

# Reporting nad podnikovými daty jako nástroj pro manažerské rozhodování

## Diplomová práce

*Studijní program:*

N6208 Ekonomika a management

*Studijní obor:*

Podniková ekonomika

*Autor práce:*

**Bc. Tadeáš Grošup**

*Vedoucí práce:*

Ing. Jan Öhm, Ph.D.

Katedra ekonomické statistiky





## Zadání diplomové práce

# Reporting nad podnikovými daty jako nástroj pro manažerské rozhodování

*Jméno a příjmení:* **Bc. Tadeáš Grošup**  
*Osobní číslo:* E18000245  
*Studijní program:* N6208 Ekonomika a management  
*Studijní obor:* Podniková ekonomika  
*Zadávací katedra:* Katedra ekonomické statistiky  
*Akademický rok:* **2019/2020**

### Zásady pro vypracování:

1. Stanovení cílů
2. Teoretická východiska práce
3. Zpracování vybraného manažerského informačního systému
4. Možnosti aplikace statistických metod ve vybraném manažerském informačním systému
5. Vyhodnocení získaných výstupů

Rozsah grafických prací:  
Rozsah pracovní zprávy:  
Forma zpracování práce:  
Jazyk práce:

65 normostran  
tištěná/elektronická  
Čeština



### Seznam odborné literatury:

- HENDL, Jan. 2012. *Přehled statistických metod: analýza a metaanalýza dat*. 4. vydání. Praha: Portál. ISBN 978-80-262-0200-4.
- KIMBALL, Ralph a Margy ROSS. 2013. *The data warehouse toolkit: The definitive guide to dimensional modeling*. 3rd ed. Indianapolis: John Wiley & Sons. ISBN: 978-1-118-53080-1.
- SYNEK, Miloslav. 2011. *Manažerská ekonomika*. 5. vydání. Praha: Grada. ISBN 978-80-247-3494-1.
- WEGNER, Trevor. 2013. *Applied Business Statistics: Methods and Excel-based Applications*. 3rd ed. South Africa: Juta and Company. ISBN 978-0-7021-7774-3.
- PROQUEST. 2019. *Databáze článků ProQuest* [online]. Ann Arbor, MI, USA: ProQuest. [cit. 2019-10-12]. Dostupné z <http://knihovna.tul.cz/>

Konzultant: Ing. Zdeněk Fikar

Vedoucí práce:

Ing. Jan Öhm, Ph.D.  
Katedra ekonomické statistiky

Datum zadání práce:

31. října 2019

Předpokládaný termín odevzdání:

31. srpna 2021

prof. Ing. Miroslav Žižka, Ph.D.  
děkan

L.S.

Ing. Jan Öhm, Ph.D.  
vedoucí katedry

V Liberci dne 31. října 2019

## Prohlášení

Prohlašuji, že svou diplomovou práci jsem vypracoval samostatně jako původní dílo s použitím uvedené literatury a na základě konzultací s vedoucím mé diplomové práce a konzultantem.

Jsem si vědom toho, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci nezasahuje do mých autorských práv užitím mé diplomové práce pro vnitřní potřebu Technické univerzity v Liberci.

Užiji-li diplomovou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti Technickou univerzitu v Liberci; v tomto případě má Technická univerzita v Liberci právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Současně čestně prohlašuji, že text elektronické podoby práce vložený do IS/STAG se shoduje s textem tištěné podoby práce.

Beru na vědomí, že má diplomová práce bude zveřejněna Technickou univerzitou v Liberci v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů.

Jsem si vědom následků, které podle zákona o vysokých školách mohou vyplývat z porušení tohoto prohlášení.

20. července 2020

Bc. Tadeáš Grošup

## **Anotace**

Hlavním cílem diplomové práce je rozšíření manažerského informačního systému společnosti ABC o Manažerskou tabulku – nástroj sledování a hodnocení plnění plánu tržeb v konsolidaci. Dále provedení statistické analýzy vybraných charakteristik a ukazatelů v manažerském informačním systému ABC s účelem zjištění jejich vlivu na úspěšnost projektů společnosti. Úvodní část práce je věnována teoretickým východiskům k problematice Business Intelligence. Je v ní vysvětlen vývoj, charakteristika a architektura Business Intelligence. Druhá kapitola pojednává o specifikách systému SAP BW. Následně je zpracován manažerský informační systém ABC. Nejprve jsou popsány skupiny reportů, které v něm byly původně zahrnuty. Poté je představeno původní zpracování Manažerské tabulky a její zavedení do datového toku v systému BW. V poslední kapitole jsou vysvětleny vybrané statistické metody, které jsou následně aplikovány na data projektů a příležitostí v manažerském informačním systému ABC. Je vytvořeno několik modelů logistické regrese, které predikují procentuální úspěšnost projektů. Jednotlivé modely jsou mezi sebou porovnány a je navržen postup aplikace výsledků do manažerského informačního systému ABC.

## **Klíčová slova**

Business Intelligence, reporting, SAP BW, statistická analýza, logistická regrese.

# **Annotation**

*Reporting on company data as a tool for managerial decision making*

The main goal of the diploma thesis is adding a Managerial table – a tool for monitoring and evaluating the fulfillment of the sales plan in consolidation – to management information system of company ABC. Furthermore, a statistical analysis of selected characteristics and indicators in management information system ABC in order to determine their impact on the success of the company's projects. The introductory part of the thesis is devoted to the theoretical basis for Business Intelligence. It explains the development, characteristics and architecture of Business Intelligence. The second chapter deals with the specifics of the SAP BW system. Subsequently, the management information system ABC is processed. First, the groups of reports that were originally included are described. Then the original processing of the Managerial Table and its introduction into the data flow in the BW system is presented. The last chapter explains selected statistical methods, which are then applied to project and opportunity data in management information system ABC. Several models of logistic regression are created, which predict the percentage success of the projects. The individual models are compared with each other and the procedure for applying the results to management information system ABC is proposed.

## **Keywords**

Business Intelligence, reporting, SAP BW, statistical analysis, logistic regression.

## **Poděkování**

Velice děkuji Ing. Janu Öhmovi, Ph.D. za odborné vedení práce, jeho rady a podněty. Dále děkuji vedení společnosti Pregis a.s. za umožnění vypracování diplomové práce s použitím jejich interních dat. Děkuji Ing. Zdeňkovi Fikarovi za jeho vstřícnost a ochotu při konzultaci práce.





# Obsah

Seznam ilustrací.....	13
Seznam tabulek.....	14
Seznam použitých zkratk a značek .....	15
Úvod .....	16
1. Business Inteligence .....	17
1.1 Data.....	17
1.2 Vývoj BI .....	18
1.3 Charakteristika BI.....	19
1.4 Architektura BI .....	20
1.4.1 Zdrojové systémy .....	21
1.4.2 Datová transformace.....	22
1.4.3 Datový sklad.....	24
1.4.4 Datové tržiště.....	25
1.4.5 Dočasné a operativní uložení.....	25
1.4.6 Multidimenzionalita .....	26
1.4.7 OLAP.....	28
1.4.8 Dolování dat .....	29
1.4.9 Reporting .....	30
1.4.10 Metadata .....	31
1.4.11 Kmenová data.....	31
2. SAP BW .....	33
2.1.1 Datový tok .....	33
2.1.2 Enterprise Data Warehousing.....	37
3. Zpracování vybraného MIS .....	39
3.1 Výchozí situace.....	39

3.1.1	Skupina reportů Pipeline .....	41
3.1.2	Skupina reportů Mastertable .....	42
3.1.3	Skupina reportů Forecasting.....	43
3.2	Manažerská tabulka .....	44
3.3	Zavedení Manažerské tabulky do BW.....	47
3.3.1	Algoritmus plnění charakteristik.....	48
3.3.2	Datový tok pro Manažerskou tabulku v BW.....	51
3.3.3	Řetězec procesů.....	56
3.3.4	Vytvoření queries .....	57
3.3.5	Vytvoření koncového reportu.....	60
3.3.6	Přístup uživatelů do reportu Manažerská tabulka .....	63
3.3.7	Produktivní využívání reportu.....	63
4.	Aplikace statistických metod v MIS ABC .....	65
4.1	Využití statistické metody.....	65
4.2	Predikce úspěšnosti projektů .....	68
4.2.1	Rating příležitostí .....	69
4.2.2	Overall project rating .....	71
4.2.3	Faktory Overall project ratingu .....	73
4.2.4	Porovnání modelů .....	76
4.2.5	Modely s přidanými proměnnými.....	78
4.3	Aplikace výsledků do datového toku.....	79
	Závěr.....	81
	Seznam použité literatury.....	83
	Seznam příloh.....	84

## Seznam ilustrací

Obrázek 1: Data, informace, znalosti a moudrost .....	17
Obrázek 2: Obecná koncepce architektury BI .....	21
Obrázek 3: Použití EAI platformy .....	24
Obrázek 4: Star schema .....	27
Obrázek 5: OLAP kostka.....	28
Obrázek 6: Diagram datového toku v BW .....	33
Obrázek 7: Extended star schema.....	35
Obrázek 8: Samostatné datové tržiště.....	37
Obrázek 9: Společný datový sklad .....	38
Obrázek 10: Vstupní obrazovka MIS ABC .....	39
Obrázek 11: Diagram datového toku.....	40
Obrázek 12: Report Pipeline.....	42
Obrázek 13: Report Mastertable.....	43
Obrázek 14: Report Forecasting .....	44
Obrázek 15: Kontingenční tabulky pro MT .....	46
Obrázek 16: Algoritmus zpracování obchodní jednotky.....	49
Obrázek 17: Algoritmus zpracování názvu skupiny materiálu .....	51
Obrázek 18: Datový tok pro Manažerskou tabulku v BW .....	52
Obrázek 19: DSO Nákladové ceny.....	53
Obrázek 20: DSO Partnerské role .....	53
Obrázek 21: Kmenová data charakteristiky specifying unit.....	56
Obrázek 22: Vymezení charakteristik pro Manažerskou tabulku .....	58
Obrázek 23: Query_1 pro Manažerskou tabulku.....	59
Obrázek 24: Query_5 pro Manažerskou tabulku.....	60
Obrázek 25: Manažerská tabulka, úvodní obrazovka.....	61
Obrázek 26: Manažerská tabulka, základní pohled .....	62
Obrázek 27: Manažerská tabulka, menu .....	62
Obrázek 28: Nová vstupní obrazovka MIS ABC .....	63
Obrázek 29: Úspěšnost příležitostí dle ratingu.....	70
Obrázek 30: Krabicový graf (Overall proj. rating a úspěšnost) .....	72

## Seznam tabulek

Tabulka 1: Druhy cen faktur .....	55
Tabulka 2: Kontingenční tabulka .....	65
Tabulka 3: Kontingenční tabulka (úspěšnost a rating příležitostí).....	70
Tabulka 4: Overall project rating (charakteristika) .....	71
Tabulka 5: Inverzní předpověď pro Overall project rating .....	72
Tabulka 6: Testy poměrem věrohodností pro faktory .....	75
Tabulka 7: Regresní model (5 faktorů) .....	76
Tabulka 8: Porovnání úspěšnosti odhadů modely A a B .....	77
Tabulka 9: Klasifikační tabulka (predikce modely A a B, mezní hodnota 0,5).....	78
Tabulka 10: Klasifikační tabulka (predikce modelem C, mezní hodnota 0,5).....	79

## Seznam použitých zkratek a značek

$\chi^2$	chí-kvadrát rozdělení
$\alpha, \beta$	regresní koeficienty v modelu logistické regrese
<i>BI</i>	Business Intelligence
<i>BW</i>	Business Warehouse
<i>CRM</i>	Customer Relationship Management
<i>DMA</i>	Data Mart
<i>DSO</i>	Data storage object (objekt datové schránky)
<i>DWH</i>	Data Warehouse
<i>ERP</i>	Enterprise Resource Planning
$H_0$	nulová hypotéza
$H_1$	alternativní hypotéza
<i>MIS</i>	Manažerský informační systém
$n$	rozsah výběrového souboru
$n_i$	absolutní četnost
$n_{i.}, n_{.j}$	marginální četnost
$n'_{ij}$	očekávaná četnost
$p_i$	relativní četnost
<i>SQL</i>	Structured Query Language

# Úvod

Obsah této diplomové práce je zaměřen v první řadě na zpracování manažerského informačního systému společnosti ABC. Ten původně obsahoval tři skupiny reportů. Zákazník si přál do MIS přidat nástroj sledování a hodnocení plnění plánu tržeb v konsolidaci. Jedná se o tzv. Manažerskou tabulku. Hlavním cílem této práce je automatizace postupu plnění Manažerské tabulky a její realizace v rámci systému BW. Nejprve je podrobně popsán postup, dle kterého oddělení controllingu společnosti ABC tuto tabulku manuálně každý měsíc sestavovalo. Před samotným zpracováním firemních údajů je představen vývoj a architektura Business Intelligence se zaměřením na specifika SAP BW.

Práce se věnuje algoritmům plnění vybraných charakteristik a ukazatelů, znázornění datového toku pro Manažerskou tabulku v systému BW, vytvoření jednotlivých queries, zpracování podoby koncového reportu a jeho začlenění do MIS ABC. Cílem je ušetřit časových kapacit odborných zaměstnanců společnosti ABC automatizací procesu sestavování Manažerské tabulky. Zároveň uživatelům MIS poskytnout nejaktuálnější data, které půjdou jednoduše filtrovat dle vybraných charakteristik. Tím vytvořit vynikající podklad manažerům pro dodatečné analýzy.

Dalším cílem je vytvoření nástroje, který bude predikovat procentuální šanci na úspěch projektů a příležitostí společnosti ABC. A to prostřednictvím statistické analýzy dat obsažených v manažerském informačním systému ABC. Konkrétně zkoumáním závislosti vybraných charakteristik a ukazatelů na úspěšnost těchto projektů. Využitým statistickým metodám je věnována vlastní kapitola. Nejprve je analyzována charakteristika Rating příležitosti a ukazatel Overall project rating. Následně jsou podrobně rozebrány pomocí chí-kvadrát testů nezávislosti a testů lineárního trendu v kontingenčních tabulkách jednotlivé faktory tohoto ukazatele. Na základě těchto faktorů a dalších charakteristik je vytvořeno několik modelů logistické regrese. V práci jsou jednotlivé modely porovnány. Cílem je nalézt model s co nejvhodnější kombinací charakteristik a ukazatelů, který bude nejpřesněji identifikovat úspěšné příležitosti a projekty. Následně představit, jak výsledky této analýzy aplikovat do datového toku v BW.

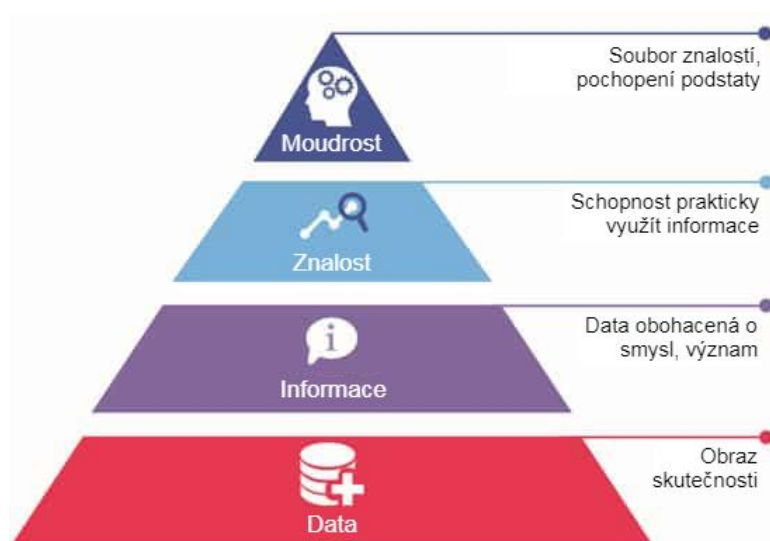
# 1. Business Intelligence

V první části diplomové práce je vysvětlena problematika dat, popsán vývoj Business Intelligence, jeho charakteristika a architektura. Tato kapitola slouží jako teoretický základ, který bude využitý pro zpracování praktické části práce.

## 1.1 Data

V dnešní době je sbíráno obrovské množství dat z různých zdrojů. Mluví se o fenoménu *Big Data*. Tyto velké objemy se nedají jednoduše metodami interpretovat pro koncového uživatele a při jejich zpracování je zapotřebí využít specializovaných nástrojů. Právě systémy Business Intelligence se zabývají získáváním agregovaných netriviálních informací z velkého objemu zdrojových dat. Pro lepší porozumění je tedy vhodné nejprve objasnit základní pojmy týkající se problematiky dat (údajů).

