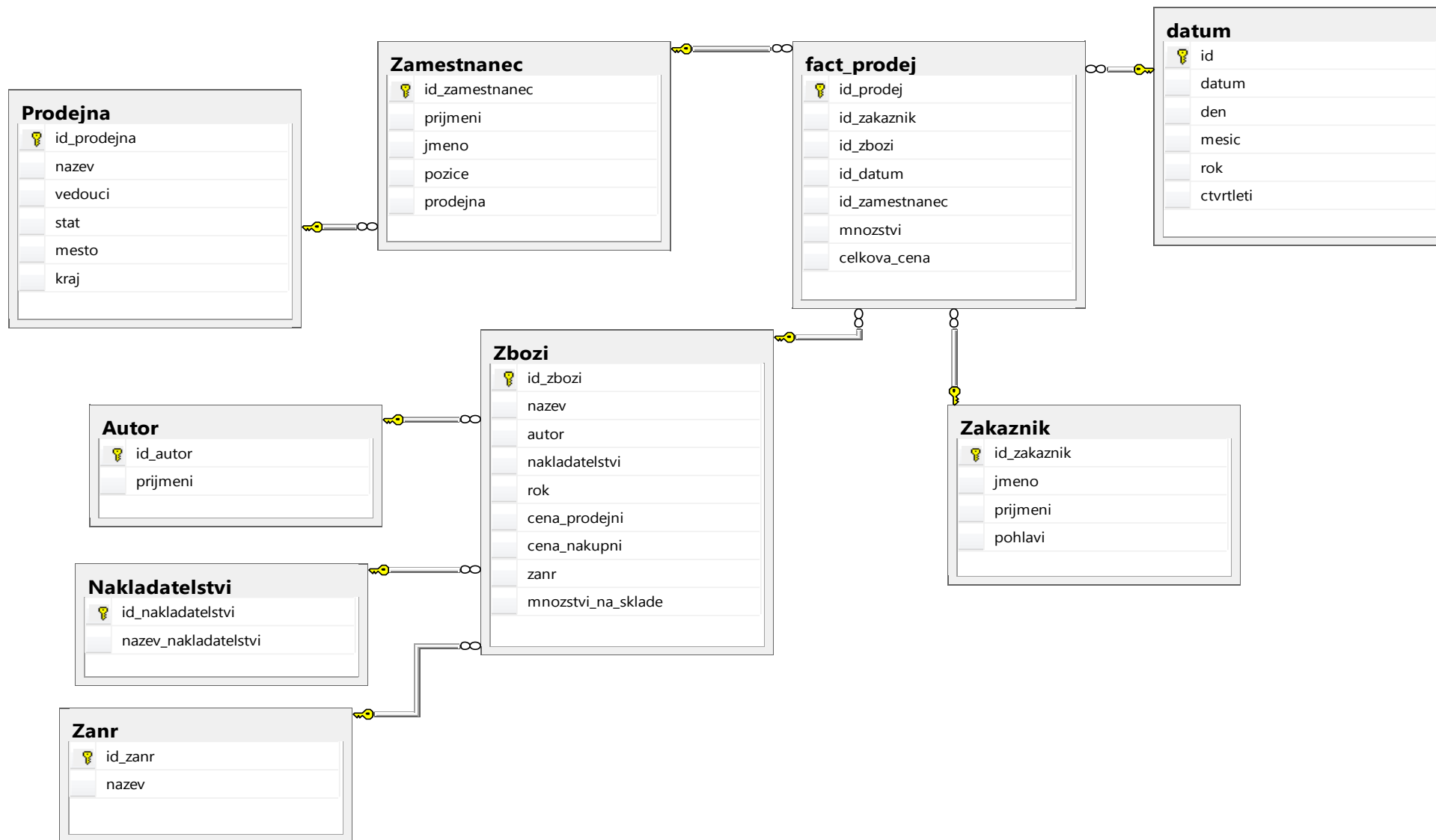
Tabulka 2 Rozdíly mezi modelem hvězdy a sněhové vločky

Hvězda	Sněhová vločka
denormalizované tabulky dimenzí	normalizované tabulky dimenzí
denormalizovaná tabulka faktů	denormalizovaná tabulka faktů
větší velikost datového skladu, neboť u denormalizovaného modelu zpravidla dochází k výskytu redundantních dat	nižší velikost datového skladu
nízké množství operací spojování (join) při tvorbě dotazů, zpracování dotazů je tedy rychlejší	vysoké množství operací spojování (join) při tvorbě dotazů, zpracování dotazů je tedy pomalejší
nejsou zde rodičovské tabulky dimenzí	obsahuje alespoň jednu rodičovskou tabulku dimenzí
model vhodný pro použití v datových tržištích s jednoduchými vazbami (1:1 nebo 1:M)	model vhodný pro použití v datových skladech s komplexními vazbami (M:N)

zdroj: autorka, „Star schema vs. snowflake“,

Joe Oates (2007) uvádí následující příklad jako ukázkou, kdy je vhodné využít model sněhové vločky: Pokud společnost nabízí velké množství typů rozdílných produktů, je vhodné vytvořit základní dimenzi Produkt, která bude obsahovat základní atributy, které jsou všem skupinám produktů společné. Následně se vytvoří podtabulky, které budou obsahovat atributy společné pro určitou skupinu produktů.

Obrázek 8 Schéma databáze ProdejDW



4.2. SQL Server Data Tools

Tato část obsahuje řešené úlohy transformace dat, práce s multidimenzionální kostkou a reportování, Pro vypracování jednotlivých úloh byly využity následující softwarové nástroje: SQL Server Data Tools a MS SQL Server 2012.

4.2.1. ETL

Tato část popisuje proces nahrání zdrojových dat, jejich úpravu a následné nahrání do datového skladu.

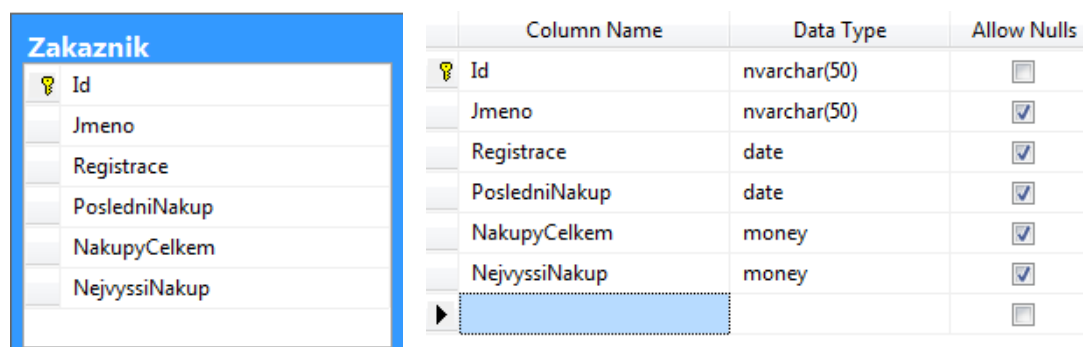
1. Založení datového skladu a souboru se zdrojovými daty

Prvním krokem je založení datového skladu, které proběhne následovně:

- 1) Spustíme SQL Server Management Studio a rozbalíme nabídku daného serveru.
- 2) Klikneme pravým tlačítkem myši na Databases, vybereme možnost New Database a zvolíme jí jméno. V mém případě ZakaznikDW.
- 3) V této databázi vytvoříme tabulku, do které se načtou upravená data tak, že rozbalíme nabídku ZakaznikDW, klikneme pravým tlačítkem myši na Tables a vybereme možnost New Table.

Tabulka může být vytvořena v návrháři tabulek či pomocí SQL dotazu.

Obrázek 9 Tabulka Zakaznik



The image shows two views of the 'Zakaznik' table structure. On the left is the 'Table Designer' view, and on the right is the 'Columns in Table' grid.

Column Name	Data Type	Allow Nulls
Id	nvarchar(50)	<input type="checkbox"/>
Jmeno	nvarchar(50)	<input checked="" type="checkbox"/>
Registrace	date	<input checked="" type="checkbox"/>
PosledniNakup	date	<input checked="" type="checkbox"/>
NakupyCelkem	money	<input checked="" type="checkbox"/>
NejvyssiNakup	money	<input checked="" type="checkbox"/>
		<input type="checkbox"/>

zdroj: autorka

```

CREATE TABLE Zakaznik(
  Id nvarchar(50),
  Jmeno nvarchar(50),
  Registrace date,
  PosledniNakup date,
  NakupyCelkem money,
  NejvyssiNakup money,
  PRIMARY KEY(Id)
);

```

- 4) Vytvoříme soubor se zdrojovými daty s příponou .csv. Soubor může být vytvořen v poznámkovém bloku či v Excelu. Obsah mého souboru zdroj_dat.csv viz Obrázek 16. Je důležité, aby po posledním údaji na posledním řádku nenásledovaly prázdné řádky, neboť ty by byly ve fázi zpracování dat vyhodnoceny jako chyba.

2. Transformace dat v SQL Server Data Tools

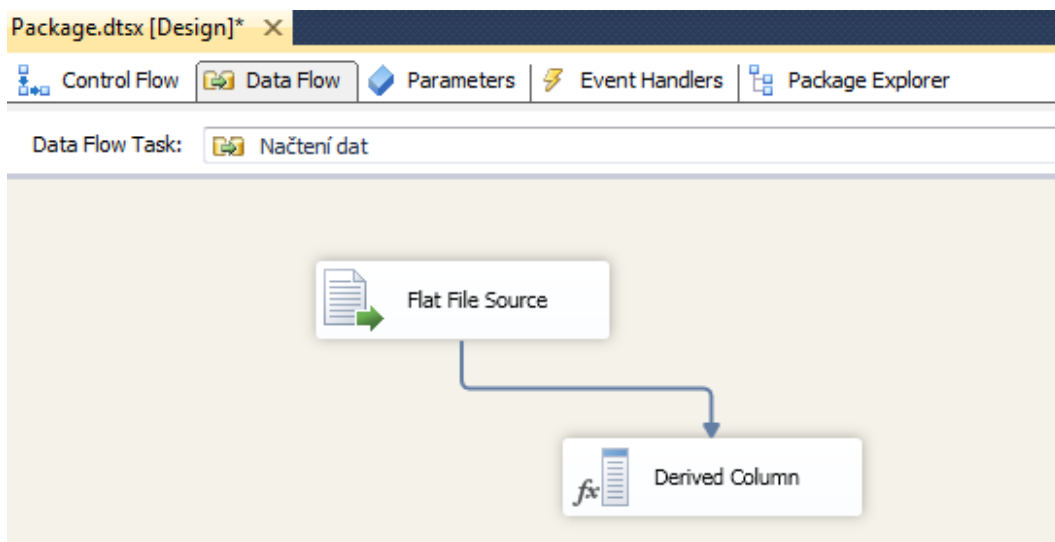
Dalším krokem je vytvoření projektu v SQL Server Data Tools a následná transformace zdrojových dat.

