



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INFORMAČNÍCH SYSTÉMŮ

DEPARTMENT OF INFORMATION SYSTEMS

OLAP ANALÝZA ROZSÁHLÝCH DAT

OLAP ANALYSIS OF BIG DATA

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

DALIBOR KRÁLÍK

VEDOUCÍ PRÁCE

SUPERVISOR

prof. Ing. TOMÁŠ HRUŠKA, CSc.

BRNO 2023

Zadání bakalářské práce



143929

Ústav: Ústav informačních systémů (UIFS)
Student: **Králik Dalibor**
Program: Informační technologie
Specializace: Informační technologie
Název: **OLAP analýza rozsáhlých dat**
Kategorie: Informační systémy
Akademický rok: 2022/23

Zadání:

1. Prostudujte teoretické základy teorie datových kostek. Seznamte se systémy SAP BW/Analysis.
2. Navrhněte zdroj dat, postup ETL a datovou kostku pro vybranou oblast dat VUT.
3. Implementujte návrh struktur zdrojových dat, proces ETL a strukturu datové kostky k bodu 2.
4. Proveďte ověření funkčnosti datové kostky pro zvolené analytické výstupy a proveďte zhodnocení výsledků.

Literatura:

- Přednášky Pokročilé informační systémy z magisterského studia (Hruška)
- Webinář na téma Datové kostky (Hruška)
- Manuály SAP pro SAP BW a SAP Analysis

Při obhajobě semestrální části projektu je požadováno:
Body 1. a 2.

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Hruška Tomáš, prof. Ing., CSc.**
Vedoucí ústavu: Kolář Dušan, doc. Dr. Ing.
Datum zadání: 1.11.2022
Termín pro odevzdání: 10.5.2023
Datum schválení: 21.10.2022

Abstrakt

Cielom tejto bakalárskej práce je vytvorenie troch OLAP dátových kociek, ktoré budú riešiť problematiku zahraničných ciest študentov a zamestnancov VUT. Primárnym účelom dátových kociek je rýchle poskytnutie správnych dát pre analytikov VUT. Na ich modelovanie a implementáciu je použitý softvér SAP Business Warehouse a na testovanie softvér SAP Analysis for Microsoft Excel. Po úspešnom návrhu, implementácii a testovaní disponuje VUT týmito tromi OLAP dátovými kockami na produkčnom a na vývojovom serveri SAP Business Warehouse, ku ktorým majú prístup analytici VUT. Vďaka tejto práci a dátovým kockám zameraným na zahraničné cesty majú analytici a pracovníci VUT zastávajúci manažérske role získať rýchly prístup k dátam a analýzám, ktoré slúžia na podporu správnych rozhodnutí spojených s fungovaním VUT.

Abstract

The goal of this bachelor's thesis is the creation of three OLAP data cubes, which will solve the problem of foreign trips of students and employees of BUT. The primary purpose of data cubes is to quickly provide correct data for BUT analysts. SAP Business Warehouse software is used for their modeling and implementation, and SAP Analysis for Microsoft Excel software is used for testing. After successful design, implementation and testing, BUT has these three OLAP data cubes on the production and development servers of SAP Business Warehouse, to which BUT analysts have access. Thanks to this work and data cubes focused on foreign trips, analysts and BUT employees in managerial roles have the opportunity to quickly access data and analyzes that serve to support correct decisions related to the functioning of BUT.

Klíčové slová

OLAP, ETL, dátové sklady, SAP, SAP Business Warehouse, analýza, SAP Analysis for Microsoft Office, dátové kocky, extrakcia, transformácia, načítanie dát, databáza, zahraničné cesty, databázový pohľad, relačná databáza, multidimenzionálna databáza, dimenzia

Keywords

OLAP, ETL, data warehouse, SAP, SAP Business Warehouse, analysis, SAP Analysis for Microsoft Office, data cubes, extraction, transformation, data loading, database, work trips abroad, database view, relational database, multidimensional database, dimension

Citácia

KRÁLÍK, Dalibor. *OLAP analýza rozsáhlých dat*. Brno, 2023. Bakalárska práca. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce prof. Ing. Tomáš Hruška, CSc.

OLAP analýza rozsáhlých dat

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením pána prof. Ing. Tomáša Hrušku, CSc.

Uviedol som všetky literárne pramene, publikácie a ďalšie zdroje, z ktorých som čerpal.

.....
Dalibor Králik
12. apríla 2023

Podakovanie

Moje podakovanie patrí vedúcemu bakalárskej práce pánovi prof. Ing. Tomášovi Hruškovi, CSc. za odbornú pomoc, vedenie a správne smerovanie pri vypracovaní práce. Podakovanie taktiež patrí Centru výpočetných a informačných služieb VUT za možnosť vypracovať bakalársku prácu so softvérom a dátami, ktoré mi poskytli a pánovi doc. Ing. Radkovi Kubáskovi Ph.D. za pomoc pri tvorbe analýz.

Obsah

1	Úvod	6
2	BI – Business Intelligence	7
2.1	Informačné systémy z pohľadu úrovne rozhodovania	7
2.2	Manažér vs. analytik	8
3	Dátové sklady	9
3.1	Vlastnosti dát v dátových skladoch	9
3.2	Dátový trh	10
3.3	Architektúra dátového skladu	10
3.4	ETL – extract, transform, load	11
3.4.1	Extrakcia	11
3.4.2	Transformácia	11
3.4.3	Nahrávanie	11
4	OLAP – Online Analytical Processing	12
4.1	MIS – manažérske informačné systémy	12
4.2	Tvorba analýzy pomocou OLAP	12
4.3	OLTP – Online Transaction Processing	13
4.4	Rozdiel medzi OLAP a OLTP	13
4.5	Multidimenzionálne objekty	14
4.5.1	Dimenzie	15
4.5.2	Ukazovatele	15
4.6	Dátová kocka	16
4.7	OLAP operácie	16
4.8	SAP Business Warehouse	18
4.9	SAP Analysis for Microsoft Office	19
4.10	SAP Business Objects	19
5	Modelovanie dátových skladov v SAP Business Warehouse	20
5.1	Tok vývoja dátových skladov a dátových kociek	20
5.2	Dátový zdroj – DataSource Object	23
5.2.1	PSA – Persistent Staging Area	23
5.3	Charakteristika	23
5.4	Ukazovateľ – key figure	23
5.5	Tvorba dátovej kocky	24
5.6	Transformácia	24
5.6.1	Transformácia v SAP BW	26

5.7	Nahrávanie dát do dátovej kocky	26
6	Problematika zahraničných ciest na VUT	27
6.1	Súčasný stav	27
6.2	Riešenie	28
6.2.1	Tri OLAP dátové kocky	29
6.3	Výročná správa o činnosti školy od Ministerstva školstva, mládeže a telovýchovy Českej republiky	30
6.3.1	Zapojenie vysokej školy do programov medzinárodnej spolupráce	30
6.3.2	Mobilita študentov, akademických a ostatných pracovníkov podľa krajiny	31
6.3.3	Mobilita absolventov	32
7	Návrh riešenia	33
7.1	Databázové pohľady	33
7.1.1	Pohľad pre kocku <i>Zahraniční studenti Příjezd</i>	33
7.1.2	Pohľad pre kocku <i>Zahraniční cesty zaměstnanci Příjezd</i>	34
7.1.3	Pohľad pre kocku <i>Zahraniční cesty Výjezd</i>	34
7.2	Návrh dátových kociek	36
7.2.1	Návrh charakteristík	37
7.2.2	Návrh ukazovateľov	37
7.3	Návrh transformácie	38
8	Implementácia	39
8.1	Tvorba dátového zdroja – DataSource	39
8.1.1	Dátový zdroj <i>Zahraniční cesty zaměstnanci Příjezd</i>	39
8.1.2	Dátový zdroj <i>Zahraniční cesty Výjezd</i>	39
8.1.3	Dátový zdroj <i>Zahraniční studenti Příjezd</i>	40
8.1.4	Dátový zdroj <i>Zahraniční cesty typ</i>	40
8.2	Nahrávanie a správa dát v PSA	40
8.3	Tvorba charakteristík a ukazovateľov	40
8.3.1	Charakteristiky pre kocku <i>Zahraniční cesty zaměstnanci Příjezd</i>	41
8.3.2	Charakteristiky pre kocku <i>Zahraniční cesty Výjezd</i>	42
8.3.3	Charakteristiky pre kocku <i>Zahraniční studenti Příjezd</i>	43
8.3.4	Tvorba atribútov a číselníkov charakteristík	43
8.3.5	Tvorba ukazovateľov	44
8.4	Modelovanie dátových kociek	44
8.4.1	Modelovanie dátovej kocky <i>Zahraniční cesty zaměstnanci Příjezd</i>	44
8.4.2	Modelovanie dátovej kocky <i>Zahraniční cesty Výjezd</i>	44
8.4.3	Modelovanie dátovej kocky <i>Zahraniční studenti Příjezd</i>	46
8.5	Implementácia transformácie	46
8.5.1	Transformácia kocky <i>Zahraniční cesty zaměstnanci Příjezd</i>	46
8.5.2	Transformácia kocky <i>Zahraniční cesty Výjezd</i>	47
8.5.3	Transformácia kocky <i>Zahraniční studenti Příjezd</i>	48
8.5.4	Transformácia atribútov charakteristiky <i>Typ zahraničnej cesty</i>	49
8.5.5	Transformácia číselníkov	50
8.6	Nahrávanie dát z PSA do dátových kociek	50
8.7	SAP BW Query pre dátové kocky	51

9 Testovanie	52
9.1 Analysis for MS Excel	52
9.2 Testovanie kocky <i>Zahraniční cesty zaměstnanci Příjezd</i>	54
9.3 Testovanie kocky <i>Zahraniční cesty Výjezd</i>	56
9.4 Testovanie kocky <i>Zahraniční studenti Příjezd</i>	58
10 Závěr	58
Literatúra	59

Zoznam obrázkov

2.1	Pyramídová schéma informačného systému (zdroj: [1, sl. 38])	7
3.1	Typická architektúra dátového skladu (zdroj: [9])	10
4.1	Princíp fungovania OLAP technológie (zdroj: [7])	13
4.2	3-dimenzionálna dátová kocka	14
4.3	Slice operácia nad dátovou kockou	17
4.4	Dice operácia nad dátovou kockou	18
5.1	Tok vývoja v SAP BW	22
5.2	Ikonka priameho priradenia v transformácii v SAP BW	25
5.3	Ikonka priradenia pomocou konštanty v transformácii v SAP BW	25
5.4	Ikonka priradenia pomocou vzorca v transformácii v SAP BW	25
5.5	Ikonka priradenia pomocou rutiny v transformácii v SAP BW	26
5.6	Tok vývoja v SAP BW	26
6.1	Tabuľka zapojenia vysokej školy do programov medzinárodnej spolupráce .	30
6.2	Záhlavie tabuľky mobility študentov, akademických a ostatných pracovníkov podľa krajiny	31
6.3	Tabuľka mobility absolventov	32
8.1	Transformácia kocky <i>Zahraniční cesty zaměstnanci Příjezd</i>	47
8.2	Transformácia kocky <i>Zahraniční cesty Výjezd</i>	48
8.3	Transformácia kocky <i>Zahraniční studenti Příjezd</i>	49
8.4	Transformácia atribútov do charakteristiky <i>Typ zahraniční cesty</i>	49
8.5	Transformácia číselníka charakteristiky <i>Typ stipendia ID</i>	50
8.6	Transformácia číselníka charakteristiky <i>Stav stipendia ID</i>	50
8.7	Transformácia číselníka charakteristiky <i>Druh stipendia ID</i>	50
9.1	Analýza príchodu zahraničných zamestnancov	53
9.2	Analýza všetkých vyslaných študentov VUT do zahraničia	55
9.3	Analýza absolventských stáží študentov VUT vyslaných do zahraničia . . .	56
9.4	Analýza prichádzajúcich študentov zo zahraničia	57

Zoznam tabuliek

7.1	Návrh pohľadu pre kocku <i>Zahraniční studenti Příjezd</i>	33
7.2	Návrh pohľadu pre kocku <i>Zahraniční cesty zaměstnanci Příjezd</i>	34
7.3	Návrh pohľadu pre kocku <i>Zahraniční cesty Výjezd</i>	36
8.1	Charakteristiky pre kocku <i>Zahraniční cesty zaměstnanci Příjezd</i>	41
8.2	Charakteristiky pre kocku <i>Zahraniční cesty Výjezd</i>	42
8.3	Charakteristiky pre kocku <i>Zahraniční studenti Příjezd</i>	43
8.4	Atribúty charakteristiky <i>Typ zahraničnej cesty</i>	43
8.5	Ukazovatele	44
8.6	Charakteristiky, ktoré patria do zoskupenia <i>Studenti</i>	45
8.7	Charakteristiky, ktoré patria do zoskupenia <i>Zaměstnanci</i>	45
8.8	Charakteristiky, ktoré patria do všeobecného zoskupenia	46

Kapitola 1

Úvod

Cieľom tejto bakalárskej práce je spracovanie veľkého rozsahu dát pomocou OLAP dátových skladov a dátových kociek. Po uvedení do problematiky sa v tejto práci zaoberám teoretickými a praktickými aspektmi tvorby dátových kociek pomocou softvéru SAP Business Warehouse (SAP BW) od spoločnosti SAP, ktorý je dostupný na oddelení Centra výpočetných a informačných služieb VUT (CVIS).

Svoj návrh a riešenie som realizoval s dátami zahraničných ciest akademických a technických pracovníkov VUT a študentov VUT. Tieto dáta mi poskytol CVIS pre realizáciu tejto práce. Výstupom práce budú tri dátové kocky, ktoré sa budú nachádzať na vývojovom (VBW) a aj produktívnom systéme (VBP) SAP BW na CVISe. Tieto dátové kocky budú následne používať analytici VUT pri vyplňaní výročnej správy o činnosti školy od Ministerstva školstva, mládeže a telovýchovy České republiky (MŠMT ČR). Po transporte dátových kociek na produktívny systém budú mať analytici možnosť dané kocky využívať a zobrazíť si v nich dáta pomocou zobrazovacieho softvéru *SAP BusinessObjects* alebo rozšírenia do programu MS Excel s názvom *SAP Analysis for Microsoft Office*.

Vzhľadom na charakter uloženia dát v centrálnej databáze VUT som bol nútený vytvoriť tri dátové kocky. Každá z nich bude obsahovať iný druh dát. Povaha dát v jednotlivých kockách je nasledovná:

- Dáta prichádzajúcich zahraničných študentov
- Dáta prichádzajúcich zahraničných akademických a technických pracovníkov
- Dáta vyslaných zahraničných akademických, technických pracovníkov a študentov

Pri vývoji dátových kociek a spracovaní dát som dbal na „good practices“ modelovania ETL pomocou softvéru SAP BW. Celý výskum a vývoj je popísaný v kapitolách tejto práce spoločne aj s výsledkami testovania a popisom validácie spracovaných dát.

Kapitola 2

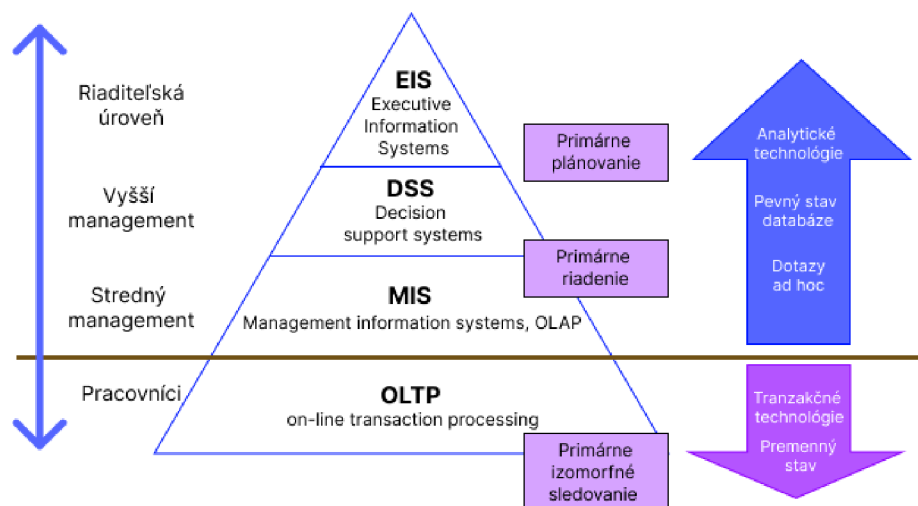
BI – Business Intelligence

Business intelligence (BI) je zastrešujúci termín, ktorý sa vzťahuje k znalostiam, procesom, technológiám, aplikáciám a postupom, ktoré uľahčujú podnikové rozhodovanie. Technológia *Business intelligence* pracuje s použitými (historickými) dátami v požadovanom kontexte a pomáha prijímať podnikové rozhodnutia pre budúcnosť. [6, strana 24]

Zameriava sa hlavne na interné informácie, ktoré súvisia so strategickým a taktickým plánovaním. *Business intelligence* pomáha previesť nespracované (raw) dáta do spracovanej a zmyslupnejšej podoby tak, aby sa nad dátami dali robiť efektívnejšie analýzy podnikových dát. Po spracovaní poskytujú výstupy analytikom rôzne možnosti pohľadu na dáta pomocou agregácií, operácií nad dátovými skladmi (data warehouse) a podobne.

2.1 Informačné systémy z pohľadu úrovne rozhodovania

Informačné systémy sú rozdelené do viacerých úrovní. Na každej z týchto úrovní sa nachádzajú iné formy dát, iné procesy, sledujú sa iné dátové ukazovatele a vykonávajú sa iné úkony. Najvýstižnejšie tieto úrovne rozhodovania znázorňuje Pyramídová schéma informačného systému 2.1:



Obr. 2.1: Pyramídová schéma informačného systému (zdroj: [1, sl. 38])

Vzhľadom na tému tejto práce sa budem zaoberať predovšetkým úrovňou MIS – manažerských informačných systémov, ktorých súčasťou je aj technológia OLAP.

2.2 Manažér vs. analytik

Pre správne pochopenie tvorby a využitia *Business Intelligence* v podniku a dátových skladov je dôležité, aby sme si objasnili role osôb, ktoré budú spracované dáta využívať vo svojej náplni práci a pri rozhodovaní podnikového riadenia.

V podnikových informačných systémoch máme prevažne dve role, ktoré využívajú výstupy BI – analytik a manažér.

Manažér sleduje dané výsledky, zhodnocuje na základe získaných dát stav podniku, analyzuje, či sa podnik svojimi výsledkami blíži k splneniu stanovených cieľov, rozhoduje o budúcich krokoch podniku, získava prehľad o dianí.

Analytik analyzuje dáta a pripravuje dátové výkazy pre manažéra. [3, sl. 3]

Kapitola 3

Dátové sklady

Jedným zo základných komponentov *Business Intelligence* je dátový sklad. Dátový sklad (Data Warehouse) je systém pre dátovú analýzu a reporting, ktorý umožňuje zhromažďovať, organizovať, uchovávať a zdieľať historické dáta pozri [6].

Dáta do systémov dátových skladov sú mnohokrát získané z rôznych dátových zdrojov, ako napríklad z databázy, externých zdrojov a logovacích súborov, následne sú transformované, spracované – ETL (extract, transform, load) a uložené v dátových skladoch. Na tieto dátové sklady sa potom užívatelia (analytici) dotazujú a dátový sklad im poskytne požadované dáta.

3.1 Vlastnosti dát v dátových skladoch

Vzhľadom na charakter systému je zrejmé, že Data Warehouse systémy pracujú s veľkým množstvom dát. Zjednodušene je možné uvažovať, že dátový sklad je uložená kolekcia spracovaných dát. [9]

Podľa odbornej literatúry [9] majú dáta v dátovom sklade nasledujúce vlastnosti: dáta sú subjektovo orientované, integrované, časovo premenné a nevolatilné.

- **Subjektovo orientované** – dáta zamerané na analytické potreby rôznych oblastí v podniku. Tieto oblasti sa líšia v závislosti od druhu činností, ktoré organizácia vykonáva.
- **Integrované** – dáta získané z viacerých zdrojov musia byť spojené dohromady, čo znamená riešenie problémov spojených s rôznou definíciou a obsahom dátových polí z rôznych zdrojov. Hlavným problémom sú rozdiely v dátových typoch, kódovaní, v možnosti existencie synonymických dát, ako napríklad polí s rozdielnym menom, ale s rovnakým obsahom dát, mnohonásobným výskytom rovnakých dát a podobne.
- **Časovo premenné** – dáta sa v rámci času môžu v dátových skladoch meniť, to znamená, že sa dáta z hľadiska času v dátových skladoch ukladajú sekvenčne, a tak reprezentujú určitý časový úsek. Dobrým príkladom je ukladanie dát o výške priemerného mesačného zostatku klientov v banke. Údaje v tomto príklade reprezentujú časový úsek jedného roka, kedy napríklad klient s číslom ID 68748 mal priemerný mesačný zostatok za rok 2022 vo výške 4 587 €.
- **Nevolatilné** – trvanlivosť dát. Je zabezpečená zakázaním modifikácie a mazaním dát. To znamená, že staršie, neaktuálne dáta sa z dátového skladu neodstraňujú ani

nemenia, ale nachádzajú sa v dátovom sklade stále. Toto umožňuje nahliadať na dáta aj z historického hľadiska, tvoriť analýzy, v ktorých je možné porovnať súčasné výsledky s tými historickými, určiť napredovanie podniku a zhodnotiť súčasné, ale aj historické výkony.