*Data, informace, znalosti, moudrost* a jejich vzájemnou souvislost zachycuje obrázek č. 1. *Data* jsou získávána různými způsoby, jako je pozorování, měření, zápis. Zachycují jistý obraz skutečnosti, který lze uložit v různých podobách. Údaje mohou mít vysokou hodnotu a na základě toho mohou být využita pro jejich obchodování. Tato skutečná hodnota se však neprojeví do té doby, než dojde k jejich zpracování. *Data* sama o sobě nedisponují sémantickou rovinou (samotná nemají význam). Tento význam je jim třeba přiřadit. Ten je přiřazován buď při jejich zaznamenávání, nebo později, za využití znalostí (Sklenák, 2001).



Obrázek 1 – *Data, informace, znalosti a moudrost*  
Zdroj: vlastní zpracování.

Rozlišujeme dva základní druhy dat, z pohledu jejich zpracování. Prvním jsou *strukturovaná data*. Ta zachycují explicitně objekty, fakta a atributy. Uložena bývají v relačním databázovém systému, který využívá hierarchii jejich elementů (pole, záznam, relace, databáze). S takto organizovanými údaji se dobře pracuje a získávají se relevantní data. Oproti tomu *nestrukturovaná data* jsou tokem bitů, který neobsahuje další specifikaci. Příkladem může být zvuková nahrávka, textový dokument nebo obrázek.

Přiřazením určitého významu nebo smyslu datům, se z nich stávají informace. Stávají se využitelná a srozumitelná a získávají hodnotu. Ta přímo nesouvisí s jejich cenou, protože se jedná o hodnotu subjektivní. Kvalita zdrojových dat, proces přeměny a potřeby koncového uživatele jsou hlavními faktory, které determinují hodnotu informací. V případě, že by uživatel informace nedokázal interpretovat, tak pro něj ztrácí cenu. Schopnost interpretace a využití informací je klíčová právě pro koncové uživatele systémů Business Intelligence. Ti musí být schopni nejen zobrazené výsledky v systému interpretovat, ale také disponovat dostatečnou informační a počítačovou gramotností.

O znalosti se jedná v tom případě, když je informace možné prakticky využít. Hrají důležitou roli v průběhu zpracování dat a interpretací informací. Shrnují poznatky získané učením a zkušenostmi. Takto získané znalosti se vždy vztahují ke konkrétním účelům, případně ke konkrétní oblasti problematiky. Oproti informacím, které jsou časově pomíjivé, jsou znalosti časově invariantní (trvalé). Znalosti lze rozdělit na *explicitní* (kodifikované, možné uchovat a předávat dál – fakta, teorie) a *tacitní* (založené na osobních zkušenostech, těžce přenositelné – hodnoty, způsoby chování). Na vrcholu pyramidy stojí moudrost, což je soubor znalostí, který vychází z pochopení podstaty problematiky v daných souvislostech. Je nejvyšší úrovní vědění a poznání člověka, které vychází z celoživotního získávání zkušeností a učení (Sklenák, 2001).

## 1.2 Vývoj BI

Pojmenování Business Intelligence je možné poprvé dohledat v článku z roku 1958 od pracovníka IBM, Hanse Petera Luhna (1958, s. 314). Luhn toto slovní spojení popsal jako „*Schopnost automatizovaně vnímat a pochopit vzájemné vztahy prezentovaných skutečností z průmyslových, vědeckých, nebo vládních organizací takovým způsobem, aby se činnost ubírala směrem k požadovanému cíli.*“ V té době se jednalo však pouze



o teoretické základy, jelikož zatím neexistovaly odpovídající technologie, které by je dokázaly přenést do praxe.

Ke konci sedmdesátých let 20. století se začala objevovat první řešení, které sloužila k podpoře analytických a manažerských úloh v podnikovém řízení. Právě v této době totiž docházelo k rozvoji on-line zpracování dat. S těmito prvotními pokusy je spojována americká firma Lockheed. Komerční produkty, založené na multidimenzionálním uložení a zpracování dat, se na americký trh dostaly poprvé v druhé polovině osmdesátých let zásluhou firem Comshare a Pilot. Tyto produkty byly označovány jako EIS (Executive Information System) a DSS (Decision Support System) (Novotný, 2005).

Se zlepšujícím se výpočetním výkonem a datovou základnou společností v devadesátých letech začalo docházet k hojnému nasazování systémů Business Intelligence. Současně se začaly objevovat trendy datových skladů (Data Warehouse) a datových tržišť (Data Marts) a s tím spojené nástroje dolování dat (Data Mining), které se pomocí statistických a matematických metod snažily o hlubší analýzy dat. O rozvoj těchto technologií se v té době zasloužili obzvláště Bill Inmon a Ralph Kimball. V 21. století se BI v oblasti podnikové informatiky stalo jedno z nejrychleji rozvíjejících se součástí. Firmy zjišťují, že konkurenční výhody a vyšší konkurenceschopnosti, mohou lépe dosáhnout díky konsolidaci firemních dat a jejich následným analyzováním (Novotný, 2005).

### **1.3 Charakteristika BI**

Jak se oblast BI vyvíjela, tak se upravovala i jeho definice z původního znění z roku 1958 od Hanse Luhna. Analytik z firmy Gartner Group Howard J. Dresner v roce 1989 BI popsal jako „*sadu konceptů a metod určených pro zkvalitnění rozhodnutí firmy*“ (Novotný, 2005, s. 18). Dále zdůraznil důležitost datové analýzy, dotazovacích nástrojů a reportingu, které umožňují uživatelům získat užitečné informace z velkého množství dostupných dat.

Se stoupající popularitou BI v 21. století vznikaly i další definice a neexistuje dnes tudíž žádná jednotná. Jako jednu ze současných lze jmenovat například definici Carla Vercelliho z jeho knihy Business Intelligence (2009, s. 3): „*Business Intelligence lze definovat jako soubor matematických modelů a analytických metodologií, které využívají dostupná data k získávání informací a užitečných znalostí pro komplexní proces rozhodování*“.

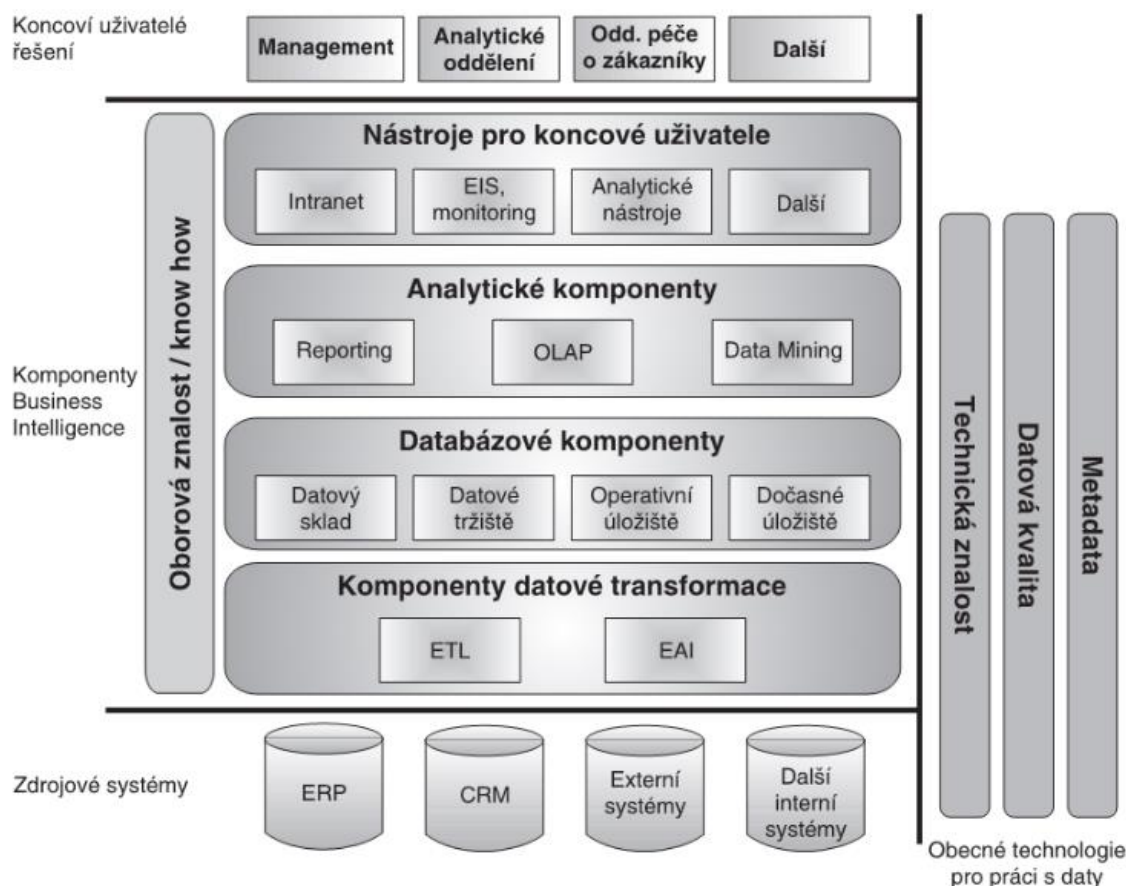
Za zmínění také stojí definice České společnosti pro systémovou integraci, která se zabývá výměnou názorů a informací v oblasti IS. Charakterizuje BI jako „*sadu procesů, aplikací a technologií, jejichž cílem je účinně a účelně podporovat rozhodovací procesy ve firmě. Podporují analytické a plánovací činnosti podniků a organizací a jsou postaveny na principech multidimenzionálních pohledů na podniková data*“ (Novotný, 2005, s. 19). Aplikace Business Intelligence podporují plánovací a analytické funkce většiny oddělení podnikového řízení (prodej, marketing, nákup, controlling, HR, výroby, finance apod.).

Na základě předchozích definice lze formulovat základní požadavky na systémy BI. Zaprvé by měly zajistit jednoduchou přístupnost uživatelů k informacím. Tyto informace musí být zobrazovány konzistentně, včas a být důvěryhodné. Systémy BI se musí adaptovat na změny (obchodní podmínky, data, technologie) a udržovat informace v utajení a chránit je před neoprávněnými osobami. Aby zapojení systému Business Intelligence v podniku fungovalo, je podstatné, aby ho uživatelé považovaly za autoritativní a důvěryhodný základ pro zlepšené rozhodování (Kimball, 2013).

## 1.4 Architektura BI

Základní uspořádání komponent v systému BI znázorňuje obrázek č. 2. Je to pouze jeden z možných návrhů. Uspořádání jednotlivých BI systémů se mohou vzájemně lišit na základě různých faktorů (požadavky zákazníků, podniku, okolního prostředí společnosti apod.). Tato řešení se mezi sebou odlišují svou komplexností, pracností, nákladností, zacílením a obtížností technologického zpracování.

Ze schématu lze vyčíst 5 základních vrstev. Na nejnižší úrovni sedí komponenty datové transformace. Tato vrstva slouží pro extrakci, transformaci, očištění a nahrávání dat z firemních zdrojových systémů. Zahrnuje **ETL** (Extract, Transform, Load) a **EAI** (Enterprise Application Integration) systémy. Data jsou poslána do následující vrstvy pro ukládání dat (databázové komponenty). Ta zahrnuje procesy, při kterých se data ukládají, aktualizují a spravují. Sem spadají komponenty **Data Warehouse** (datový sklad), **Data Mart** (datové tržiště), **ODS** (Operational Data Store – operativní datové uložení) a **DSA** (Data Staging Areas – dočasná uložení dat).



Obrázek 2 – Obecná koncepce architektury BI

Zdroj: NOVOTNÝ, Ota 2005, Business Intelligence: Jak využít bohatství ve vašich datech, s. 27.

Ve třetí vrstvě se nacházejí analytické komponenty, které zajišťují zpřístupňování těchto dat a jejich analýzu. Spadá sem **Reporting**, systémy **OLAP** (On-Line Analytical Processing) a **Data Mining** (dolování dat). Tato analyzovaná data musí být dostupná koncovým uživatelům. Právě k tomu slouží další vrstva, která zahrnuje portálové aplikace (webové technologie), systémy Executive Information Systems (**EIS**) a jiné analytické aplikace. Dohromady jsou tyto 4 vrstvy zastřešovány oborovou znalostí, neboli „best practices“ využívaných při nasazování řešení Business Intelligence pro jednotlivé situace v organizaci. V dalších podkapitolách budou přiblíženy jmenované součásti jednotlivých vrstev.

### 1.4.1 Zdrojové systémy

Existují dva základní typy informací, se kterými mohou informační systémy pracovat. Jsou jimi operativní a analytické informace. Operativní zobrazují současný stav podniku, během dne se mění a bývají uloženy v relačních databázích (jedná se například o účetnictví). Zpracování operativních informací v reálném čase provádějí tzv. **OLTP** (On Line

Transaction Processing) systémy. Data v těchto transakčních systémech jsou označována jako zdrojová, primární nebo produkční. Z nich čerpají aplikace BI data a samy o sobě tedy nepatří do skupiny Business Intelligence aplikací (Novotný, 2005).

Tyto zdrojové systémy jsou specifické tím, že jejich architektura umožňuje práci s daty (modifikace, ukládání) v reálném čase. Oproti tomu nejsou stavěny pro provádění analytických úloh. Produkční databáze jsou realizovány prostřednictvím databázových systémů, jako jsou MS SQL Server, ORACLE apod. Mezi systémy OLTP řadíme například systémy:

- **CRM** – Customer Relationship Management (aplikace pro správu vztahu se zákazníky),
- **ERP** – Enterprise Resource Planning (aplikace pro plánování podnikových zdrojů),
- **SCM** – Supply Chain Management (aplikace pro řízení dodavatelských vztahů).

Zdrojem dat ale nemusí být pouze tyto velké databáze. Může jím být i běžný tabulkový kalkulátor (MS Excel), soubor aplikace databázového typu (MS Access) nebo textové soubory s danou strukturou (tzv. flat files). Čerpat data je možné nejen z podnikových zdrojů, ale také z veřejných databází (statistické úřady, vzorníky, vládní instituce). Pro jakýkoliv zdroj platí, že na jeho kvalitě závisí následné výsledky v systému BI.

### 1.4.2 Datová transformace

Na tuto část řešení BI je kladen veliký důraz. Zajišťuje analýzu zdrojů pro potřeby konkrétního projektu a vybírá relevantní data, které následně posílá pro integraci do BI systému. Na kvalitě zpracovaných dat prostřednictvím datové transformace samozřejmě závisí kvalita všech analýz, reportů a dalších výstupů. V případě, že by byla poslána nesprávná nebo špatně interpretovaná data, bude výrazně ovlivněna rozhodovací schopnost uživatelů výstupů. Proto bývá tato část z pohledu Business Intelligence finančně a časově nejnáročnější (může se jednat až o 80 % časové náročnosti celého projektu) (Pour, 2012).

Získávat data z produkčních systémů je možná v intervalech, nebo v reálném čase. V prvním případě se mluví o tzv. **ETL** procesu (*Extract, Transform, Load*) a v druhém o Real-Time Data Warehouse.

Proces ETL se rovněž označuje jako **datová pumpa**. Jak již napovídá název, tak prvním úkolem procesu je data extrahovat (neboli získat a vybrat z primárních systémů). Není to

jednorázová akce, ale pravidelná činnost, prováděná v delším časovém období. Data jsou přenášena v tzv. batch (dávkovém) režimu zpravidla v denních, týdenních nebo ročních intervalech. *Extract* bývá označováno jako nejpodstatnější část ELT procesu, jelikož na správném získání dat závisí všechny následující procesy.

Druhým úkolem je transformace (*Transform*) extrahovaných dat. To představuje úpravu a čištění těchto dat pomocí specifických nástrojů a docílení jejich převodu do požadované podoby. Mezi tyto nástroje řadíme agregace, třídění, normalizace, filtrování, slučování tabulek, apod. Transformace dat vede k zajištění jejich úplnosti, správnosti a jednotnosti.

*Load* je poslední částí ELT procesu, která se zabývá nahrávání transformovaných dat do specifických datových struktur (datových schémat, datového skladu). Tyto formáty struktur musí být předem navrženy tak, aby vyhovovaly potřebám řízení podniku.