- 1) Vybereme možnost New Project a potom Integration Services Project. Tím nám vznikne tzv. balíček (package), který obsahuje jednotlivé části projektu. Balíček se skládá z pěti částí: Control Flow, Data Flow, Parametres, Event Handlers, Package Explorer.
- 2) Do první části, Control Flow, vložíme z postranního boxu (SSIS Toolbox) Data Flow Task. V této úloze jsou zahrnuty nástroje pro transformaci, čištění a úpravu dat během jejich přenosu z počátečního do cílového úložiště. Data Flow Task lze přejmenovat, což je vhodné, jelikož v reálných případech dochází k načtení z různých zdrojů a mohou vyvstat různé požadavky na ETL proces.
- 3) Přepneme se do části Data Flow. Je zde předem vybrán vytvořený Data Flow Task. Pokud by jich bylo vytvořeno více, lze se mezi nimi přepnout pomocí rozbalovacího menu pod názvy jednotlivých částí balíčku.
- 4) Z SSIS Toolboxu vybereme Flat File Source z nabídky Other Sources. Z této položky vedou 2 šipky: modrá představuje výstupní data a červená případné chyby. Vedle názvu položky je také křížek v červeném kolečku. To znamená, že

tento prvek ještě nebyl nastaven. Nastavení provedeme přes pravé tlačítko myši a možnost Edit.

- 5) V zobrazeném průvodci dáme vytvořit nové spojení (Connection) pomocí tlačítka New. Dále vyhledáme soubor se vstupními daty, v tomto případě zdroj_dat.csv.
- 6) Jako další krok je vhodné zkontrolovat v sekci Columns, zda-li je vybrán jako oddělovač stejný znak, který jsme použili, tedy středník. Pokud není, lze ho změnit v nabídce Column Delimiter. A dáme OK.
- 7) Dalším krokem je vytvoření položky Derived Column z nabídky Common (v SSIS Toolboxu). Tento prvek slouží pro transformaci vstupních dat podle vybraných pravidel a vytváří nový sloupec či přepíše některý ze vstupních sloupců dat.
- 8) Nyní propojíme obě vytvořené komponenty natažením modré šipky z Data Flow Task do Derived Column.

Obrázek 10 Propojené prvky Data Flow Task a Derived Column



zdroj: autorka

- 9) Nyní je potřeba nastavit pravidla transformace. Cílem bude nahradit nevyhovující údaje ve sloupci Pohlavi, který obsahuje různá označení pro muže (muz, m, MUZ) i pro ženu (z, Zena, Z). Výsledný sloupec by měl obsahovat pouze označení muz a zena. Prvním úkolem tedy bude převést velká písmena na malá a dále proběhne nahrazení nesprávných tvarů.

- 10) Klikneme pravým tlačítkem myši na Derived Column a vybereme Edit. V pravé části nového okna vybereme z nabídky String Functions možnost LOWER a přetáhneme ji do spodního okna. Pro lepší orientaci si můžeme nový sloupec pojmenovat. V mém případě se jmenuje „zmenseni“. Jako argument této funkce zvolíme Pohlavi z levé nabídky Columns a přetáhneme ho do spodního okna na pozici character expression. Viz obrázek 11. Stiskneme OK.
- 11) Dalším krokem bude nahrazení výrazu „m“ za „muz“. Vytvoříme tedy nový Derived Column, ve kterém použijeme funkci REPLACE a podmínku, kterou použijeme pro omezení délky nahrazovaného pojmu na 1 znak. Kdybychom toto omezení nepoužili, tak by došlo k nahrazení daného znaku i uvnitř slov. Podmínka má následující syntaxi:

„boolean_expression?expression1:expression2“

(? : (Conditional) (SSIS Expression))

- boolean_expression = výraz, o jehož pravdivosti se rozhoduje,
- expression1 = akce, která se provede, pokud bude výraz vyhodnocen jako pravdivý,
- expression2 = akce, která se provede, pokud bude výraz vyhodnocen jako nepravdivý

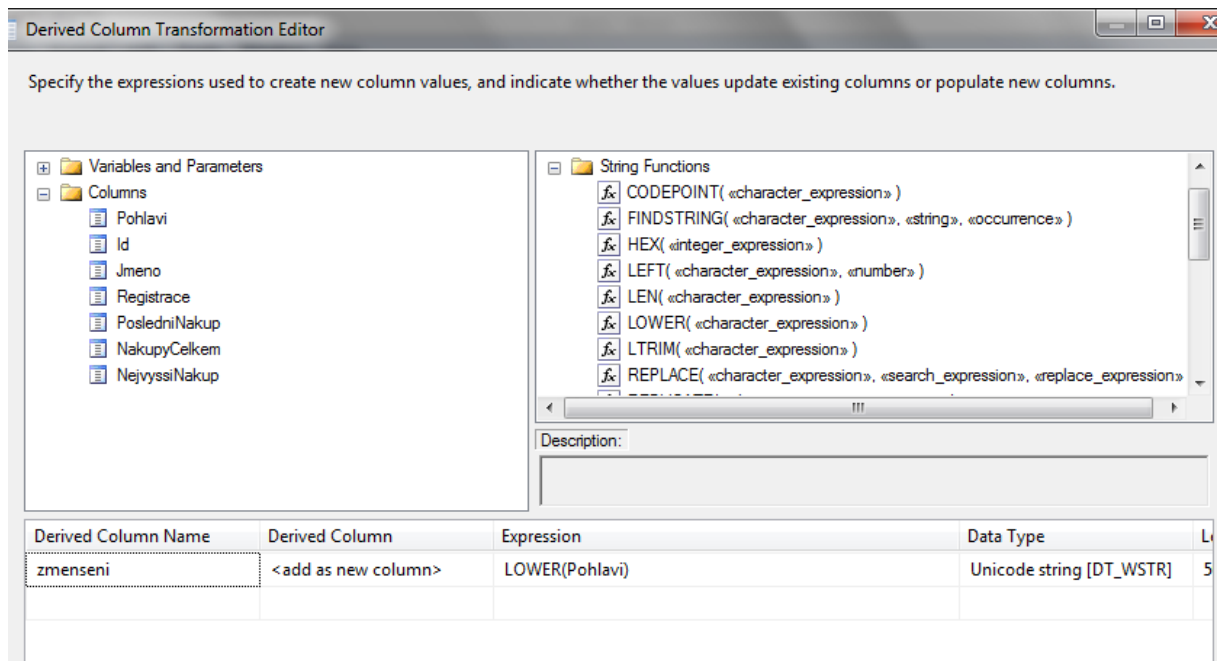
V našem případě bude zápis podmínky vypadat následovně:

`LEN(zmenseni) == 1 ? REPLACE(zmenseni,"m","muz")` : Derived Column 1

Tento výraz znamená, že pokud bude délka (LEN) výrazu rovna 1, dojde k nahrazení, jinak se přepíše původní nenahrazený výraz.

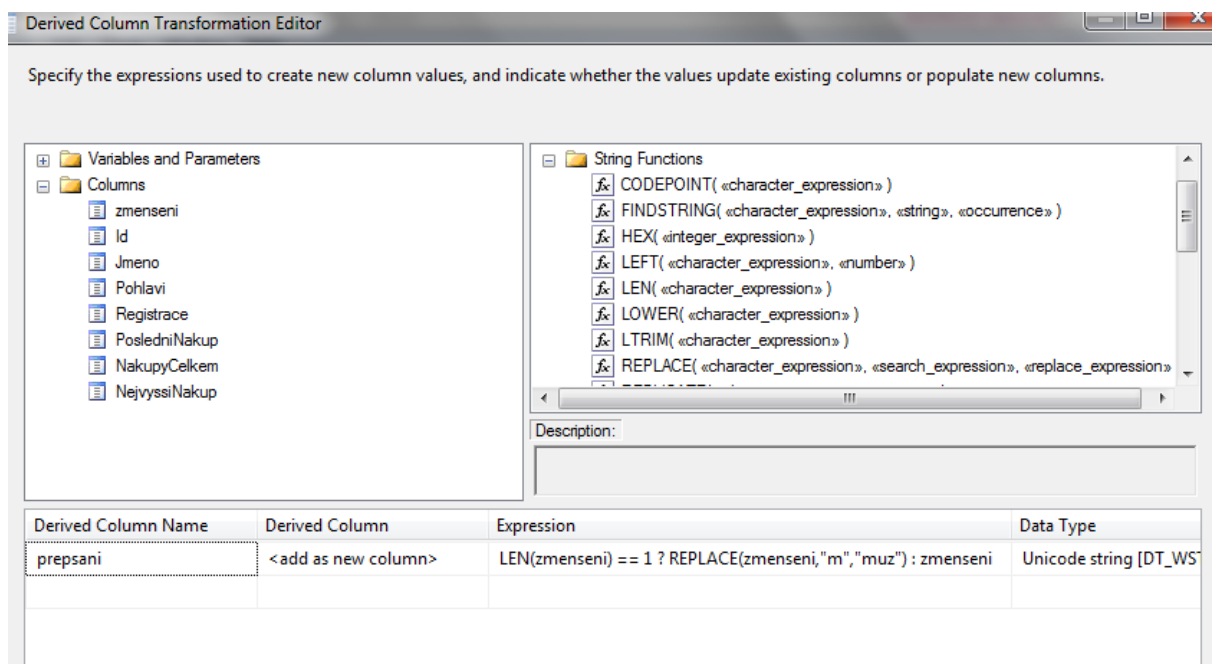
Jako argumenty funkce REPLACE zvolíme „zmenseni“ jako character expression, „m“ bude na pozici search expression a „muz“ bude v části replace expression. Zmenseni je nově vzniklý sloupec z předchozí úpravy. V dolní části okna ve sloupci Derived Column zvolíme „add as a new column“. Viz Obrázek 12.