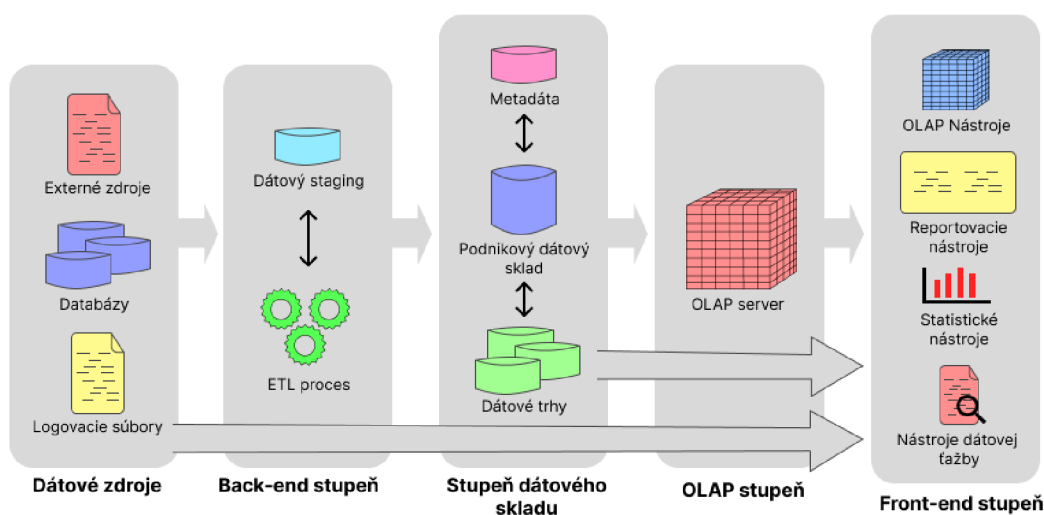
3.2 Dátový trh

Dátový trh je v podstate jednoduchšia forma dátového skladu zameraná na jedno konkrétne oddelenie v podniku. Dá sa povedať, že dátový trh je podmnožinou dátového skladu.

Pomocou dátových trhov môžu tímy analytikov pristupovať ku konkrétnym požadovaným dátam rýchlejšie vďaka rozdeleniu podľa oddelení podniku a nemusia tak stráviť čas hľadaním požadovaných dát v komplexných dátových skladoch, ktoré obsahujú dáta celého podniku.

3.3 Architektúra dátového skladu

Typická architektúra systému dátového skladu podľa odbornej literatúry [9, str. 77] zahrňuje 4 stupne, ktoré na seba nadväzujú.



Obr. 3.1: Typická architektúra dátového skladu (zdroj: [9])

1. **Back-end stupeň** – je stupeň, v ktorom prebieha **ETL** – **extract, transform, load** proces (sekcia 3.4). Tento stupeň slúži na naplnenie dát do dátového skladu z rôznych zdrojov. Obsahuje taktiež data staging oblasť, ktorá slúži ako prechodná databáza, v ktorej prebiehajú procesy transformácie a integrácie pred vkladaním dát do dátových skladov. [9, str. 76]
2. **Stupeň dátového skladu** – tento stupeň znázorňuje uloženie dát do podnikového **dátového skladu** a do viacerých **dátových trhov**. Ako už bolo vyššie spomenuté, podnikový dátový sklad obsahuje a reprezentuje všetky dáta podniku. Dátový trh je podmnožinou dátového skladu a obsahuje dáta zamerané len na určitú oblasť podniku – je teda podmnožinou dátového skladu.

Posledným komponentom tohto stupňa sú **metadáta**. Všeobecne je známe, že metadáta sú dáta o dátach. Tieto metadáta sa v Data Warehouse rozdeľujú ešte na dva druhy. Prvým druhom sú **business metadáta**, ktoré opisujú význam alebo sémantiku dát, väzby medzi dátami a podobne. Ďalším druhom sú **technické metadáta**, ktoré opisujú to, ako sú dáta štruktúrované a ukladané, a popisujú aplikácie a procesy, ktoré s takýmito dátami manipulujú [9, str. 78].

3. **OLAP stupeň** – pozostáva z OLAP servera, ktorý užívateľom v podniku (väčšinou analytikom a manažérom) sprístupňuje **multidimenzionálne objekty** (sekcia 4.5) s dátami z dátových skladov alebo dátových trhov.
4. **Front-end stupeň** – nachádzajú sa tu klientske nástroje, ktoré umožňujú užívateľom využívať obsah dátových skladov a dátových trhov. Tieto klientske nástroje zahŕňajú nástroje, ako napríklad **OLAP nástroje**, ktoré povolia využívať a manipulovať s dátovým skladom pomocou OLAP technológie, **reportovacie nástroje**, ktoré slúžia na vytvorenie reportov a záverečných správ s dátami podniku pre vyšší manažment podniku, **štatistické nástroje** používané hlavne analytikmi pri tvorbe analýz podniku a **nástroje na ťažbu dát**, ktoré sa využívajú hlavne na objavenie hodnotných informácií, ako napríklad trendov.

3.4 ETL – extract, transform, load

ETL je skratka pre dátovú extrakciu (extract), transformáciu dát (transform) a nahrávanie dát (load). Ako je znázornené na obrázku č. 3.1, ETL procesy sa nachádzajú v back-end stupni architektúry systému dátového skladu. Celou úlohou ETL je zozbieranie, transformovanie a nahrávanie dát z rôznych dátových zdrojov a ich uloženie do jedného dátového skladu (viacerých dátových trhov) [10].

3.4.1 Extrakcia

Tento proces zhromažďuje dáta z rôznych rozdielnych zdrojov dát, ako napríklad logovacie súbory, databáza, externé zdroje a podobne. Mnohokrát majú dáta rozdielny formát.

3.4.2 Transformácia

Počas procesu transformácie sa modifikuje formát dát z dátových zdrojov do formátu dátového skladu. Tento proces zahŕňa aj rôzne operácie s dátami, ako napríklad *čistenie dát* – odstraňovanie nepožadovaných dát, *konverzia dátových typov*, *konverzia formátu dát*, *integrácia dát* a *agregácia dát* – spracovanie dát získaných z dátových zdrojov na základe úrovne zanorenia v dátovom sklade.

3.4.3 Nahrávanie

Napĺňa dátový sklad (dátové trhy) spracovanými, transformovanými dátami.

Kapitola 4

OLAP – Online Analytical Processing

OLAP – Online Analytical Processing systémy umožňujú užívateľom interaktívne žiadať a automaticky využívať agregačné operácie s dátami v dátových skladoch.

Ako uvádza odborná literatúra, týmto spôsobom môžu osoby s rozhodovacou právomocou ľahko pristupovať k požadovaným informáciám a analyzovať ich na rôznych úrovniach detailov. [9]

OLAP technológia dokáže pomôcť užívateľom rýchlo a jednoducho analyzovať veľké množstvo dát z rôznych dátových zdrojov, ako napríklad finančné dáta, dáta o predajoch, tržové dáta pomocou multidimenzionálnych objektov OLAP, ako sú napríklad dátové kocky.

4.1 MIS – manažérske informačné systémy

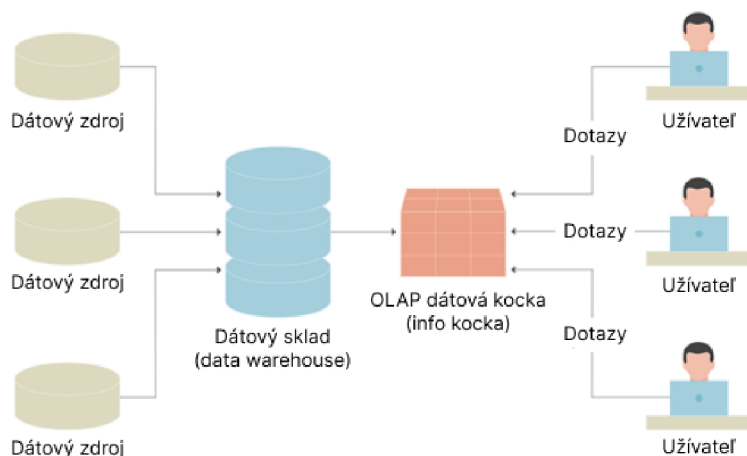
Manažérske informačné systémy sa používajú na rozhodovanie, koordináciu, kontrolu, analýzu a vizualizáciu informácií v podniku alebo organizácii. Úlohou je poskytnúť informácie pre vedúcich pracovníkov, ktorí vďaka dátam v spracovanej forme dokážu analyzovať z rôznych pohľadov situáciu, či už ekonomickú, ale aj personálnu, a rozhodnúť o budúcnosti daného sektoru. Vďaka dátovým skladoch a dátovým kockám si dokážu dáta zobrazit z rôznych pohľadov a vďaka širokej a hlbkej analýze vytvorit čo najoptimálnejšie riešenie alebo rozhodnutie.

4.2 Tvorba analýzy pomocou OLAP

Pre vytvorenie analýzy OLAP technológiou je potrebné, aby sme mali k dispozícii dáta, poznali ich význam, aby sme dáta, s ktorými budeme vytvárať multidimenzionálny objekt, mali uložené v dátovom sklade, aby sme mali vytvorený multidimenzionálny objekt, k dispozícii zobrazovací softvér OLAP objektov a na záver potrebujeme znalosť, ako spraviť analýzu dát.

V praxi a aj pre potreby tejto práce nedokáže všetky tieto zložky zabezpečiť jeden jediný človek. Dáta pre analýzu všeobecne pripravuje databázista, ktorý dané dáta uloží do databázového view. Následne sa dát ujme vývojár, ktorý vytvorí dátový sklad, multidimenzionálny objekt (prevažne dátovú kocku), a na záver prichádza analytik alebo manažér, ktorý si dátovú kocku zobrazí v zobrazovacom softvéri a vytvorí požadovanú analýzu.

Zjednodušene je princíp fungovania OLAP technológie z technického aspektu znázornený na obrázku 4.1.



Obr. 4.1: Princíp fungovania OLAP technológie (zdroj: [7])

4.3 OLTP – Online Transaction Processing

OLTP – Online Transaction Processing je technológia, ktorá umožňuje spracovávať veľké množstvo databázových transakcií veľkým množstvom ľudí v reálnom čase [4]. Preto sú OLTP systémy zamerané na obsluhu veľkého množstva databázových požiadaviek pre veľké množstvo ľudí v reálnom čase. OLTP systémy sa využívajú napríklad pri nakupovaní v online obchodoch, v internetovom bankovníctve, pri transakciách v ATM bankomatoch, pri rezervácii ubytovania v hoteloch alebo pri rezervácii leteniek.

Z podstaty OLTP systémov teda vyplýva, že sa využívajú najmä na finančné transakcie alebo na transakcie, ktoré sa vykonávajú v jednom okamihu mnohokrát rôznymi užívateľmi.

Vzhľadom na povahu OLTP systémov obsahujú tieto systémy dáta v tzv. *raw* forme, teda neupravenej, nespracovanej forme, a mnohokrát sa poskytujú užívateľom bez zmien tak, ako sú uložené v databáze. Pri OLTP sa nepoužíva spracovanie dát, ako napríklad dátové sklady (data warehouse), agregácie alebo dátové kocky z toho dôvodu, že by to systém spomaľovalo. Keďže cieľom OLTP je poskytnúť veľké množstvo dát veľkému množstvu užívateľov v reálnom čase, dáta nie sú upravované a transformované do zložitejších štruktúr ako napríklad dátové kocky. [8]

4.4 Rozdiel medzi OLAP a OLTP

Hoci to nie je na prvý pohľad zrejmé, medzi OLAP a OLTP technológiami je veľmi veľký rozdiel.

Rozdielom je už len minimálne to, že pri OLAP systémoch sa používajú nielen aktuálne, ale aj historické dáta a analýzy sa tvoria z pohľadu celej histórie podniku. Pri OLTP sa nepoužívajú historické dáta, ale len dáta aktuálne, súčasné.

Pri OLTP sú záznamy dát v tzv. transakciách – ide o transakčné spracovanie, pri OLAP sú dáta transformované do dátových skladov, agregované a uložené do OLAP multidimenzionálnych objektov. To znamená, že dáta v OLAP sú upravené, očistené a všeobecne

spracované. Pri OLTP to však neplatí, dáta sa používajú väčšinou v rovnakej forme, ako do OLTP systému prišli. Táto skutočnosť má za následok to, že dáta pri OLTP sa využívajú na každodenné operácie, ako napríklad v bankovníctve pri presunoch financií. Dáta sú spracované veľmi rýchlo.

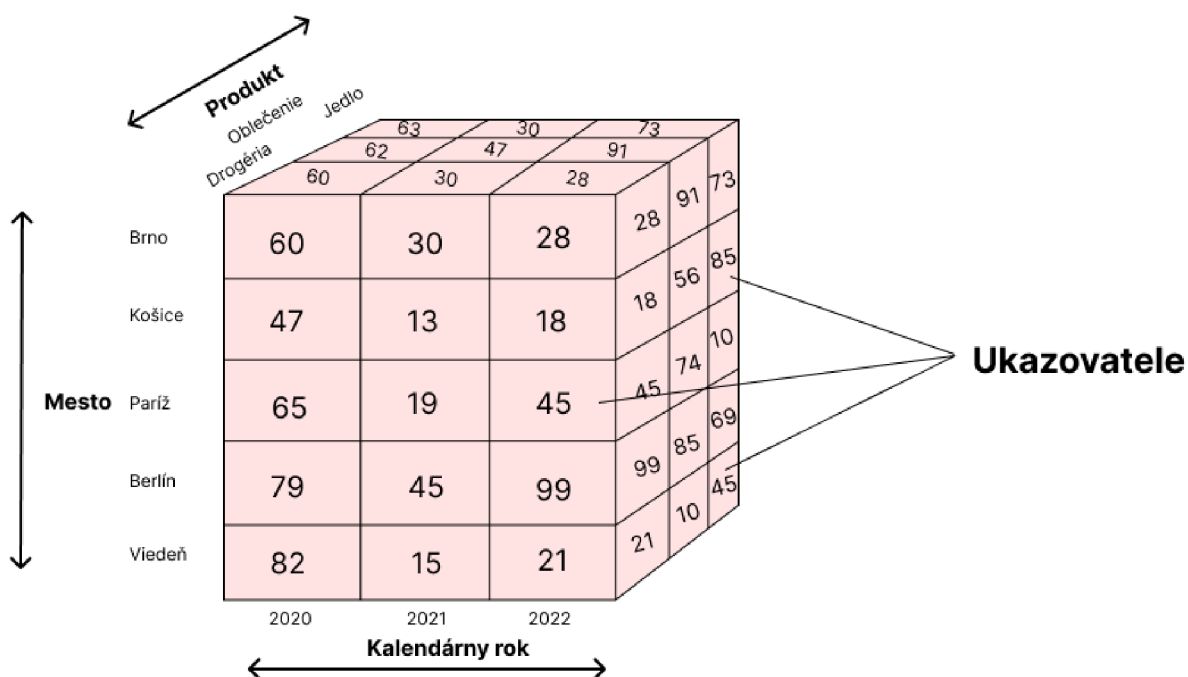
Avšak pri OLAP je to naopak. Tým, že sú dáta spracované a že sú medzi nimi vytvorené väzby, ich časová náročnosť spracovania je vyššia ako pri OLTP. Avšak OLAP nám poskytuje komplexnejší pohľad na dáta. Cieľom OLAP je poskytnúť dáta vo vhodnej forme pre tvorbu analýz potrebných na rozhodovanie a podporiť tak tvorbu správnych rozhodnutí manažérov podnikov.

4.5 Multidimenzionálne objekty

V OLAP sú dátové sklady založené na multidimenzionálnom modeli, kde sú dáta reprezentované OLAP objektami, ako sú napríklad dátové kocky. Tieto objekty tvoria potom výstupy, ktoré využívajú užívatelia (analytici a manažéri) pri zobrazení dát a pri tvorbách analýz. [9, str. 53]

Multidimenzionálne objekty sú všeobecne *n-dimenzionálne objekty*, ktoré poskytujú užívateľom na stupni front-endu možnosť nahliadať na dáta a operovať s nimi, filtrovať ich, vytvárať rôzne pohľady na ne, agregovať číselné výstupy a všeobecne sprístupniť dáta, vďaka ktorým je podporená tvorba správnych rozhodnutí manažérov v podnikoch.

V tejto práci budem pre čo najjednoduchšie pochopenie multidimenzionálnych OLAP objektov používať príklad dátovej kocky, ktorá bude zobrazovať počet predaných produktov v konkrétnych rokoch a v mestách, kde sa nachádzajú predajne obchodného reťazca.



Obr. 4.2: 3-dimenzionálna dátová kocka

4.5.1 Dimenzie

Dimenzie (v SAP Business Warehouse nazývané *charakteristiky*) sú kategórie dát, ktoré môžu byť využité pri analýze a tvorbe reportov. Tieto dimenzie sú používané pre operácie s dátami kocky. Vďaka dimenziám sa dokážeme na dáta v dátovej kocke pozerat z iného uhla pohľadu, filtrovať dáta, tvoriť podkocky s menšou množinou dát alebo pridaním dimenzií tvoriť hĺbkovú analýzu dát.

Na účely tejto práce som vytvoril ukážkovú dátovú kocku, ktorá sa nachádza na obrázku 4.2. Táto dátová kocka je vytvorená z troch dimenzií, v ktorých sa nachádzajú dáta. Tieto dimenzie charakterizujú: *Kalendárny rok*, *Mesto*, *Produkt*. Vďaka dátovým skladom a OLAP technológii sú vytvorené väzby medzi týmito dimenziami a poskytujú užívateľom agregovaný číselný údaj, ktorý sa nazýva **ukazovateľ**, po anglicky *measure* (v SAP Business Warehouse nazývaný *key figure*). Z dátovej kocky 4.2 je teda možné vyčítať, že počet predaných kusov produktov v kategórii *Drogéria* bol v predajni v meste *Paríž* v roku *2022* vo výške 45.

Atribúty charakteristík

Dáta v dimenziách (charakteristikách) môžu mať svoje vlastnosti, ktoré sa taktiež modelujú a zahŕňajú do dátových skladov. Tieto vlastnosti dát v charakteristikách sa nazývajú *atribúty charakteristík*.

Atribúty by mali mať povahu unikátnosti pre daný údaj v charakteristike – tieto dáta by mali byť pre daný údaj charakteristiky jedinečné. Napríklad pre charakteristiku *Mesto* môžu byť atribúty: počet obyvateľov daného mesta, štát, v ktorom sa mesto nachádza, rozloha v km^2 alebo meno starostu. Ako už je z povahy dát zrejmé, sú jedinečné, teda napríklad mesto Berlín nebude mať 2 a viac štátov, v ktorých sa nachádza, alebo súčasne viac hodnôt počtu obyvateľov mesta.

Hierarchie

Dáta v dimenziách môžu byť ďalej členené na menšie zložky. Pri dátových kockách sa tomu hovorí *hierarchia charakteristiky*. Vďaka hierarchiám dokážeme vytvoriť dátové sklady, v ktorých sme schopní dáta so závislosťou viacej členiť. Pri analýzach sa hierarchie veľmi často využívajú.

Príkladom môže byť hierarchia pre charakteristiku *Kalendárny rok*. Tu môžeme členiť kalendárny rok na kvartály, mesiace a ak to má zmysel, tak aj na dni. V praxi sa hierarchie používajú napríklad pri členení oddelení podniku, kde každé oddelenie má svoje pododdelenie a podobne.

4.5.2 Ukazovatele

Ukazovateľ, po anglicky *measure* (v SAP Business Warehouse terminológii *key figure*), je číselný údaj, ktorý reprezentuje závislosť dimenzií v danom prieniku dimenzií. Analytik tak dostáva číselný údaj o závislosti, väzbe n-dimenzií.

Mnohokrát sa ako ukazovateľ berú finančné údaje, ako napríklad číselná suma, počet záznamov v dátovom sklade, podiel kontribúcie pri tvorbe kníh a článkov, počet zamestnancov a podobne.

Každý ukazovateľ má priradenú svoju agregačnú funkciu, vo väčšine prípadov to býva agregačná funkcia SUM, teda sumarizácia. V prípade SAP Business Warehouse máme k dispozícii aj funkcie ako MAX, MIN a NO2 (maximum, minimum a bez agregácie).

Vzhľadom na fakt, že ukazovateľ je číselná hodnota, ktorá má svoju agregáčnú funkciu, nemá ukazovateľ atribúty ani hierarchie. Z pohľadu významu dát to nedáva zmysel.

Dobrym príkladom ukazovateľa sú číselné hodnoty v dátovej kocke na obrázku 4.2, ktoré sa nachádzajú na spojnici dimenzií. Reprezentujú tak počet predaných produktov v konkrétnom roku, v konkrétnom meste podľa druhu produktov.