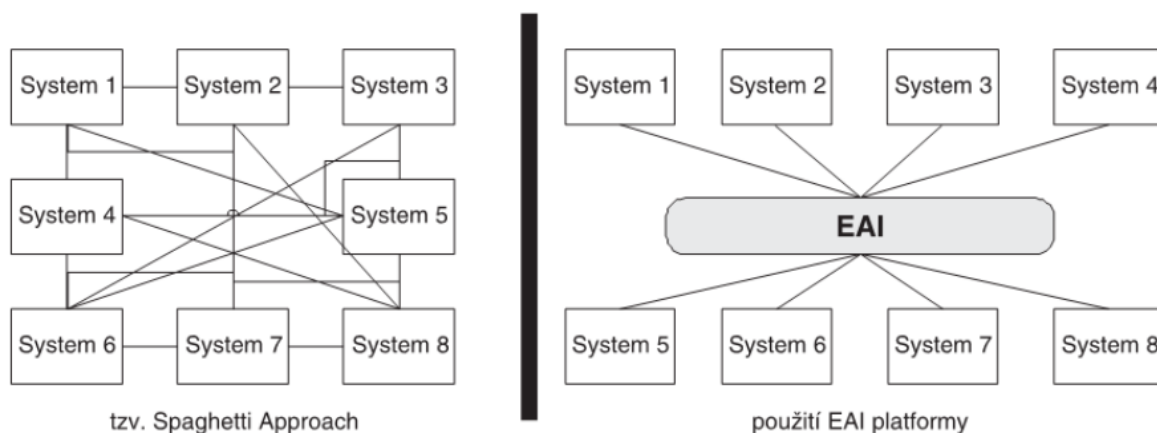
Proces datové pumpy zkráceně slouží k přenosu a úpravě dat z různorodých databázových architektur. Data nemusí být nutně transformována přímo do datového skladu, ale mohou využívat nejprve dočasné databáze. V ní jsou provedeny operace *Transform* popsané výše a následně přenesena do datového skladu (Pour, 2012).

### **ETL vs. ELT**

Vedle nástroje Extract, Transform, Load (**ETL**) se začal používat i Extract, Load, Transform (**ELT**). V druhém případě jsou data nahrávána do uložště před tím, než dojde k jejich transformaci. Data nejsou zpracovávána prostřednictvím transformačních nástrojů, nýbrž je využíváno dotazovacího jazyka v samotných databázích, kam se data ukládají. Vede to k snížení času, který je nutný pro přenos dat z produkčních systémů. Ty pak mohou být transformována později, až to bude potřeba. Nástroje ELT mají i svá negativa. Nelze například manipulovat s daty, která se nachází mimo databázi, což činí celý proces méně flexibilní. ELT oproti ETL také vyžaduje dočasné tabulky, aby byla umožněna následná transformace (Pour, 2012).

Jako alternativu k ETL lze použít proces **EAI** (Enterprise Application Integration). Na rozdíl od ETL funguje v reálném čase (viz zmíněný Real-Time Data Warehouse) a zaměřuje se na integraci dat. To má za cíl propojit dva a více systémů při zachování co nejvyšší vzájemné nezávislosti. Redukuje se tím počet vzájemných rozhraní

zdrojových podnikových systémů. Graficky znázorňuje EAI platformu v porovnání s tzv. Spaghetti přístupem obrázek č. 3.



Obrázek 3 – Použití EAI platformy

Zdroj: NOVOTNÝ, Ota 2005, Business Intelligence: Jak využít bohatství ve vašich datech, s. 30.

### 1.4.3 Datový sklad

Bill Inmon definuje **DWH** (Data Warehouse – datový sklad) jako „*integrováný, subjektivě orientovaný, stálý a časově rozlišený souhrn dat, uspořádaný pro podporu rozhodování managementu*“ (Novotný, 2005, s. 32). Jedná se o základní databázovou komponentu v architektuře BI, která primárně slouží jako podklad pro reporting a analýzy. Je to jedno místo, kde jsou shromážděna a uložena data ze všech zdrojů, ve kterých by se mohla vyskytovat relevantní informace pro koncového uživatele.

Z definice lze stanovit, že DWH by měl být:

- *integrováný* – ukládání dat probíhá v rámci podniku jako celku;
- *konsolidováný* – data z různých zdrojů musí být v DWH převedeny do jedné formy;
- *subjektivě orientováný* – rozdělení dat je na základě jejich typu, nikoliv na základě místa, kde vznikly;
- *stálý* – data v DWH nevznikají manuálním porízením a nelze je nikterak upravovat, slouží pouze pro čtení;
- *časově rozlišený* – DWH musí obsahovat dimenzi času, aby mohla být prováděna analýza za určité období a uložena i historie dat.

Do datového skladu tedy data natékají v daných časových intervalech (dávkově) a zpravidla nedochází mezi intervaly k jejich změně nebo přidávání. Odstraňování

záznamů také není většinou žádoucí, aby byly zachovány historické informace a bylo možno sledovat změny.

V dnešní době je běžnější, že se data v DWH vyskytují v normalizované formě a denormalizace (ukládání do databázových schémat) se provádí až na úrovni datových tržišť. Zatímco dříve k normalizaci (například na základě STAR schémat) docházelo již na úrovni DWH (Pour, 2012).

#### **1.4.4 Datové tržiště**

Podstata **DMA** (Data Mart – datových tržišť) je podobná datovým skladům. Hlavní odlišností je, že DMA jsou určena pro užší okruh uživatelů. Může tím být jen jedno oddělení, divize, závod, pobočka apod. Jinými slovy se jedná o decentralizované datové sklady, které se postupem času mohou integrovat do celopodnikového řešení (Novotný, 2005).

DMA bývají většinou právě tím místem, z kterého čerpají data analytické komponenty architektury BI. A jelikož jsou datová tržiště stavěna na základě potřeb jednotlivých okruhů uživatelů společnosti, liší se mezi sebou tím, jaká data vyžadují, jak dlouhou historii potřebují, jak často musí být data aktualizována.

Lze rozlišovat dva přístupy k budování DMA a DWH. Autorem prvního je R. Kimball (zakladatel multidimenzionálního modelování), který Data Warehouse popsal jako prosté sjednocení datových tržišť. Budují se tedy nejprve jednotlivé DMA na základě požadavků jednotlivých skupin uživatelů. Tato tržiště jsou zakládána postupně, dle potřeb uživatelů a jsou mezi sebou nezávislá. Sjednocením všech těchto Data Marts dostaneme Data Warehouse. Tento princip izolovaných DMA se také označuje jako *dvouvrstvá architektura*. Oproti tomu existuje také přístup *třívrstvé architektury*, s kterou přišel W. Inmon. Zde se data ze zdrojových systémů nahrávají do centrálního datového skladu. Nad tímto centralizovaným DWH se následně vytvářejí jednotlivé datové tržiště. Tento přístup je nákladnější a časově náročnější na realizaci. Naopak v něm však oproti prvnímu způsobu nedochází výskytu redundantních dat (Pour, 2013).

#### **1.4.5 Dočasné a operativní uložení**

**DSA** (Data Staging Area – dočasné uložení) je nepovinná součást databázových komponent v architektuře BI. Slouží k prvotnímu uložení netransformovaných dat

z produkčních systémů a podporuje rychlou a kvalitní extrakci dat do datového skladu. V DSA se nachází data, která jsou neagregovaná, nekonzistentní, detailní, neobsahující historii, měnící se a v identické struktuře, jako jsou uložena v primárních systémech. V dočasném uložišti jsou pouze aktuální data, která se po zpracování a přenosu do DWH (nebo DMA) z DSA odstraní (Novotný, 2005).

**ODS** (Operational Data Store – operativní uložště dat) je také komponent, který není nezbytný v architektuře Business Intelligence. Je to uložště, které slouží jako databáze podporující proces analýzy. Mají do něj přístup i koncoví uživatelé a ostatní systémy, což je jeden z rozdílů oproti DSA. Dalším je ten fakt, že v ODS se nachází data konsolidovaná, konzistentní, subjektivě orientovaná a někdy i doplněná o agregace. Operativní uložště dat lze tedy definovat jako databázi, která podporuje jednodušší dotazy nad menším množstvím aktuálních analytických dat. Příkladem může být referenční databáze zákazníků nebo produktů (Novotný, 2005).

#### **1.4.6 Multidimenzionalita**

Aplikace Business Intelligence jsou založeny na principu několikadimenzionální tabulky, která umožňuje pružně a velmi rychle jednotlivé dimenze měnit. Tím uživateli nabízí různé pohledy na danou ekonomickou realitu. Takový model používá struktury optimalizované pro dotazy koncových uživatelů a nástroje pro ukládání dat.

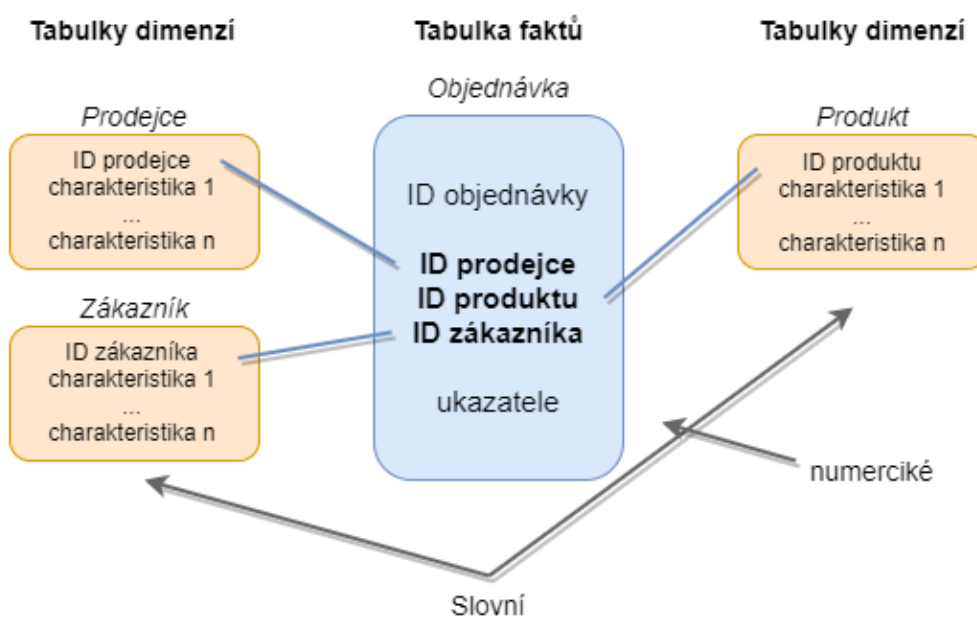
Multidimenzionalitu dat lze implementovat na úrovni relační databáze. V takové případě tyto modely rozlišují dva typy relací – *tabulky faktů* a *tabulky dimenzí*. Tabulka faktů v sobě udržuje ukazatele, které je žádoucí analyzovat (jedná se většinou o ekonomický ukazatel, který je numerický). Úroveň míry detailů v tabulce faktů představuje pojem *granularita*. Čím méně jsou data podrobná, tím je její míra nižší a naopak (Pour, 2013).

Tabulky dimenzí obsahují popisné atributy (slovního charakteru), které dávají význam konkrétnímu faktu. Na základě těchto atributů lze manipulovat a třídit data v tabulkách faktů. Vzájemně jsou propojené pomocí tzv. cizích klíčů. Běžně může dimenzionální tabulka obsahovat až několik desítek atributů. Z některých těchto atributů lze v rámci dimenze definovat tzv. *hierarchie* (jsou uspořádány v hierarchické struktuře). Prakticky to znamená, že prvky dimenzí se rozdělují na skupiny prvků, podskupiny až na jednotlivé prvky (Novotný, 2005). Ty slouží k vytváření agregací. Na základě toho lze odlišit dva druhy dimenzionálních atributů – úroňové atributy (určující agregační úroň hierarchií)



a popisné atributy (pouze specifikující danou úroveň v hierarchii). Pouze pomocí úrovnových atributů lze provádět operace v OLAP kostce (viz následující kapitola). Popisné atributy pouze doplňují informace k dané úrovni a operace pomocí nich nelze.

Databáze OLTP (transakčních) systémů jsou zpravidla modelovány v **3NF** (třetí normální forma). Tato úprava umožňuje rychlé a jednoduché ukládání dat při současné optimalizaci velikosti této databáze. Tabulky dimenzí však nemusejí 3NF dodržovat. V takovém případě nesou název de-normalizované a dodržující **2NF** (druhou normální formu).



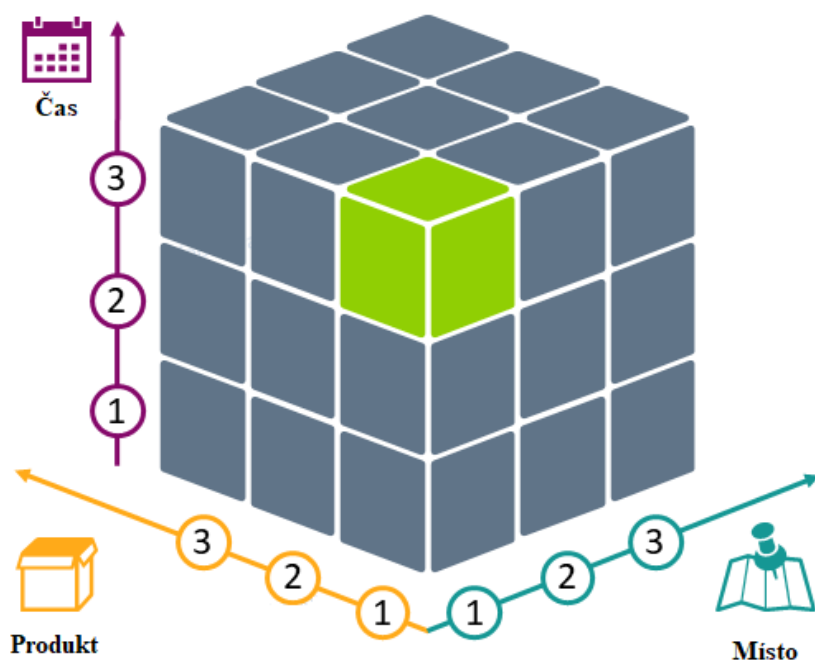
Obrázek 4 – Star schema

Zdroj: vlastní zpracování dle INMON, Bill 2002, Building the Data Warehouse, s. 140.

Datové modely produkčních systémů často obsahují mnoho tabulek a jsou komplexní. Je tedy nutné využití relačních dimenzionálních modelů. Nejzákladnější z nich – *Star schema* (multidimenzionální model typu hvězda) je znázorněn na obrázku číslo 4. Model obsahuje jednu tabulku faktů, která je obklopena několika de-normalizovanými tabulkami dimenzí. Mezi sebou jsou spojeny prostřednictvím identifikátorů a vytváří podobu hvězdy. V případě, že by byla alespoň jedna z dimenzí normalizovaná do dílčích tabulek, jednalo by se o model sněhové vločky (*Snowflake schema*). Dá se tedy tvrdit, že Star schema je specifickým typem Snowflake schema, které má pouze jednu úroveň hierarchie. Model, který je kombinací několika schémat hvězdy (obsahuje více tabulek faktů) se nazývá *Constellation schema* (souhvězdí) (Kimball, 2013).

## 1.4.7 OLAP

**OLAP** (Online Analytical Processing) je informační technologie, která je založená na koncepci multidimenzionálních databází. V architektuře BI je tvořena datovými OLAP kostkami (OLAP Cubes), jejichž základem jsou tabulky faktů a dimenzí. Dimenze přitom mají většinou hierarchickou strukturu. Jedna nebo i více těchto souvisejících kostek dohromady představuje OLAP databázi. V ní se nachází data nenormalizovaná – obsahuje tedy předem zpracované agregace dat dle definovaných struktur a jejich kombinací. OLAP databáze je stěžejní analytický komponent celé architektury Business Intelligence, jelikož uživatelů přehledně zpracovává a rychle zpřístupňuje velké objemy dat z různých pohledů (Novotný, 2005).



Obrázek 5 – OLAP kostka

Zdroj: vlastní zpracování.

Na obrázku č. 5 je možné vidět OLAP kostku zobrazující agregovaná data podle následujících faktů: produkt, čas a místo (každá představuje jednu dimenzi). Každý jeden průsečík všech dimenzí představuje právě jednu konkrétní hodnotu (prvek multidimenzionální databáze). Každý tento prvek může obsahovat algoritmy (předpisy) pro jejich transformace.

Existuje několik základních operací, pomocí kterých se rozšiřuje pohled na tato agregovaná data. Mezi operace, které toto podrobnější procházení dat v OLAP kostce umožňují, patří:

- *Slicing* (omezení jedné dimenze – například zobrazení jen jednoho roku),
- *Dicing* (omezení prvky více dimenzí, dělení na menší kostky – například zobrazení jednoho roku a jednoho produktu),
- *Drill-down* (posun níže v hierarchii o úroveň – například u dimenze místo to může být přechod z Evropy na Německo),
- *Drill-up* (opak předchozí operace),
- *Pivoting* (změna pohledu – otáčení kostky) (Kimball, 2013).

V porovnání s OLTP systémy, které uchovávají data na nejvyšší úrovni detailu, ukládají OLAP databáze jen data, která jsou relevantní pro analýzy. Jsou buď agregovaná na vyšší úroveň než jednotlivá transakce, nebo zahrnují jen nějaké její atributy. Dalším velkým rozdílem mezi těmito systémy je ten fakt, že do OLTP systémů jsou data pořizována v reálném čase (dochází ke kontinuálnímu zatěžování), zatímco u OLAP databáze jsou data aktualizovaná v daných intervalech (nepravidelná zátěž systému) (Novotný, 2005).