Obrázek 11 Zmenšení všech písmen



zdroj: autorka

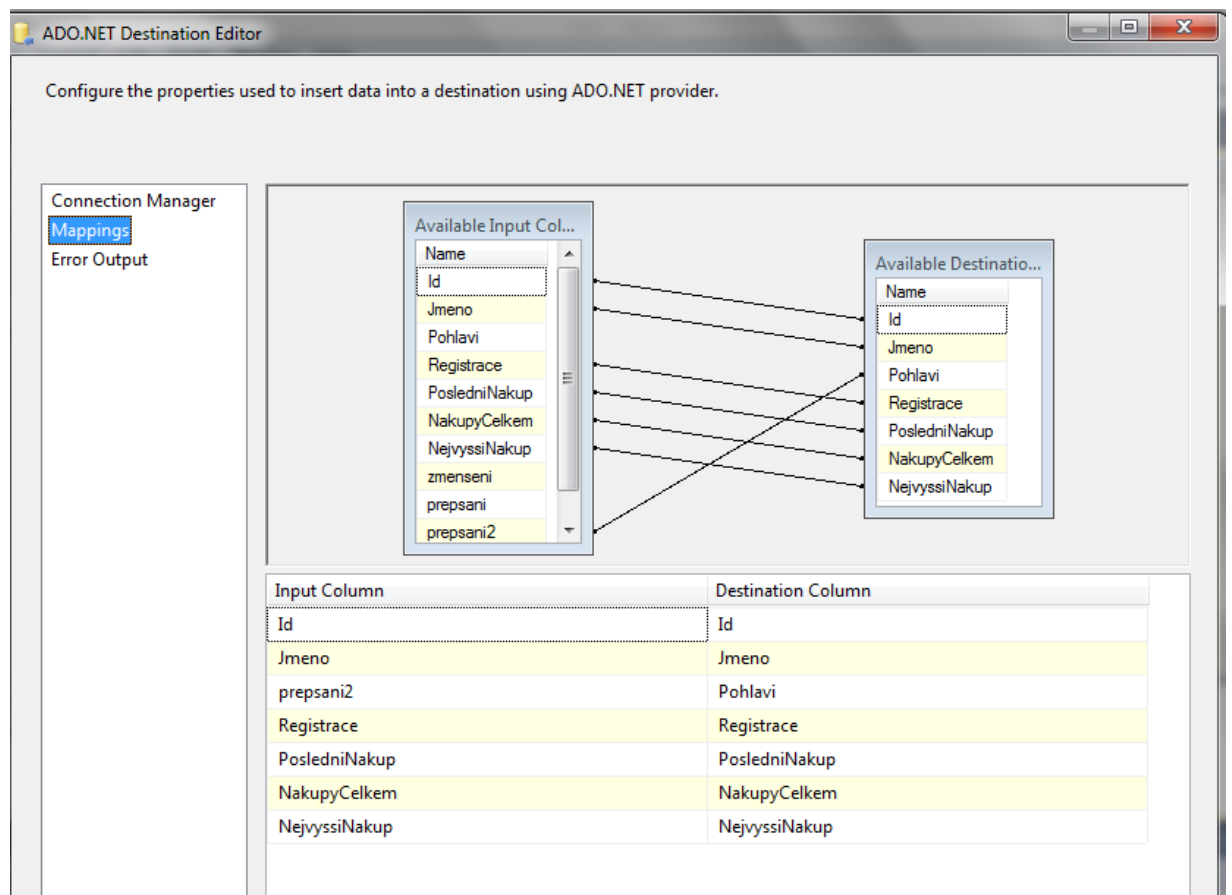
Obrázek 12 Nahrazení vybraného výrazu



zdroj: autorka

- 12) Nahrazení výrazu „z“ za „zena“ proběhne analogicky. Výrazy MUZ a Zena již není třeba nahrazovat, neboť k jejich opravě došlo již při první úpravě.
- 13) Nyní dojde k nahrání upravených dat do datového skladu. Z SSIS Toolboxu přidámě nový prvek, kterým je ADO NET Destination z nabídky Other Destinations. I u tohoto prvku bude potřeba provést nastavení, takže opět zvolíme možnost Edit z nabídky pravého tlačítka myši.
- 14) V novém okně vybereme vytvořit nový Connection manager. Jako jméno serveru uvedeme název SQL Serveru a zadáme jméno používané databáze v tomto případě ZakaznikDW. V Connection manageru vybereme také tabulku, do které se nahrají upravená data.

Obrázek 13 Mappings

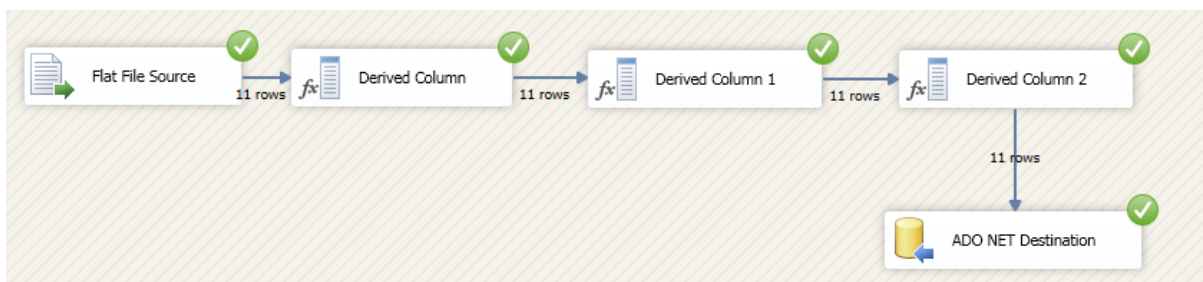


zdroj: autorka

15) Dalším krokem je zkontrolování a upravení položky Mappings, která zobrazuje spojení sloupců, mezi kterými dojde k nahrání dat. Pokud některé spojení není správné, lze ho vymazat a pouhým tažením myši vytvořit spojení nové. Výsledné nastavení viz obrázek 13.

16) Nyní je vše připraveno pro spuštění transformace dat. Samotné spuštění proběhne stisknutím tlačítka Start. Nyní by Data Flow mělo vypadat následovně, viz obrázek 14.

Obrázek 14 Hotové Data Flow



zdroj: autorka

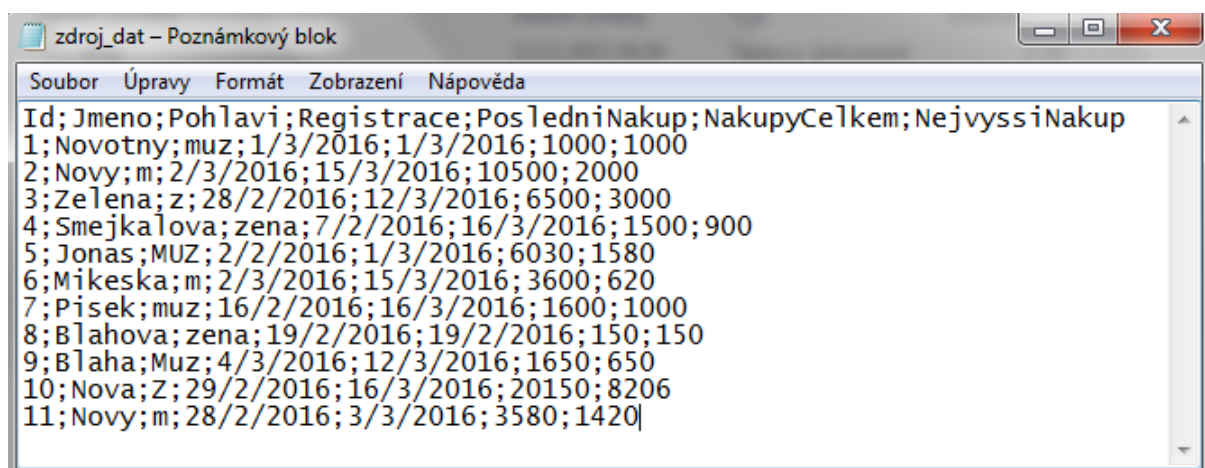
17) Výsledek můžeme zkontrolovat v datovém skladu, kde po vypsání tabulky Zakaznik vidíme, že je naplněna upravenými daty.

Obrázek 15 Naplněná tabulka Zakaznik

	Id	Jmeno	Pohlavi	Registrace	PosledniNakup	NakupyCelkem	NejvyssiNakup
1	1	Novotny	muz	2016-03-01	2016-03-01	1000,00	1000,00
2	10	Nova	zena	2016-02-29	2016-03-16	20150,00	8206,00
3	11	Novy	muz	2016-02-28	2016-03-03	3580,00	1420,00
4	2	Novy	muz	2016-03-02	2016-03-15	10500,00	2000,00
5	3	Zelena	zena	2016-02-28	2016-03-12	6500,00	3000,00
6	4	Smejkalova	zena	2016-02-07	2016-03-16	1500,00	900,00
7	5	Jonas	muz	2016-02-02	2016-03-01	6030,00	1580,00
8	6	Mikeska	muz	2016-03-02	2016-03-15	3600,00	620,00
9	7	Pisek	muz	2016-02-16	2016-03-16	1600,00	1000,00
10	8	Blahova	zena	2016-02-19	2016-02-19	150,00	150,00
11	9	Blaha	muz	2016-03-04	2016-03-12	1650,00	650,00

zdroj: autorka

Obrázek 16 Původní data



zdroj: autorka

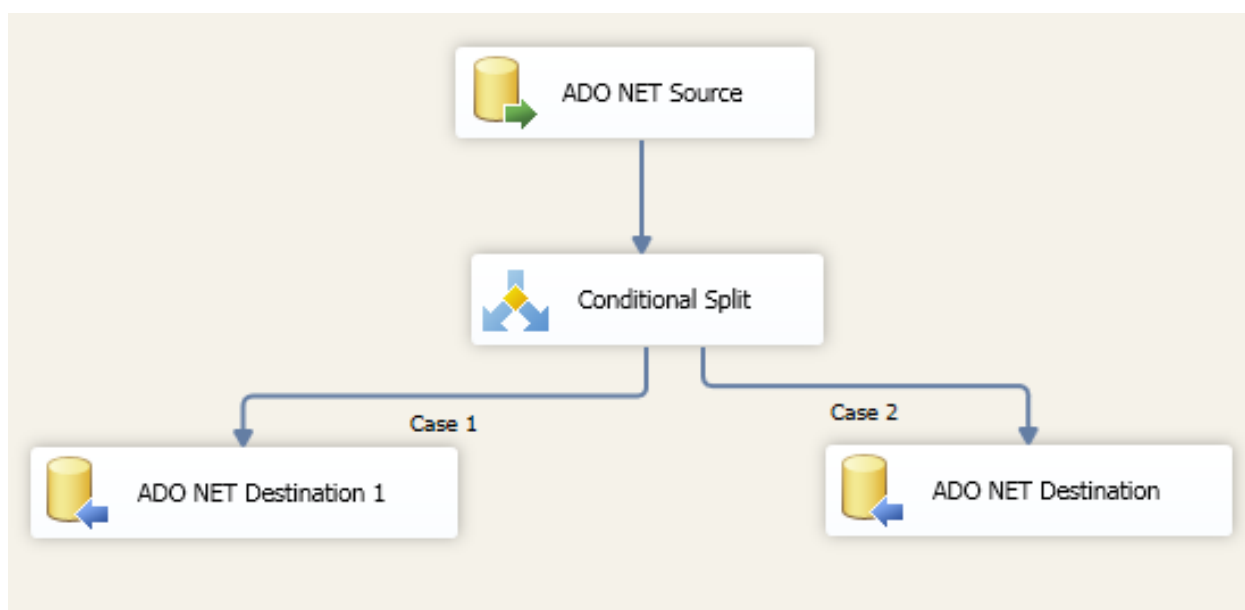
3. Rozdělení

Pro jednotlivé záznamy lze vybrat různá umístění na základě hodnot, kterých nabývají vybrané sloupce. V našem případě rozdělíme záznamy z tabulky Zakaznik do tabulek ZakaznikM, pokud je hodnota ve sloupci NakupyCelkem menší než 5000, a ZakaznikV, pokud je hodnota ve sloupci NakupyCelkem větší nebo rovna 5000.