4.6 Dátová kocka

Dátová kocka (v SAP Business Warehouse terminológii nazývaná aj ako *info kocka*) je n-dimenzionálny OLAP objekt, ktorý je postavený nad multidimenzionálnou databázou, v ktorej sú uložené dáta. Dátová kocka sa skladá z dimenzií (charakteristík) a ukazovateľov, ktoré túto kocku definujú. Dátová kocka musí obsahovať minimálne jeden ukazovateľ. Bez neho sa nedá dátová kocka namodelovať.

V dátovej kocke sú všetky dimenzie (charakteristiky) nastavené ako kľúčové. To znamená, že prienik dát dimenzií je vždy unikátny, jedinečný – každému prieniku dimenzií zodpovedá ukazovateľ a je jasne definovaný výsledok prieniku. Napríklad pre produkty *Oblečenie* v roku 2022 existuje v dátovej kocke 4.2 rozdielna hodnota ukazovateľa v rôznych mestách.

Pomocou tejto vlastnosti dokáže zobrazovací softvér zobraziť správne relácie medzi dátami v dimenziách bez straty dát.

Ako už bolo spomenuté vyššie, dátovú kocku najskôr vyvíjajú programátori, ktorí vytvoria ETL, dátový sklad, namodelujú dátovú kocku a potom do nej nahrajú dáta. Následne sú analytici schopní na úrovni front-endu dátovú kocku zobraziť, operovať s dátami pomocou OLAP operácií opísaných v sekcii 4.7 a tvoriť analýzy s dátami v dátovej kocke.

Príkladom dátovej kocky je obrázok 4.2 v sekcii 4.5, ktorý znázorňuje 3-dimenzionálnu dátovú kocku.

4.7 OLAP operácie

Ako bolo už spomenuté v sekcii 4.2, výstupom OLAP sú multidimenzionálne objekty, prevažne dátové kocky. Dáta týchto objektov sú uložené v multidimenzionálnych databázach, ktoré sú práve optimalizované pre účel OLAP analýzy a pre uloženie veľkého množstva dát do dátových skladov. Tento druh databáz poskytuje možnosť rýchleho spracovania dát a generovania rýchlych odpovedí.

Nad multidimenzionálnymi databázami existujú operácie, ktoré dokážu pracovať veľmi rýchlo a efektívne s dátami uloženými v nich. Tieto operácie vykonávajú na úrovni front-endu užívateľa (analytici a manažéri) manipulujúci s OLAP objektmi za účelom analýzy.

Základné OLAP analytické operácie sú nasledovné:

- **Roll-up** – taktiež známa ako *drill-up* alebo tzv. vynorenie, táto operácia sumarizuje (agreguje) údaje pozdĺž dimenzie. Zmenšuje sa teda počet aktívnych dimenzií (charakteristík), zväčšuje sa agregácia a zmenšuje počet medzivýsledkov. [3]
- **Drill-down** – tzv. zavíranie, pomocou tejto operácie sa analytik „ponorí“ hlbšie do dát v dátovej kocke. Získava tak detailnejší pohľad na dáta. Pri tejto operácii sa zväčšuje počet aktívnych dimenzií (charakteristík) dátovej kocky – pridávajú sa dimenzie z neaktívnych dimenzií. Zmenšuje sa agregácia a zväčšuje sa počet medzivýsledkov.

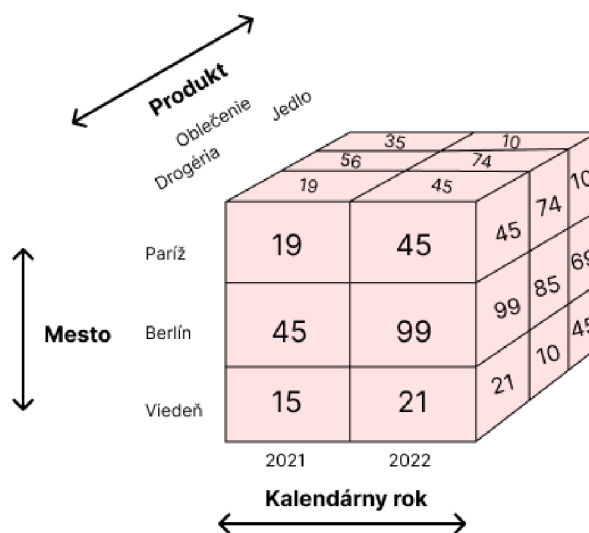
- **Slice** – inak nazvané filtrovanie jednou hodnotou. Touto operáciou dokáže analytik vyfiltrovať dáta jednej dimenzie, a tak aj „odstrániť“ dimenziu dátovej kocky avšak takým spôsobom, že dáta v kocke budú podliehať kontextu vyfiltrovannej dimenzie tejto operácie. Napríklad, ak sa v dátovej kocke nachádza aktívna dimenzia s názvom *Produkt*, užívateľ môže vyfiltrovať produkt *Drogéria*, dimenzia sa zo zoznamu aktívnych dimenzií vymaže, ale nastaví sa kontext dát len pre záznamy s kategóriou produktu *Drogéria*. Následne budú užívateľovi v analýze zobrazené len dáta tejto kategórie. Vďaka tejto operácii sú užívatelia schopní analyzovať konkrétne dáta, na ktoré sa potrebujú zamerať, a vytvoriť tak hĺbkovú analýzu. Treba si však uvedomiť, že operácia *Slice* je filtrovanie iba jednej dimenzie dátovej kocky, v našom príklade to bola dimenzia *Produkt*. Táto operácia nepodlieha „filtrovaníu“ viacerých dimenzií naraz. Prostredníctvom tejto operácie vzniká podkocka, v ktorej sú dáta v kontexte vyfiltrovannej dimenzie.

Kontext: Produkt = "Drogéria"

Mesto	Brno	60	30	28
	Košice	47	13	18
	Paríž	65	19	45
	Berlín	79	45	99
	Viedeň	82	15	21
		2020	2021	2022
		Kalendárny rok		

Obr. 4.3: Slice operácia nad dátovou kockou

- **Dice** – touto operáciou užívateľ vytvorí podkocku z celej OLAP kocky prostredníctvom výberu dvoch alebo viacerých dimenzií. Týmto dimenziám môže nastaviť aj dodatočné filtre. V príklade si zvolíme aktívne dimenzie *Kalendárny rok*, *Mesto* a *Produkt*. Pri dimenzii *Produkt* si nenastavíme žiaden filter a necháme v dátovej kocke všetky dáta tejto dimenzie. V dimenzii s názvom *Mesto* sa môže v dátovej kocke nastaviť filter na mestá *Berlín*, *Viedeň* a *Paríž*, a pri dimenzii *Kalendárny rok* sa nastaví filter na roky *2021* a *2022*. Nastaví sa tak kontext dát a zobrazí sa nám tzv. podkocka – podmnožina dát celej OLAP dátovej kocky.



Obr. 4.4: Dice operácia nad dátovou kockou

- **Pivot** – rotovanie osí dátovej kocky za účelom získania nového pohľadu nad rovnakými dátami. Táto operácia dodá analytikovi možnosť nahliadať na dáta z iného uhla pohľadu a dodá mu tak flexibilnejšie možnosti analýzy.

Podľa odborných článkov uverejnených na internete poznáme aj ďalšie operácie, ako napríklad Scoping, Screening, Drill across, Drill through, Sort, Add measure, Drop measure, Union a Difference. [2]

4.8 SAP Business Warehouse

SAP Business Warehouse (SAP BW) je softvér od spoločnosti SAP vytvorený najmä na tvorbu dátových skladov a správu dát využívaných na podporu rozhodovania v podnikoch. Zbiera, transformuje a ukladá dáta generované v SAP alebo non-SAP aplikáciách a robí ich dostupné pomocou vstavaných reportingových, analytických, business intelligence nástrojov a taktiež aj prostredníctvom nástrojov tretích strán. SAP BW je založený na SAP NetWeaver platforme. [5]

SAP BW je taktiež platforma pre programátorov, ktorí v nej môžu vytvárať dátové sklady, OLAP objekty, vykonávať úlohy pre správu dát, ako napríklad automatické nahrávanie nových dát do dátových skladov, generovať reporty a vytvárať analytické aplikácie. Poskytuje tiež niekoľko nástrojov na podporu extrakcie dát, dátovej ťažby, ukladania dát, správu dát a podobne.

Na VUT nahráva SAP BW dáta z rôznych dátových zdrojov. Jedným z najpoužívanejších je Centrálna databáza VUT (CDB), ktorá je v podstate Oracle databáza. Ďalším významným zdrojom dát sú ekonomické dáta, ktoré sú uložené v systémoch SAP.

Na účely tejto záverečnej práce budem používať softvér SAP Business Warehouse pre tvorbu dátového skladu, tvorbu dátového zdroja, modelovanie a tvorbu dátovej kocky, dimenzií (charakteristík) a celkovo OLAP objektov. Následne mi SAP BW ako výstup poskytne dátové kocky, ktoré je možné načítať a zobraziť v zobrazovacom softvéri SAP Analysis for Microsoft Office, ktorý je opísaný v sekcii 4.9 nižšie.

4.9 SAP Analysis for Microsoft Office

SAP Analysis for Microsoft Office je softvérový nástroj od spoločnosti SAP, ktorý umožňuje užívateľom zobraziť, analyzovať a vizualizovať podnikové dáta uložené v SAP Business Warehouse (SAP BW) systéme prostredníctvom programov Microsoft Excel a Microsoft PowerPoint. Taktiež umožňuje zobrazenie OLAP dátových kociek a Query objektov. Počas analýzy je užívateľ schopný využívať vstavané funkcie programov Microsoft Excel a PowerPoint. Užívateľ má možnosť vďaka *SAP Analysis for Microsoft Office* využiť rozšírené funkcionality tohto softvéru na tvorbu analýz.

V tejto práci budem používať *SAP Analysis for Microsoft Office* na zobrazenie dát OLAP dátových kociek, ktoré som vytvoril. Dokážem tak vytvoriť potrebné analýzy, validovať namodelované a vytvorené dátové kocky a poskytnúť tieto dátové kocky ďalším analytikom VUT, ktorí ich dátový obsah budú používať v budúcnosti.

4.10 SAP Business Objects

SAP Business Objects je *Business Intelligence* softvérová platforma, ktorá sa používa na zobrazovanie, tvorbu a doručenie reportov, analýz a dátových vizualizácií. Služi ako webová aplikácia a nachádza sa na úrovni front-endu z modelu 3.1.

SAP Business Objects obsahuje aplikácie ako Crystal Reports, Web Intelligence alebo Xcelsius, ktoré slúžia na vyššie spomenuté prípady použitia.

Kapitola 5

Modelovanie dátových skladov v SAP Business Warehouse

Softvér SAP Business Warehouse poskytuje nástroje na návrh a vývoj OLAP objektov, načítanie dát z rôznych zdrojov a možnosť vložiť tieto dáta do OLAP objektov.

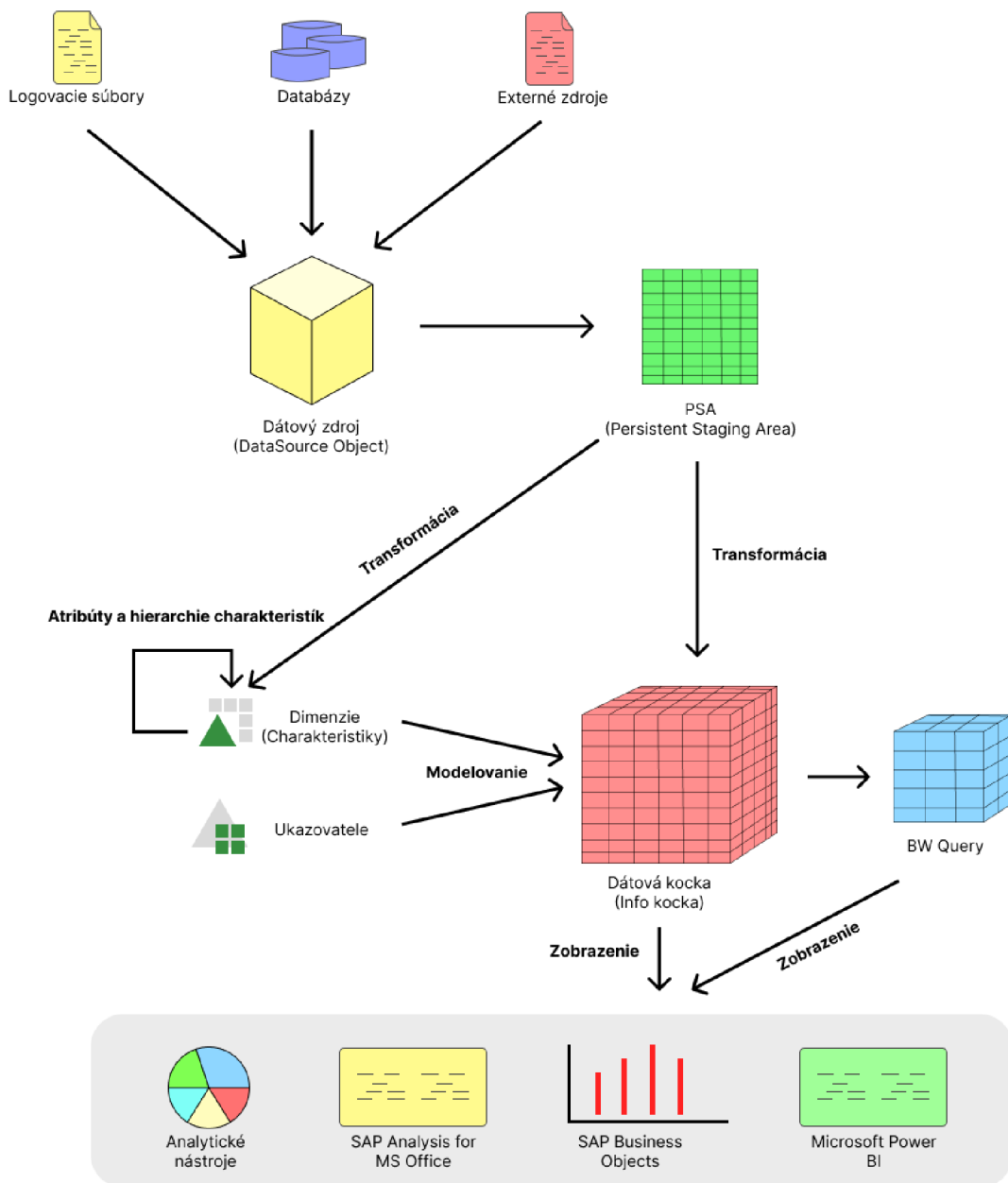
5.1 Tok vývoja dátových skladov a dátových kociek

Vývoj dátových skladov a dátových kociek v SAP BW pozostáva z viacerých štádií. V prvom rade je potrebné, aby sme vedeli, nad akými dátami chceme vytvoriť OLAP dátovú kocku, pretože budeme vytvárať charakteristiky (dimenzie), ktoré budú uchovávať dáta. Vývoj môžeme rozdeliť do 10 krokov:

1. Získanie zdroja dát, ktorý máme uložený lokálne (na zariadení, na ktorom budeme vykonávať nahrávanie dát do dátovej kocky) alebo vzdialene (externe napojená databáza, ktorá obsahuje databázové view s požadovanými dátami).
2. Vytvorenie objektu *Dátového zdroja (DataSource Object)* a prepojenie s lokálnym alebo externým zdrojom dát.
3. Vytvorenie *InfoPacketu*, ktorý slúži na počiatočné nahranie dát z *DataSource* objektu do *Persistent Staging Area (PSA)*.
4. Vytvorenie charakteristík a ukazovateľov správneho dátového typu na základe dát uložených v PSA a požadovaného koncového tvaru.
5. Vytvorenie atribútov charakteristík. Vytvorenie transformácie a nahranie dát do týchto atribútov.
6. Vytvorenie dátovej kocky pomocou charakteristík a ukazovateľov.
7. Vytvorenie transformácie, vytvorenie a programovanie namapovania a formátovania dát z PSA do dátovej kocky.
8. Vytvorenie objektu *Procesu prenosu dát*, ktorý je súčasťou dátovej kocky a slúži na aktiváciu transformácie dát z PSA do dátovej kocky.

9. Zobrazenie dát v zobrazovacom softvéri, ako je napríklad *SAP Analysis for MS Office* alebo *SAP Business Objects*.
10. V prípade nutnosti je možné vytvoriť nad štruktúrou dátovej kocky dodatočné *BW Query*, ktoré slúži ako nadstavba nad dátovou kockou. Slúži primárne na obmedzenie charakteristík a ukazovateľov na základe užívateľských práv alebo na vytvorenie nových vypočítaných ukazovateľov už zo súčasných dostupných charakteristík. Príkladom, kedy je vhodné použiť *BW Query*, je prípad, kedy ekonóm z Fakulty informačných technológií žiada o dáta nachádzajúce sa v dátovej kocke, ktorá obsahuje ekonomické údaje o celom VUT. Tento ekonóm nemá práva vidieť ekonomické dáta celého VUT, preto mu SAP BW vývojár obmedzí dáta len na Fakultu informačných technológií.

Celá schéma toku vývoja sa nachádza na obrázku 5.1, kde sú vyobrazené prepojenia medzi objektami aj s popisom objektov a prechodov.



Obr. 5.1: Tok vývoja v SAP BW

5.2 Dátový zdroj – DataSource Object

Tento objekt slúži na nastavenie prepojenia zo zdrojov dát (lokálnym alebo externým). Umožňuje nahliadnuť na názov polí v zdroji dát (názov databázových domén v databázovom view), zobrazí navrhované dátové typy a počet požadovaných znakov pre tieto polia a umožní nahliadnuť na dáta, ktoré sa nachádzajú v zdroji dát. Povoľuje taktiež vykonať počítačové spracovanie vstupných dát pomocou vykonania konverznej rutiny.

5.2.1 PSA – Persistent Staging Area

Je to databázová tabuľka, ktorá slúži ako dočasné úložisko pre dáta, ktoré boli nahrané z externého prostredia do systému SAP BW. Používa sa na dočasné uloženie dát pred začatím procesu transformácie dát do OLAP objektov, v prípade tejto práce do dátových kociek. Jednou z funkcií PSA je zabezpečenie perzistentnosti zdrojových dát v systéme, ktoré zostanú v SAP BW systéme uložené aj po odpojení externého dátového zdroja (po odpojení databázy, logovacieho súboru, CSV a podobne).

5.3 Charakteristika

Charakteristika je základným článkom OLAP dátovej kocky. Reprezentuje dimenziu dátovej kocky, v ktorej sa nachádzajú dáta pre analýzy. V SAP BW je tvorba charakteristík veľmi jednoduchá. Na základe dát, ktoré sa v charakteristike budú nachádzať, zvolíme vhodný dátový typ, dĺžku (maximálny počet znakov) a voliteľnú konverznú rutinu. Pri voľbe dátového typu máme na výber 4 typy:

- **CHAR** – znakové reťazce
- **NUMC** – numerický text, podpora len celých čísel
- **DATS** – dátum vo formáte RRRRMMDD
- **TIMS** – čas vo formáte HHMMSS

SAP BW poskytuje aj ďalšie možnosti nastavenia charakteristík, ako sú napríklad atribúty charakteristík (k ID osoby je napríklad priradený login osoby spolu s krstným menom a priezviskom), kmeňové dáta (inak povedané aj číselníky), ktoré je vhodné použiť v kombinácii s číslom ID a názvom nejakého objektu, napríklad fakulty univerzity, a na záver hierarchie, ktoré sú bližšie popísané v sekcii 4.5.1.

5.4 Ukazovateľ – key figure

Keďže ukazovatele zohrávajú inú úlohu ako charakteristiky, ich tvorba je mierne odlišná. Pri ukazovateľoch máme na výber z viacerých typov, ktoré SAP BW ponúka:

- čiastka s nastaviteľnou menou
- číslo
- dátum
- množstvo s nastaviteľnou jednotkou

- integer
- čas

Pri niektorých dátových typoch sa dá nastaviť aj pevne stanovený počet desatinných miest, ktorý sa mnohokrát využíva. Posledným a zároveň takmer najdôležitejším nastavením pri ukazovateľoch je agregáčna funkcia, ktorá sa aplikuje na daný ukazovateľ. SAP BW podporuje 4 agregáčne funkcie:

- **SUM** – suma hodnôt
- **MAX** – maximálna hodnota
- **MIN** – minimálna hodnota
- **Bez agregácie**

5.5 Tvorba dátovej kocky

Tvorba dátovej kocky v SAP BW je vo všeobecnosti veľmi jednoduchá záležitosť, pretože štruktúra dátovej kocky pozostáva z ukazovateľov a charakteristík, ktoré je potrebné vytvoriť separátne mimo modelovania dátovej kocky. Pri vytváraní dátovej kocky v podstate len vložíme požadované charakteristiky a ukazovatele do štruktúry dátovej kocky a uložíme dátovú kocku. Pri vytváraní je užívateľ schopný prideliť dátovej kocke dodatočné metadáta, teda popisy štruktúry a popisy prípadov použitia. Taktiež má vývojár možnosť zoskupovať charakteristiky na základe sémantiky. Tieto zoskupenia slúžia pri analýzach na lepšiu užívateľskú orientáciu v tom, čo dátová kocka ponúka.