#### **1.4.8 Dolování dat**

*Data Mining* (dolování dat) je jedním z analytických komponent systému Business Intelligence. W. Inmon ho ve své knize (2002, s. 389) definuje jako „*proces analyzování velkého množství dat sloužící k odhalení dříve neznámých obchodních souvislostí*“. Zjednodušeně jde o objevení strategických informací v datech pomocí speciálních automatických algoritmů. Základem jsou kvalitní data, která pochází ze správně postaveného datového skladu.

U procesu data miningu se jedná nejen o zobrazení deskriptivních informací, ale především prediktivních. Manažeři tyto informace využívají pro objevování nových skutečností o činnostech společnosti, testování hypotéz, odhalování skrytých závislostí. K dolování dat a získávání cenných informací je využívána řada statistických a matematických modelů. Samotný manažer již však nemusí být specialistou na statistiku, aby těmito informací porozuměl, jelikož mu jsou přehledně a srozumitelně prezentována prostřednictvím BI nástrojů pro koncové uživatele.

### 1.4.9 Reporting

Pojem *Reporting* v oblasti BI zahrnuje reportingové nástroje, které slouží k přehlednému zobrazení podstatných informací pro konkrétní potřebu koncového uživatele. Tyto nástroje jsou postavené na dotazování do databází. Zdroje těchto dat mohou být transakční databáze, data warehouse, ale v rámci správné architektury BI to bývá především OLAP databáze. K získávání dat se využívá *Structured Query Language* (SQL) dotazů do těchto databází. Následně jsou tato data načtena do aplikací pro tvorbu reportů (Novotný, 2005).

Aplikace pro tvorbu reportů slouží k tabulkovému nebo grafickému zobrazení (mluvíme také o statickém a dynamickém obsahu) získaných dat. Kvalitní report poskytuje rozšířené nastavení, které uživatelům umožňuje zobrazit více podrobností, data filtrovat nebo je ukázat z jiného úhlu pohledu. Po prvotním vytvoření podoby reportu již bývají jeho aktualizace zcela automatizované.

Report poté může být zasílán uživatelům v pravidelných intervalech na email nebo k němu mohou mít řídicí pracovníci přístup například skrz firemní intranet. Zobrazení prostřednictvím webového rozhraní má nespornou výhodu v tom, že jsou všechny informace na jednom místě (serveru). Koncoví uživatelé nemusejí řešit samotnou strukturu Business Intelligence, ani jednotlivé komponenty. Pouze si nechají zobrazit informace, které jsou pro ně relevantní. Přístup k nim navíc mohou mít i přes mobilní aplikace a tedy odkudkoliv a kdykoliv.

Speciální typem reportu je tzv. *Dashboard* (palubní deska). Ten je zpravidla složen z několika reportů. Pohled na celý systém je přehledně zobrazen na jedné stránce a obsahuje interaktivní prvky. Klade přitom důraz na grafické zobrazení informací (indikátory, budíky apod.). Příkladem zobrazovaných informací může být denní objem prodeje, aktuální stav cash-flow, plnění měsíčního plánu, porovnání současného stavu ukazatele oproti předchozímu časovému období, ...

Reporting se dělí na *standartní* a *ad hoc*. Standartní je právě ten, který využívá předpřipravené dotazy, které v daných časových intervalech spouští. Ad hoc reporting pokrývá aktuální požadavky uživatelů, které nezahrnuje standartní reporting. Jeho základem jsou specifické, jednorázové dotazy, které jsou explicitně vytvořené uživatelem (Novotný, 2005).

### 1.4.10 Metadata

Důležitou součástí architektury BI jsou i tzv. *metadata* (metaúdaje). Zjednodušeně řečeno to jsou data o datech. Metadata byla součástí různých programů a dat již několik desetiletí, v rámci Business Intelligence však dosáhly další úrovně významu. V datovém skladu neobsahujícím metadata, by bylo pro uživatele nebo analytika problematické determinovat, kde začít s analýzou. Musel by totiž podrobně probírat datový sklad a zkoumat, jaká data tam jsou a jaká ne, což by bylo časově náročné. S pomocí metadat však může konečný uživatel rychle přejít na potřebná data nebo zjistit, že tam nejsou (Inmon, 2002).

V kontextu Business Intelligence se metadata chovají jako indexy, které stojí nad celým datovým skladem a udržují přehled o tom, kde se co nachází. Jednotlivými položkami, které metadata v datovém skladu obsahují, mohou být: struktura dat známá programátorovi/analytikovi, zdroj dat do DWH, transformace dat procházejících DWH (transformační pravidla), datový model, vztah mezi datovým modelem a datovým skladem, historie extrakcí.

V prostředí BI lze rozlišovat následující druhy metadat:

- *metadata zdrojových systémů* – popisují zdrojová data (identifikace a vzájemné vztahy),
- *metadata datových pump* – popisují původ dat,
- *metadata databázové vrstvy* – popisují jednotlivé sloupce a řádky databází a také obsah samotných dat nebo objektů (slouží jako technická dokumentace, optimalizují pohledy do databáze),
- *metadata uživatelské vrstvy* – slouží k pochopení vztahů mezi informacemi pro koncové uživatele a poskytují zpětný pohled na původ dat (Pour, 2012).

### 1.4.11 Kmenová data

Kmenová data zahrnují údaje o jednotlivých objektech (zákazníci, závody, dodavatelé, ...), zdrojích (majetek, finance) a produktech (materiál, projekt, ...). Jejich řízením se zabývá tzv. *Master Data Management* (MDM). Důležitou roli hraje v tom případě, když jsou do datového skladu transportována data z několika odlišných provozních systémů. Každý z nich může pracovat například s vlastním seznamem materiálů. MDM má zabránit tomu, aby se v DWH objevily duplicity a chybné informace, pokud by byl nový materiál zaveden do jednoho systému, ale v ostatních by scházel. Kmenová data musí být jediným zdrojem

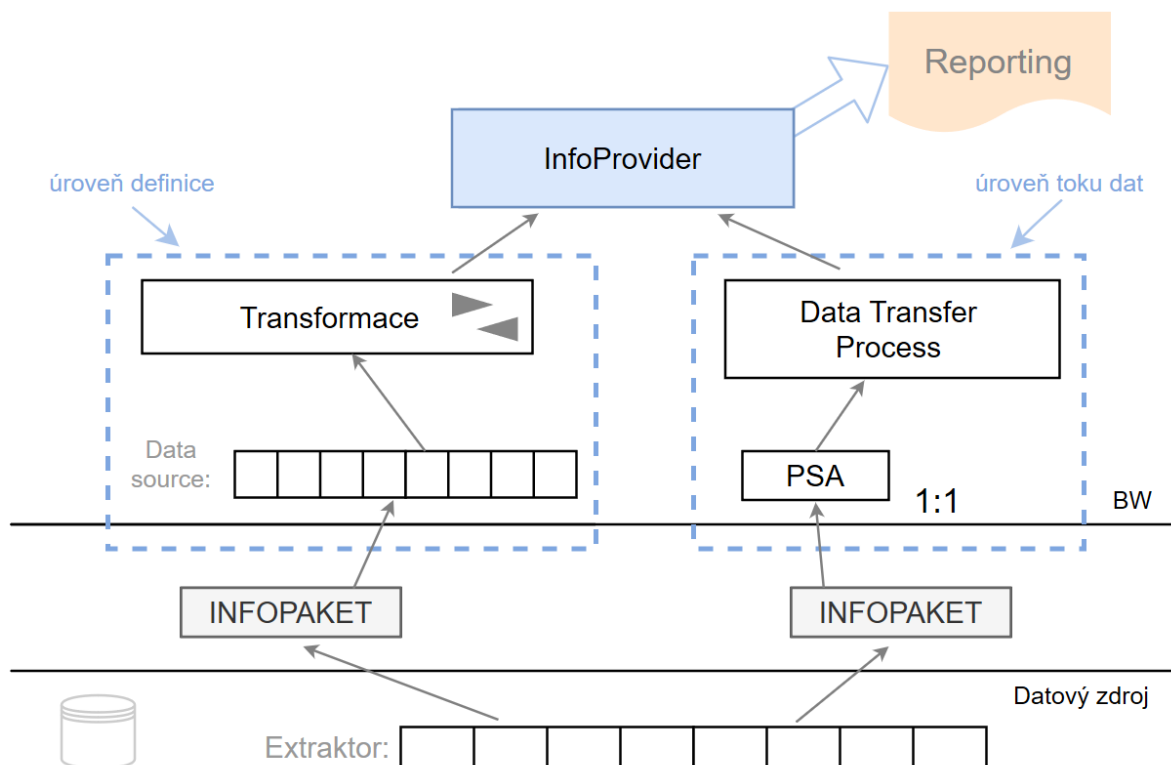
pravdy a podle toho být ukládána a řízena. MDM zajišťuje jejich synchronizaci napříč systémy a i uvnitř systému BI. Když jsou podrobná data (například o materiálu) uložena v master databázi, projeví se každá jejich změna ve všech ostatních systémech, což vede k jejich jednodušší správě (Pour, 2012).

## 2. SAP BW

Pro zpracování praktické části této diplomové práce byl využit systém SAP BW (Business Warehouse). Pro snadnější porozumění praktické části budou některá specifika jeho architektury popsána v této kapitole.

### 2.1.1 Datový tok

Obecný diagram datového toku v SAP BW znázorňuje obrázek č. 6. V tomto zjednodušeném modelu se nahrávají data ze zdrojového systému pomocí infopaketů sestavených nad extraktorem do PSA (Persistent Staging Area = uložisko pro dosud nezměněná data ze zdrojového systému). Odtud se data přes DTP (Data Transfer Process = transformace vstupních dat) dostávají do InfoProvideru podle definice zanesené v předpisu zvaném Transformace (pozn.: datový tok zde probíhá na dvou úrovních – jedna definuje, jakým způsobem budou data protékat (Transformace), a druhá obsahuje samotný tok dat (DTP)). Nad InfoProviderem jsou postaveny tzv. Query, pomocí kterých se sestavují konkrétní dotazy na data. Tyto dotazy slouží jako podklad pro reportovací nástroje.



Obrázek 6 – Diagram datového toku v BW  
Zdroj: vlastní zpracování.

Datový zdroj neukládá žádná data, pouze definuje, jaká pole se budou jakým způsobem přenášet. V SAP BW je vrstva získávání dat reprezentována PSA. Jedná se o povinnou součást toku dat a slouží pouze k dočasnému uložení dat. PSA je příchozí úložiště v SAP BW pro data ze zdrojových systémů. Požadovaná data jsou uložena beze změny ze zdrojového systému (1:1) v databázových tabulkách, které odpovídají struktuře polí v datovém zdroji. Formát dat zůstává nezměněn, což znamená, že nedochází k shrnutí ani transformaci. Lze nahrávat data ze SAP systémů, externích databází i souborů. Aby bylo možné nahrávat data ze SAP systému, musí být propojen s BW systémem – komunikovat s bází.

Datové zdroje ve zdrojovém systému mohou být:

a) Předdefinované extraktory SAP:

- extraktory SAP Business contentu vytvořené společností SAP a připravené pro obecné použití v rámci BW;
- do této skupiny patří i LISovské datové zdroje postavené nad logistickými daty SAP.

b) Vlastní extraktory:

- extraktory vyvinuté dle konkrétních potřeb zákazníka.

Extraktory mohou být do BW nahrávány 2 odlišnými způsoby:

1. Full load ~ pokaždé jsou do BW nahrávány všechny záznamy ze zdrojového systému (není vyžadován Init load).
2. Delta load ~ první load by měl být Init, kdy dojde k jednorázovému načtení všech dat do BW; následně je zahájeno sledování změn/ nových záznamů na zdrojovém systému tak, aby pouze tyto byly v dalším běhu načteny do BW jako datový přírůstek (delta).

Existují různé možnosti nastavení pravidel transformace dat poskytnutých extraktorem do InfoProvideru:

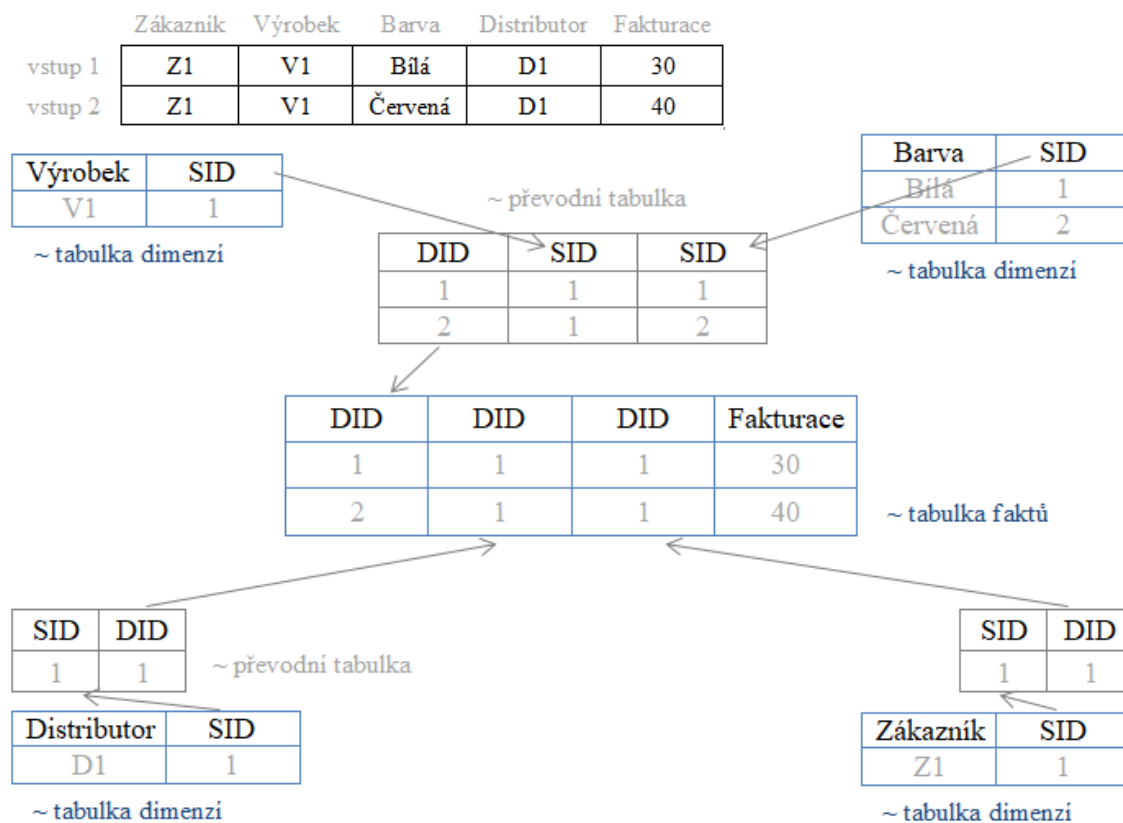
- pravidlo postavené na datech datového zdroje (spouštěcí rutina),
- pravidlo v procesu převodu,
- pravidlo na konci objektu (koncová rutina).



Výstupem datového toku jsou datové objekty (InfoProviders), přičemž rozlišujeme následující typy:

- Kostky = InfoCubes – datová logika založena na „star schema“,
- DSO (Data storage Object, Objekt datové schránky) – datové uspořádání představuje klasickou tabulku s klíčem (jednoznačným identifikátorem) a ostatními datovými poli,
- MultiProvider – datový objekt, který kombinuje data z několika jiných datových objektů pomocí operace „UNION“ a následně je poskytuje k dalším analytickým účelům; multiprovider sám o sobě nikdy neobsahuje data, data jsou stále uložena v datových objektech, nad kterými je MP prostaven.

Do InfoKostky se data pouze vkládají jako nové záznamy, není zde umožněna jejich částečná aktualizace (např. pokud se změní status obchodní příležitosti, není možné v datech aktualizovat stávající datovou větu změnou jednoho pole, ale je nutné původní datovou větu smazat a nahradit ji pomocí nové datové věty. Potencionální duplicity proto v datech musí být ošetřeny dříve, než se začnou data do kostky nahrávat. InfoCubes jsou založeny na Extended Star Schema (znázorněno na obrázku č. 7).



Obrázek 7 – Extended star schema  
Zdroj: vlastní zpracování.