- 1) Z SSIS Toolboxu vložíme ADO NET Source, protože budeme používat tabulku, která je již uložena na SQL Serveru a ne CSV soubor.
- 2) Vložíme položku Conditional Split.
- 3) Nastavíme podmínky rozdělení:
 - Case 1 – NakupyCelkem < 5000,
 - Case 2 – NakupyCelkem >= 5000.
- 4) Vytvoříme dvě ADO NET Destination pro obě cílové tabulky, které spojíme s Conditional Split.
- 5) Jedna ADO NET Destination bude mít přiřazen Case 1 a cílovou tabulku ZakaznikM a druhá bude mít přiřazen Case 2 a cílovou tabulku ZakaznikV.

Výsledné zapojení viz Obrázek 17.

Obrázek 17 Conditional Split



zdroj: autorka

4.2.2. Multidimenzionální kostka

Tato část obsahuje postup vytvoření jednoduché multidimenzionální kostky.

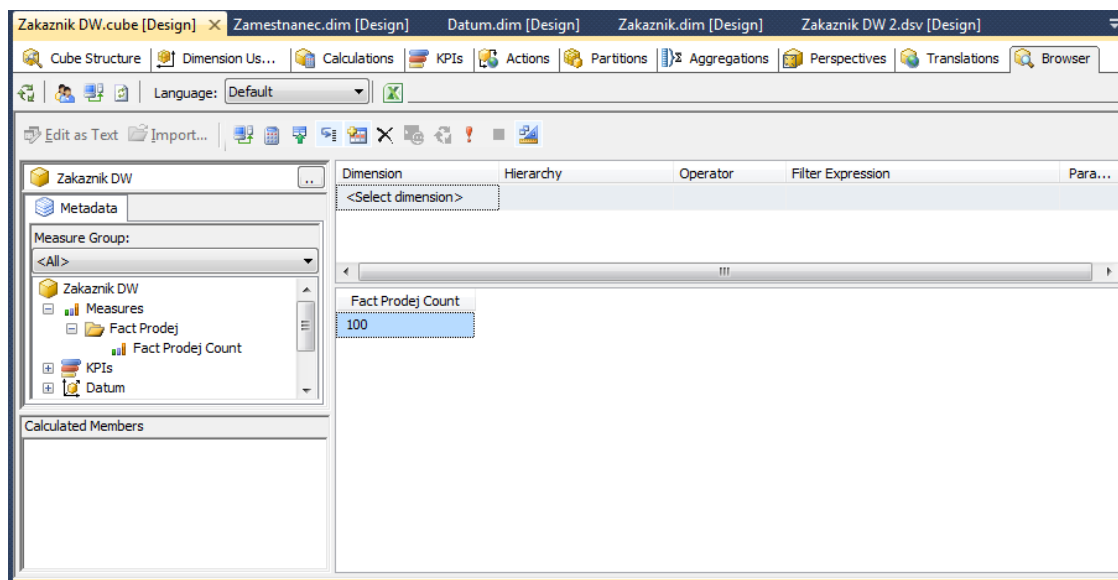
- 1) Zvolíme New Project – Analysis Services – Analysis Services Multidimensional and Data Mining Project a dáme OK.
- 2) V Solution Explorer klikneme pravým tlačítkem na položku Data Sources a zvolíme New Data Source.
- 3) V průvodci nastavíme jako zdroj databázi ProdejDW.
- 4) Nyní je potřeba nastavit Data Source View. Toto nastavení proběhne obdobně jako u zdroje, pouze je nutné vybrat tabulky, které budeme potřebovat: všechny tabulky kromě systémové tabulky sysdiagrams.
- 5) Nyní je nutné vytvořit dimenze. Vytvoření proběhne podobně jako v předchozích bodech.
- 6) V průvodci vybereme Use existing table. Zvolíme tabulku, ze které vytvoříme dimenzi, a zvolíme atributy, které budeme chtít použít, v našem případě všechny.
- 7) U dimenzí lze nastavit hierarchie. Například u data můžeme zvolit hierarchii: Rok – Čtvrtletí – Měsíc.

- 8) Nyní je nutné dimenzi zpracovat. Toho dosáhneme tak, že pravým tlačítkem myši klikneme na vytvořenou dimenzi v Solution Exploreru a vybereme možnost Process.²
- 9) Stejným způsobem vytvoříme i ostatní dimenze.
- 10) Nyní vytvoříme kostku. Průvodce vytvořením opět vyvoláme přes pravé tlačítko myši kliknutím na položku Cubes. V průvodci vše odsouhlasíme, pouze je nutné vybrat tabulku faktů. Tou bude v našem případě fact_prodej.
- 11) Kostku necháme zpracovat kliknutím na tlačítko Process.
- 12) Nyní můžeme na záložce Browser pracovat s kostkou.

Vybrání konkrétních dat proběhne pouhým přetažením metrik a dimenzí do pracovního pole:

- 1) Vypsání počtu všech nákupů
 - a. Z tabulky Measure Group přetáhneme do pole položku Fact Prodej Count. Výsledek viz obrázek 18.

Obrázek 18 Vypsání počtu všech nákupů



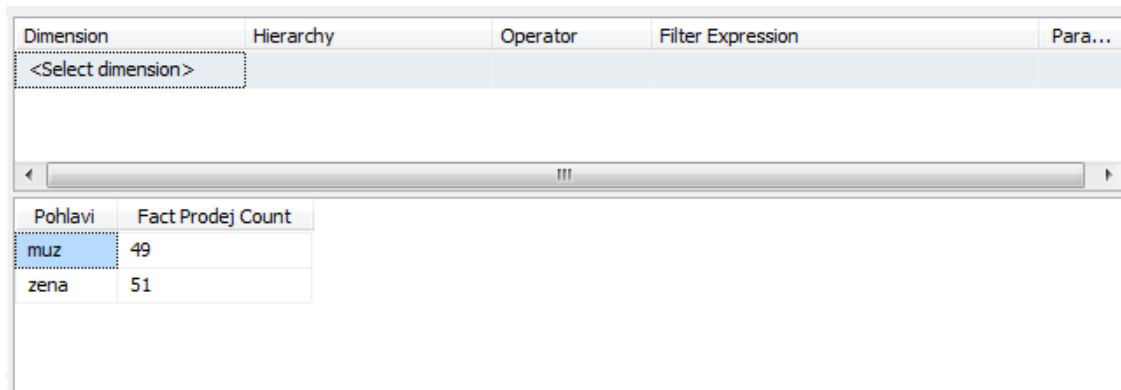
zdroj: autorka

² Je nutné, aby služba SQL Server Analysis Services využívala účet s dostatečně vysokým oprávněním. Pokud tomu tak není, lze vybrat jiný účet v SQL Server Configuration Manager, který je dostupný z nabídky Start – Microsoft SQL Server 2012 – Configuration Tools.

2) Rozdělení prodejů podle pohlaví zákazníka

- a. Do pracovního pole přetáhneme položku pohlaví z dimenze Zakaznik.

Obrázek 19 Rozdělení prodejů podle pohlaví zákazníka



Dimension	Hierarchy	Operator	Filter Expression	Para...
<Select dimension>				

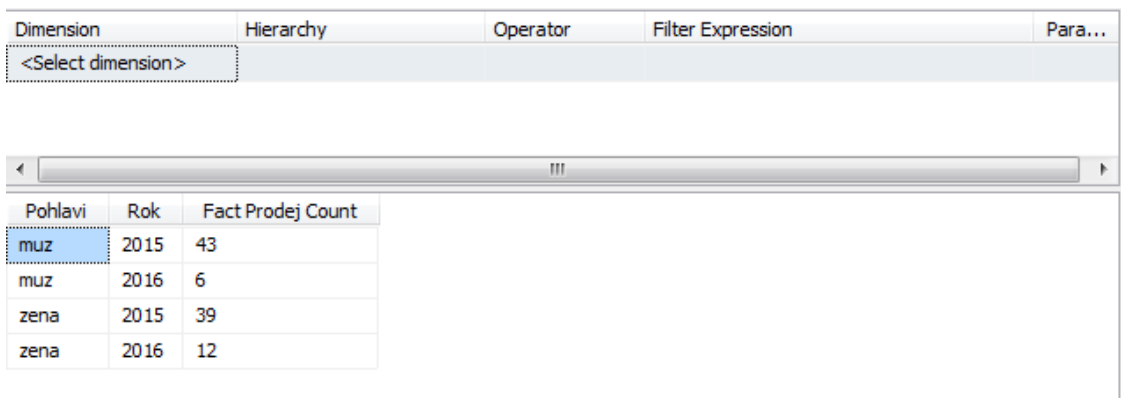
Pohlavi	Fact Prodej Count
muz	49
zena	51

zdroj: autorka

3) Rozdělení prodejů podle pohlaví zákazníka a roku uskutečnění prodeje

- a. Do pracovního pole přetáhneme položku rok z dimenze Datum.

Obrázek 20 Rozdělení prodejů podle pohlaví zákazníka a roku



Dimension	Hierarchy	Operator	Filter Expression	Para...
<Select dimension>				

Pohlavi	Rok	Fact Prodej Count
muz	2015	43
muz	2016	6
zena	2015	39
zena	2016	12

zdroj: autorka

Vypsaná data lze též filtrovat na základě hodnot, kterých nabývají vybrané sloupce z dimenzí.