5.6 Transformácia

Transformácia ako objekt je v SAP BW komponent dátovej kocky, vďaka ktorému vývojár vytvára transformáciu (mapovanie) dát z PSA do dátovej kocky. Je to proces, ktorý je súčasťou ETL, ktorá je bližšie popísaná v sekcii 3.4. Jednou z hlavných úloh transformácie je preformátovanie vstupných dát a uloženie týchto dát do charakteristík, ktoré budú pri analýzach tieto dáta reprezentovať. Transformácia je tiež miesto, kde vznikajú väzby medzi jednotlivými dátami – vytvoria sa závislosti medzi charakteristikami (dimenziami) a vypočítajú sa presne príslušné ukazovatele na základe priradených agregáčnych funkcií. Na tomto mieste prebieha celá „*mágia*“ OLAP objektov, pretože vďaka naprogramovanej transformácii SAP BW vie, ako má uložiť dáta do multidimenzionálnej databázy tak, aby zachoval vzájomné vzťahy medzi jednotlivými charakteristikami a vypočítal hodnoty ukazovateľov.

Pri priradovaní (mapovaní) dát z PSA do dátovej kocky sa používajú 4 základné spôsoby priradenia (mapovania):

- Priame priradenie – priradenie dát v rovnakom tvare, ako sa nachádzajú v PSA (Persistent Staging Area).



Obr. 5.2: Ikonka priameho priradenia v transformácii v SAP BW

- Priradenie konštanty – priradenie konštantnej hodnoty. Vo väčšine prípadov sa toto priradenie používa, ak vytvárame ukazovateľ, ktorý počíta počet databázových záznamov, ktoré sa pri transformácii do dátovej kocky prenesú.



Obr. 5.3: Ikonka priradenia pomocou konštanty v transformácii v SAP BW

- Priradenie vzorcom – SAP BW ponúka vstavané funkcie a priradenie vzorcom využíva práve tieto funkcie. Ide napríklad o konverzné, formátovacie, dátumové a číselné funkcie, ktoré ponúkajú možnosti, ako formátovať dáta vstupujúce do charakteristík.



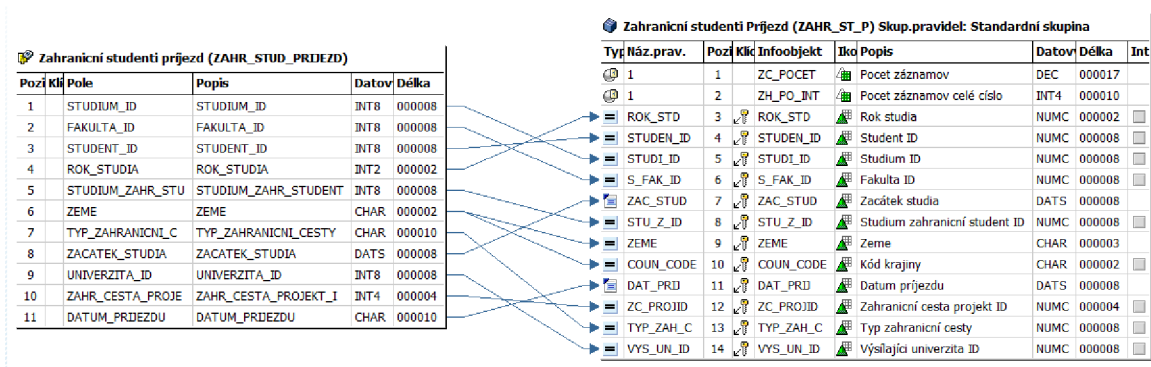
Obr. 5.4: Ikonka priradenia pomocou vzorca v transformácii v SAP BW

- Priradenie pomocou rutiny – rutina je program, ktorý zo vstupných polí dáta spracuje, naformátuje a uloží do výstupného poľa, ktoré reprezentuje výslednú charakteristiku. Rutiny v SAP BW sa programujú v jazyku ABAP, ktorý vytvoril SAP. ABAP je procedurálny programovací jazyk, ktorý podporuje objektovo orientované programovanie. Rutiny som v tejto práci použil hlavne na zabezpečenie správneho spracovania nepodporovaných znakov v znakových reťazcoch a na konverziu dátumov z externého formátu do interného podporovaného formátu SAPu.

Okrem týchto základných spôsobov existujú aj ďalšie, ktoré sa však bežne používajú len zriedka.



Obr. 5.5: Ikonka priradenia pomocou rutiny v transformácii v SAP BW



Obr. 5.6: Tok vývoja v SAP BW

5.6.1 Transformácia v SAP BW

V SAP BW vyzerá objekt transformácie nasledovne:

Na ľavej strane sa nachádza štruktúra polí v *DataSource* objekte a na pravej strane sa nachádza štruktúra dátovej kocky.

Šípky znázorňujú mapovanie (priradenie) dát z *DataSource* objektu do dátovej kocky. PSA slúži len ako databázová tabuľka, ktorá uchováva dáta *DataSource* objektu.

5.7 Nahrávanie dát do dátovej kocky

Na spustenie nahrávania dát z PSA do dátovej kocky slúži objekt *Proces prenosu dát*, ktorý je súčasťou dátovej kocky a má vytvorenú väzbu na transformáciu a *DataSource* objekt. Poskytuje nám možnosti, ako spustiť transformáciu a prenos dát. Dáta transformuje prostredníctvom paketov (súbory záznamov dát). Vývojár je schopný nastaviť veľkosť paketov, paralelné alebo sériové transformovanie dát a filtre, ktoré sa aplikujú na dáta uložené v databázovej tabuľke *DataSource* objektu – v PSA.

Kapitola 6

Problematika zahraničných ciest na VUT

Ako každá univerzita, aj VUT (Vysoké učení technické) vysiela svojich zamestnancov a študentov na zahraničné cesty a vzdelávacie pobyty. Taktiež sa angažuje a prijíma zahraničných pracovníkov a študentov, ktorí vykonávajú vzdelávaciu alebo pracovnú činnosť pod záštitou VUT. Vďaka každému druhu zahraničných ciest ponúka univerzita svojim pracovníkom a študentom možnosť realizovať sa, vzdelávať sa alebo reprezentovať univerzitu nielen v Českej republike, ale aj v zahraničí.

Pod zahraničné cesty spadajú rôzne programy, ako napríklad Erasmus+, Nórske fondy, Leonardo da Vinci a CEEPUS, ktoré sú mnohokrát zamerané práve na vzdelávanie a výskumnú činnosť.

Dáta o všetkých druhoch zahraničných ciest má VUT uložené v centrálnej databáze (CDB). Jednotlivé fakulty univerzity majú svoje oddelenia zahraničných ciest, ktoré vysielať, prijímajú pracovníkov a študentov a spravujú zahraničné cesty jednotlivých fakúlt.

V tejto práci sa zameriam na spracovanie dát pomocou OLAP technológie nad všetkými dátami zahraničných ciest. Spadajú sem dáta o prijatých a vyslaných akademických a technických pracovníkoch, ale aj dáta o vyslaných a prijatých študentoch. Budem teda pracovať s rozsiahlym počtom dát.

6.1 Súčasný stav

Aktuálne sa k dátam zahraničných ciest dokážu zamestnanci a analytici dostať iba jediným spôsobom – pomocou informačného systému Apollo, ktorý slúži ako interný informačný systém VUT.

V IS Apollo existujú moduly pre zahraničné cesty, ktoré obsahujú dáta práve o zahraničných cestách. Taktiež v module *Přehledy a statistiky* existuje sekcia zameraná na zahraničné cesty. Toto je jediné možné miesto, ako sa užívatelia mimo databázistov dokážu dostať k dátam o zahraničných cestách.

V bežných prípadoch je súčasné riešenie postačujúce, ale v prípade špecifickejších požiadaviek na dáta alebo hĺbkových analýz je problematika zahraničných ciest v IS Apollo vyriešená nedostatočne.

Jednou z týchto situácií je výročná správa o činnosti školy od MŠMT, ktorú musia vyplniť všetky univerzity v Českej republike. Jednou časťou tejto výročnej správy sú aj údaje a štatistiky o zahraničných cestách.

V minulosti sa získavali dáta o zahraničných cestách pomocou SQL dotazov vykonávaných nad relačnými databázovými tabuľkami v centrálnej databáze VUT. Tento proces nebol praktický a mohol viesť k nesprávnemu poskytnutiu dát. Stačila mierna chyba v SQL dotaze a vo výročnej správe sa mohli nachádzať nesprávne dáta, ktoré by s realitou nemuseli mať nič spoločné.

Pomocou získavania dát z SQL dotazov nemali analytici možnosť vytvárať s dátami profesionálne analýzy a aktívne sa vracieť k dátam v priebehu niekoľkých minút – vždy bolo nutné vykonať SQL dotaz a zatažovať tak ľudské zdroje VUT.

Ďalšou možnosťou bolo vykonávať analýzy v module *Přehledy a statistiky*. Analytik mal možnosť zobraziť si dáta na základe filtrov a exportovať dáta do MS Excelu. Tento modul neponúka žiadne pokročilé možnosti analýz, ktoré by analytikom poskytli flexibilitu. Analytik je schopný maximálne vytvoriť všeobecné grafy a exportovať ich.

6.2 Riešenie

Pomocou praktickej časti tejto práce sa pokúsim vyriešiť danú problematiku s dátami zahraničných ciest a poskytnúť tak analytikom a vedúcim nielen zahraničných oddelení flexibilnejšie prostredie, v ktorom môžu vykonávať hĺbkové analýzy s dátami zahraničných ciest.

Pomocou môjho riešenia tak ušetrim aj ľudské zdroje VUT, ktoré boli potrebné na poskytovanie špecifickjších dát pre ojedinelé úlohy a požiadavky. Analytici nebudú nútení zatažovať požiadavkami databázistov, ktorí by následne museli vykonávať komplikované SQL dotazy nad centrálnou databázou VUT.

Mojím riešením budú 3 OLAP dátové kocky zahraničných ciest vytvorené pomocou OLAP technológie a softvéru SAP Business Warehouse. Pôjde o tieto dátové kocky:

- **Zahraníční cesty Výjezd** – Dátová kocka obsahujúca dáta vyslaných akademických a technických pracovníkov VUT spoločne s vyslanými študentmi VUT. Dátová kocka obsahuje dáta aj o výške grantov (štipendií pre zahraničné pobyty) pridelených k danej zahraničnej ceste. Pomocou kocky je možné sledovať nielen ekonomickú stránku zahraničných ciest vyslaných účastníkov, ale aj údaje o tom, do akej krajiny účastník zahraničnej cesty cestuje, na akú univerzitu, na ako dlho a za akým účelom.
- **Zahraníční cesty zamestnanci Příjezd** – Obsahuje dáta o prichádzajúcich zahraničných akademických a technických pracovníkoch na vzdelávacie účely, ale aj na účely výskumnej činnosti. V dátovej kocke sa nachádzajú dáta o ich vysielajúcej inštitúcii, krajine, z ktorej prišli, dátumy príchodu aj odchodu a typ ich zahraničnej cesty.
- **Zahraníční študenti Příjezd** – Nachádzajú sa tu dáta o prichádzajúcich študentoch zo zahraničných škôl. Sú tu k dispozícii informácie o krajine, z ktorej pochádzajú, univerzity, z ktorej boli vyslaní, na akú fakultu VUT prišli študovať, ich dátum príchodu a začiatku štúdia.

Tieto dátové kocky budú dostupné pre celé VUT a budú mať k nim prístup zamestnanci VUT, ktorí budú mať potrebné oprávnenia na zobrazenie týchto dát. Keďže ide o citlivé dáta, v ktorých sa nachádzajú aj ekonomické dáta ohľadom grantov a dotácií v kombinácii s dátami osobných údajov jednotlivých osôb, je potrebné, aby prístup mali len oprávnení zamestnanci.

6.2.1 Tri OLAP dátové kocky

Všetky tri dátové kocky sú zamerané na zahraničné cesty, avšak každá z nich na iný sektor zahraničných ciest. Ich štruktúra má teda podobné črty – nachádzajú sa tu charakteristiky so sémanticky rovnakým významom dát.

Je nutné uviesť dôvod, prečo je potrebné vytvoriť tri OLAP dátové kocky a nie je možné vytvoriť jednu veľkú komplexnú dátovú kocku, ktorá by poskytovala dáta o všetkých typoch zahraničných ciest akademických a technických pracovníkov VUT a študentov VUT.

Tým dôvodom je spôsob uloženia dát v tabuľkách relačnej databázy v centrálnej databáze VUT. Pre získanie všetkých požadovaných dát o jednotlivých sektoroch zahraničných ciest som musel študovať uloženie dát v centrálnej databáze a nájsť v ERD (entity-relationship diagram) schémach VUT správne zdrojové tabuľky, ktoré obsahovali mnou požadované dáta. Problém však ale nastal vo väzbách a prepojení jednotlivých tabuliek, pretože v niektorých prípadoch neexistovali väzby medzi tabuľkami, ktoré obsahovali špecifickejšie dáta.

Preto som v spolupráci s databázovým oddelením VUT prišiel k záveru, že nebude možné vytvoriť jednu komplexnú dátovú kocku, pretože mi to štruktúra databázy nedovoľuje, ale budem musieť navrhnúť a vytvoriť 3 databázové kocky, kde každá z kociek bude mať ako zdroj dát rozdielny databázový pohľad.

Po preštudovaní ERD schém, ktoré má vytvorené VUT práve na účely dokumentácie štruktúry centrálnej databázy, som navrhol tri databázové pohľady. Následne som požiadal databázové oddelenie VUT, aby mi z môjho návrhu vytvorilo práve tieto tri databázové pohľady (každý pohľad zameraný na jednotlivú časť zahraničných ciest). Pre každý databázový pohľad som vytvoril *DataSource object* v SAP BW a využil tieto pohľady ako externý zdroj dát. Spôsob fungovania *DataSource object* je popísaný v sekcii 5.2.

6.3 Výročná správa o činnosti školy od Ministerstva školstva, mládeže a telovýchovy Českej republiky

Ako už bolo spomenuté, hlavným dôvodom spracovania problematiky zahraničných ciest a vytvorenie OLAP dátových kociek je výročná správa o činnosti školy od Ministerstva školstva, mládeže a telovýchovy Českej republiky, ktorá sa taktiež zaoberá štatistikou spojenou so zahraničnými cestami VUT. Túto správu musia vyplňať všetky univerzity v Českej republike. Zo súčasného stavu vyplýva, že tento proces získavania dát nie je automatizovaný a každý rok stojí zamestnancov VUT desiatky hodín práce. OLAP dátové kocky dokážu tento proces zefektívniť a ušetriť ľudské zdroje VUT.

6.3.1 Zapojenie vysokej školy do programov medzinárodnej spolupráce

V tejto tabuľke sa zisťuje počet osôb vyslaných a prijatých univerzitou v závislosti od programu vrátane finančných dotácií a počet projektov univerzity. Vo výročnej správe o činnosti táto tabuľka zodpovedá tabuľke číslo 7.1.

Vysoké učení technické v Brně	H2022/ 7. rámcový program EK			CELKEM
	CELKEM	Z toho Marie- Curie Actions	Ostatní	
Počet projektů*				0
Počet vyslaných studentů**				0
Počet přijatých studentů***				0
Počet vyslaných akademických a vědeckých pracovníků****				0
Počet přijatých akademických a vědeckých pracovníků*****				0
Dotace v tis. Kč*****				0

Obr. 6.1: Tabuľka zapojenia vysokej školy do programov medzinárodnej spolupráce

Pre tabuľku platia dodatočné poznámky, ktoré bližšie špecifikujú povahu dát v jednotlivých poliach. Tieto poznámky sú naviazané na hviezdičky v bunkách tabuľky. Poznámky sú doslovne prevzaté z výročnej správy o činnosti:

- * – Jedná se o v daném roce probíhající projekty.
- ** – Vyjíždějící studenti (tj. počty výjezdů) – kteří v roce 2022 absolvovali zahraniční pobyt; započítávají se i ti studenti, jejichž pobyt začal v roce 2021. Započítávají se pouze studenti, jejichž pobyt trval více než 4 týdny (28 dní). Pokud VŠ uvádí i jinak dlouhé výjezdy, uvede to v poznámce k tabulce.

- *** – Přijíždějící studenti (tj. počty příjezdů) – kteří přijeli v roce 2022; započítávají se i ti studenti, jejichž pobyt začal v roce 2021. Započítávají se pouze studenti, jejichž pobyt trval více než 4 týdny (28 dní). Pokud VŠ uvádí i jinak dlouhé výjezdy, uvede to v poznámce k tabulce.
- **** – Vyjíždějící akademičtí pracovníci (tj. počty výjezdů) – kteří v roce 2022 absolvovali zahraniční pobyt; započítávají se i ti pracovníci, jejichž pobyt začal v roce 2021.
- ***** – Přijíždějící akademičtí pracovníci (tj. počty příjezdů) – kteří přijeli v roce 2022; započítávají se i ti pracovníci, jejichž pobyt začal v roce 2021.
- ***** – Uvedené částky představují celkové finanční zdroje projektů, včetně spolufinancování MŠMT.

6.3.2 Mobilita studentů, akademických a ostatních pracovníků podle krajiny

Do tabulky sa uvádza štatistika počtu vyslaných a prijatých študentov, akademických a ostatných pracovníkov VUT vzhľadom na krajiny, do ktorých boli vyslaní alebo z ktorých boli prijatí. Vo výročnej správe o činnosti táto tabuľka zodpovedá tabuľke číslo 7.2.

Vysoké učení technické v Brně	Počet vyslaných studentů*			Počet přijatých studentů**		Počet vyslaných akademických pracovníků***	Počet přijatých akademických pracovníků****	Počet vyslaných ostatních pracovníků***	Počet přijatých ostatních pracovníků****	CELKEM za zemi
	Celkem	Absolventské stáže (z celkem)*****	Virtuálně ¹ (z celkem)	Virtuálně ¹ (z celkem)	Celkem					
Země										

Obr. 6.2: Záhľad tabuľky mobility študentov, akademických a ostatných pracovníkov podľa krajiny

Tak ako pri tabuľke v sekcii 6.3.1, aj pri tejto tabuľke sa nachádzali dodatočné poznámky, ktoré špecifikovali konkrétnejšie dáta v poliach:

- * – Vyjíždějící studenti (tj. počty výjezdů) – studenti, kteří v roce 2022 absolvovali (ukončili) zahraniční pobyt; započítávají se i ti studenti, jejichž pobyt začal v roce 2021. Započítávají se pouze studenti, jejichž pobyt trval alespoň 2 týdny (14 dní).
- ** – Přijíždějící studenti (tj. počty příjezdů) – studenti, kteří přijeli v roce 2022; započítávají se i ti studenti, jejichž pobyt začal v roce 2021. Započítávají se pouze studenti, jejichž pobyt trval alespoň 2 týdny (14 dní).
- *** – Vyjíždějící akademičtí/ostatní pracovníci (tj. počty výjezdů) – pracovníci, kteří v roce 2022 absolvovali (ukončili) zahraniční pobyt; započítávají se i ti pracovníci, jejichž pobyt začal v roce 2021. Započítávají se pouze pracovníci, jejichž pobyt trval alespoň 5 dní.
- **** – Přijíždějící akademičtí/ostatní pracovníci (tj. počty příjezdů) – pracovníci, kteří přijeli v roce 2022; započítávají se i ti pracovníci, jejichž pobyt začal v roce 2021. Započítávají se pouze pracovníci, jejichž pobyt trval alespoň 5 dní.
- ***** – V tabulce 7.2 Mobilita studentů a akademických a ostatních pracovníků podle zemí je uveden výčet všech zemí; účelem je usnadnění zpracování získaných

údajů MŠMT. Současně by neměl představovat dodatečnou zátěž pro vysoké školy při vyplňování. V případě neexistence mobility z dané země nevyplňujte prosím buňku.

- ***** – Absolventskou stáží se rozumí praktická stáž v zahraničním podniku nebo organizaci v délce 2-12 měsíců, která je započatá po úspěšném absolvování studia a ukončená do jednoho roku od absolvování studia. Absolventská stáž je realizována na základě trojstranné dohody mezi studentem, vysílající vysokoškolskou institucí a přijímající organizací, institucí, podnikem.
- ¹ – Pokud se virtuální mobilita na VŠ nerealizovala, uveďte nulu. Pokud se realizovala, ale nejsou dostupné údaje, uveďte kvalifikovaný odhad a buňku/buňky/sloupec okomentujte (např. kvalifikovaný odhad). Pokud není možné uvést kvalifikovaný odhad, nechte buňku prázdnou a a buňku/buňky/sloupec okomentujte (např. n/a).

6.3.3 Mobilita absolventov

Tabuľka vyžaduje dáta o podiele a počte absolventov jednotlivých stupňov štúdia, ktorí absolvovali počas svojho štúdia zahraničný pobyt s dĺžkou minimálne 14 dní. Vo výročnej správe o činnosti táto tabuľka zodpovedá tabuľke číslo 7.3.