To se skládá z tabulky faktů, která je obklopena tabulkami dimenzí. Každá dimenze je reprezentována svou tabulkou. Primární klíč (SID) z tabulky dimenzí odpovídá cizímu klíči (DID) z tabulky faktů. Pokaždé, když je do tabulky dimenzí přidána hodnota, která v ní ještě není, systém ji automaticky přiřadí SID. Tabulka faktů a dimenze jsou propojeny identifikací abstraktních čísel (ID), které jsou umístěny v klíčové části příslušné databázové tabulky. Používají se k propojení klíče InfoKostky s charakteristikami dimenze. Charakteristiky nastavují granularitu, v níž jsou udržovány klíčové údaje pro InfoKostku. Distributor a zákazník jsou separátní dimenze – jednoduché převodní tabulky. Výrobek a barva tvoří jednu dimenzi – DID vytvořeno pro jedinečnou kombinaci SID. Může existovat maximálně 16 dimenzí.

V DSO se rozlišují pouze klíčová (mají na sebe navázaný index) a datová pole. Používá se pro ukládání konsolidovaných a vyčištěných transakčních dat a kmenových dat. Klasicky obsahuje více granulární data než InfoKostka, která jsou uložena v transparentních, plochých tabulkách. Nad DSO lze vytvářet indexy, což usnadňuje výběr dat. Rychlost zpracování v analytických reportech je horší v porovnání s kostkou. To se projeví obzvláště při velkém množství dat. V DSO lze použít funkcionalitu prepisování záznamů. Znamená to, že při aktualizaci dat v DSO dochází k přepsání charakteristik, které nejsou součástí klíče a zvolené aktualizaci ukazatelů. Oproti tomu InfoCubes vytvoří vždy nový záznam, pokud charakteristiky ve dvou různých záznamech nejsou úplně stejné.

Každý InfoProvider obsahuje InfoObjekty:

- a) Charakteristiky,
- b) Ukazatele (Key Figures).

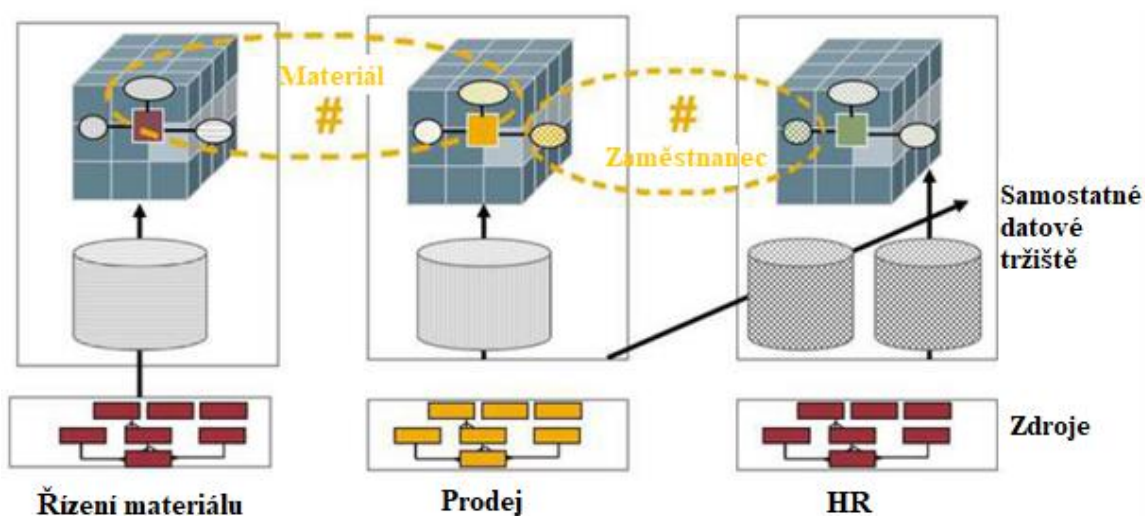
Charakteristika může být jak jednoduché pole (číslo domu), tak i sofistikovaná tabulka kmenových dat (adresa). V takovém případě představují jednoduché charakteristiky atributy složených charakteristik. Atributy mohou být definovány jako navigační (lze pomocí nich vyhledávat/filtrovat) a zobrazovací (slouží pouze k zobrazení). Vedle atributů se rozlišují ještě 2 další typy kmenových dat – texty a hierarchie. Texty mohou obsahovat časově a jazykově závislé popisy charakteristiky. Ukazatele jsou ústřední „fakta“ uložená v systému BW a představují odpověď, kterou uživatel obvykle hledá. Základním

ukazatelem bývají čísla jako hodnota prodejů, náklady, počet zaměstnanců, vyrobené množství.

## 2.1.2 Enterprise Data Warehousing

EDW (Enterprise Data Warehousing) je souhrn všech rozhodnutí učiněných pro celou společnost k vytvoření platného a stabilního řešení podnikového datového skladu, které splňuje všechny požadavky na integrované a důsledně strukturované informace.

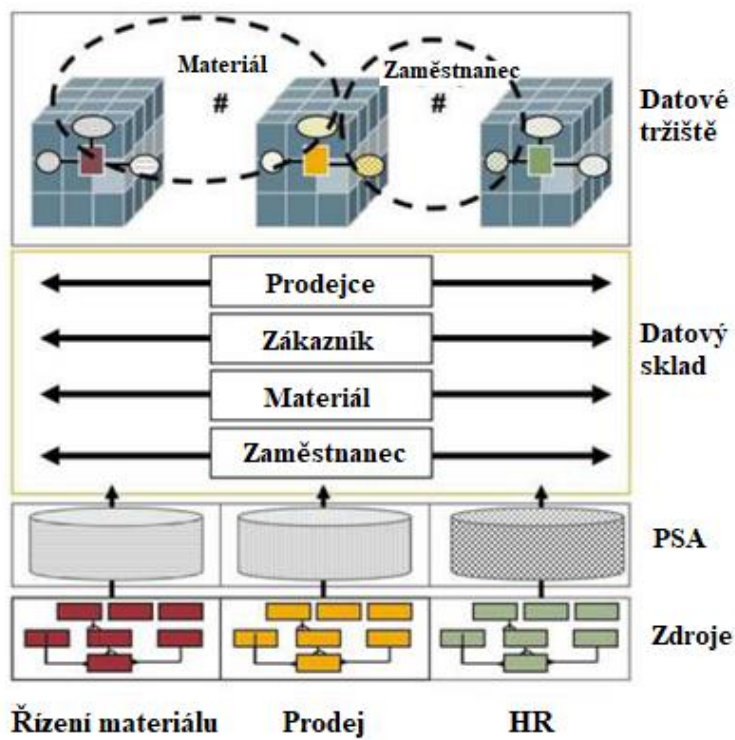
Existují dva základní přístupy k architektuře v SAP BW. Jsou znázorněny na obrázku č. 8 a č. 9. Na prvním se nachází samostatné datové tržiště. Neexistuje zde EDW a tím pádem jsou datové modely vystaveny nezávisle na sobě. Z hlediska dobré praxe se jedná o špatné řešení. Rozdílné datové modely totiž způsobují nekonzistence v sémantice i datech – ve dvou datových modelech budou například hodnoty zaměstnanců s odlišnými daty. Pro různé datové modely existuje také vlastní datový management a je zapotřebí vícenásobné extrakce.



Obrázek 8 – Samostatné datové tržiště  
Zdroj: vlastní zpracování dle interní dokumentace.

V druhém případě je integrovaným základem pro jednotlivá datová tržiště společný datový sklad. SAP BW datové modely jsou vystaveny na společném Data Warehouse. Než dojde k načtení dat do datových tržišť (obecně do InfoKostek) jsou data nahrána do sémanticky jednotného Datového skladu a hodnoty jsou zde jednotně očištěny. Výhodou tohoto přístupu je standardizace extrakce a nahrávání dat, integrace dat s ohledem na jednotnou sémantiku a jednotné hodnoty. Jsou zde vypočteny a skladovány společné (jednotné)

mezivýsledky. Dochází také ke snížení (nebo odstranění) redundancí v datových modelech i datech. Mluví se o tzv. single version of truth.



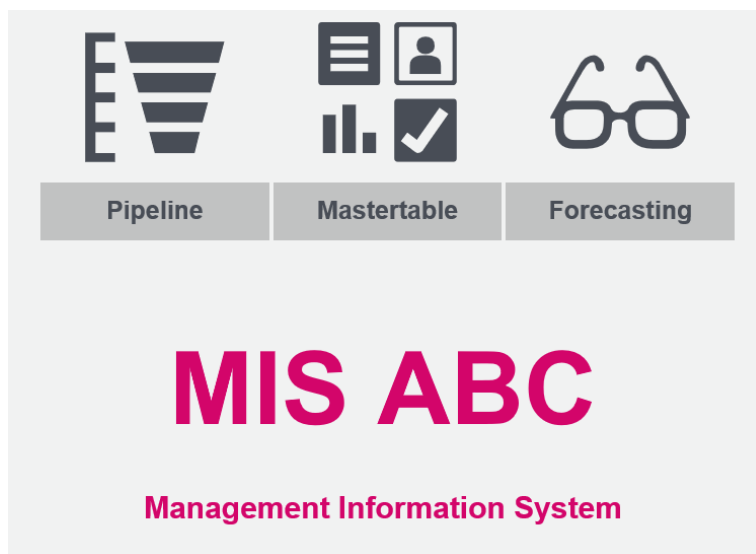
Obrázek 9 – Společný datový sklad  
Zdroj: vlastní zpracování dle interní dokumentace.

### 3. Zpracování vybraného MIS

Praktická část této diplomové práce se věnuje zpracování vybraného manažerského informačního systému. Vybraný MIS vznikl ve firmě Pregis, která poskytuje služby v oblasti vývoje, implementace a podpory informačních systémů SAP. Vzhledem k citlivosti dat, která jsou v koncových reportech obsažena, je zákazník, pro kterého byl informační systém zpracován, utajen a označován jako společnost ABC. Hodnoty údajů jsou také pro účely utajení upraveny nebo cenzurovány.

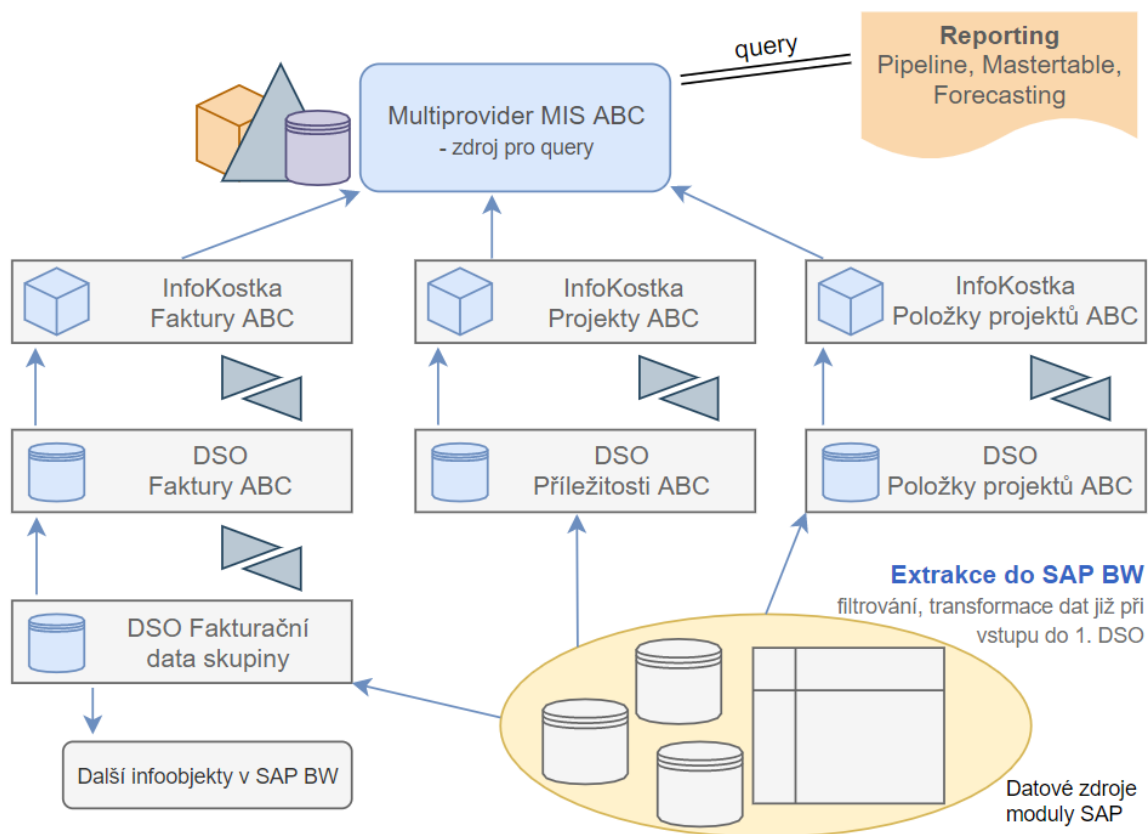
#### 3.1 Výchozí situace

Manažerský informační systém pro společnost ABC původně obsahoval tři reporty. Vstupní obrazovka MIS v prostředí webového prohlížeče je znázorněna na obrázku číslo Obrázek 10. Kliknutím na jednu z ikon uživatel vstoupil do požadovaného reportu. Jednotlivé skupiny reportů jsou označovány názvy Pipeline, Mastertable a Forecasting a každý z nich byl vytvořený dle specifických požadavků a potřeb zákazníka.



Obrázek 10 – Vstupní obrazovka MIS ABC  
Zdroj: interní zdroje společnosti.

Zjednodušený diagram datového toku v SAP BW pro zmíněné reporty znázorňuje obrázek číslo Obrázek 11. Program, který vytváří výsledný report, se dotazuje pomocí queries do multiprovideru MIS ABC. Ten čerpá data ze tří různých InfoCubes – Faktury ABC, Projekty ABC a Položky projektů ABC.



Obrázek 11 – Diagram datového toku  
Zdroj: vlastní zpracování.

Nejtěžejnější a nejrozsáhlejším datovým objektem je kostka Faktury ABC. Obsahuje klíčové pole číslo faktury a položka faktury. V datových polích se nachází charakteristiky jako je prodejní doklad, plátce, příjemce faktury, cesta odbytu, materiál, číslo operace, odpovědný pracovník a mnoho dalších. Mezi ukazatele v kostce patří například fakturované množství, fakturace netto v CZK a v USD, hodnota netto položky faktury. Do fakturační kostky tečou data ze stejnojmenného data storage object. Pod ním stojí DSO s fakturačními daty celé skupiny (dále jen DSO FAK). Extraktor do DSO FAK nahrává z modulu SAP fakturační data různých společností skupiny. DSO Faktury ABC si denní delta extrakcí vybírá a filtruje automaticky jen údaje podstatné pro společnost ABC. Jedná se o přístup, kdy jsou datové modely vystaveny na společném Data Warehouse. Zajišťuje se zde single version of truth a snižuje se redundance v datových modelech i datech.

Data do objektů pod kostkami Projekty ABC a Položky projektů ABC jsou nahrávána pomocí jiného extraktoru a nevyužívají stejný datový zdroj, jako datový sklad s fakturačními údaji. Klíčovým polem zde je číslo operace (tj. číslo projektu), které se také nachází mezi datovými poli ve fakturační kostce. V multiprovideru a na něj navazujících

query je proto nutné pomocí tohoto pole hodnoty propojit. InfoCube Projekty ABC dále obsahuje datová pole ohledně statusu příležitosti, časové podrobnosti o projektu, pravděpodobnost, očekávaný obrat v CZK a USD a další. Kostka Položky projektů ABC zahrnuje charakteristiky jako je požadovaný, dohodnutý a skutečný termín projektu, potvrzení do výroby, datum kompletizace, rating příležitosti apod.

Je také nutné zmínit, že data jsou v procesu datového toku ze zdroje transformována. K nejrozsáhlejší transformaci dochází nejčastěji při toku dat do DSO. Zde je hlavně využíváno koncové rutiny pro čištění a sjednocení formátu dat, ale také pro plnění nových datových polí pomocí SQL dotazů do jiných infoobjektů, které v rámci BW systému existují a obsahují údaje relevantní pro potřebu daného DSO. Dále je zde například využíváno nahrávání kmenových dat charakteristiky materiálu. Tyto transformace, ale také i nastavení extraktorů, se mohou v průběhu času upravit, pokud se změní požadavky zákazníka na koncový report.

### **3.1.1 Skupina reportů Pipeline**

První a také nejrozsáhlejší je skupina reportů Pipeline. Poskytuje zákazníkovi přehled o obchodních příležitostech a běžících projektech z pohledu množství, peněz, času a úspěšnosti získávání. Zpravidla každý z těchto grafů a tabulek je zastoupen svou vlastní query. Jeden z pohledů (očekávaný obrat) je znázorněn na obrázku číslo Obrázek 12. Jedná se o přehledný graf s očekávanou hodnotou příležitostí a projektů v jednotlivých měsících. Tyto projekty a příležitosti jsou barevně rozlišeny dle fáze, v které se nachází. Projekty a obchodní příležitosti mohou být Open (otevřené), InProcess (právě zpracovávány), Confirmed (potvrzené), Won (vyhrané), Realized (realizované), nebo Lost (ztracené). Report nabízí jednotlivé grafy pro projekty a příležitosti v každé fázi. U projektů Won a Lost lze zobrazit procentuální zastoupení důvodů, které stály za tím, že projekt byl vyhraný, nebo ztracený. Záložka Success rate obsahuje bližší pohled na úspěšnost získávání příležitostí za posledních 12 měsíců, tzv. WinRate, který je definován jako podíl získaných příležitostí oproti všem rozhodnutým (zisk/ ztráta) příležitostem za sledované období. Úspěšnost získávání příležitostí je možné sledovat jak za celou společnost, tak za jednotlivá obchodní oddělení, případně v jednotlivých prodejních programech či ratingových skupinách



Obrázek 12 – Report Pipeline  
Zdroj: interní zdroje společnosti.