1) Vyfiltrování prodejů, které proběhly v roce 2015

- a. V menu nad pracovním polem vybereme dimenzi Datum, v hierarchii zvolíme Rok, Operator bude Equal a Filter Expression bude 2015.

Obrázek 21 Vyfiltrování prodejů, které proběhly v roce 2015

Dimension	Hierarchy	Operator	Filter Expression	Para...
Datum	Rok	Equal	{ 2015 }	<input type="checkbox"/>
<Select dimension>				

Pohlavi	Fact Prodej Count
muz	43
zena	39

zdroj: autorka

- 2) Vyfiltrování prodejů, které uskutečnily ženy a které proběhly v 2. čtvrtletí roku 2015
 - a. Přidáme další filtr. Vybereme dimenzi Datum, v hierarchii zvolíme Ctvrtletí, Operator bude Equal a Filter Expression bude 2.
 - b. Přidáme ještě jeden filtr. Vybereme dimenzi Zakaznik, v hierarchii zvolíme Pohlavi, Operator bude Equal a Filter Expression bude žena.

Obrázek 22 Vyfiltrování prodejů, které uskutečnily ženy a které proběhly v 2. čtvrtletí roku 2015

Dimension	Hierarchy	Operator	Filter Expression	Para...
Datum	Rok	Equal	{ 2015 }	<input type="checkbox"/>
Datum	Ctvrtletí	Equal	{ 2 }	<input type="checkbox"/>
Zakaznik	Pohlavi	Equal	{ žena }	<input type="checkbox"/>

Pohlavi	Fact Prodej Count
zena	11

zdroj: autorka

4.2.3. Reportování

K vytvoření reportu využijeme nástroj SQL Server Data Tools a databázi ProdejDW viz obrázek 8. Report lze vytvořit bez nebo s pomocí průvodce.

Tvorba reportu bez průvodce

- 1) Zvolíme New Project – Reporting Services – Report Server Project a dáme OK.
- 2) V Solution Explorer klikneme pravým tlačítkem na položku Reports a zvolíme možnost Add New Report. Tím nám vznikne prázdný report, který je možné libovolně naplnit.
- 3) Nyní je potřeba nastavit zdroj dat, takže Toolbox po levé straně nastavíme na Report Data a nastavíme položku Data Sources -> Add Data Source.
- 4) Nyní je potřeba nastavit zdroj dat, takže stiskneme tlačítko Edit. Objeví se Connection Properties, kde nastavíme jméno serveru a vybereme databázi. Dáme OK a klikneme na Next.
- 5) Dále nastavíme Datasets, kde bude vytvořen dotaz pro vybrání dat. Dotaz lze začít psát přímo nebo využít Query Builder.
- 6) V tomto reportu budou vypsána příjmení autorů, názvy jejich knih, nakladatelství, která je vydala a prodejní ceny. Dotaz bude vypadat následovně:

```
SELECT      Autor.prijmeni, Zbozi.nazev, Zbozi.cena_prodejni,  
            Nakladatelstvi.nazev_nakladatelstvi
```

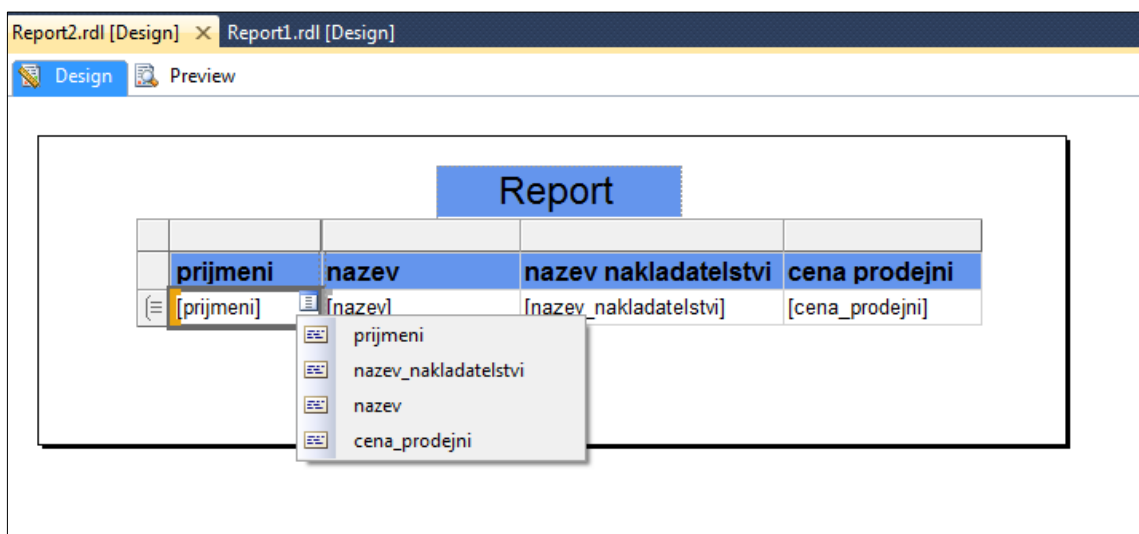
```
FROM        Autor INNER JOIN
```

```
            Zbozi ON Autor.id_autor = Zbozi.autor INNER JOIN
```

```
            Nakladatelstvi ON Zbozi.nakladatelstvi = Nakladatelstvi.id_nakladatelstvi
```

- 7) Nyní vytvoříme vzhled reportu. Pro nadpis použijeme Text Box, který pouze přetáhneme z Toolboxu.
- 8) Z Toolboxu přetáhneme tabulku (Table). Vytvořená tabulka má tři sloupce, tento počet lze zvýšit označením sloupce a zvolením Insert Column z nabídky pravého tlačítka myši.
- 9) Nyní naplníme jednotlivé sloupce tabulky, viz obrázek 23.

Obrázek 23 Naplnění tabulky daty



zdroj: autorka

10) Hotový report je nyní potřeba nahrát na Report Server. Toho lze docílit dvěma způsoby:

- a. V SQL Server Data Tools vybereme možnost Build – Deploy Report Project, následně spustíme prohlížeč s adresou <http://localhost/ReportServer/> a klikneme na odkaz na daný report.
- b. Spustíme v internetovém prohlížeči stránku <http://localhost/Reports/Pages/Folder.aspx>. Zvolíme možnost Upload a vybereme soubor s reportem manuálně. Tato možnost též umožňuje vytvářet složky a reporty do nich zařazovat.

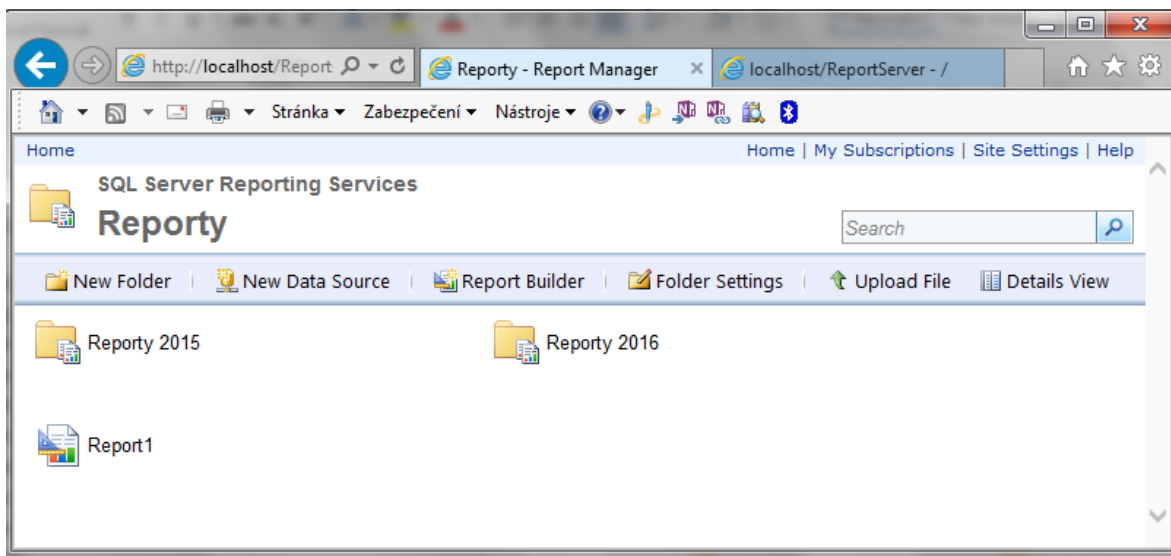
11) Obrazovka s nahráním reportu viz obrázek 24.

12) Výsledný report by měl vypadat následovně, viz obrázek 25

Za účelem příjemnějšího uživatelského zážitku a větší přehlednosti se příjmení autora nebude opakovat, ale bude pouze u první knihy a zbývající buňky se sloučí do jedné.

- 1) Označíme buňky nazev, nazev_nakladatelstvi, cena_prodejni a z nabídky pravého tlačítka myši zvolíme Add Group -> Parent Group -> Group By Prijmeni.
- 2) Výsledné uspořádání, viz obrázek 26.

Obrázek 24 SQL Server Reporting Services



zdroj: autorka

Obrázek 25 Tvorba reportu bez průvodce 1

Report			
prijmeni	nazev	nazev nakladatelstvi	cena prodejni
Tolkien	Hobit	Mladá fronta	199,0000
Tolkien	Pán prstenů 1	Mladá fronta	239,0000
Tolkien	Pán prstenů 2	Mladá fronta	239,0000
Tolkien	Pán prstenů 3	Mladá fronta	239,0000
Pratchett	Barva kouzel	Talpress	199,0000
Pratchett	Maškaráda	Talpress	199,0000
Pratchett	Zaslaná pošta	Talpress	199,0000
Pratchett	Podivný regiment	Talpress	210,0000
Pratchett	Buch	Talpress	210,0000
Rowling	Harry Potter 1	Albatros	259,0000
Rowling	Harry Potter 2	Albatros	259,0000
Rowling	Harry Potter 3	Albatros	259,0000
Rowling	Harry Potter 7	Albatros	259,0000
Rowling	Prázdné místo	Albatros	249,0000
Flanagan	Hraničářův učeň	Egmont	199,0000
Flanagan	Hraničářův učeň 2	Egmont	199,0000
Deary	Děsivé dějiny 1	Egmont	129,0000
Deary	Děsivé dějiny 2	Egmont	129,0000
Asimov	Nadace	Mladá fronta	239,0000
Asimov	Nadace 2	Mladá fronta	239,0000

zdroj: autorka

Obrázek 26 Tvorba reportu bez průvodce 1 - upraveno

Report			
prijmeni	nazev	nazev nakladatelstvi	cena prodejni
Asimov	Nadace	Mladá fronta	239,0000
	Nadace 2	Mladá fronta	239,0000
Deary	Děsivé dějiny 1	Egmont	129,0000
	Děsivé dějiny 2	Egmont	129,0000
Flanagan	Hraničářův učeň	Egmont	199,0000
	Hraničářův učeň 2	Egmont	199,0000
Pratchett	Barva kouzel	Talpress	199,0000
	Maškaráda	Talpress	199,0000
	Zaslaná pošta	Talpress	199,0000
	Podivný regiment	Talpress	210,0000
	Buch	Talpress	210,0000
Rowling	Harry Potter 1	Albatros	259,0000
	Harry Potter 2	Albatros	259,0000
	Harry Potter 3	Albatros	259,0000
	Harry Potter 7	Albatros	259,0000
	Prázdné místo	Albatros	249,0000
Tolkien	Hobit	Mladá fronta	199,0000
	Pán prstenů 1	Mladá fronta	239,0000
	Pán prstenů 2	Mladá fronta	239,0000
	Pán prstenů 3	Mladá fronta	239,0000