Vysoká škola (název)	Bakalárske štúdium		Magisterské štúdium		studium		Doktorské štúdium		CELKEM	
	podíl	počet	podíl	počet	podíl	počet	podíl	počet	podíl	počet
Fakulta 1 (název)										
Podíl [%] a počet absolventů, kteří během svého studia vyjeli na zahraniční pobyt v délce alespoň 14 dní										
Podíl [%] a počet absolventů doktorského studia, u nichž délka zahraničního pobytu dosáhla alespoň 1 měsíc (tj. 30 dní)									0,0%	0,0
Fakulta 2 (název)										
Podíl [%] a počet absolventů, kteří během svého studia vyjeli na zahraniční pobyt v délce alespoň 14 dní										
Podíl [%] a počet absolventů doktorského studia, u nichž délka zahraničního pobytu dosáhla alespoň 1 měsíc (tj. 30 dní)									0,0%	0,0
Vysoká škola (název)										
Podíl [%] a počet absolventů, kteří během svého studia vyjeli na zahraniční pobyt v délce alespoň 14 dní										
Podíl [%] a počet absolventů doktorského studia, u nichž délka zahraničního pobytu dosáhla alespoň 1 měsíc (tj. 30 dní)									0,0%	0,0
Vysoká škola (název)	0,0%	0,0	0,0%	0,0	0,0%	0,0	0,0%	0,0	0,0%	0,0

Obr. 6.3: Tabuľka mobility absolventov

Kapitola 7

Návrh riešenia

Po tom, ako sme v spolupráci s databázovým oddelením VUT usúdili, že nebude možné vytvoriť jedno komplexné databázové view, na ktoré by sme vytvorili jednu komplexnú databázovú kocku, som začal vytvárať tri návrhy databázových pohľadov.

7.1 Databázové pohľady

Vďaka získaniu potrebných práv v IS Apollo mi bol sprístupnený modul *Dokumentace DB*, v ktorom sa nachádzajú ERD schémy, ktoré reprezentujú štruktúry centrálnej databázy VUT. Tieto ERD schémy obsahujú tabuľky modelu ERD (entity-relationship diagram) a znázorňujú vzájomné väzby medzi tabuľkami. Na základe týchto ERD schém som navrhol databázové pohľady, ktoré mi potom databázové oddelenie VUT vytvorilo a sprístupnilo v testovacej, ale aj ostrej centrálnej databáze VUT.

Je potrebné spomenúť, že v stĺpci *Tabuľka v ERD schéme* som zmenil presný názov tabuliek v centrálnej databáze VUT z dôvodu bezpečnosti. Avšak názvy som zmenil tak, aby názov stále zodpovedal sémantike dát, ktoré táto tabuľka obsahuje.

7.1.1 Pohľad pre kocku *Zahraniční studenti Příjezd*

Tento databázový pohľad obsahuje dáta o zahraničných študentoch, ktorí prišli študovať na VUT.

Atribút	Typ v kocke	Tabuľka v ERD schéme
STUDIUM_ID	charakteristika	STUDIUM
FAKULTA_ID	charakteristika	STUDIUM
STUDENT_ID	charakteristika	STUDIUM
ROK_STUDIA	charakteristika	STUDIUM
ZACATEK_STUDIA	charakteristika	STUDIUM
STUDIUM_ZAHR_STUDENT_ID	charakteristika	STUDIUM_ZAHR_STUDENT
ZEME	charakteristika	STUDIUM_ZAHR_STUDENT
TYP_ZAHRANICNI_CESTY_ID	charakteristika	STUDIUM_ZAHR_STUDENT
UNIVERZITA_ID	charakteristika	STUDIUM_ZAHR_STUDENT
ZAHR_CESTA_PROJEKT_ID	charakteristika	STUDIUM_ZAHR_STUDENT
DATUM_PRIJEZDU	charakteristika	STUDIUM_ZAHR_STUDENT

Tabuľka 7.1: Návrh pohľadu pre kocku *Zahraniční studenti Příjezd*

7.1.2 Pohľad pre kocku *Zahraniční cesty zaměstnanci Příjezd*

Pri tvorbe tohto databázového pohľadu som sa zameril na ERD schému, ktorá sa venovala práve dátam s prichádzajúcimi pracovníkmi, ktorí prišli na zahraničnú cestu na VUT ako zahraniční pracovníci.

Atribút	Typ v kocke	Tabuľka v ERD schéme
TYP_ZAHRANICNI_CESTY_ID	charakteristika	TYP_ZAHR_CESTY
JE_ERASMUS	charakteristika	TYP_ZAHR_CESTY
TYP_SE_STUD	charakteristika	TYP_ZAHR_CESTY
TYP_ZAMESTNANCI	charakteristika	TYP_ZAHR_CESTY
TYP_SE_STAZ	charakteristika	TYP_ZAHR_CESTY
TYP_STUD	charakteristika	TYP_ZAHR_CESTY
TYP_STAZ	charakteristika	TYP_ZAHR_CESTY
ZKRATKA	charakteristika	TYP_ZAHR_CESTY
POPIS	charakteristika	TYP_ZAHR_CESTY
ZAHR_CESTA_PRIJEZD_ID	charakteristika	PRIJEZD_ZAHR_CESTA
JMENO	charakteristika	PRIJEZD_ZAHR_CESTA
PRIJIMENI	charakteristika	PRIJEZD_ZAHR_CESTA
POHLAVI	charakteristika	PRIJEZD_ZAHR_CESTA
STATNI_PRISLUSNOST	charakteristika	PRIJEZD_ZAHR_CESTA
STAT_UNIVERZITY	charakteristika	PRIJEZD_ZAHR_CESTA
FAKULTA_VUT_ID	charakteristika	PRIJEZD_ZAHR_CESTA
ZAMESTNANI_NA_VUT	charakteristika	PRIJEZD_ZAHR_CESTA
KONTAKTNI_OSOBA_VUT_ID	charakteristika	PRIJEZD_ZAHR_CESTA
DATUM_PRIJEZDU	charakteristika	PRIJEZD_ZAHR_CESTA
DATUM_ODJEZDU	charakteristika	PRIJEZD_ZAHR_CESTA
ZAHR_CESTA_PROJEKT_ID	charakteristika	PRIJEZD_ZAHR_CESTA
VYSILAJICI_UNIVERZITA_ID	charakteristika	PRIJEZD_ZAHR_CESTA
ZAHR_CEST_PRIJ_DUVOD_ID	charakteristika	ZAHR_CEST_DUVOD
NAZEV	charakteristika	ZAHR_CEST_DUVOD

Tabuľka 7.2: Návrh pohľadu pre kocku *Zahraniční cesty zaměstnanci Příjezd*

7.1.3 Pohľad pre kocku *Zahraniční cesty Výjezd*

Tým, že sa v tomto databázovom pohľade nachádzajú dáta o akademických a technických pracovníkoch, ktorých VUT vyslalo na zahraničnú cestu, ale aj o študentoch VUT, ktorí boli taktiež vyslaní VUT na študijný pobyt, nachádza sa v tomto pohľade najviac atribútov tabuliek z centrálnej databázy. ERD schéma bola veľmi rozsiahla, ale na základe požiadaviek analytikov a samotnej výročnej správy o činnosti školy pre MŠMT, som vytvoril návrh databázového pohľadu, ktorý obsahuje všetky potrebné dáta pre dostatočné pokrytie problematiky zahraničných ciest, na ktoré VUT vysiela svojich pracovníkov a študentov. Zoznam atribútov, typ atribútov v budúcej OLAP dátovej kocke a sémantika tabuliek, z ktorých som atribúty získal, sa nachádza v tabuľke 7.3.

Atribút	Typ v kočke	Tabuľka v ERD schéme
ZAHR_CESTA_PROJEKT_ID	charakteristika	ZAHR_CESTA_PROJEKT
ZAHRANICNI_CESTA_ID	charakteristika	ZAHRANICNI_CESTA
STUDIUM_ID	charakteristika	ZAHRANICNI_CESTA
TYP_ZAHRANICNI_CESTY_ID	charakteristika	ZAHRANICNI_CESTA
UNIVERZITA_ID	charakteristika	ZAHRANICNI_CESTA
AKADEMICKY_ROK	charakteristika	ZAHRANICNI_CESTA
COUNTRY_CODE	charakteristika	ZAHRANICNI_CESTA
ZISKANE_KREDITY	charakteristika	ZAHRANICNI_CESTA
PER_ID (ID zamestnance)	charakteristika	ZAHRANICNI_CESTA
ORGUNITID	charakteristika	ZAHRANICNI_CESTA
NAKLAD_CESTA_ID	charakteristika	ZAHRANICNI_CESTA
POCET_PRACOVNICH_DNU	charakteristika	ZAHRANICNI_CESTA
OD	charakteristika	ZAHRANICNI_CESTA
DO	charakteristika	ZAHRANICNI_CESTA
DELKA_POBYTU	charakteristika	ZAHRANICNI_CESTA
CELKEM_OD	charakteristika	ZAHRANICNI_CESTA
CELKEM_DO	charakteristika	ZAHRANICNI_CESTA
CELKOVA_DELKA_POBYTU	charakteristika	ZAHRANICNI_CESTA
ZRUSENA	charakteristika	ZAHRANICNI_CESTA
ZAHR_CESTA_STIPENDIUM_ID	charakteristika	STIPEND_ZAHR_CESTA
ZAHRANICNI_CESTA_ID	charakteristika	STIPEND_ZAHR_CESTA
STIPENDIUM_ID	charakteristika	STIPEND_ZAHR_CESTA
TYP_ZAHRANICNI_CESTY_ID	charakteristika	TYP_ZAHR_CESTY
ZKRATKA	charakteristika	TYP_ZAHR_CESTY
POPIS	charakteristika	TYP_ZAHR_CESTY
JE_ERASMUS	charakteristika	TYP_ZAHR_CESTY
TYP_SE_STUD	charakteristika	TYP_ZAHR_CESTY
TYP_ZAMESTNANCI	charakteristika	TYP_ZAHR_CESTY
TYP_SE_STAZ	charakteristika	TYP_ZAHR_CESTY
TYP_STUD	charakteristika	TYP_ZAHR_CESTY
TYP_STAZ	charakteristika	TYP_ZAHR_CESTY
FAKULTA_ID	charakteristika	STUDIUM
STUDENT_ID	charakteristika	STUDIUM
ROK_STUDIA	charakteristika	STUDIUM
ZACATEK_STUDIA	charakteristika	STUDIUM
CASTKA	ukazovateľ	STIPENDIUM
MENA	ukazovateľ	STIPENDIUM
STIPENDIUM_ID	charakteristika	STIPENDIUM
TYP_STIPENDIA_ID	charakteristika	STIPENDIUM
STAV_STIPENDIA_ID	charakteristika	STIPENDIUM
SKOLNI_ROK	charakteristika	STIPENDIUM
SPP_PRVEK	charakteristika	STIPENDIUM
UCETNI_ROK	charakteristika	STIPENDIUM
DRUH_STIPENDIA_ID	charakteristika	DRUH_STIPENDIA

NAZEV	charakteristika	DRUH_STIPENDIA
DRUH_STIP_ZKRATKA	charakteristika	DRUH_STIPENDIA
DRUH_STIP_POPIIS	charakteristika	DRUH_STIPENDIA
TYP_STIP_ZKRATKA	charakteristika	TYP_STIPENDIA
TYP_STIP_POPIIS	charakteristika	TYP_STIPENDIA
STAV_STIP_STAV	charakteristika	STAV_STIPENDIA
STAV_STIP_POPIIS	charakteristika	STAV_STIPENDIA

Tabuľka 7.3: Návrh pohľadu pre kocku *Zahraniční cesty Výjezd*

7.2 Návrh dátových kociek

Pre správny návrh OLAP dátových kociek je veľmi dôležité poznať význam dát, ktoré nám poskytol dátový zdroj. Je to dôležité práve z dôvodu tvorby charakteristík a ukazovateľov, z ktorých sa následne dátová kocka vytvorí.

V dátovej kocke potom môžeme charakteristiky zoskupovať podľa sémantického významu alebo podľa väzby na dôležitú charakteristiku v dátovej kocke. Zoskupenie charakteristík môžeme vykonať napríklad s charakteristikami *Získané kredity* a *Rok štúdia*, ktoré umiestnime spoločne s charakteristikou *Student ID* do zoskupenia *Student*, pretože sa týkajú iba študentov a nie zamestnancov. To isté by sa dalo spraviť pre charakteristiky *Vedúci*, *Pracovné oddelenie* a *Zamestnanec ID*, ktoré by sme umiestnili do zoskupenia *Zamestnanec*. Tieto zoskupenia pomáhajú hlavne analytikom rýchlejšie sa zorientovať a intuitívnejšie pracovať s dátovou kockou.

Pri dátových kockách *Zahraniční študenti Príjezd* a *Zahraniční cesty zamestnanci Príjezd* bol návrh dátovej kocky veľmi jednoduchý. V oboch prípadoch sú charakteristiky zamerané buď na zamestnancov, alebo na študentov, takže nebolo potrebné vytvárať sémantické zoskupenia charakteristík. Stačilo vytvorené charakteristiky a ukazovatele vložiť do dátovej kocky, a tak dátovú kocku vytvoriť.

V prípade dátovej kocky *Zahraniční cesty Výjezd* je situácia iná. V dátovej kocke sa miešajú charakteristiky akademických a technických pracovníkov VUT s charakteristikami naviazanými iba na študentov VUT. Navyše ešte existovali charakteristiky, ktoré boli spoločné aj pre pracovníkov VUT, aj pre študentov VUT. Tieto skutočnosti som pri návrhu dátovej kocky zohľadnil a vytvoril som 3 zoskupenia charakteristík, aby som analytikom uľahčil tvorbu analýz a oddelil tak charakteristiky, ktoré nemajú medzi sebou väzby a nevytvárajú žiaden rozumný výstup pri tvorbe analýz. Tieto tri zoskupenia sú:

- **Studenti** – obsahuje charakteristiky, ktoré majú väzby len na charakteristiky *Stúdium ID* a *Student ID*, a nie na charakteristiku *Zamestnanec ID*.
- **Zamestnanci** – obsahuje charakteristiky, ktoré majú väzbu len na charakteristiku *Zamestnanec ID* a nemajú väzbu na charakteristiky *Stúdium ID* a *Student ID*.
- **Obecné** – obsahuje všetky ostatné charakteristiky, teda tie, ktoré majú väzby na všetky tri charakteristiky *Stúdium ID*, *Student ID* a *Zamestnanec ID*.

7.2.1 Návrh charakteristík

Pri návrhu charakteristík je potrebné opäť dobre poznať sémantiku dát z viacerých dôvodov:

1. pridanie správneho názvu a popisu charakteristiky,
2. tvorba atribútov, hierarchií a číselníka charakteristiky v prípade, ak tieto dáta máme k dispozícii.

Správny dátový typ dokáže vývojár vyčítať už pri napojení externého zdroja (databázového pohľadu) a vytvorení *DataSource objektu*. V *DataSource objekte* SAP Business Warehouse navrhne optimálny dátový typ so znakovou dĺžkou poľa.

Dátový typ sa medzi charakteristikou a polom v *DataSource objekte* môže však líšiť v prípade, ak chceme zmeniť formát dát pomocou transformácie do tvaru, v ako dáta nie sú uložené v zdrojovom systéme – v prípade tejto práce v databázovom pohľade.

Túto skutočnosť som musel pri návrhu zohľadniť hlavne pri charakteristikách, ktoré mali charakter dátumu vo formáte **DD.MM.YYYY**. Pre správne spracovanie a formátovanie dátumov bolo potrebné v *DataSource objekte* pretypovať tieto vstupné polia na dátový typ CHAR s dĺžkou 10 znakov, pri transformácii použiť transformačnú rutinu v kombinácii s ABAP konverznou funkciou `'CONVERT_DATE_TO_INTERNAL'` a uložiť výsledok do charakteristiky s dátovým typom DATS s dĺžkou 8 znakov.

V návrhu charakteristík som zohľadnil taktiež možnosť číselníkov. Pre charakteristiky *Stav stipendia ID*, *Druh stipendia ID* a *Typ stipendia ID* som navrhol číselníky, v ktorých ku konkrétnemu ID naviazem aj jeho skratku a popis. Pomocou číselníkov zaistím v dátových kockách jednoduchšie porozumenie číslam ID, pretože číslo ID má pre neznalého užívateľa nulovú výpovednú hodnotu, ale popis daného čísla ID už neznalému užívateľovi poskytne informáciu o tom, čo toho číslo ID znamená.

7.2.2 Návrh ukazovateľov

Už pri návrhu databázového pohľadu som si uvedomil, že v dátových kockách *Zahraniční studenti Příjezd* a *Zahraniční cesty zaměstnanci Příjezd* nebude ukazovateľ niť hlbší sémantický význam, ako napríklad výška dotácie v eurách, ale ukazovateľ bude mať v transformácii priradenú konštantu 1 a bude reprezentovať počet záznamov, ktoré boli do dátovej kocky vložené. Táto skutočnosť vznikla preto, lebo v dátových pohľadoch sa nenachádzal numerický údaj, ktorý by mal vzťah k hlavným charakteristikám danej dátovej kocky.

Avšak ukazovateľ v dátovej kocke musí existovať, pretože vytvára agregáčnú číselnú závislosť medzi charakteristikami a je základnou zložkou OLAP dátovej kocky.

Pri dátovej kocke *Zahraniční cesty Výjezd* je situácia odlišná. Už z počiatkových požiadaviek, z výročnej správy o činnosti školy od MŠMT a tabuľky zobrazenej na obrázku 6.1 bolo zrejmé, že bude potrebné získať z centrálnej databázy VUT číselný údaj o *výške Dotácie v tis. Kč* práve pre zahraničné cesty, na ktoré VUT vyslalo svojich pracovníkov a študentov. S touto skutočnosťou bolo nutné počítať už pri návrhu databázového pohľadu. Preto som v návrhu štruktúry dátovej kocky navrhol ukazovateľ s dátovým typom **CURR**, ktorý mal k sebe priradený aj údaj o mene, v akej bola dotácia zahraničnej ceste udelená. Údaj o mene bolo potrebné zahrnúť z toho dôvodu, že v centrálnej databáze boli pridelené dotácie v rôznych menách (vo väčšine prípadoch išlo o eurá alebo o české koruny). Rozhodol som sa avšak, že pre dátovú kocku vytvorím ďalší ukazovateľ, ktorý bude ukazovať počet záznamov vložených do dátovej kocky – rovnaký ukazovateľ, aký som navrhol pre dátové kocky

Zahraniční studenti Přijezd a Zahraniční cesty zaměstnanci Přijezd. Z toho teda vyplýva, že dátová kocka *Zahraniční cesty Výjezd* bude mať vo finálnom stave až dva ukazovatele.

7.3 Návrh transformácie

Pri návrhu transformácie v SAP BW je cieľom prepojiť správne polia z *DataSource* objektu do charakteristík a ukazovateľov v dátovej kocke. V procese návrhu som preto pomenoval charakteristiky kociek sémanticky rovnako ako dátové polia v *DataSource* objekte. Uľahčilo mi to potom pri reálnej tvorbe transformácie prepojenie a mapovanie polí medzi *DataSource* objektmi a poliami v dátových kockách.

Návrh transformácie v SAP BW nemá zmysel robiť dopredu veľmi podrobne, pretože až pri implementácii transformácie vznikajú situácie, ako napríklad nesprávne formátovanie dát alebo problémy s mapovaním dát do polí v dátových kockách. Tieto situácie sa nedajú ľahko ošetriť ešte pred implementáciou, pretože sú závislé od skutočného uloženia a formátu dát v dátovom zdroji.

Kapitola 8

Implementácia

Pri implementácii a tvorbe dátových kociek som postupoval podľa toku vývoja v SAP BW zobrazenom na obrázku 5.1. Využil som primárne návrhy, ktoré som zrealizoval v počiatočnej fáze tejto práce.

8.1 Tvorba dátového zdroja – DataSource

Po úspešnom návrhu databázových pohľadov a poskytnutia databázových pohľadov od databázového oddelenia CVIS VUT bolo všetko pripravené na to, aby som mohol vytvoriť *DataSource* objekty, ktoré slúžia na prepojenie externého zdroja dát do systému SAP BW. Podrobnejší popis *DataSource* objektu je v sekcii 5.2.

8.1.1 Dátový zdroj *Zahraniční cesty zaměstnanci Příjezd*

Tento dátový zdroj slúži na napojenie databázového pohľadu pre dátovú kocku *Zahraniční cesty zaměstnanci Příjezd*, ktorý bol navrhnutý a popísaný v sekcii 7.1.2.

Po vytvorení objektu *DataSource* som v nastavenia napojil poskytnutý databázový pohľad **ST01.V_SAPBW_ZAHR_CESTY_PRIJEZD**.

Po úspešnom napojení mi SAP BW poskytol náhľad na polia, ktoré sa v databázovom pohľade nachádzali spolu s dátovými typmi a dĺžkami jednotlivých polí. Keďže sa v databázovom pohľade nachádzali polia s dátovým typom *DATS* (polia, ktoré obsahujú záznamy vo formáte dátumu), podľa predchádzajúcich skúseností som vedel, že ich musím pretypovať do dátového typu *CHAR* a v transformácii pomocou transformačnej rutiny preformátovať do interného formátu SAP BW. Toto je jediný spôsob, ako z formátu dátumu z centrálnej databázy VUT môže vývojár získať správny interný formát dátumu pre SAP BW.