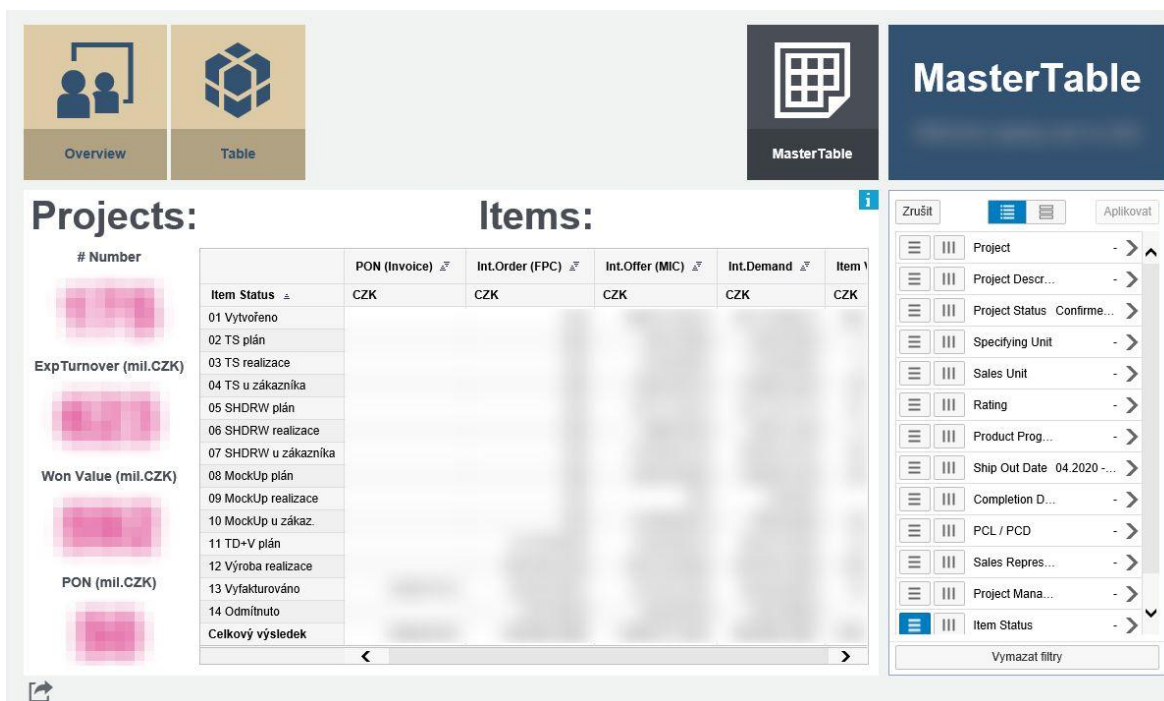
Celá skupina reportů Pipeline je vybavena sofistikovaným filtrem. Aplikováním filtru na čtyři hlavní charakteristiky (Prodejní jednotka, Obchodní zástupce, Produktový program a Rating) se promítne na všechny dílčí reporty. Pokud tedy do záložky Pipeline vstoupí například manažer specifické prodejní jednotky, stačí, že filtr aplikuje na začátku (na prvním reportu). Jak se bude proklikávat dalšími reporty, uvidí pouze hodnoty týkající se jeho prodejní jednotky. Každý report má poté vlastní podrobnější filtr s charakteristikami, u kterých není žádoucí, aby se jejich omezení promítalo mezi záložkami. Výstupní hodnoty reportů je možné jedním kliknutím přepínat mezi početním a peněžním vyjádřením a dále také mezi měnou v CZK a USD. Většina grafů lze zobrazit i v tabulkovém vyjádření a následně přidávat více charakteristik a ukazatelů pro podrobnější pohled. Tento tabulkový výstup je možné jednoduše exportovat do formátu xlsx. Uživatel poté může s exportovanými údaji provádět podrobnější analýzu například v Excelu.

### 3.1.2 Skupina reportů Mastertable

Skupina reportů Mastertable ve firemním procesu navazuje na reporty Pipeliny, když sleduje životní cyklus potvrzených (Confirmed) a získaných (Won) příležitostí, které jsou aktuálně firmou realizovány (v tuto chvíli se příležitosti mění v projekty). Základní pohled do záložky Mastertable je vidět na obrázku číslo 13. Zde se nachází na hlavní straně



přehledná tabulka, v které jsou položky u projektů klasifikovány dle 14 různých stavů. Ke stanovení statusu položky dochází na úrovni BW komplexním výpočtem na základě informací z různých systémů společnosti (CRM, aplikace pro projektové manažery, fakturační data SAP, objednávkový systém). Každá položka se nachází právě v jednom stavu a k ní se váží vybrané ukazatele jako například fakturovaná částka, cena objednávky a cena nabídky. V levé části je číselně vyjádřen počet projektů ve zvoleném období, z nich plynoucí očekávaný obrat, hodnota za kterou byly projekty vysoutěženy (Won Value) a hodnota přímých obchodních nákladů (PON). Stejně jako u předchozího reportu, lze výstup zobrazit v podobnější tabulce, hodnoty filtrovat a exportovat pro podrobnější analýzu v nástroji MS Excel.

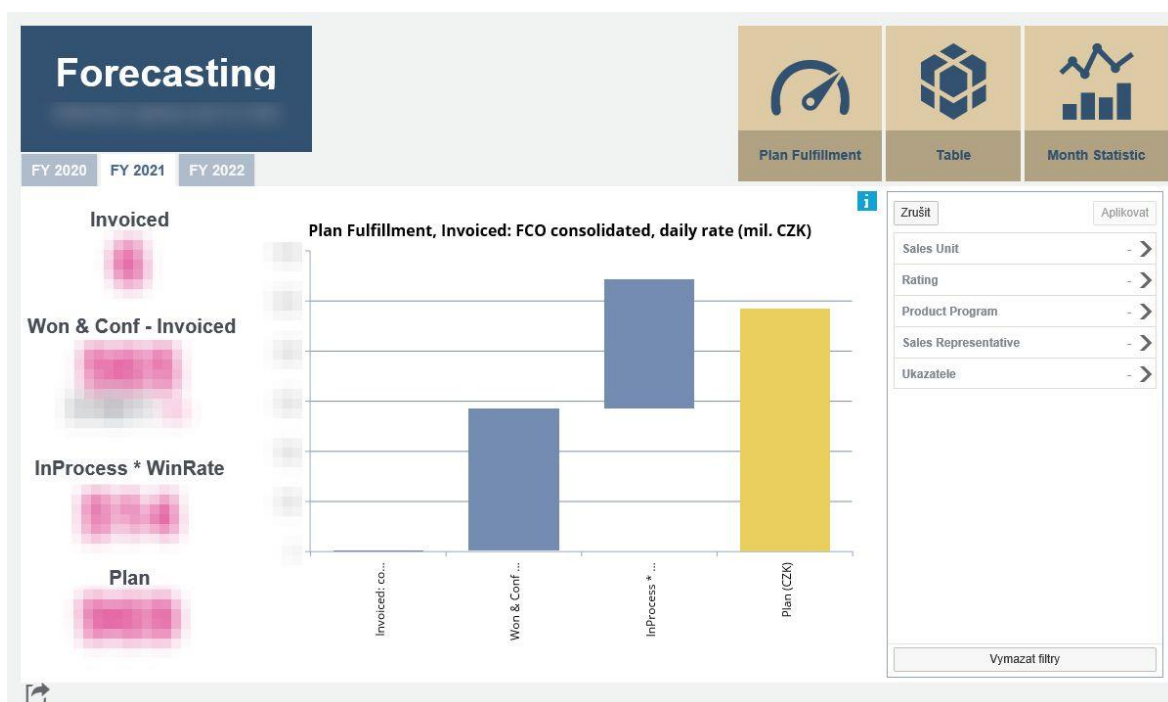


Obrázek 13 – Report Mastertable  
Zdroj: interní zdroje společnosti.

### 3.1.3 Skupina reportů Forecasting

Skupina reportů Forecasting slouží manažerům společnosti ABC pro výhled do budoucna. Pro každý fiskální rok mají interně stanovenou plánovanou peněžní hodnotu vyfakturovaných projektů. Tento report poskytuje přehledný pohled na skutečnost prostřednictvím grafu typu vodopád (viz obrázek číslo 14). Levý sloupec představuje vyfakturované projekty, následující sloupec Won a Confirmed hodnotu aktuálně realizovaných projektů očištěnou o již vyfakturované položky (aby nedocházelo

k duplikaci s daty v prvním sloupci) a třetí sloupec znázorňuje potenciální příležitosti (Open, In Process) s datem realizace v aktuálním fiskálním roce. Hodnota těchto potenciálních příležitostí je ještě ponížena WinRate koeficientem, aby představovala reálný odhad toho, co je společnost schopna na trhu získat. Pokud součet těchto tří sloupců převyšuje poslední (plán), znamená to, že je společnost na dobré cestě, aby plán v daném fiskálním roce dodržela. Report také nabízí pohled měsíčních statistik plánu oproti skutečnosti. Také zde zůstává základní funkčnost jako u předchozích reportů co se týče filtrování a exportu.



Obrázek 14 – Report Forecasting  
Zdroj: interní zdroje společnosti.

### 3.2 Manažerská tabulka

V průběhu roku 2019, kdy již delší dobu byly úspěšně zavedeny předchozí zmíněné reporty, přišel zákazník s požadavkem na rozšíření MIS o další report. Zástupci oddělení controllingu společnosti ABC každý měsíc vytváří tzv. Manažerskou tabulku a měli by zájem o automatizaci jejího sestavování.

Manažerská tabulka je nástroj ke sledování a hodnocení plnění plánu tržeb v konsolidaci. Vykazují se v ní tržby za skupiny materiálu a afilace ve shodě s finančním plánem společnosti ABC. Výstup je provázaný s finančním účetnictvím. Kontrolní mechanismy zajišťují kompatibilitu údajů manažerské tabulky a FÚ. Jedná se o kontrolu faktur

nepropadlých do účetnictví, porovnání výstupu s účetnictvím a kontrolu zaúčtovaných nevyfakturovaných dodávek.

## **Východisko - Původní (ruční) zpracování Manažerské tabulky**

Jak již bylo zmíněno, manažerská tabulka byla manuálně zpracovávána každý měsíc dle následujícího postupu. V prvním kroku byla získána data ze SAP pomocí transakce k exportu seznamu faktur za sledovaný kalendářní měsíc. Před provedením transakce bylo nutné správně zaškrtnout doplňující atributy a omezit filtr, aby výstup nezahrnoval jednotky náležící nežádoucí prodejní kanceláři nebo s nechtěnou cestou odbytu. Následně byl výstup ze SAPu uložený do xlsx formátu.

Další úpravy byly prováděny v Excelu. Z výstupu bylo vyloučit jednotky s takovým číslem materiálu, který značí, že se jedná o službu nebo o zástupný materiál. Následně se do tabulky vložil sloupec *Název skupiny materiálu* a *Název obchodní jednotky*. Jejich správné manuální naplnění je stěžejní pro správnost výstupu. Jsou rozlišovány následující názvy skupiny materiálu:

- **VA (vybavovací akce)** – jedná se o zakázky, které pro společnost ABC představují kompletní výrobní projekt. Příslušnost k této skupině se zjišťuje dle *cesty odbytu*. Dále se dělí na VA maintenance a VA collections a to dle hodnoty pole *typ produktu*.
- **Zákaznický ceník** - zjišťuje se dle hodnoty pole *úroveň prodeje* a *čísla příjemce*, nebo zadavatele zakázky;
- **Collections** – zjišťuje se dle hodnoty pole *úroveň prodeje*;
- **Výprodej** – zjišťuje se dle hodnoty pole *úroveň prodeje*;
- **Old catalogue** - zjišťuje se dle seznamu číselných *skupiny materiálů*;
- **Promo** – jedná se o katalogy a propagační materiály (vybrané číselné skupiny materiálu);
- **Ostatní** - všechny skupiny materiálu mimo výše uvedené.

Při doplňování skupin bylo nutné dodržet správné pořadí, jelikož například jednotka, která odpovídá zařazení do skupiny Výprodej, zároveň může být vybavovací akcí. V tom případě je žádoucí, aby byla zařazena do skupiny VA.

Do sloupce *Název obchodní jednotky* byly doplněny odpovídající afilace. Ty se většinou zjišťovaly dle hodnoty uvedené v charakteristice *prodejní kancelář*. Některé obchodní

jednotky jsou však přiřazovány dle *zadavatele zakázky*, nebo dle hodnoty, která se nachází se v *čísle projektu*. Stejně jako u přiřazování názvu skupiny materiálu, i zde bylo nutné dodržet správné pořadí. Také bylo třeba kontrolovat, zda obchodník, který vystavil fakturu (tj. *zadavatel zakázky*), opravdu patří k dané prodejní kanceláři. Někdy se stává, že obchodník fakturuje pod nesprávnou prodejní kanceláři. Správné přiřazení pak probíhá manuálně prostým přepsáním správné prodejní kanceláře v Excelu. Pokud se obchodník dopustí uvedené chyby, ve fakturačním procesu SAP již není možné tuto chybu jednoduše zpětně opravit.

V dalším kroku se doplnil ukazatel *cena FCO ve fixním kurzu*. FCO zde znamená tzv. franko. Je to označení pro dodávky, u kterých náklady platí odesílatel. Její hodnota ve fixním kurzu se v manažerské tabulce získávala přepočtem hodnoty sloupce *cena FCO* v měně, která je na faktuře. Pro přepočet je využíván fixní kurz jednotlivých měn vůči CZK vyhlášený společností pro na daný fiskální rok.

Po roztřídění dat do správných skupin a k odpovídajícím obchodním jednotkám se vytvořila kontingenční a kontrolní kontingenční tabulka (příklad na obrázku číslo 15). První obsahuje výsledky za jednotlivé obchodní jednotky a druhá za názvy skupin materiálů. Výsledné číslo a součty pro jednotlivé skupiny se musí v obou tabulkách shodovat.

Popisky řádků	Součet z Cena FCO v CZK	Součet z Fixní kurz			Popisky řádků	Součet z Cena FCO v CZK	Součet z Fixní kurz
▢ Londýn					collections		
ostatní					ostatní		
VA					promo		
▢ MAP					VA		
ostatní					výprodej		
promo					<b>Celkový součet</b>		
▢ OJ1							
collections							
ostatní							
promo							
VA							
výprodej							
▢ OJ2							
collections							
ostatní							
VA							
výprodej							
▢ OJ3							
collections							
ostatní							
výprodej							
▢ PAE							

Obrázek 15 – Kontingenční tabulky pro MT  
Zdroj: interní zdroje společnosti.

Po odstranění chyb a nesrovnalostí se správné výsledky kontingence doplnily k příslušným afilacím do souboru s manažerskou tabulkou, která je uložena na firemním intranetu.

Zpracovaná data musí být do manažerské tabulky doplněna na začátku měsíce a následně je nutné provést kontrolu fakturace s účetnictvím. U případných neshod je hledána příčina a následně dojde k úpravě. Ještě k jedné kontrole dojde přibližně v polovině měsíce, kdy se uzavírá účetnictví. Je třeba zkontrolovat, zda někdo proti pravidlům něco nevyfaktoval. V případě, že ano, musí být realizace o tuto částku upravena. 20. kalendářní den je doplněna vedoucím správy zahraničních afilací konsolidovaná realizace afilací a tím je proces pro daný měsíc ukončen.

### **3.3 Zavedení Manažerské tabulky do BW**

Pouze na základě návodky s původním zpracováním manažerské tabulky by bylo složité provést její realizaci v rámci systému BW. Lze vidět, že některé kroky při zpracování jsou činěny dle lidského úsudku, a může si lišit případ od případu. Proto bylo zprvu nutné se zákazníkem problematiku body postupu prodiskutovat a vyjasnit. Některé kroky bylo potřeba zjednodušit a umožnit zákazníkovi případné úpravy provádět i na výstupu z BW. Přípravná fáze zavedení byla časově náročná a vyžadovala několik konzultací se zákazníkem pro nalezení nejefektivnějšího řešení pro obě strany. V systému Business Intelligence lze ke stejnému problému přistoupit z několika úhlů a použít různé postupy. Ty se mezi sebou liší tím, jak jednoduše do těchto postupů lze v budoucnu zasahovat a upravovat je, a také v první řadě jejich časovou náročností – což znamená i finanční nákladností pro zákazníka.