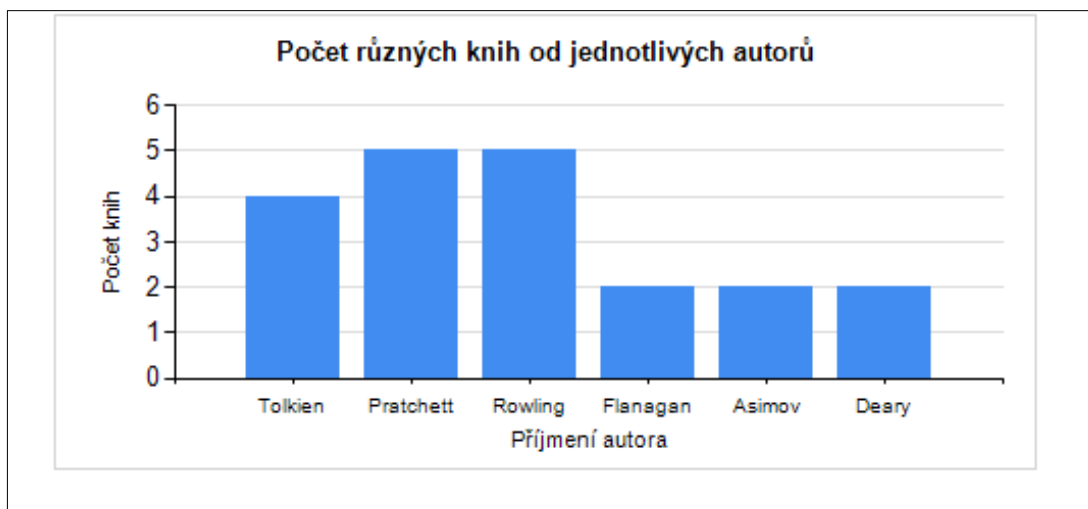
zdroj: autorka

Pro report lze vybrat různá rozložení, místo tabulky můžeme zvolit např. diagram. Cílem bude vytvořit diagram, který zobrazuje, kolik je máme různých knih od jednotlivých autorů.

- 1) Diagram vložíme přetažením položky Chart, kde zvolíme odpovídající typ diagramu.
- 2) Pro dataset budou vybrána data pomocí dotazu:

```
SELECT      Zbozi.autor, COUNT(*) AS Expr1, Autor.prijmeni
FROM        Zbozi INNER JOIN
            Autor ON Zbozi.autor = Autor.id_autor
GROUP BY Zbozi.autor, Autor.prijmeni
```
- 3) U diagramu je potřeba nastavit Values, kterým přiřadíme sloupec Expr1 (nově vypočítané hodnoty), a Category Groups, které budou obsahovat sloupec prijmeni.
- 4) Report opět nahrajeme na Report Server, který spustíme v prohlížeči. Výsledný report může vypadat následovně, viz obrázek 27.

Obrázek 27 Tvorba reportu bez průvodce - diagram



zdroj: autorka

Tvorba reportu s průvodcem

- 1) Po spuštění SQL Server Data Tools zvolíme New Project – Reporting Services – Report Server Project Wizard a dáme OK.
- 2) Nyní je potřeba nastavit zdroj dat, takže stiskneme tlačítko Edit. Objeví se Connection Properties, kde nastavíme jméno serveru a vybereme databázi. Dáme OK a klikneme na Next.
- 3) Nyní je potřeba zadat dotaz, podle kterého bude vygenerován report. Dotaz lze začít psát přímo do nebo využít Query Builder.
- 4) V tomto reportu budou vypsána příjmení autorů, názvy jejich knih, nakladatelství, která je vydala a prodejní ceny. Dotaz bude vypadat následovně:

```
SELECT    Autor.prijmeni, Zbozi.nazev, Zbozi.cena_prodejni,  
Nakladatelstvi.nazev_nakladatelstvi
```

```
FROM      Autor INNER JOIN
```

```
          Zbozi ON Autor.id_autor = Zbozi.autor INNER JOIN
```

```
          Nakladatelstvi ON Zbozi.nakladatelstvi =
```

```
Nakladatelstvi.id_nakladatelstvi
```

- 5) Klikneme Next, vybereme požadované rozložení dat a znovu klikneme Next.

- 6) Nyní je potřeba zvolit hierarchii dat. V tomto případě můžeme příjmení autora zvolit jako nadřazený údaj a ostatní údaje jako podřízené.
- 7) U reportu lze zatrhnout možnost drill-down, která vytvoří rozbalovací nabídky.
- 8) Dále je možné nastavit požadovanou barvu výsledného reportu.
- 9) Následuje zvolení serveru, na který bude report nasazen a kliknutím na tlačítko Finish je report vytvořen.
- 10) Výsledný report si lze prohlédnout v záložce Preview.
- 11) Hotový report je nyní potřeba nahrát na Report Server.
- 12) Výsledný report by měl vypadat následovně, viz obrázek 28.

Obrázek 28 Report 1

Report1			
prijmeni	nazev	nazev nakladatels tví	cena prodejní
<input type="checkbox"/> Asimov			
	Nadace	Mladá fronta	239.0000
	Nadace 2	Mladá fronta	239.0000
<input type="checkbox"/> Deary			
	Děsivé dějiny 1	Egmont	129.0000
	Děsivé dějiny 2	Egmont	129.0000
<input type="checkbox"/> Flanagan			
<input type="checkbox"/> Pratchett			
	Barva kouzel	Talpress	199.0000
	Maškaráda	Talpress	199.0000
	Zaslaná pošta	Talpress	199.0000
	Podivný regiment	Talpress	210.0000
	Buch	Talpress	210.0000
<input type="checkbox"/> Rowling			
<input type="checkbox"/> Tolkien			
	Hobit	Mladá fronta	199.0000
	Pán prstenů 1	Mladá fronta	239.0000
	Pán prstenů 2	Mladá fronta	239.0000
	Pán prstenů 3	Mladá fronta	239.0000

zdroj: autorka

Úprava reportu

Výsledný report (viz obrázek 28) můžeme ještě dále upravit:

- 1) Přidání možnosti seřazení dat podle vybraného sloupce,
- 2) Podmíněné formátování.

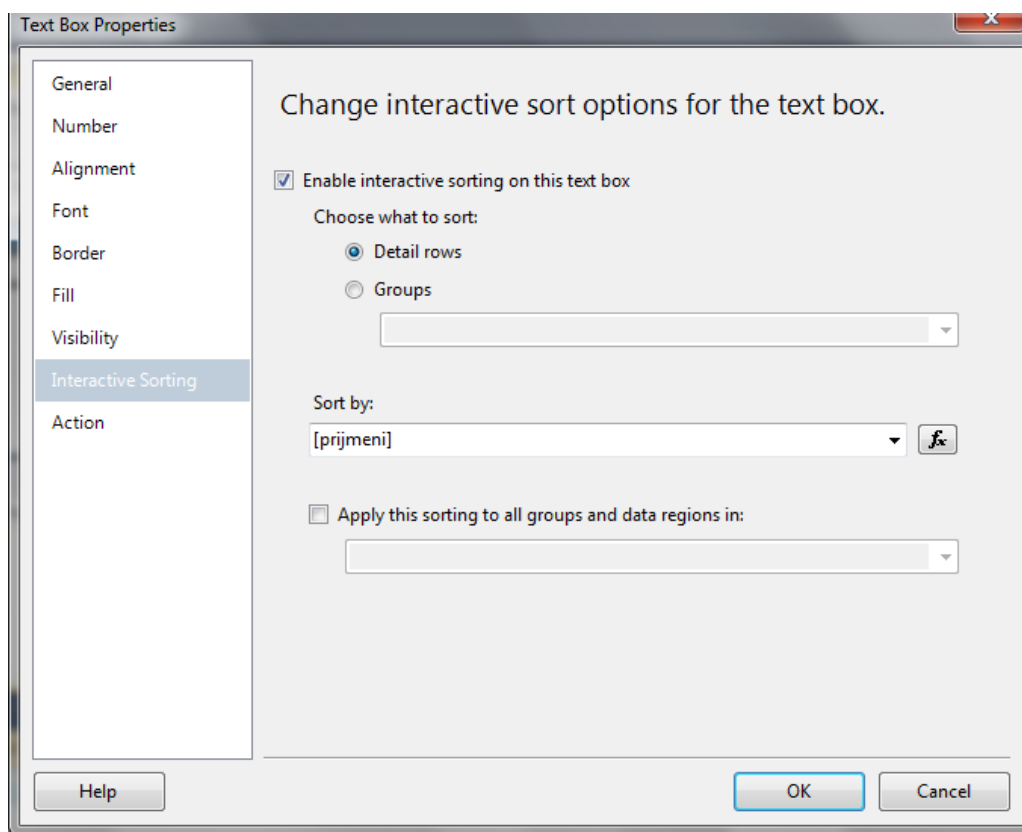
Přidání možnosti seřazení dat podle vybraného sloupce

- 1) Pomocí průvodce vytvoříme nový report.
- 2) Vybereme následující data: název knihy, příjmení autora, žánr, nakladatelství, rok a prodejní cenu. Použijeme následující dotaz:

```
SELECT      Autor.prijmeni, Nakladatelstvi.nazev_nakladatelstvi, Zanr.nazev
AS zanr, Zbozi.nazev, Zbozi.rok, Zbozi.cena_prodejni
FROM        Nakladatelstvi INNER JOIN
            Autor INNER JOIN
            Zbozi ON Autor.id_autor = Zbozi.autor ON
Nakladatelstvi.id_nakladatelstvi = Zbozi.nakladatelstvi INNER JOIN
            Zanr ON Zbozi.zanr = Zanr.id_zanr
```

- 3) Pravým tlačítkem myši klikneme na prijmeni ve vytvořené tabulce a zvolíme Text Box Properties.
- 4) Vybereme záložku Interactive Sorting, povolíme řazení a vybereme, podle čeho se budou data v tabulce řadit. V tomto případě půjde o hodnotu prijmeni. Dáme OK. Viz obrázek 29.
- 5) Tuto akci zopakujeme pro všechny ostatní sloupce.
- 6) Nyní lze seřadit řádky v tabulce podle jakéhokoli sloupce.