Po správnej konfigurácii vstupných polí v *DataSource* objekte som objekt uložil a pripravil na použitie.

8.1.2 Dátový zdroj *Zahraniční cesty Výjezd*

Dátový zdroj slúži na napojenie databázového pohľadu pre dátovú kocku *Zahraniční cesty Výjezd*, v ktorého nastaveniach bolo potrebné poskytnutý databázový pohľad od databázového oddelenia VUT s názvom **ST01.V_SAPBW_ZAHR_CESTY_STIP** napojiť.

Tak, ako v prípade dátového zdroja *Zahraniční cesty zaměstnanci Příjezd* popísaného v sekcii 8.1.1, musel som tiež v *DataSource* objekte pretypovať polia, ktoré reprezentovali dáta v dátumovom formáte z dátového typu *DATS* na dátový typ *CHAR*.

8.1.3 Dátový zdroj *Zahraniční studenti Příjezd*

Dátový zdroj slúži na poskytnutie dát pre dátovú kocku *Zahraniční studenti Příjezd*. Dáta sa nachádzajú v databázovom pohľade **ST01.V_SAPBW_ZAHR_STUD_PRIJEZD**, ktorý som v nastaveniach napojil na tento *DataSource* objekt.

Opätovne som vykonal rovnaké úpravy polí v *DataSource* objekte ako pri predchádzajúcich *DataSource* objektoch popísaných v sekciách 8.1.1 a 8.1.2.

8.1.4 Dátový zdroj *Zahraniční cesty typ*

Tento dátový zdroj sa v celku líši od predchádzajúcich troch. Je špeciálny v tom, že nebude slúžiť na poskytnutie dát priamo do dátovej kocky, ale bude slúžiť ako zdroj dát pre atribúty charakteristiky **Typ zahraniční cesty**.

Vo všetkých predchádzajúcich troch zdrojoch sa nachádza pole, ktoré reprezentuje dáta *typu zahraničnej cesty ID*, teda dáta sú v tvare identifikačného čísla. Každé identifikačné číslo má v systéme VUT inú sémantiku a poskytuje informáciu o tom, o aký typ zahraničnej cesty ide.

Ku každému identifikačnému číslu typu zahraničnej cesty je priradený popis, popis v anglickom jazyku, skratka a skratka v anglickom jazyku.

Databázový pohľad s názvom **ST01.V_SAPBW_ZAHRANICNI_CESTY_TYP** od databázového oddelenia VUT obsahuje práve tieto štyri polia, ktoré v budúcom vývoji použijem ako zdroj dát pre atribúty charakteristiky **Typ zahraniční cesty**.

8.2 Nahrávanie a správa dát v PSA

PSA (Persistent Staging Area) je interná databázová tabuľka, ktorá slúži ako dočasné úložisko dát. Podrobnejšie som sa problematike PSA venoval v sekcii 5.2.1.

Pre nahratie dát z dátového zdroja do PSA je potrebné vytvoriť *InfoPacket*, ktorý toto nahratie do PSA vykoná.

Pri vypracovaní tejto práce som pre každý *DataSource* objekt vytvoril vlastný *InfoPacket* a pomocou neho nahral dáta do samostatného PSA – pre každý dátový zdroj existuje samostatné PSA.

8.3 Tvorba charakteristík a ukazovateľov

Ako bolo už v tejto práci spomenuté, tvorba charakteristík sa odvíja najmä od zdroja dát a samozrejme od požadovaného výstupu, ktorý chceme tvorbou OLAP dátových kociek dosiahnuť.

Pri tvorbe charakteristík pre dátové kocky, ktoré sú výstupom tejto práce, som využil rovnaký spôsob, a teda na základe dátových zdrojov, návrhu a požadovaného výsledku som vytvoril potrebné charakteristiky s požadovaným dátovým typom tak, aby som bol schopný vykonať úspešnú transformáciu dát z dátového zdroja (v tejto fáze už transformáciou dát uložených v PSA) do dátových kociek a získal v charakteristikách dátových kociek požadovaný výstupný formát týchto dát.

8.3.1 Charakteristiky pre kocku *Zahraniční cesty zaměstnanci Příjezd*

Technické meno	Popis	Dátový typ	Délka
JEERASMUS	Je Erasmus+	NUMC	2
JE_STAZ	Je Stáž	NUMC	2
JE_STUD	Je Studium	NUMC	2
KONTAKT_O	Kontaktní osoba VUT	NUMC	8
POHLAVI	Pohlaví	CHAR	2
STAT_PRIS	Státní příslušnost	CHAR	2
STAT_UNIV	Stát univerzity	CHAR	2
TYP_SE_ST	Je Socrates/Erasmus+ studium	NUMC	2
TYP_ZAH_C	Typ zahraniční cesty	NUMC	8
TYP_ZAMES	Je zaměstnanec	NUMC	2
T_SE_STAZ	Je Socrates/Erasmus+ stáž	NUMC	2
ZAM_NAVUT	Je zaměstnaní na VUT	NUMC	2
ZC_FAK_ID	Fakulta / součást	NUMC	8
ZC_JMENO	Jméno	CHAR	100
ZC_PRIJME	Příjmení	CHAR	100
ZA_CESTA	Zahraniční cesta příjezd ID	NUMC	8
VYS_UN_ID	Vysílající univerzita	NUMC	8
DAT_PRIJ	Datum příjezdu	DATS	8
ZDAT_OD	Datum odjezdu	DATS	8
ZDEL_POBY	Délka zahraniční cesty	CHAR	5

Tabulka 8.1: Charakteristiky pre kocku *Zahraniční cesty zaměstnanci Příjezd*

8.3.2 Charakteristiky pre kocku *Zahraniční cesty Výjezd*

Technické meno	Popis	Dátový typ	Dĺžka
ZCELDELP	Celková dĺžka pobytu	CHAR	10
ZDELKA_P	Dĺžka pobytu	CHAR	10
ZDRUH_SID	Druh stipendia ID	NUMC	2
ZKREDITY	Získané kredity	CHAR	7
ZSTAVSTID	Stav stipendia ID	NUMC	8
ZTSTIPID	Typ stipendia ID	NUMC	8
ZUCETROK	Účetní rok	CHAR	6
STUDI_ID	Studium ID	NUMC	8
STP_ID	Stipendium ID	NUMC	10
JE_STAZ	Je Stáž	NUMC	2
JE_STUD	Je Studium	NUMC	2
TYP_SE_ST	Je Socrates/Erasmus+ studium	NUMC	2
T_SE_STAZ	Je Socrates/Erasmus+ stáž	NUMC	2
S_FAK_ID	Fakulta / součást	NUMC	8
STUDEN_ID	Student ID	NUMC	8
ROK_STD	Rok studia	NUMC	2
SPP_PRVEK	SPP prvek	CHAR	24
ZDS_POPIŠ	Druh stipendia popis	CHAR	250
ZCASTKAC	Částka	CHAR	10
ZPOCP_DNI	Počet pracovních dní	NUMC	6
ZZAMESTID	Zaměstnanec ID	NUMC	8
ZORGUNID	Vysílající útvar	NUMC	8
ZAHR_C_ID	Zahraniční cesta ID	NUMC	8
ZCELDO	Cesta do celkem	DATS	8
ZCELKOD	Cesta od celkem	DATS	8
JEERASMUS	Je Erasmus+	NUMC	2
TYP_ZAMES	Je zaměstnanec	NUMC	2
ZZAHR_DO	Cesta do	DATS	8
ZZAHR_OD	Cesta od	DATS	8
ZZRUSENA	Je zrušena	CHAR	2
ZUNIV_ID	Univerzita ID	NUMC	8
TYP_ZAH_C	Typ zahraniční cesty	NUMC	8
ZEME	Stát	CHAR	3
Z_CDELDNI	Délka cesty dní celkem	CHAR	5
Z_DELPDNI	Délka cesty dní	CHAR	5
ZAKROK	Akademický rok	CHAR	7
ZSKROK	Školní rok	CHAR	7

Tabuľka 8.2: Charakteristiky pre kocku *Zahraniční cesty Výjezd*

8.3.3 Charakteristiky pre kocku *Zahraniční studenti Příjezd*

Technické meno	Popis	Dátový typ	Dĺžka
ROK_STD	Rok studia	NUMC	2
STUDEN_ID	Student ID	NUMC	8
STUDI_ID	Studium ID	NUMC	8
S_FAK_ID	Fakulta / součást	NUMC	8
ZAC_STUD	Začátek studia	DATS	8
STU_Z_ID	Studium zahraniční student ID	NUMC	8
ZEME	Stát	CHAR	3
COUN_CODE	Kód krajiny	CHAR	2
DAT_PRIJ	Datum příjezdu	DATS	8
ZC_PROJID	Zahraniční cesta projekt ID	NUMC	4
TYP_ZAH_C	Typ zahraniční cesty	NUMC	8
VYS_UN_ID	Vysílající univerzita ID	NUMC	8

Tabuľka 8.3: Charakteristiky pre kocku *Zahraniční studenti Příjezd*

8.3.4 Tvorba atribútov a číselníkov charakteristík

Pri vypracovaní tejto práce som využil atribúty charakteristík, pri charakteristikách ktorých dáta boli identifikačné čísla (ID), napríklad ID osôb, ID zahraničných univerzít, ID fakúlt VUT a pri ID typoch zahraničných ciest.

Pri tvorbe charakteristík pre identifikačné čísla osôb, zahraničných univerzít a fakúlt VUT som vytvoril charakteristiky ako referencie na otestované a používané charakteristiky v systéme SAP BW. Mnou vytvorené charakteristiky prevzali atribúty z referenčných charakteristík a boli pripravené na použitie.

Pri tvorbe charakteristiky *Typ zahraničnej cesty* som nemohol zvoliť rovnaký postup, pretože táto charakteristika v systéme SAP BW ešte neexistovala. Dáta pre atribúty tejto charakteristiky som získal pomocou *DataSource* objektu **Zahraniční cesty typ**, ktorému som sa podrobnejšie venoval v sekcii 8.1.4. Dátový zdroj obsahoval dáta s významom: skratka typu zahraničnej cesty, skratka typu zahraničnej cesty v anglickom jazyku, popis typu zahraničnej cesty, popis typu zahraničnej cesty v anglickom jazyku. Pre tieto dáta bolo potrebné vytvoriť separátne charakteristiky, ktoré potom budú priradené ako atribúty charakteristiky *Typ zahraničnej cesty*. Do atribútov charakteristiky je však potrebné nahráť dáta z *DataSource* objektu rovnakým spôsobom ako do dátovej kocky – pomocou transformácie. Implementácii tejto transformácie sa venujem v sekcii 8.5.4.

Technické meno	Popis	Dátový typ	Dĺžka
ZPOPISTYP	Popis	CHAR	250
ZPOPISTYE_ID	Popis EN	CHAR	250
ZZKRATKAT_ID	Zkratka	CHAR	25
ZZKRATKAE_ID	Zkratka EN	CHAR	25

Tabuľka 8.4: Atribúty charakteristiky *Typ zahraničnej cesty*

Pre charakteristiky *Stav stipendia ID*, *Druh stipendia ID* a *Typ stipendia ID* som navrhol číselníky, ktoré som aj implementoval. Ako zdroj dát číselníkov som použil *DataSource*

objekt pre kocku *Zahraniční cesty Výjezd*, v ktorej som práve navrhol zahrnutie týchto dát z centrálnej databázy VUT. V SAP BW sa číselníky nahrávajú ako tzv. texty charakteristiky, do ktorých som nahrával dáta, ktoré mali povahu stredne dlhého textu a extra dlhého textu. Do stredne dlhých textov som nahral skratky týchto identifikačných čísel a do extra dlhých textov som nahral popis týchto identifikačných čísel. Dáta do textov charakteristiky sa rovnako ako pri atribútoch charakteristiky nahrávajú pomocou transformácie. Tieto transformácie som musel pre každú charakteristiku implementovať zvlášť. Implementácii sa venujem v sekcii 8.5.5.

8.3.5 Tvorba ukazovateľov

Pri všetkých troch dátových kockách som zvolil dva rovnaké ukazovatele, ktoré oba sémanticky reprezentovali počet záznamov, jeden ukazovateľ v celých číslach, druhý v desatinných číslach kvôli kontrole agregácií a zabezpečenia správnosti. Pre dátovú kocku *Zahraniční cesty Výjezd* som vytvoril aj dodatočný ukazovateľ s názvom **Částka** a reprezentoval výšku grantu alebo dotácie pridelenú jednotlivých zahraničných cestám.

Technické meno	Popis	Dátový typ
ZC_PO CET	Počet záznamov	DEC
ZH_PO_INT	Počet záznamov celé číslo	INT4
ZCASTKA	Částka	CURR

Tabuľka 8.5: Ukazovatele

8.4 Modelovanie dátových kociek

Samotné modelovanie dátových kociek nie je veľmi náročná vývojárska činnosť. Stačí založiť nový objekt dátovej kocky, vytvoriť takzvané zoskupenia, ktoré zoskupujú charakteristiky, ktoré majú na seba väzbu alebo ktoré majú sémanticky podobný význam a zároveň sa ich význam líši od charakteristík v ďalších zoskupeniach a na záver do zoskupení vložiť už vytvorené charakteristiky.

8.4.1 Modelovanie dátovej kocky *Zahraniční cesty zaměstnanci Příjezd*

Vzhľadom na fakt, že sa v dátovej kocke nachádzajú charakteristiky len o prichádzajúcich zamestnancoch na VUT, nebolo potrebné pre lepšiu orientáciu v dátovej kocke vytvárať zoskupenia charakteristík. Všetky charakteristiky pripravené pre dátovú kocku *Zahraniční cesty zaměstnanci Příjezd* som teda vložil do jedného zoskupenia nazvaného *Obecné*.

8.4.2 Modelovanie dátovej kocky *Zahraniční cesty Výjezd*

Dátová kocka *Zahraniční cesty Výjezd* je najkomplexnejšia zo všetkých troch kociek vytvorených v tejto práci, pretože obsahuje dáta o odchádzajúcich študentoch, akademických a technických pracovníkov VUT na zahraničné cesty či už na študijné pobyty, ale aj vykonávať pracovné a výskumné činnosti. Obsahuje taktiež najviac dát zo všetkých troch dátových kociek.

Z návrhu kocky a databázového pohľadu je jasné, že obsahuje veľké množstvo rozdielnych charakteristík, ktoré nie každá medzi sebou súvisí. Preto som sa pri modelovaní tejto dátovej kocky rozhodol vytvoriť tri zoskupenia charakteristík, do ktorých som uložil sémanticky

vzájomne súvisiace charakteristiky a oddelil tie, ktoré nemali medzi sebou žiadnu vzájomnú väzbu či už sémantickú, ale aj z pohľadu závislostí v dátovom sklade. Vytvorené zoskupenia:

- **Studenti** – sémantická skupina obsahuje charakteristiky súvisiace so študentmi, ktorí boli vyslaní na zahraničné cesty, vo väčšine prípadov študijné pobyty. To znamená, že ide o charakteristiky, ktorých dáta majú v databázovom pohľade priamy vzťah len s dátami charakteristiky *Student ID* (identifikačné číslo študenta) a nemajú vzťah s dátami charakteristiky *Zaměstnanec ID*.

Zoznam charakteristík, ktoré patria do tohto zoskupenie:

Technické meno	Popis
ZCELDELP	Celková dĺžka pobytu
ZDELKA_P	Dĺžka pobytu
ZDRUH_SID	Druh stipendia ID
S_FAK_ID	Fakulta / súčasť
ZKREDITY	Získané kredity
ROK_STD	Rok Studia
STUDEN_ID	Student ID
JE_STAZ	Je Stáž
JE_STUD	Je Studium
TYP_SE_ST	Je Socrates/Erasmus+ studium
T_SE_STAZ	Je Socrates/Erasmus+ stáž
ZSTAVSTID	Stav stipendia ID
ZTSTIPID	Typ stipendia ID
ZUCETROK	Účetní rok
SPP_PRVEK	SPP prvek
STUDI_ID	Studium ID
STIP_ID	Stipendium ID
ZDS_POPIS	Druh stipendia popis
ZCASTKAC	Částka
ZSKROK	Akademický rok stip.

Tabuľka 8.6: Charakteristiky, ktoré patria do zoskupenia *Studenti*

- **Zaměstnanci** – skupina obsahuje charakteristiky, ktorých dáta majú vzťah v databázovom pohľade len s dátami charakteristiky *Zaměstnanec ID* a nie s dátami charakteristiky *Student ID*.

Zoznam charakteristík, ktoré patria do tohto zoskupenie:

Technické meno	Popis
ZPOCP_DNI	Počet pracovných dní
ZZAMESTID	Zaměstnanec ID
ZORGUNID	Vyslající útvar

Tabuľka 8.7: Charakteristiky, ktoré patria do zoskupenia *Zaměstnanci*

- **Obecné** – toto zoskupenie je univerzálne a boli sem pridané charakteristiky, ktoré nebolo možné pridať do predchádzajúcich zoskupení. Ide teda o charakteristiky, ktoré majú vzťah aj na charakteristiku *Student ID*, aj na charakteristiku *Zaměstnanec ID*.

Zoznam charakteristík, ktoré patria do všeobecného zoskupenia:

Technické meno	Popis
ZAHR_C_ID	Zahraniční cesta ID
ZCELDO	Cesta do celkem
ZCELKOD	Cesta od celkem
JEERASMUS	Je Erasmus+
TYP_ZAMES	Je zaměstnanec
ZZAHR_DO	Cesta do
ZZAHR_OD	Cesta od
ZZRUSENA	Je zrušena
ZUNIV_ID	Univerzita ID
TYP_ZAH_C	Typ zahraniční cesty
ZEME	Stát
Z_CDELDNI	Délka cesty dní celkem
Z_DELPDNI	Délka cesty Dní
ZAKROK	Akademický rok

Tabulka 8.8: Charakteristiky, ktoré patria do všeobecného zoskupenia

8.4.3 Modelovanie dátovej kocky *Zahraniční studenti Příjezd*

Pri modelovaní dátovej kocky *Zahraniční studenti Příjezd* nastala rovnaká situácia, ako pri dátovej kocke *Zahraniční cesty zaměstnanci Příjezd*, teda nebolo potrebné vytvárať zoskupenia charakteristík, pretože dátová kocka obsahovala dáta zamerané len na prichádzajúcich zahraničných študentov. Vytvoril som teda všeobecné zoskupenie s menom *Obecné*, do ktorého som vložil všetky navrhnuté charakteristiky pre túto dátovú kocku.

8.5 Implementácia transformácie

Transformácia je jednou z najdôležitejších častí celého procesu vývoja dátovej kocky, pretože vďaka nej transformujeme dáta z dátového zdroja, dáta spracujeme a preformátujeme do tvaru, ktorý chceme mať vo výslednej dátovej kocke. Bližší popis transformácie sa nachádza v sekcii 5.6.

V SAP BW transformácia figuruje ako objekt, ktorý je priradený dátovej kocke. Po vytvorení objektu transformácie sa zobrazia nastavenia danej transformácie. Štruktúra transformácie pozostáva z dvoch častí. Na ľavej strane sa nachádza napojený dátový zdroj (*Data-Source* objekt) a na pravej strane sa nachádza namodelovaná dátová kocka. Úlohou vývojára je prepojiť polia *DataSource* objektu a využiť jednu z transformačných funkcií (spôsobov priradenia) popísaných v sekcii 5.6.

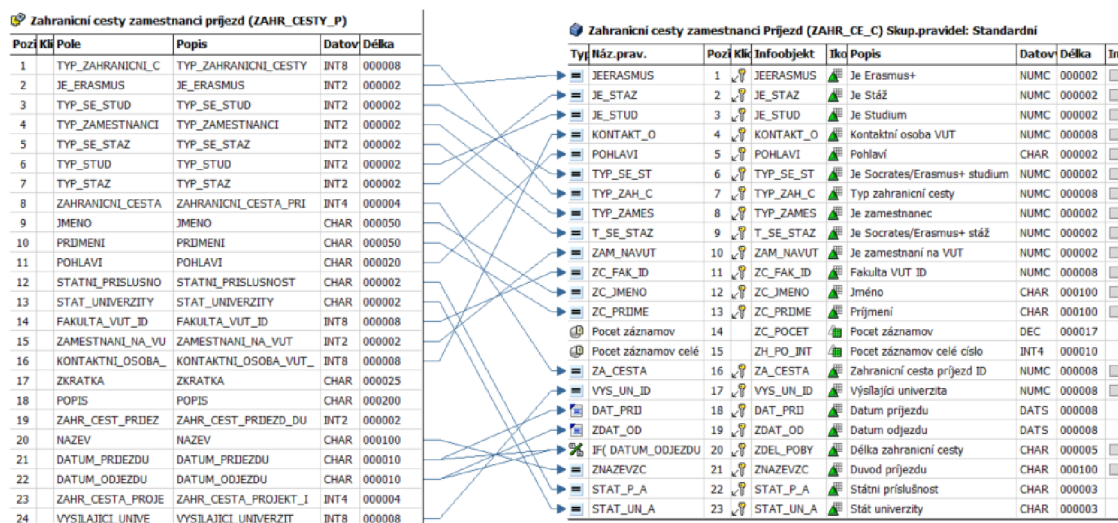
8.5.1 Transformácia kocky *Zahraniční cesty zaměstnanci Příjezd*

Z obrázku 8.1 môžeme vidieť, že v tejto transformácii som vo väčšine prípadov využil metódu priameho priradenia polí z *DataSource* objektu do charakteristík dátovej kocky. Dôvodom je skutočnosť, že dáta v dátovom zdroji – databázovej tabuľke boli už v tvare rovnakom, aký bol požadovaný v dátovej kocke.