Stěžejní pro celé zavedení bylo zajištění, aby se všechna data, která jsou vyžadována pro sestavení manažerské tabulky, nacházela v BW systému. Data ze SAP transakce seznamu faktur se již vyskytovala ve fakturačním datovém skladu DSO FAK (viz obrázek číslo 11). Bylo zde nutné zkontrolovat extraktor ze zdrojového systému a zjistit, zda poskytoval všechny potřebné charakteristiky a ukazatele a že na něm není zaveden žádný filtr, který by vylučoval data, která jsou potřebná pro manažerskou tabulku. Údaje do fakturačního DSO FAK jsou nahrávána v podstatně nezměněné podobě. Veškerá transformace se děje až v dalších infoobjektech.

Multiprovider MIS ABC již obsahoval fakturační data společnosti ABC, která byla využívána pro předchozí skupiny reportů MasterTable a Forecasting. Jak DSO, tak infokostka Faktury ABC obsahovaly však jen podmnožinu fakturačních údajů, které vyžaduje manažerská tabulka. Doposud se v těchto objektech nacházely pouze fakturační

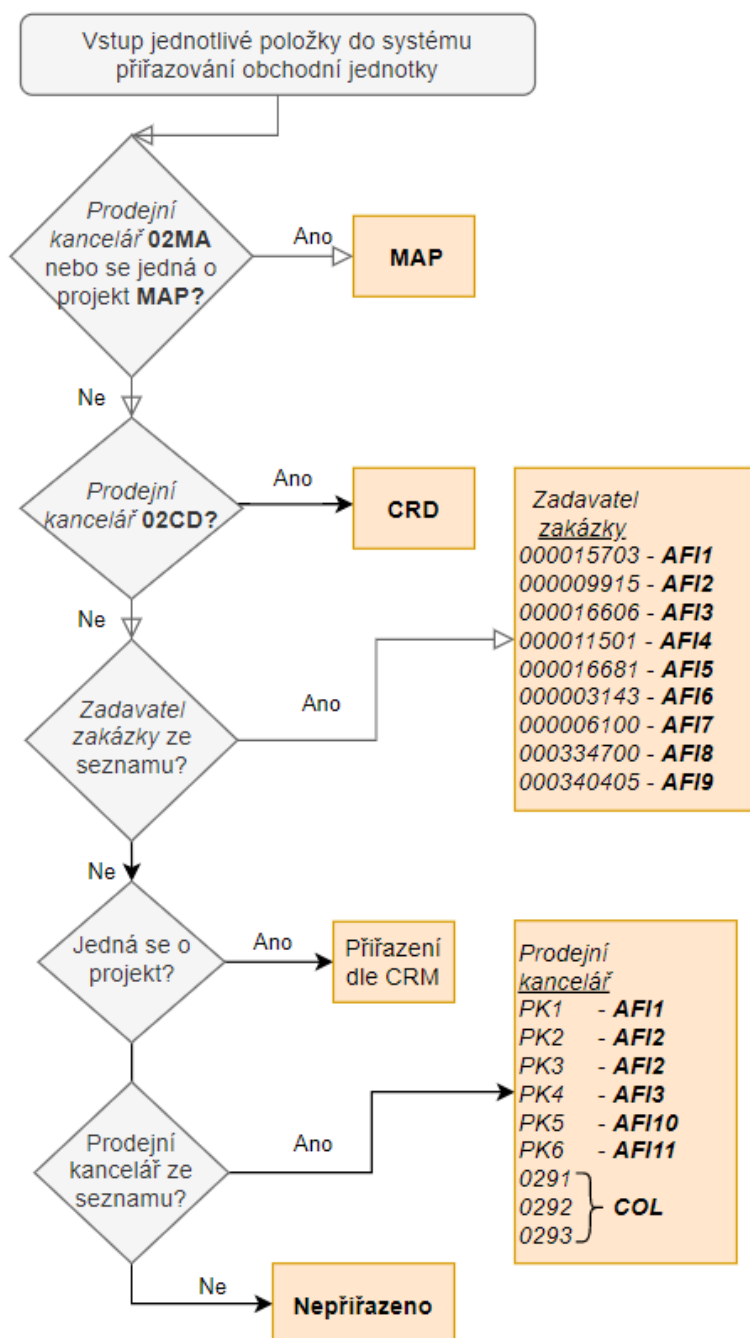
hodnoty zakázek, které jsou projekty (jedná se o vybavovací akce). Pro manažerskou tabulku je to však jedna z podkategorií, kterou sleduje. Zde bylo tedy nutné učinit rozhodnutí, zda vytvářet nové infoobjekty napojené na DSO FAK, které by obsahovaly všechna fakturační data relevantní pro manažerskou tabulku, nebo zda provádět úpravy na stávajícím DSO a infokostce. Lepším řešením se nakonec ukázala druhá možnost a to jak z důvodu nižší časové náročnosti, tak i dodržení dobrých zásad při výstavbě datového skladu. Pokud navíc v budoucnu dojde ke strukturálním změnám ve fakturačních datech, je spíše pravděpodobné, že se to bude týkat faktur celé společnosti, nejen projektů. Proto je lepší všechny tyto údaje mít v jednom infoobjektu, aby bylo možné úpravy provádět na jednom místě. Pro skupiny reportů MasterTable a Forecasting bylo možné omezit fakturační data pouze na vybavovací akce na úrovni jednotlivých query a datové rozšíření stávající DSO tak na ně nemělo žádný negativní dopad.

### 3.3.1 Algoritmus plnění charakteristik

Kritickým krokem pro funkčnost procesu a správnost výstupu bylo dobré sestavení algoritmu, podle kterého byly provádět úpravy, které byly dříve ručně prováděny v nástroji MS Excel. Vyloučení nežádoucích jednotek (s číslem materiálu, který značí, že se jedná o službu nebo o zástupný materiál) šlo nastavit poměrně jednoduše. O mnoho složitější bylo vložení a správné naplnění nových charakteristik *Název skupiny materiálu* a *Název obchodní jednotky*. Byl vytvořen vývojový diagram (viz obr. č. 16 a č. 17), který problematiku naplnění těchto polí zachycuje (skutečná jména afilací (obchodních jednotek) jsou v diagramu nahrazena zkratkami AFI a pořadovým číslem). Přehledné zobrazení se ukázalo jako velmi cenné při komunikaci se zákazníkem. Pomocí graficky vyobrazeného rozhodovacího stromu bylo možné zákazníkovi přiblížit fungování algoritmu a společně tak objevit místa, která bylo potřeba ve zdrojovém kódu dále upravit. Po implementaci se toto zobrazení dále použilo při diskusi nad novými zákaznickými požadavky na další úpravy v naplňování těchto polí.

Podle původní návodky se *obchodní jednotka* přiřazovala tak, že se nejprve zjistilo, zda zadavatel zakázky je interní zákazník. Jedná se o 9 konkrétních hodnot, které určují, jaká obchodní jednotka bude položce přiřazena. Po prvotním sestavení automatizované manažerské tabulky v BW a následné analýze odchylek výstupu oproti skutečnosti bylo zjištěno, že dochází k nesrovnalostem u dvou jednotek, na které nebyl při sestavování původní návodky kladen příliš velký důraz, jelikož se jedná o méně důležité kategorie

zboží (marketingové produkty). Při původním ručním sestavování je zaměstnanec zpracovávající tabulku jednoduše rozeznal dle zkušenosti. V automatizované verzi však tyto položky propadávaly mezi jiné obchodní jednotky. S pomocí zákazníka byly přidány dva kroky na počátek algoritmu, aby k těmto nesrovnalostem nedocházelo. Pokud tedy položka s prodejní kanceláří 02MA nebo 02CD nebo projekt s označením MAP obsahuje v poli zadavatel zakázky jedno z 9 hodnot, již se nedostane do odpovídající, nesprávné skupiny (obchodní jednotky).



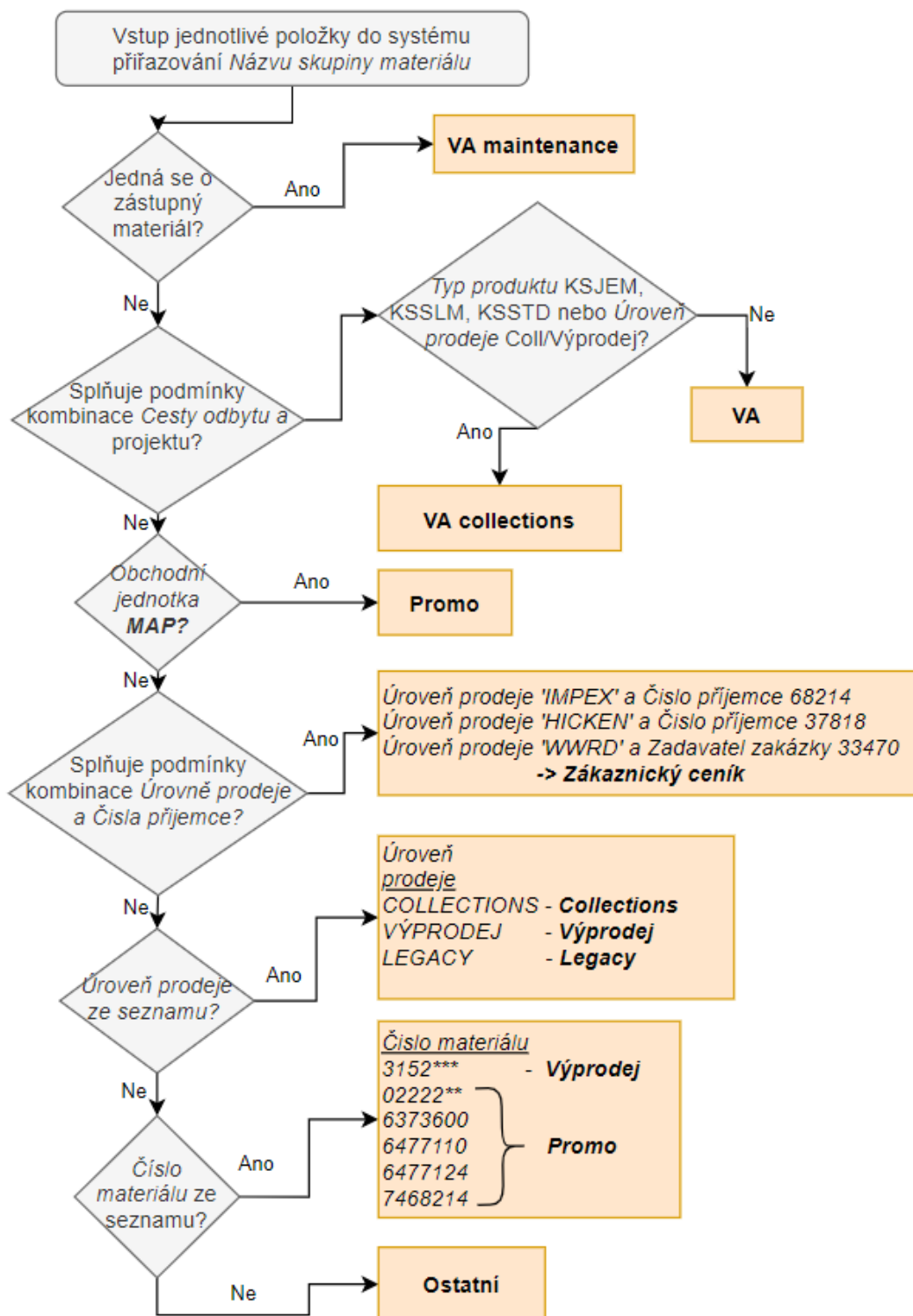
Obrázek 16 – Algoritmus zpracování obchodní jednotky  
Zdroj: vlastní zpracování.

V dalším kroku, pokud se nejedná o interního zákazníka, algoritmus zjišťuje, zda se jedná o projekt (má číslo operace). Pokud ano, tak lze obchodní jednotku přiřadit stejně, jako k tomu docházelo již v původních 3 skupinách reportů, které už BW úspěšně fungovaly. Informaci o obchodní jednotce projektu lze tudíž snadno pomocí SQL dotazu do DSO příležitosti ABC získat. V posledním kroku, pokud se nejedná ani o interního zákazníka, ani o projekt, přiřazuje se obchodní jednotka na základě prodejní kanceláře. Jak již bylo zmíněno, pořadí v tomto postupu je důležité. Pokud by se například obchodní jednotka nejprve zjišťovala dle prodejní kanceláře a poté až dle zadavatele zakázky, docházelo by k velké chybovosti.

Postup pro automatické naplnění pole *název skupiny materiálu* je komplikovanější. Stejně jako u obchodní jednotky, také zde muselo dojít k úpravám oproti původní návodce. Nejprve se přiřadí vybavovací akce do 3 kategorií. Pokud se jedná o zástupný materiál ZAST001, tak se přiřadí skupina VA maintenance. Klasickou vybavovací akcí je každý projekt. Výjimka nastává v případě, že položka má v poli cesta odbytu hodnotu 50 a jedná se o specifickou skupinu prodeje. Mezi VA patří také všechny položky s cestou odbytu 80. Zda se jedná o VA collections se rozhoduje dle pole *typ produktu* a *úroveň prodeje*. V následujícím kroku bylo nutné přidat podmínku pro zachycení marketingových produktů (stejně jako u předchozího algoritmu). Zde už lze využít informace získané z prvního algoritmu a nastavit přiřazení tak, že pokud obchodní jednotka je MAP, tak název skupiny materiálu bude Promo. V dalších dvou krocích se znovu využívá pole *úroveň prodeje* k zařazení položky do skupin *Zákaznický ceník*, *Výprodej* a *Legacy*. V případě, že doposud nebyla položka přiřazena do žádné skupiny, podívá se algoritmus na její číslo materiálu a hledá jeho shodu se seznamem materiálů pro *Výprodej* a *Promo*.

Během několika měsíců vývoje manažerské tabulky v BW byly tento algoritmy mnohokrát upravovány. I v současnosti, při běžném chodu, je to nejčastější důvod zásahu do procesu automatického sestavování výstupu. Důvodem pro to je, že obchodní jednotky (afilace) se v průběhu času mohou měnit. Společnost otvírá a zavírá po světě pobočky nebo dochází k restrukturalizaci. Stejně tak dochází ke změnám u názvů skupin materiálů. Mohou vznikat nová čísla materiálů, kategorie nebo i přesuny mezi jednotlivými skupinami.



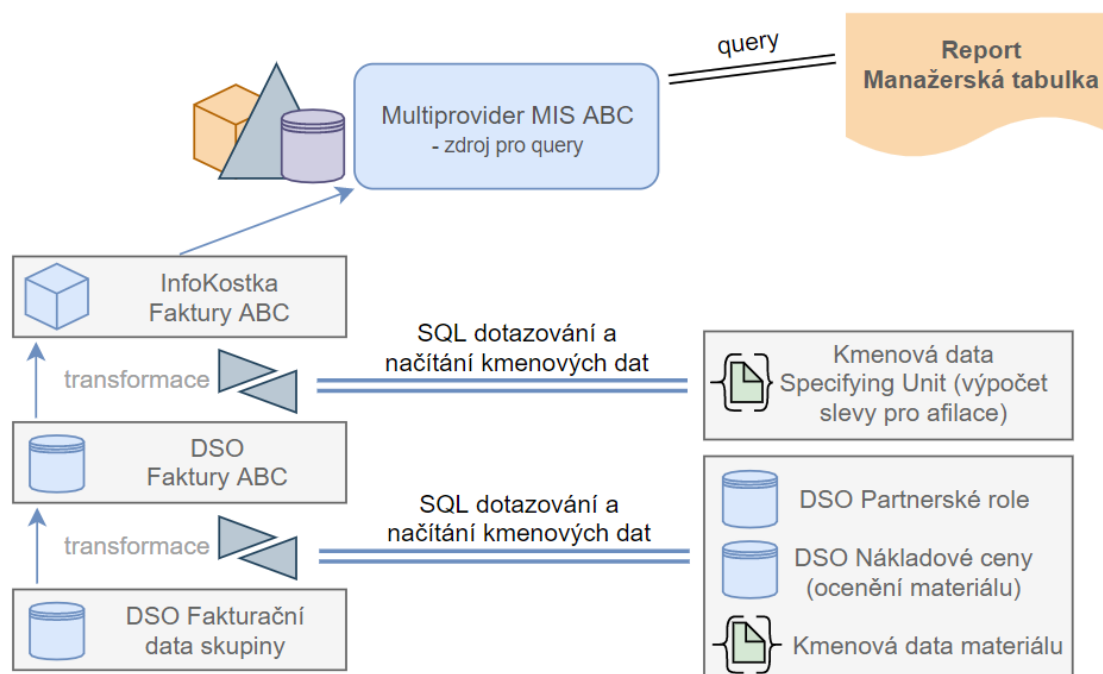


Obrázek 17 – Algoritmus zpracování názvu skupiny materiálu  
Zdroj: vlastní zpracování.