Obrázek 29 Interactive Sorting



zdroj: autorka

Podmíněné formátování

Pokud chceme některá data pro koncového uživatele zvýraznit, může využít podmíněného formátování.

- 1) Využijeme předchozí report, doplníme však do tabulky sloupec `mnozstvi_na_sklade`.
- 2) V reportu označíme celý bílý řádek, do kterého se načítají hodnoty z databáze.
- 3) V pravém menu Properties klikneme na položku `BackgroundColor` a zvolíme `Expression`.
- 4) Zde je potřeba zadat výraz, podle kterého se bude měnit formát buněk. Naším cílem bude změnit pozadí buněk, pokud je množství knihy na skladě rovno nule, na červené, a pokud je větší než pět na fialové.
- 5) Výraz bude vypadat následovně:

=Iif(Fields!mnozstvi_na_sklade.Value>5,"Plum",Iif(Fields!mnozstvi_na_sklade.Value >0,"Transparent","Red"))

Iff znamená podmínku, argumenty v závorce znamenají:

- Fields!mnozstvi_na_sklade.Value>5 – pole a jeho hodnota, která bude určovat formátování,
- "Plum" – barva, která bude použita, pokud bude podmínka splněna,
- třetí část představuje akci, která se provede, nebude-li podmínka splněna, v tomto případě je zde vložena druhá podmínka, neboť je zde více intervalů, které se porovnávají.

6) Výsledný report viz obrázek 30.

Obrázek 30 Podmíněné formátování

Report6

<div style="display: flex; justify-content: space-around; margin-bottom: 5px;"> 0 ks 1 - 5 ks >5 ks </div>					
příjmení	název	žánr	název nakladatelství	rok	množství na sklade
Tolkien	Hobit	fantasy	Mladá fronta	2002	5
Tolkien	Pán prstenů 1	fantasy	Mladá fronta	2001	6
Tolkien	Pán prstenů 2	fantasy	Mladá fronta	2001	8
Tolkien	Pán prstenů 3	fantasy	Mladá fronta	2001	9
Pratchett	Barva kouzel	fantasy	Talpress	1999	3
Pratchett	Maškaráda	fantasy	Talpress	2001	5
Pratchett	Zaslaná pošta	fantasy	Talpress	2002	3
Pratchett	Podivný regiment	fantasy	Talpress	2003	15
Pratchett	Buch	fantasy	Talpress	2002	0
Rowling	Harry Potter 1	fantasy	Albatros	2016	3
Rowling	Harry Potter 2	fantasy	Albatros	2016	2
Rowling	Harry Potter 3	fantasy	Albatros	2016	6
Rowling	Harry Potter 7	fantasy	Albatros	2016	2
Rowling	Prázdné místo	fantasy	Albatros	2016	10
Flanagan	Hraničářův učeň	fantasy	Egmont	2015	2
Flanagan	Hraničářův učeň 2	fantasy	Egmont	2015	6
Deary	Děsivé dějiny 1	naucny	Egmont	2001	1
Deary	Děsivé dějiny 2	naucny	Egmont	2002	3
Asimov	Nadace	sci-fi	Mladá fronta	2014	0
Asimov	Nadace 2	sci-fi	Mladá fronta	2014	6

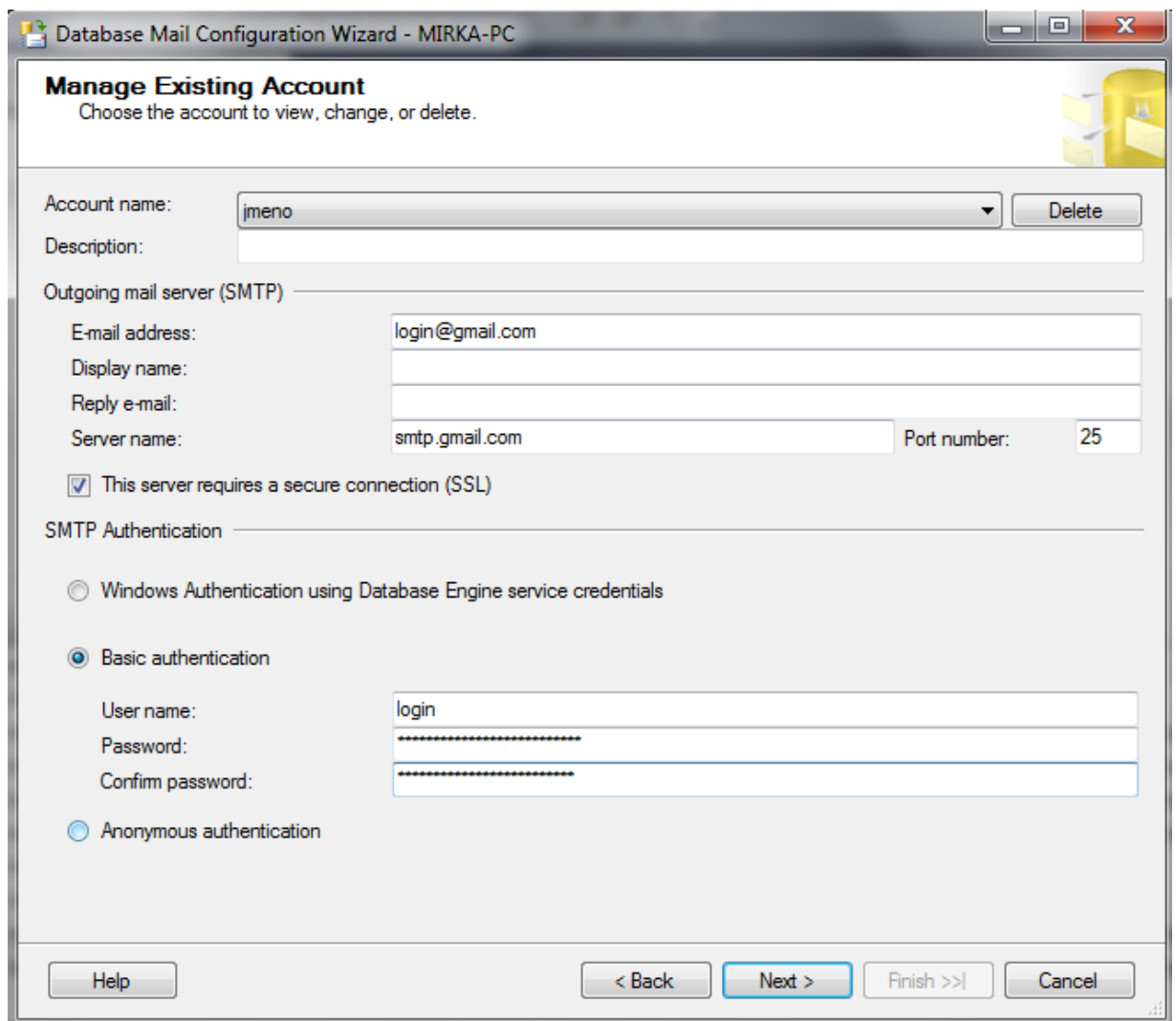
zdroj: autorka

4.3. Email

Pracovníci mohou potřebovat upozornění, že určitá úloha byla dokončena, proto je užitečné nastavit profily v SQL Server Management Studiu pro zasílání emailů. Nastavení proběhne přes:

1. Kliknutí na položku Management pravým tlačítkem,
2. kliknutí na položku Database Mail pravým tlačítkem,
3. kliknutí na Configure Database Mail a dojde ke spuštění průvodce.
4. Zde se ponechá zaškrtnutá možnost Set Up Database Mail.
5. Vyplní se jméno profilu a klikne se na tlačítko Add.
6. Objeví se následující formulář:

Obrázek 31 Průvodce konfigurací emailu



The screenshot shows the 'Manage Existing Account' dialog box in the Database Mail Configuration Wizard. The window title is 'Database Mail Configuration Wizard - MIRKA-PC'. The main heading is 'Manage Existing Account' with the instruction 'Choose the account to view, change, or delete.' The dialog contains several input fields and options:

- Account name:** A dropdown menu with the value 'jmeno' and a 'Delete' button.
- Description:** An empty text input field.
- Outgoing mail server (SMTP):**
 - E-mail address:** 'login@gmail.com'
 - Display name:** Empty
 - Reply e-mail:** Empty
 - Server name:** 'smtp.gmail.com'
 - Port number:** '25'
 - This server requires a secure connection (SSL)
- SMTP Authentication:**
 - Windows Authentication using Database Engine service credentials
 - Basic authentication
 - User name:** 'login'
 - Password:** Masked with dots
 - Confirm password:** Masked with dots
 - Anonymous authentication

At the bottom, there are navigation buttons: 'Help', '< Back', 'Next >', 'Finish >>', and 'Cancel'.

zdroj: autorka

7. Vyplní se jméno profilu.
8. Dále se vyplní emailová adresa a jméno serveru, které SQL server využije pro odeslání emailu.³
9. Zaškrtně se, že server požaduje bezpečné spojení.
10. V Basic authentication se jako uživatelské jméno uvede login a heslo emailového účtu.
11. Nový profil se objeví v nabídce, v následujícím okně se označí jako Public.
12. Následující okno obsahuje přehled nastavených omezení, např. maximální povolená velikost přílohy, zakázané přípony souborů v příloze, atd.
13. Po kliknutí na tlačítko Next následuje okno se souhrnem informací.
14. Kliknutí na tlačítko Finish; nenastanou-li nějaké nečekané komplikace, je profil založen.

Je-li potřeba některé údaje opravit, tak po krocích 1-3 následuje:

4. Vybrání možnosti Manage Database Mail accounts and profiles.
5. Vybrání možnosti View, change, or delete an existing account.
6. Zvolení profilu, který je třeba upravit a samotná úprava.

4.3.1. Odeslání emailu

Odeslání emailu proběhne následovně:

1. Kliknutí na položku Database Mail pravým tlačítkem,
2. Kliknutí na Send Test E-mail,
3. V následující tabulce se vyplní emailová adresa a klikne se na tlačítko Send.