V transformácii som využil dvakrát priradenie konštantou, a to pri ukazovateľoch v dátovej kocke. Priradil som vždy číslo jeden ako konštantu za účelom plnenia funkcie počítadla nahraných záznamov do dátovej kocky.

Pri transformácii dát do charakteristík **DAT__PRIJ** (Datum příjezdu) a **ZDAT__OD** (Datum odjezdu) som kvôli potrebným konverziám dátumov do interného formátu SAP musel použiť transformáciu pomocou rutiny, a teda musel som naprogramovať krátky program v jazyku ABAP, ktorý vykoná túto konverziu do interného formátu SAP BW.

Na záver som pri transformácii dát do charakteristiky **ZDEL__POBY** (Délka zahraniční cesty) využil priradenie pomocou vzorca. Vytvoril som teda pomocou vzorca veľmi krátky program, ktorý odčítal *Dátum příjezdu*, *Dátum odjezdu*, skontroloval dáta v týchto poliach a na záver previedol výsledok do interného formátu a uložil do charakteristiky.



Obr. 8.1: Transformácia kocky *Zahraniční cesty zaměstnanci Příjezd*

8.5.2 Transformácia kocky *Zahraniční cesty Výjezd*

Táto transformácia bola najzložitejšia zo všetkých, pretože dátová kocka obsahovala najviac charakteristík a je najkomplexnejšia.

V transformácii som opäť vo väčšine prípadov použil priame priradenie, pretože formát dát z databázy bol vhodný a nebolo potrebné dáta ďalej upravovať.

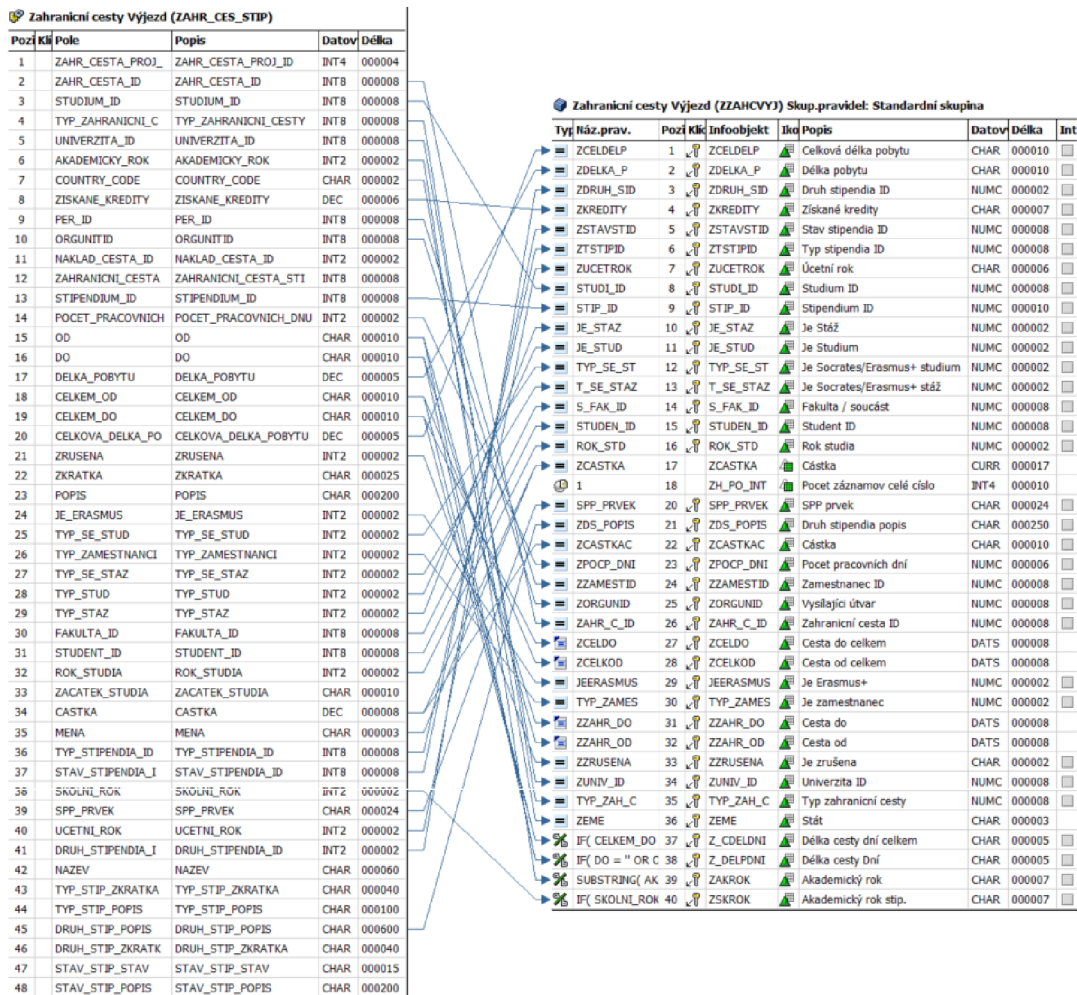
Tak ako v predchádzajúcej dátovej kocke, tak aj v tejto sa nachádza ukazovateľ, ktorého dáta reprezentujú počet nahraných záznamov v dátovej kocke – znova som použil priradenie konštantou, a teda číslom jeden.

V dátovej kocke sa nachádzajú aj charakteristiky, ktoré obsahujú dáta v dátumovom formáte. Ide o charakteristiky **ZCELKOD**, **ZCELKOD**, **ZZAHR__DO** a **ZZAHR__OD**. Pre konverziu týchto dát do interného dátumového formátu SAP BW bolo teda potrebné vykonať priradenie pomocou transformačnej rutiny a naprogramovať krátky program v jazyku ABAP.

Pri posledných štyroch charakteristikách v transformácii v dátovej kocke som sa rozhodol využiť priradenie pomocou vzorca, ktorý mi poskytol najjednoduchší spôsob, ako získať požadovaný formát dát. Pri charakteristikách **Z_CDELDNI** (Délka cesty dní celkem) a **Z_DELPDNI** (Délka cesty Dní) som potreboval získať rozdiel dátumov, a teda počet dní medzi dátumami *OD* a *DO*. Jednoduchým programom vo vzorci sa mi to podarilo

docieľiť a aj ošetriť prípadné nevhodné dáta vo vstupných poliach a zachovať tak správnosť nových dát v týchto charakteristikách.

Pri charakteristikách **ZAKROK** () a **ZSKROK** () som potreboval získať výsledný formát dát v tvare YYYY/YY, teda napríklad pre akademický rok 2020 vo vstupnom poli som vo výstupnej charakteristike potreboval získať záznam s hodnotou 2020/21.



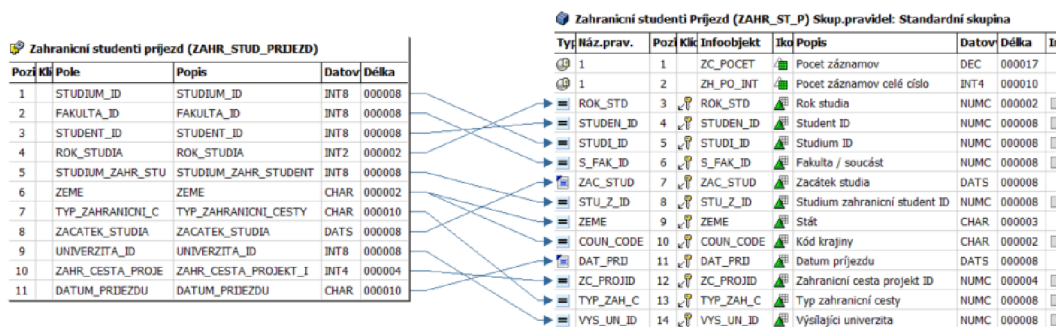
Obr. 8.2: Transformácia kocky *Zahranční cesty Výjezd*

8.5.3 Transformácia kocky *Zahranční studenti Příjezd*

Keďže dátová kocka *Zahranční studenti Příjezd* obsahuje ukazovatele, ktorých úlohou je uchovávať údaje o počte nahraných záznamov do dátovej kocky, využil som opäť priradenie pomocou konštanty tak, ako pri predchádzajúcich dátových kockách.

Nachádzajú sa tu tiež dve charakteristiky, ktoré obsahujú dátumy, takže je potrebné ich pomocou transformačnej rutiny prekonvertovať do interného formátu SAP BW.

Pri ostatných charakteristikách som využil priame priradenie, pretože nebolo nutné meniť ich formát alebo štruktúru.

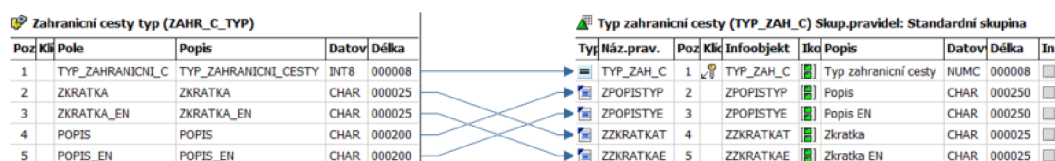


Obr. 8.3: Transformácia kocky *Zahraniční studenti Příjezd*

8.5.4 Transformácia atribútov charakteristiky *Typ zahraničnej cesty*

Transformácia dát do atribútov charakteristiky je bežná vec vývoja v SAP BW, pokiaľ chceme dosiahnuť práve použitie atribútov. Atribúty sú v podstate charakteristiky, ktoré sú priradené materskej charakteristike, takže do nich musia byť nahrané dáta rovnako, ako do bežných charakteristik.

Pri transformácii som vytvoril prepojenie medzi poliami *DataSource* objektu na ľavej strane s atribútmi na pravej strane. Z obrázku je vidieť, že som pri prepojení použil transformačné rutiny. Všetky rutiny sú rovnaké a slúžia na ošetrovanie nepovolených znakov v systéme SAP BW. Volá sa v nich funkcia, ktorá nepovolené znaky odstráni a nahradí za validné znaky. Zabezpečí tak úspešný priebeh transformácie dát z *DataSource* objektu do atribútov charakteristiky.



Obr. 8.4: Transformácia atribútov do charakteristiky *Typ zahraniční cesty*

8.5.5 Transformácia číselníkov

Ako už bolo spomenutú v návrhu charakteristík a implementácii číselníkov charakteristík, v tejto práci som implementoval tri charakteristiky, ktoré využívajú číselníky.

Číselník súži na naviazanie významu identifikačnému číslu (ID) prostredníctvom textu alebo popisu. Z tejto skutočnosti teda vyplýva, že je potrebná transformácia dát, a teda nahranie a prepojenie popisov s identifikačnými číslami.

Transformáciu číselníkov som teda implementoval pre charakteristiky **ZTSTIPID** (Typ stipendia ID), **ZSTAVSTID** (Stav stipendia ID) a **ZDRUH_SID** (Druh stipendia ID).

Pozí	Klíč	Pole	Popis	Datov	Délka
36		TYP_STIPENDIA_ID	TYP_STIPENDIA_ID	INT8	000008
37		STAV_STIPENDIA_I	STAV_STIPENDIA_ID	INT8	000008
38		SKOLNI_ROK	SKOLNI_ROK	INT2	000002
39		SPP_PRVEK	SPP_PRVEK	CHAR	000024
40		UCETNI_ROK	UCETNI_ROK	INT2	000002
41		DRUH_STIPENDIA_I	DRUH_STIPENDIA_ID	INT2	000002
42		NAZEV	NAZEV	CHAR	000060
43		TYP_STIP_ZKRATKA	TYP_STIP_ZKRATKA	CHAR	000040
44		TYP_STIP_POPIIS	TYP_STIP_POPIIS	CHAR	000100

Typ	Náz.prav.	Pozí	Klíč	Infoobjekt	Ikq	Popis	Datov	Délka	Int
ZTSTIPID	1		ZTSTIPID	ZTSTIPID		Typ stipendia ID	NUMC	000008	
0TXMTD	2		0TXMTD	0TXMTD		Strední popis	CHAR	000040	
0TXTXL	3		0TXTXL	0TXTXL		Extra dlhý popis	SSTR	001333	

Obr. 8.5: Transformácia číselníka charakteristiky *Typ stipendia ID*

Pozí	Klíč	Pole	Popis	Datov	Délka
37		STAV_STIPENDIA_I	STAV_STIPENDIA_ID	INT8	000008
38		SKOLNI_ROK	SKOLNI_ROK	INT2	000002
39		SPP_PRVEK	SPP_PRVEK	CHAR	000024
40		UCETNI_ROK	UCETNI_ROK	INT2	000002
41		DRUH_STIPENDIA_I	DRUH_STIPENDIA_ID	INT2	000002
42		NAZEV	NAZEV	CHAR	000060
43		TYP_STIP_ZKRATKA	TYP_STIP_ZKRATKA	CHAR	000040
44		TYP_STIP_POPIIS	TYP_STIP_POPIIS	CHAR	000100
45		DRUH_STIP_POPIIS	DRUH_STIP_POPIIS	CHAR	000600
46		DRUH_STIP_ZKRATK	DRUH_STIP_ZKRATKA	CHAR	000040
47		STAV_STIP_STAV	STAV_STIP_STAV	CHAR	000015
48		STAV_STIP_POPIIS	STAV_STIP_POPIIS	CHAR	000200

Typ	Náz.prav.	Pozí	Klíč	Infoobjekt	Ikq	Popis	Datov	Délka	Int
ZSTAVSTID	1		ZSTAVSTID	ZSTAVSTID		Stav stipendia ID	NUMC	000008	
0TXMTD	2		0TXMTD	0TXMTD		Strední popis	CHAR	000040	
0TXTXL	3		0TXTXL	0TXTXL		Extra dlhý popis	SSTR	001333	

Obr. 8.6: Transformácia číselníka charakteristiky *Stav stipendia ID*

Pozí	Klíč	Pole	Popis	Datov	Délka
40		UCETNI_ROK	UCETNI_ROK	INT2	000002
41		DRUH_STIPENDIA_I	DRUH_STIPENDIA_ID	INT2	000002
42		NAZEV	NAZEV	CHAR	000060
43		TYP_STIP_ZKRATKA	TYP_STIP_ZKRATKA	CHAR	000040
44		TYP_STIP_POPIIS	TYP_STIP_POPIIS	CHAR	000100
45		DRUH_STIP_POPIIS	DRUH_STIP_POPIIS	CHAR	000600
46		DRUH_STIP_ZKRATK	DRUH_STIP_ZKRATKA	CHAR	000040

Typ	Náz.prav.	Pozí	Klíč	Infoobjekt	Ikq	Popis	Datov	Délka	Int
ZDRUH_SID	1		ZDRUH_SID	ZDRUH_SID		Druh stipendia ID	NUMC	000002	
0TXTSH	2		0TXTSH	0TXTSH		Krátký popis	CHAR	000020	
0TXMTD	3		0TXMTD	0TXMTD		Strední popis	CHAR	000040	
0TXTXL	4		0TXTXL	0TXTXL		Extra dlhý popis	SSTR	001333	

Obr. 8.7: Transformácia číselníka charakteristiky *Druh stipendia ID*

8.6 Nahrávanie dát z PSA do dátových kociek

Každá dátová kocka má transformáciu, ktorá je napojená na *DataSource* objekt a prepája tento objekt s charakteristikami dátovej kocky. Dáta z *DataSource* objektu sa nachádzajú v PSA. PSA som sa venoval v sekcii 5.2.1.

Je potrebné mať však objekt, ktorý dokáže zahájiť prenos dát z PSA, implementovať vytvorenú transformáciu a nahráť tak dáta do dátovej kocky. Na to slúži objekt s názvom *Proces prenosu dát* a má väzbu na dátovú kocku, transformáciu a DataSource objekt. Slúži na zahájenie procesu prenosu dát z PSA do dátovej kocky a poskytuje vývojárovi rôzne možnosti tohto prenosu.

Pre každú transformáciu bolo teda nutné vytvoriť objekt *Proces prenosu dát* a pomocou neho spustiť dátový prenos do dátových kociek. To isté sa týka aj atribútov a číselníkov charakteristík.

Pre všetky tieto transformácie som v mojej práci použil rovnaké nastavenia objektov *Procesu prenosu dát*. Pri prenose dát som zvolil paralelnú extrakciu s 50 000 záznamami v jednom dátovom pakete (balíček záznamov nahrávaný do dátovej kocky).

Vďaka týmto objektom sa mi podarilo nahráť dáta do dátových kociek a charakteristík a ukončiť tak úspešne tvorbu OLAP dátových kociek.

8.7 SAP BW Query pre dátové kocky

Ako bolo popísané v sekcii 5.1 v bode 10, nad dátovou kockou sa dá vytvoriť aj tzv. *SAP BW Query*, ktoré môže obsahovať len niektoré charakteristiky a ukazovatele počiatočnej kocky, avšak môže obsahovať aj dodatočne vytvorené vypočítané ukazovatele, ktoré sa využívajú ako doplnujúce dáta k analýzam. Pre systém SAP BusinessWarehouse prebieha vytváranie *Query* v softvéri **Query Designer**. Tento softvér slúži na tvorbu a administráciu *SAP BW Queries*.

Pre niektoré potreby analytikov je potrebné vytvárať nad dátovými kockami aj *SAP BW Queries*, ktoré obsahujú už spomenuté vypočítané ukazovatele. Rovnako to bolo aj pri týchto dátových kockách zahraničných ciest. Vzhľadom na fakt, že súčasťou mojej roly vývojára dátových kociek nie je tvorba *SAP BW Queries*, vytvoril ich ďalší člen nášho tímu, pán doc. Ing. Radek Kubásek Ph.D., ktorého náplňou práce je aj tvorba týchto *Queries* na základe potrieb analytikov.

Pri dátovej kocke *Zahraniční cesty Výjezd* vytvoril pán doc. Ing. Radek Kubásek Ph.D. tri vypočítané ukazovatele:

- **Počet cest**
- **Počet studentů**
- **Počet zaměstnanců**

Pri dátovej kocke *Zahraniční cesty zaměstnanci Příjezd* vytvoril iba jeden dodatočný vypočítaný ukazovateľ:

- **Počet cest**

Pri dátovej kocke *Zahraniční studenti Příjezd* vytvoril dva dodatočné vypočítané ukazovatele:

- **Počet studií**
- **Počet studentů**

Je potrebné brať túto skutočnosť do úvahy, pretože záverečné analýzy, ktoré budú aj opísané v tejto práci, budú vytvorené nad *SAP BW Queries* v takom stave, ako budú použité pri získavaní dát pre výročnú správu o činnosti školy pre MŠMT.

Kapitola 9

Testovanie

Dátové kocky som testoval z 2 aspektov:

1. či sú dáta v dátových kockách správne a zodpovedajú dátam v informačnom systéme Apollo – týmto testovaním som bol schopný odhaliť nesprávny SQL dotaz pri tvorbe databázového pohľadu alebo stratu dát počas transformácie dát do dátovej kocky. V informačnom systéme Apollo som na kontrolu dát používal moduly *Přehledy a statistiky*, *Zahraniční cesty* a *Zahraniční cesty zaměstnanců*. Avšak vzhľadom na bezpečnosť dát a GDPR nemôžem v tejto práci zahrnúť snímky z týchto modulov informačného systému Apollo, pretože sa tam nachádzajú citlivé dáta, ako napríklad osobné údaje, identifikačné čísla osôb v internom systéme a ďalšie citlivé dáta.
2. či štruktúra kocky a dáta v dátovej kocke po transformácii dokážu slúžiť ako správny zdroj dát pre vyplnenie výročnej správy o činnosti pre MŠMT.

9.1 Analysis for MS Excel

Tento softvér som používal na zobrazenie dát v dátových kockách a SAP BW Queries. Bližšie som sa popisu *Analysis for MS Excel* venoval v sekcii 4.9. Taktiež všetky výstupy, ktoré budú zobrazené v tejto časti práce, budú pochádzať z *Analysis for MS Excel*.

9.2 Testovanie kocky *Zahraniční cesty zaměstnanci Přijezd*

Základná analýza dát v programe *Analysis for MS Excel* neodhalila žiadne problémy. Všetky charakteristiky obsahovali dáta a jednotlivé charakteristiky mali medzi sebou väzby, s ktorými sa pri implementácii počítalo.

Pri overovaní správnosti dát s IS (informačným systémom) Apollo sa podarilo dosiahnuť úspešný výsledok – všetky porovnané dáta boli totožné, a teda dáta medzi kockou a IS Apollo boli rovnaké.

Nasledovalo testovanie vhodnosti dát v dátovej kocke pre potreby výročnej správy o činnosti od MŠMT. V tomto testovacom prípade som použil túto dátovú kocku na vyplnenie tabuľky 7.2 „Mobilita študentov, akademických a vedeckých pracovníkov VUT“, ktorá je podrobnejšie popísaná v sekcii 6.2. Keďže dátová kocka obsahuje dáta o prichádzajúcich akademických a technických pracovníkoch na VUT, venoval som sa práve stĺpčekom *Počet prijatých akademických pracovníků* a *Počet přijatých ostatních pracovníků*.

V dátovej kocke som do riadkov vložil charakteristiku „Stát univerzity“, do stĺpčeka som vložil ukazovateľ „Počet cest“. V poznámkach k tabulke bolo k týmto prichádzajúcim pracovníkom napísané obmedzenie dát, ktoré majú byť brané do úvahy. Ide o poznámku „****“ teda „Příjíždějící akademičtí/ostatní pracovníci (tj. počty příjezdů) – pracovníci, kteří přijeli v roce 2022; započítávají se i ti pracovníci, jejichž pobyt začal v roce 2021. Započítávají se pouze pracovníci, jejichž pobyt trval alespoň 5 dní“.

Na základe týchto obmedzení sa nastavili filtre na pozadí nad dátami v charakteristikách „Datum odjezdu“ a „Délka zahraniční cesty“.

Charakteristika „Datum odjezdu“ reprezentuje dátum, kedy zahraniční pracovníci opustili VUT, teda ukončili svoju zahraničnú cestu v roku 2022 na VUT a zároveň mohli zahájiť svoju zahraničnú cestu v roku 2021. Preto filter dát nad touto charakteristikou bol nastavený na dátumy od „1.1.2022“ do „31.12.2022“.

Charakteristika „Délka zahraniční cesty“ mala nastavený filter na číslo => 5, teda zobrazili sa len záznamy o zahraničných cestách, ktoré trvali päť a viac dní.

Výsledok analýzy je zobrazený na obrázku 9.1 nižšie:

Stát univerzity	Počet cest	
AR	Argentina	2
AT	Rakousko	7
BA	Bosna-Hercegovina	5
BE	Belgie	1
BG	Bulharsko	3
CN	Čína	3
CZ	Česká republika	3
DE	Nemecko	3
DZ	Alžírsko	1
ES	Španělsko	2
FI	Fínsko	1
FR	Francie	3
GB	Velká Británie	2
GE	Gruzie	1
GH	Ghana	1
GR	Řecko	1
HU	Maďarsko	2
IN	Indie	14
IQ	Irak	1
IR	Írán	1
IT	Itálie	10
JO	Jordánsko	2
JP	Japonsko	1
KR	Jižní Korea	2
LV	Lotyšsko	1
MA	Maroko	1
MT	Malta	1
PK	Pakistán	3
PL	Polsko	12
RO	Rumunsko	1
RS	Srbsko	1
RU	Ruská federace	1
SK	Slovensko	7
TR	Turecko	2
TW	Tchaj-wan	1
UA	Ukrajina	3
US	USA	1
VN	Vietnam	1
#	Nepřirazený	18
Celkový výsledek		126

Analysis

View for: **Zahranční cesty příjezd zaměstnanců všech**

Find:

Data Source: **Zahranční cesty příjezd zar**

Columns:

- Measures
- Datum odjezdu**
- Datum příjezdu
- Délka zahr. cesty**
- Důvod příjezdu
- Fakulta VUT ID
- Je Erasmus+
- Je Soc/E+ stáž
- Je Soc/E+ studium
- Je Stáž
- Je Studium
- Je zaměstnanec
- Je zaměstnaní na VUT
- Jméno
- Kont. osoba VUT
- Pohlaví
- Příjmení
- Stát univerzity**
- Státní příslušnost
- Typ zahr. c. Popis
- Typ zahr. c. Zkratka
- Typ zahr. c. Zkratka
- Typ zahraniční cesty
- Výslající univ.
- Zahr. cesta př. ID

Rows:

- Stát univerzity

Background Filter:

- Datum odjezdu ▼
- Cell Selection
- Délka zahr. cesty ▼
- Cell Selection

Analysis

View for: **Zahranční cesty příjezd zaměstnanců všech**

Find:

Data Source: **Zahranční cesty příjezd zar**

Columns:

- Measures
- Datum odjezdu**
- Datum příjezdu
- Délka zahr. cesty**
- Důvod příjezdu
- Fakulta VUT ID
- Je Erasmus+
- Je Soc/E+ stáž
- Je Soc/E+ studium
- Je Stáž
- Je Studium
- Je zaměstnanec
- Je zaměstnaní na VUT
- Jméno
- Kont. osoba VUT
- Pohlaví
- Příjmení
- Stát univerzity**
- Státní příslušnost
- Typ zahr. c. Popis
- Typ zahr. c. Zkratka
- Typ zahr. c. Zkratka
- Typ zahraniční cesty
- Výslající univ.
- Zahr. cesta př. ID

Rows:

- Stát univerzity

Background Filter:

- Datum odjezdu ▼
- Cell Selection
- Délka zahr. cesty ▼
- Cell Selection

Obr. 9.1: Analýza príchodu zahraničných zamestnancov

9.3 Testovanie kocky *Zahraniční cesty Výjezd*

Pri základnej analýze dát v programe *Analysis for MS Excel* som narazil na charakteristiky, ktoré neobsahovali žiadne dáta – boli prázdne. Po analýze transformácie a celého procesu vývoja v SAP BW som prišiel k záveru, že na strane SAP BW nie je problém, ale problém nastal v zdroji dát – v databázovom pohľade. V spolupráci s databázovým oddelením VUT sme narazili na chybné zvolenie databázovej domény v databázovom pohľade. Problém sme vyriešili a chýbajúce dáta sa v charakteristikách zobrazili. Nasledovala úspešná kontrola správnosti dát s IS Apollo.

Pri testovaní vhodnosti štruktúry dátovej kocky pre výročnú správu o činnosti od MŠMT som si opäť zvolil tabuľku 6.2, pretože analýzy pre túto tabuľku sú ľahké na popis a intuitívne pre čitateľov tejto práce. Prvá analýza je zameraná na celkový počet vyslaných študentov VUT do zahraničia. Pre dáta v tomto stĺpčeku sa vzťahujú obmedzenia uvedené v poznámke pod tabuľkou v označení „*“ v nasledujúcom znení: „Pozn.: * = Vyjíždějící studenti (tj. počty výjezdů) – studenti, kteří v roce 2022 absolvovali (ukončili) zahraniční pobyt; započítávají se i ti studenti, jejichž pobyt začal v roce 2021. Započítávají se pouze studenti, jejichž pobyt trval alespoň 2 týdny (14 dní)“. Na základe týchto obmedzení bolo potrebné nastaviť dodatočné filtre na pozadí, ktoré filtrujú dáta.

Analýza má nasledovnú štruktúru:

- **Stĺpce** – Obsahuje iba ukazovateľ „Počet cest“.
- **Riadky** – Obsahuje charakteristiku „Stát“, ktorá reprezentuje štát, do ktorého bol študent VUT vyslaný.
- **Filtre na pozadí** – Prvý filter na pozadí je nastavený na charakteristiku „Cesta do celkem“, ktorej dáta reprezentovali dátum, kedy študenti ukončili zahraničnú cestu. Filter bol teda nastavený na dátum od „1.1.2022“ do „31.12.2022“. Druhou charakteristikou je „Délka cesty dní celkem“, ktorej dáta reprezentovali počet dní strávených na tejto zahraničnej ceste. V našom prípade mali byť zahrnuté iba zahraničné cesty, ktoré trvali aspoň 14 dní. Podľa toho bol nastavený aj tento filter. Treťou charakteristikou je „Je zaměstnanec“, ktorého hodnotu sme nastavili na hodnotu „#“, čo reprezentuje hodnotu „false“. Dosiahli sme teda to, že sa dáta v analýze týkali len študentov VUT a nie aj zamestnancov VUT. Na záver bol nastavený filter aj na charakteristiku „Je zrušena“ na hodnotu „0“, ktorá reprezentovala hodnotu „false“. Docielil som teda stav, že v analýze sú zahrnuté len dáta o zahraničných cestách, ktoré neboli zrušené.

Analýza v programe SAP Analysis for Microsoft Excel vyzerá nasledovne:

Stát		Počet cest
AE	Sp.arab.emirát	2
AL	Albánie	1
AT	Rakousko	67
BE	Belgie	13
BG	Bulharsko	2
CA	Kanada	5
CH	Švýcarsko	10
CN	Čína	1
CY	Kypr	3
DE	Německo	46
DK	Dánsko	21
EE	Estonsko	14
ES	Španělsko	30
FI	Finsko	26
FR	Francie	22
GB	Velká Británie	18
GE	Gruzie	1
GR	Řecko	12
HR	Chorvatsko	5
HU	Maďarsko	1
ID	Indonésie	2
IE	Irsko	4
IL	Izrael	1
IS	Island	9
IT	Itálie	15
JO	Jordánsko	3
JP	Japonsko	1
KR	Jižní Korea	18
LI	Lichtenštejnsko	2
LT	Litva	8
LV	Lotyšsko	2
MK	Makedonie	1
MT	Malta	6
MX	Mexiko	1
NL	Nizozemí	13
NO	Norsko	21
NP	Nepál	1
PL	Polsko	8
PT	Portugalsko	34
RO	Rumunsko	4
RS	Srbsko	3
SE	Švédsko	10
SG	Singapur	2
SI	Slovinsko	14
SK	Slovensko	17
TH	Thajsko	1
TW	Tchaj-wan	5
US	USA	11

Analysis

View for: **Zahraniční cesty výjezd všechny charakteristiky**

Find:

Data Source

- Zahraniční cesty výjezd vše
- Measures
- Akademický rok
- Akademický rok stip.
- Celková délka pobytu
- Cesta do
- Cesta do celkem**
- Cesta od
- Cesta od celkem
- Částka
- Délka cesty Dní
- Délka cesty dní celk**
- Délka pobytu
- Druh stipendia ID
- Druh stipendia popis
- Fakulta / součást
- Je Erasmus+
- Je Soc/E+ stáž
- Je Soc/E+ studium
- Je Stáž
- Je Studium
- Je zaměstnanec**
- Je zrušena**
- Ména
- Ména/jednotka
- Počet pracovních dní
- Rok studia
- SPP prvek
- Stát**
- Stav stipendia ID
- Stipendium ID
- Student E-mail
- Student fakulta ID
- Student ID
- Student login
- Student pohlaví
- Student pozice
- Student PSC tr. byd.
- Student stát tr.
- Student stát. přísl.
- Student ústav
- Studium ID
- Typ stipendia ID
- Typ zahraniční cesty
- Účetní rok
- Univerzita ID
- Vysílající útvar
- Zahraniční cesta ID
- Zaměstnanec E-mail
- Zaměstnanec fakulta
- Zaměstnanec ID

Columns

Measures

Rows

Stát

Background Filter

- Cesta do celkem
- Cell Selection
- Délka cesty dní celk
- Cell Selection
- Je zaměstnanec
- #
- Je zrušena
- 0

Obr. 9.2: Analýza všech vyslaných studentů VUT do zahraničia

Ďalšia analýza za použitia tejto dátovej kocky bola vykonaná na získanie počtu vyslaných študentov do zahraničia, ktorý absolvovali absolventskú stáž. Tieto dáta sú opäť potrebné pre tabuľku mobility študentov a pracovníkov VUT zobrazenej na obrázku 6.2.

K štruktúre analýzy vykonanej na obrázku 9.2 bol pridaný iba ďalší filter na pozadí nad dátami charakteristiky „Typ zahraniční cesty“. Tu boli vyfiltrované len zahraničné cesty, ktorých typ bol „ERASMUS PLUS absolvent“. Ostatné filtre, ukazovatele a charakteristiky zostali totožné s analýzou vykonanou na obrázku 9.2.

Výsledok analýzy dopadol nasledovne:

Stát	Počet cest	
AT	Rakousko	6
DE	Nemecko	2
EE	Estonsko	1
ES	Španielsko	2
FR	Francie	1
GB	Veľká Británie	1
IS	Island	2
IT	Itálie	1
MT	Malta	1
NO	Norsko	1
PT	Portugalsko	1
RS	Srbsko	1
SE	Švédsko	2
SK	Slovensko	1

The screenshot also shows the SAP BW Query Analysis interface with the following configuration:

- View for:** Zahraniční cesty výjezd všechny charakteristiky [DS_27]
- Data Source:** Zahraniční cesty výjezd vše
- Columns:** Measures, Rows, Stát
- Background Filter:** Cesta do celkem, Délka cesty dní celk, Je zaměstnanec, Je zrušena, Typ zahraniční cesty (ERASMUS PLUS absolvent)

Obr. 9.3: Analýza absolventských stáží študentov VUT vyslaných do zahraničia

9.4 Testovanie kocky *Zahraniční studenti Přijezd*

Pri analýze získaných dát z centrálnej databázy VUT v porovnaní s dátami v informačnom systéme Apollo som zistil, že databázový pohľad neobsahuje všetky dáta, ktoré by podľa IS Apolla mal, databázový pohľad obsahoval len zlomok dát. Tento problém sme začali riešiť s databázovým oddelením VUT a prišli sme k záveru, že problém nastal v zlej *join condition* pri spojovaní tabuliek, z ktorých sa získavali dáta do databázového pohľadu. Problém sa nám podarilo vyriešiť a dáta boli po opätovnej kontrole v poriadku.

Pri tvorbe analýzy sme použili vytvorené SAP BW Query nad touto dátovou kockou. Cieľom tejto analýzy je získanie dát o celkovom počte prijatých zahraničných študentoch na VUT.

Analýza má nasledujúcu štruktúru:

- **Stĺpce** – Obsahuje iba vypočítaný ukazovateľ „Počet štúdií“.
- **Riadky** – Obsahuje charakteristiku „Stát“, ktorá reprezentuje štát, z ktorého bol študent vyslaný.
- **Filtre na pozadí** – Filtrovanie prebiehalo len nad jednou charakteristikou, a tou je „Studium ID“. Do filtra boli vložené identifikačné čísla štúdií, ktoré prebiehali viac

ako trinásť dní a boli ukončené v roku 2022. Tieto identifikačné čísla boli získané pomocou dátovej kocky *Studia zmeny*.

Analýza v programe SAP Analysis for Microsoft Excel vyzerá nasledovne:

Stát		počet studij
AL	Albánie	1
BE	Belgie	5
BA	Bosna-Herceg.	4
BR	Brazílie	10
BG	Bulharsko	2
ME	Černá Hora	3
CZ	Česká republika	2
DK	Dánsko	3
EE	Estonsko	8
FI	Fínsko	9
FR	Francie	156
HR	Chorvatsko	7
IN	Indie	2
IR	Írán	1
IE	Irsko	2
IT	Itálie	29
IL	Izrael	2
JP	Japonsko	1
KR	Jižní Korea	5
LT	Litva	22
LV	Lotyšsko	2
LU	Lucembursko	1
MK	Makedonie	2
MT	Malta	4
MX	Mexiko	4
DE	Nemecko	13
NL	Nizozemí	6
PK	Pakistán	1
PL	Polsko	11
PT	Portugalsko	51
AT	Rakousko	10
RO	Rumunsko	11
GR	Řecko	18
SK	Slovensko	18
SI	Slovinsko	7
RS	Srbsko	3
ES	Španělsko	111
CH	Švýcarsko	1
TW	Tchaj-wan	7
TR	Turecko	62
UA	Úkrajina	14
US	USA	5
GB	Velká Británie	2
Celkový výsledek		638

Analysis

View for: zahr cety stud přijezd test1

Find:

Data Source

- zahr cety stud přijezd tes
- Measures
- Datum příjezdu
- Fakulta / součást
- Kód krajiny
- Název vysílající uni
- Rok studia
- Stát**
- Student E-mail
- Student fakulta ID
- Student ID
- Student login
- Student pohlaví
- Student pozice
- Student PSČ tr. byd.
- Student stát tr.
- Student stát. přísl.
- Student ústav
- Studium ID**
- Typ zahr. c. Popis
- Typ zahr. c. Popis E
- Typ zahr. c. Zkratka
- Typ zahr. c. Zkratka
- Typ zahraniční cesty
- Vysílající univ.
- Z. student E-mail
- Z. student fakulta
- Z. student login
- Z. student pohlaví
- Z. student pozice
- Z. student PSČ tr.
- Z. student stát tr.
- Z. student stát. př.
- Z. student ústav
- Začátek studia
- Zahr. cesta proj. ID
- Zahr. student ID

Columns

Measures

Rows

Stát

Background Filter

Studium ID

Obr. 9.4: Analýza prichádzajúcich študentov zo zahraničia

Kapitola 10

Záver

Cieľom práce bolo naprogramovať a vytvoriť tri dátové kocky s dátami zahraničných ciest (Zahranční studenti Příjezd, Zahranční cesty Výjezd, Zahranční cesty zaměstnanci Příjezd), ktoré budú slúžiť primárne na získavanie dát pre vyplnenie výročnej správy o činnosti pre Ministerstvo školstva, mládeže a telovýchovy Českej republiky. V druhom rade budú dátové kocky slúžiť analytikom VUT pre získavanie dát zahraničných ciest študentov a pracovníkov VUT. Dátové kocky som implementoval v softvéri SAP Business Warehouse. S pomocou databázového oddelenia CVIS VUT som získal dáta z centrálnej databázy VUT a pomocou ďalších osôb na VUT som vytvoril optimálny výsledok a poskytol tieto dátové kocky analytikom VUT.

Cieľ práce sa mi podarilo naplniť a úspešne vytvoriť tieto dátové kocky. Počas procesu tvorby dátových kociek vznikli určité problémy, ktoré sa mi podarilo vyriešiť a v závere nemali vplyv na celkovú úspešnosť práce.

Po úspešnom preštudovaní teórie dátových kociek a s podrobným zoznámením sa so softvéri SAP Business Warehouse a SAP Analysis for Microsoft Excel som navrhol databázové pohľady pre jednotlivé kocky. Nasledoval návrh samotných dátových kociek, ktoré som na základe toku vývoja dátových kociek v softvéri SAP Business Warehouse implementoval a otestoval. Testovanie dátových kociek prebiehalo v dvoch fázach. Môžem zhodnotiť, že všetky dátové kocky prešli úspešne každou fázou testovania a boli použité na získanie dát zahraničných ciest pre vyplnenie výročnej správy o činnosti vysoké školy pre MŠMT.

Vypracovanie tejto práce mi dalo mnoho nových vedomostí ohľadne problematiky zahraničných ciest, teórie a vývoja dátových kociek, technológie OLAP a modelovania v SAP Business Warehouse.

Problematika zahraničných ciest je len podmnožinou problematiky pracovných ciest na VUT. Preto verím, že v rámci CVIS VUT budem stáť v budúcnosti za vývojom nových dátových kociek, ktoré sa budú venovať aj celej problematike pracovných ciest na VUT. Táto problematika je veľmi dôležitá z pohľadu manažmentu VUT ako inštitúcie. Preto verím v ďalší vývoj a nutnosť spracovania tejto problematiky pomocou OLAP dátových kociek.

Literatúra

- [1] BURGET, R. *Informační systémy – Pojem informačního systému, data, informace* [Fakultná přednáška]. Brno: FIT VUT Brno, 2022 [cit. 2022-12-19]. Dostupné z: <https://www.fit.vut.cz/study/course/IIS/.cs>.
- [2] GALAKTIKASOFT. *OLAP Operations in Data Mining* [online]. Galaktikasoft, január 2018 [cit. 2022-12-25]. [vid. 2022-12-25]. Dostupné z: <https://galaktika-soft.com/blog/olap-operations-in-data-mining.html>.
- [3] HRUŠKA, T. *Manažérský informační systém – Analytický přístup k datům* [Prednáška]. Brno: FIT VUT Brno, 2020 [cit. 2022-12-19]. Dostupné z: <https://www.fit.vut.cz/study/course/PIS/.cs>.
- [4] IBM. *OLTP* [online]. IBM, 2022 [cit. 2022-12-19]. [vid. 2022-12-19]. Dostupné z: <https://www.ibm.com/cloud/learn/oltp>.
- [5] JEWETT, E. *SAP BW (Business Warehouse)* [online]. TechTarget, september 2018 [cit. 2022-12-25]. [vid. 2022-12-25]. Dostupné z: <https://www.techtarget.com/searchsap/definition/Business-Information-Warehouse>.
- [6] LABERGE, R. *Datové sklady : agilní metody a business intelligence*. 1. vyd. Brno: Computer Press, 2012 [cit. 2022-12-19]. ISBN 978-80-251-3729-1.
- [7] LUTKEVICH, B. *OLTP (online transaction processing)* [online]. TechTarget, 2021 [cit. 2022-12-19]. [vid. 2022-12-19]. Dostupné z: <https://www.techtarget.com/searchdatacenter/definition/OLTP>.
- [8] TAYLOR, D. *What is OLTP? Definition, Architecture, Example* [online]. Guru99, október 2022 [cit. 2022-12-19]. [vid. 2022-12-19]. Dostupné z: <https://www.guru99.com/what-is-oltp.html>.
- [9] VAISMAN, A. a ZIMÁNYI, E. *Data Warehouse Systems: Design and Implementation*. 2014. vyd. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2014. Data-Centric Systems and Applications. ISBN 9783642546549.
- [10] WREMBEL, R. *Data warehouses and OLAP : concepts, architectures, and solutions*. 1. Hershey: IRM Press, 2007. ISBN 1-59904364-5.