4.3.2. Seznam odeslaných emailů

Chceme-li zjistit, jaké emaily byly odeslány, lze využít následující dotaz, který z tabulky `sysmail_sentitems` vypíše všechny emaily, nebo pokud je využita klauzule `WHERE` (jako v uvedeném příkladu), pouze emaily odeslané z vybraného účtu.

³ Na obrázku je použit účet `login@gmail.com` společnosti Google a jako poštovní server `smtp.gmail.com`. Skutečná společnost by jistě měla svůj poštovní server, takže by nevyužívala těchto služeb. Ve spojení s SQL Serverem by pravděpodobně šlo o další produkt společnosti Microsoft, tedy MS Exchange Server.

```
SELECT * FROM sysmail_sentitems WHERE profile_id=1
```

Podobně lze zobrazit nepovedená odeslání, např.:

```
SELECT * FROM sysmail_event_log
```

K tomuto přehledu se lze dostat i přes nabídku: Database Mail -> View Database Mail Log.

5. Závěr

Cílem této diplomové práce bylo vytvořit sbírku řešených příkladů pro výuku předmětu Business Intelligence na Ekonomické fakultě Jihočeské Univerzity. Za tímto účelem bylo využito MS SQL Serveru a nástroje MS Server Data Tools, které jsou na této fakultě využívány.

V této práci jsem se zaměřila zejména na popis vytvoření jednotlivých částí analytického procesu. Jednalo se o proces ETL, tvorbu multidimenzionální kostky a vytváření reportu. V jednotlivých úlohách byly představeny možnosti, které MS Server Data Tools pro jednotlivé úlohy nabízí.

V praxi je proces Business Intelligence poměrně komplexní problém, a proto jsem se rozhodla využít ve své práci zjednodušený model datového skladu, který studentům umožní snadnější pochopení dané problematiky a vyzkoušení jednotlivých úloh.

Podle mého názoru jsou informační systémy a s nimi spojené nástroje Business Intelligence stále důležitější pro úspěšné fungování většiny podniků. Studenti Ekonomické fakulty by tedy měli být s touto problematikou obeznámeni a měli by být schopni vytvářet alespoň základní reporty. K osvojení těchto znalostí jim může posloužit právě má práce, která obsahuje názorné postupy.

Summary

The subject of the diploma thesis titled, "Collection of Solved Problems of Business Intelligence for Students of the Faculty of Economics" is a creation and description of a series of solved problems of Business Intelligence.

The theoretical part is focused on basic terms associated with Business Intelligence. In the practical part, there is showed the usage of MS SQL Server and MS Server Data Tools for creation different parts of the Business Intelligence process.

Keywords:

Business Intelligence, SQL Server, ETL, Reporting

Zdroje

1. AGARWAL, V. V. a HUDDLESTON J. (2009) *Databáze v C# 2008: průvodce programátora*. (1st ed.). Brno, Czechia: Computer Press.
2. GÁLA, L., POUR, J. a ŠEDIVÁ, Z. (2009) *Podniková informatika. 2., přeprac. a aktualiz. vyd.* (2nd ed.) Praha, Czechia: Grada publishing.
3. HERSCHEL, R. T. (2012) *Organizational applications of business intelligence management: emerging trends* Hershey, PA: Business Science Reference,
Dostupné z:
<https://books.google.cz/books?id=FdNJhIMUxFEC&printsec=frontcover&vq=Organizational+Applications+of+Business+Intelligence+Management:+Emerging&hl=cs#v=onepage&q&f=false>
4. HUMPHRIES, M. (2002) *Data warehousing: návrh a implementace*. (1st ed.). Praha, Czechia: Computer Press,
5. KIMBALL, R. (1998) *The data warehouse lifecycle toolkit: expert methods for designing, developing, and deploying data warehouses*. New York: Wiley.
6. KUČEROVÁ H. (2004) *Databázové systémy*. Praha, Czechia: Vyšší odborná škola informačních služeb. Retrieved from: info.sks.cz/users/ku/DOKUMENTY/das_syl.pdf
7. LABERGE, R. (2012) *Datové sklady: agilní metody a business intelligence*. (1st ed.). Brno, Czechia: Computer Press.
8. LACKO, L. (2003) *Databáze: datové sklady, OLAP a dolování dat s příklady v Microsoft SQL Serveru a Oracle: agilní metody a business intelligence*. (1st ed.) Brno, Czechia: Computer Press.
9. LANS, R. F. (2012) *Data virtualization for business intelligence systems: revolutionizing data integration for data warehouses*. Amsterdam: Elsevier, Morgan Kaufmann.
10. NOVOTNÝ, O. (2005) *Business intelligence: jak využít bohatství ve vašich datech*. (1st ed.) Praha, Czechia: Grada.
11. POKORNÝ, J. (2012) *Moderní databáze 2012 Architektura informačních systémů - NoSQL databáze - Současný stav vývoje*. Praha, Czechia: KOMIX.
12. VERCELLIS, C. (2009) *Business intelligence: data mining and optimization for decision making*. (1st ed.) Chichester, United Kingdom: Wiley,.

Elektronické zdroje:

1. Jake Freivald – *EAI – Enterprise Application Integration*. *InformationBuilders*. (©2016) Retrieved from: <http://www.informationbuilders.com/eai-enterprise-application-integration>
2. Milan Kučera – *Dva způsoby budování datového skladu*. (©2001) *SystemOnline*. Retrieved from: <http://www.systemonline.cz/clanky/dva-zpusoby-budovani-datoveho-skladu.htm>
3. Joe Oates - *When do you use a star schema and when to use a snowflake schema?* (©2007) *InformationManagement*. Retrieved from: http://www.information-management.com/news/ask_the_experts/-10000243-1.html
4. Michael Petrášek - *Business intelligence a reporting nejen pro krizové období*. (©2016) *SystemOnline*. Retrieved from: <http://www.systemonline.cz/business-intelligence/business-intelligence-a-reporting-nejen-pro-krizove-obdobi.htm>
5. Jaromír Skřivan – *Databáze a jazyk SQL*. (©2000) *Interval* Retrieved from: <https://www.interval.cz/clanky/databaze-a-jazyk-sql/>
6. *Business Intelligence Reporting (BI Reporting)*. (©2016) *Techopedia*. Retrieved from: <https://www.techopedia.com/definition/30217/business-intelligence-reporting-bi-reporting>
7. Milena Tvrđíková - *EIS - nezbytná součást business intelligence* (© 2016) *SystemOnline*, Retrieved from: <http://www.systemonline.cz/clanky/eis-nezbytna-soucast-business-intelligence.htm>
8. ? : (Conditional) (SSIS Expression). (©2016) *Microsoft*. Retrieved from: <https://msdn.microsoft.com/en-us/library/ms141680.aspx>
9. *Decision Support System (DSS)*. (©2016) *Techopedia*. Retrieved from: <https://www.techopedia.com/definition/770/decision-support-system-dss>
10. *Magic Quadrant for Business Intelligence and Analytics Platforms* (© 2015) Retrieved from: <http://www.gartner.com/technology/reprints.do?id=1-2ADAAYM&ct=150223&st=sb>
11. *OldanyGroup* – *OLTP* (©2016) Retrieved from: <http://www.oldanygroup.cz/index-stranek-115/oltp/>

12. *Příručka Začínáme s Power BI* (©2015) Retrieved from:
<https://support.office.com/cs-cz/article/Příručka-Začínáme-s-Power-BI-bd30711a-7ccf-49e8-aafa-2e8f481e675d?ui=cs-CZ&rs=cs-CZ&ad=CZ>
13. *Snowflake Schema vs. Star Schema* (©2012) SearchBusinessIntelligence. Retrieved from: <http://searchbusinessintelligence.techtarget.in/answer/Star-schema-vs-snowflake-schema-Which-is-better>
14. *Star schema vs. snowflake schema: Which is better?* Diffen. Retrieved from: http://www.diffen.com/difference/Snowflake_Schema_vs_Star_Schema
15. *WHAT IS EXPERT SYSTEM?* (©2010) World of Computing. Retrieved from: <http://intelligence.worldofcomputing.net/ai-branches/expert-systems.html#.VsweR9xsiw1>
16. *Základy normalizace databáze.* (©2015) Microsoft. Retrieved from: <https://support.microsoft.com/cs-cz/kb/283878>

Seznam obrázků

Obrázek 1 Datový model Sněhová vločka.....	10
Obrázek 2 Datový model Hvězda.....	11
Obrázek 3 Části systému Business Intelligence	13
Obrázek 4 Hlavní komponenty Business Intelligence a jejich vazby.....	14
Obrázek 5 Normalizovaný datový model	25
Obrázek 6 Schéma Hvězda	26
Obrázek 7 Schéma Sněhová vločka.....	27
Obrázek 8 Schéma databáze ProdejDW	29
Obrázek 9 Tabulka Zakaznik.....	30
Obrázek 10 Propojené prvky Data Flow Task a Derived Column	32
Obrázek 11 Zmenšení všech písmen	34
Obrázek 12 Nahrazení vybraného výrazu.....	34
Obrázek 13 Mappings.....	35
Obrázek 14 Hotové Data Flow	36
Obrázek 15 Naplněná tabulka Zakaznik.....	36
Obrázek 16 Původní data.....	37
Obrázek 17 Conditional Split	38
Obrázek 18 Vypsání počtu všech nákupů.....	39
Obrázek 19 Rozdělení prodejů podle pohlaví zákazníka.....	40
Obrázek 20 Rozdělení prodejů podle pohlaví zákazníka a roku	40
Obrázek 21 Vyfiltrování prodejů, které proběhly v roce 2015.....	41
Obrázek 22 Vyfiltrování prodejů, které uskutečnily ženy a které proběhly v 2. čtvrtletí roku 2015	41
Obrázek 23 Naplnění tabulky daty	43
Obrázek 24 SQL Server Reporting Services	44
Obrázek 25 Tvorba reportu bez průvodce 1	44
Obrázek 26 Tvorba reportu bez průvodce 1 - upraveno	45
Obrázek 27 Tvorba reportu bez průvodce - diagram.....	46
Obrázek 28 Report 1	47
Obrázek 29 Interactive Sorting.....	49
Obrázek 30 Podmíněné formátování	50
Obrázek 31 Průvodce konfigurací emailu	51

Seznam tabulek

Tabulka 1 Rozdíly mezi OLTP a OLAP systémy.....	17
Tabulka 2 Rozdíly mezi modelem hvězdy a sněhové vločky.....	28