### 3.3.2 Datový tok pro Manažerskou tabulku v BW

Ne však všechna data potřebná pro kompletní sestavení manažerské tabulky a její automatické fungování se vyskytují v DSO FAK (nedají se získat pomocí extraktoru na

něm napojeném). Pro některé specifické ukazatele a charakteristiky byl zaveden vlastní DSO nebo se vyskytují v kmenových datech jiných charakteristik. Do těchto DSO se v koncové rutině nad DSO Faktury ABC pomocí SQL dotazů data přidají. Obrázek číslo 18 graficky znázorňuje situaci.



Obrázek 18 – Datový tok pro Manažerskou tabulku v BW  
Zdroj: vlastní zpracování.

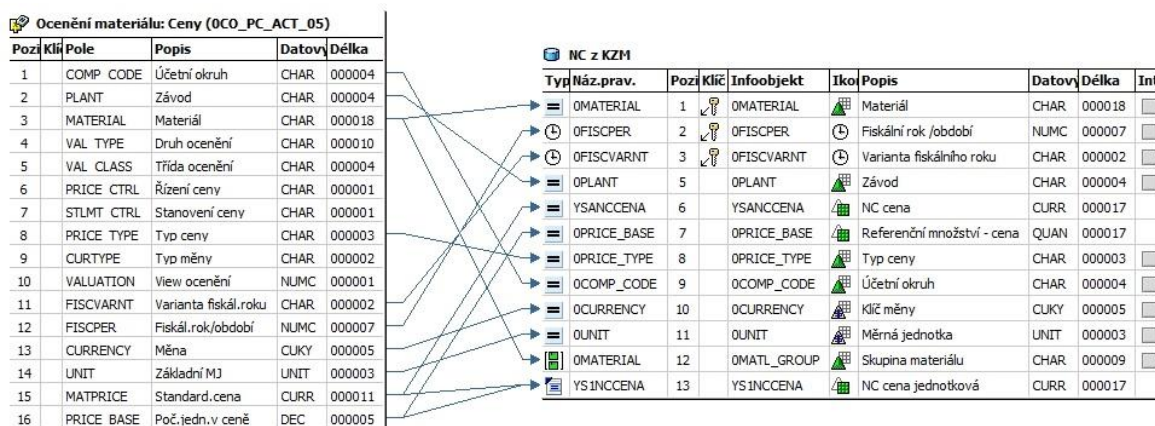
## Úroveň prodeje

V algoritmu pro přiřazování *názvu skupiny materiálu* je využíváno pole *úroveň prodeje*. To však doposud v systému BW nebylo a muselo být přidáno. Údaje o úrovni prodeje se ve zdrojovém systému nachází v klasifikaci kmenového záznamu materiálu. Ze zdrojového systému bylo toto pole dotaženo do kmenových dat materiálu v BW. V transformaci z DSO Fakturační data skupiny do DSO Faktury ABC lze tedy pole *úroveň prodeje* naplnit pomocí pole *materiál* a jeho kmenových dat. Výhodou přidání této charakteristiky do kmenových dat materiálu je to, že lze jednoduše použít i v jiných infoobjektech. Pokud by se však tato charakteristika přidala pouze do jednoho DSO, tak by bylo nutné se do něj vždy dotazovat pro její nahrání do dalších infoobjektů.

## Nákladové ceny

Pro manažerskou tabulku také v DSO Faktury ABC chyběl ukazatel *nákladové ceny*. Ta je taktéž získávána z v klasifikaci kmenového záznamu materiálu. Na obrázku číslo 19

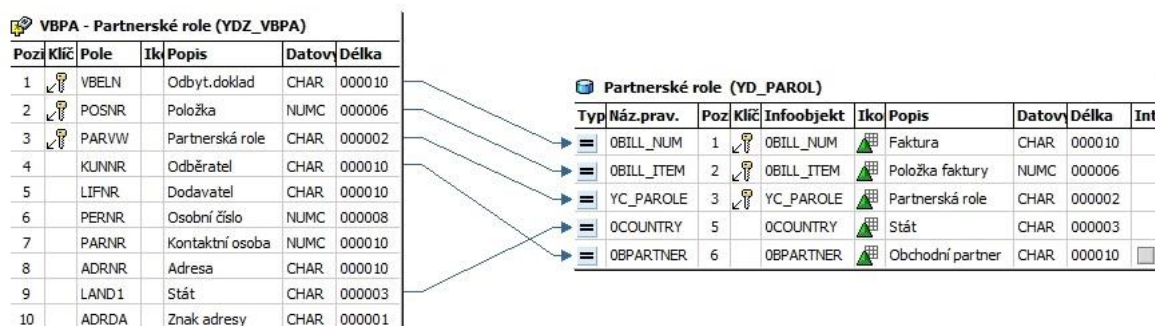
je vidět její extrakce ze zdrojového systému do vlastního DSO Nákladové ceny. Do něj jsou nahrány údaje o standartní ceně materiálu a počtu jednotek v ceně. Z těchto dvou charakteristik je vypočtena *Nákladová cena jednotková*. Právě ta je přidána do DSO Faktury ABC pomocí SQL dotazování v koncové rutině v transformaci. Pole *materiál, účetní okruh, závod, typ ceny a fiskální období* slouží k přiřazení odpovídající hodnoty. Vedle *nákladové ceny* se dotahuje i informace o měně, v které je vedena.



Obrázek 19 – DSO Nákladové ceny  
Zdroj: interní zdroje společnosti.

## Partnerské role

Podobně jsou plněny charakteristiky *zadavatel zakázky, příjemce materiálu a koncový zákazník*. Každá z nich je variací charakteristiky *obchodní partner* a vychází z ní. Na odbytovém dokladu se také nachází *partnerská* role. Je to kódové označení, které určuje, do jaké z podkategorií *obchodní partner* patří. V našem případě je AG *zadavatel zakázky*, WE *příjemce materiálu* a YK *koncový zákazník*. Tento kód spolu s číslem faktury, položkou, obchodním partnerem a jeho státem získán ze zdrojového systému a dotažen do nově založeného DSO Partnerské role (viz obr. číslo 20).



Obrázek 20 – DSO Partnerské role  
Zdroj: interní zdroje společnosti.

Existuje více partnerských rolí, než jen tyto 3, ale do systému BW jsou nahrávány pouze položky s těmito partnerskými rolemi, jelikož ostatní nejsou vyžadovány. Nahrávání všech by pouze zbytečně zatěžovalo systém a bylo by časově náročné. Do tohoto DSO si v transformaci v koncové rutině pomocí SQL dotazování sahá DSO Faktury ABC. Vybírá pole *obchodní partner* a *stát* pro stejná čísla faktury a položky na faktuře a plní je do polí *zadavatel zakázky* (*stát zadavatele zakázky*), *příjemce materiálu* (*stát příjemce materiálu*) a *koncový zákazník* (*stát koncového zákazníka*), na základě toho, jaká hodnota se nachází v poli *partnerská role*. Je také nutné poznamenat, že před zaváděním automatizované manažerské tabulky v BW se již pole *zadavatel zakázky* a *příjemce materiálu* nacházela v DSO FAK a napřímo byla plněna až do multiprovideru. Manažerská tabulka vyžaduje přidání charakteristiky *koncový zákazník* a ta musí být naplněna výše uvedeným způsobem. Pro zachování konzistence, přehlednosti a správných zásad toku dat v BW bylo předěláno i plnění dvou předchozích charakteristik na tento nový způsob.

### **Druhy cen faktur**

Dalším komplikovanějším krokem při tvorbě automatizované manažerské tabulky bylo na požadavek zákazníka zavést různé druhy cen faktur. Cena na faktuře může být sledována jako netto nebo jako FCO cena. Fakturuje se po světě ve spoustě různých měn. Je tedy nutné pomocí kurzu přepočítat částku do sledovaných měn – CZK a USD. Může však být použit dvojí kurz. První je aktuální ke dni vystavení faktury. Druhý je fixní pro přepočet do CZK, který je společností vyhlášený na dané období a během této doby se nemění. Vedle kurzu se také řeší konsolidace cen. U konsolidovaných cen se pro výpočet zohledňuje sleva z nekonsolidované ceny, která je pro jednotlivé afilace předem stanovená. Problematiku nově sledovaných cen pro reporty společnosti ABC přehledněji zobrazuje tabulka číslo 1.

Doposud se v ostatních reportech vyskytovaly pouze cena netto, FCO cena v CZK a USD nekonsolidovaná s kurzem aktuálním ke dni vystavení faktury. Vypočítávala se i konsolidovaná cena, ale pouze v omezeném rozsahu (jednotně pro vybavovací akce). První dva druhy cen z tabulky přicházejí do BW přímo ze zdrojového systému a nacházejí se tedy již v DSO YV\_FAK. Netto cena v CZK je vypočítána v transformaci do DSO Faktury ABC z netto ceny v původní měně a z příslušného aktuálního kurzu ke dni vystavení této faktury. Všech následujících 8 cen je přepočítáno až v transformaci do infokostky. Důvodem pro to je, aby všechny FCO ceny byly přehledně vypočítávány

na jednom místě a zároveň aby se nemusely zbytečně nahrávat z DSO do infokostky a tím komplikovat datový tok.

**Tabulka 1: Druhy cen faktur**

Technický název	Název	Konsolidace	Kurz
0NETVAL_INV	Netto cena v původní měně	nekonsolidovaná	původní měna
0REBATE_BAS	FCO cena v původní měně	nekonsolidovaná	původní měna
YU_FAKCZK	Netto cena v CZK	nekonsolidovaná	aktuální ke dni vystavení
YU_FCZMP	FCO cena v CZK	nekonsolidovaná	aktuální ke dni vystavení
YU_CZKFIX	FCO cena v CZK	nekonsolidovaná	fixní
YU_FAKOC	FCO cena v CZK	konsolidovaná	aktuální ke dni vystavení
YU_FAKFC	FCO cena v CZK	konsolidovaná	fixní
YU_FUSMP	FCO cena v USD	nekonsolidovaná	aktuální ke dni vystavení
YU_USDFIX	FCO cena v USD	nekonsolidovaná	fixní
YU_FAKOU	FCO cena v USD	konsolidovaná	aktuální ke dni vystavení
YU_FAKFU	FCO cena v USD	konsolidovaná	fixní

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

Jak již bylo řečeno, nejprve musí být zavedeny nekonsolidované ceny. YU\_FCZMP a YU\_FUSMP jsou vypočítány obdobně jako netto cena v CZK, akorát místo netto ceny v původní měně využívá FCO cenu v původní měně. Pro získání YU\_CZKFIX je využit funkční modul v kódu, který pro dané datum faktury hledá v systému odpovídající stanovený fixní kurz a aplikuje ho na FCO cena v původní měně. Pokud již faktura byla v CZK, tak samozřejmě k žádné změně nedojde. Jelikož jsou v systému zavedeny fixní kurzy pouze pro přepočítání do CZK, musí být YU\_USDFIX přepočítána z YU\_CZKFIX v koncové rutině. Nejprve je zkontrolováno, zda faktura již nebyla v USD, aby nedošlo ke zbytečnému přepočítávání dvakrát. Pokud je v jiné měně, pomocí funkčního modulu se získá fixní kurz z daného data pro přepočítání z USD do CZK. Tímto kurzem je vydělena hodnota YU\_CZKFIX a tím se získá YU\_USDFIX.

Vypočtení konsolidované ceny YU\_FAKOC, YU\_FAKFC, YU\_FAKOU, YU\_FAKFU vychází z odpovídajících nekonsolidovaných cen. K přepočtu je však ještě potřeba dvou hodnot slev – jedna pro skupiny Collections, Výprodej, Old catalogue, Zákaznický ceník, Legacy a jedna pro ostatní skupiny. Výše slevy pro dané období je nahráváno do kmenových dat charakteristiky *specifying unit* (viz obrázek číslo 21).

Poz	Klíč	Pole	Ik	Popis	Datov	Délka
1		TYP_PROD		Typ a produkt	CHAR	000005
2		OBCHODNI_JEDN		Obchodní jednotka	CHAR	000003
3		TYP_PODMINKY		Typ podmínky	CHAR	000004
4		PLATI_OD		Platí od	DATS	000008
5		PLATI_DO		Platí do	DATS	000008
6		ZAKLAD		Základ	DEC	000006
7		STRMOST		Strmost	DEC	000006
8		DNO		Spodní hranice	DEC	000006

Typ	Náz.prav.	Poz	Klíč	Infoobjekt	Ik	Popis	Datov	Délka	Int
	YV_RS	1		YV_RS		Specifying Unit	CHAR	000004	
	ODATETO	2		ODATETO		Platí do	DATS	000008	
	ODATEFROM	3		ODATEFROM		Platí od	DATS	000008	
	YU_SLEAF1	4		YU_SLEAF1		Sleva ařlance pro VA	FLTP	000016	
	YU_SLEAF2	5		YU_SLEAF2		Sleva ařlance pro COL, VYP, OLD	FLTP	000016	

Obrázek 21 – Kmenová data charakteristiky specifying unit

Zdroj: interní zdroje společnosti.

Jedná se také o obchodní jednotku a lze se na to pomocí SQL dotazů v koncové rutině v kostce Faktury ABC dotazovat. Sleva vychází z parametrů cenové kalkulace ve zdrojovém systému. Využívá se charakteristiky *Typ produktu*. Jednotlivé typy mají přiřazenou svoji slevu pro určené období a tyto informace jsou uloženy právě v parametrech cenové kalkulace. První zmiňované skupině náleží *typ produktu* KSSTD, druhé VANO. Jelikož pro jednu *specifying unit* existují dvě různé slevy, musí být její kmenová data nahrávána nadvakrát. Při současném nahrávání obou docházelo k tomu, že se navzájem přehrávaly.

V koncové rutině transformace infokostky se nejprve kód podívá, o jaký *název skupiny materiálu* jednotka má. Dle toho z kmenových dat vybere požadovanou hodnotu slevy pomocí *obchodní jednotky* a *data faktury*, které napojí na *specifying unit* a *platí od/platí do*. Sleva je následně aplikována na všechny nekonsolidované druhy cen a tím jsou získány všechny 4 typy konsolidovaných.

### 3.3.3 Řetězec procesů

Poté, co je datový tok ze zdrojového systému až do multiprovideru správně nastaven, je třeba stanovit čas a pravidelnost nahrávání a aktualizace těchto dat. Jedná se o automatické spuštění řetězce procesů. Aktualizace dat v BW systému je logicky poskládána tak, aby byly respektovány důležité návaznosti.

Nejdříve dochází k načtení základních dat datového skladu (kmenová data, obecná transakční data) a posléze se teprve aktualizují konkrétní koncové reporty. Kmenová data *materiálu* a *specifying unit* jsou přehrávána každý den v pozdních odpoledních hodinách. Aktuálnost těchto dat je podmínkou pro správné naplnění infoobjektů pro manažerskou tabulku. Aktualizace fakturačních dat YV\_FAK probíhá denně mezi jednou a třetí hodinou ráno. Reporting ABC pak začíná ve čtyři hodiny ráno, kdy se nejdříve aktualizují data z CRM (DSO Příležitosti ABC a infokostka Projekty ABC), pak interní nabídka s interní

poptávkou. Následně proběhne aktualizace MasterTable (DSO a infokostka Položky projektů ABC).

Až po naplnění všech zmíněných objektů přichází na řadu data pro manažerskou tabulku. Zařazení aktualizace reportu ManTab až na konec úplný konec řetězce procesů je dáno tím, že tento report vyžaduje data ze všech ostatních infoobjektů, které při jeho zpracování již musí být datově naplněny.

Řetězec procesů pro manažerskou tabulku zahrnuje v první řadě naplnění infoobjektů DSO Partnerské role a DSO Nákladové ceny, do kterých je následně v transformacích dotazováno. DSO Faktury ABC jsou nahrávány v režimu full. Dojde tedy každý den nejdříve k výmazu stávajících dat a poté jejich opětovnému přidání spolu s novými záznamy za uplynulý den (popřípadě s modifikovanými staršími jednotkami – stornované faktury apod.). Všechny tyto záznamy projdou každý den transformací. Stejný proces platí pro tok dat do infokostky Faktury ABC. Multiprovider MIS ABC nemusí být do řetězce procesů zahrnut, protože sám neobsahuje žádná dat, pouze seskupuje objekty, které mohou být využity pro reportovací dotazy (queries).

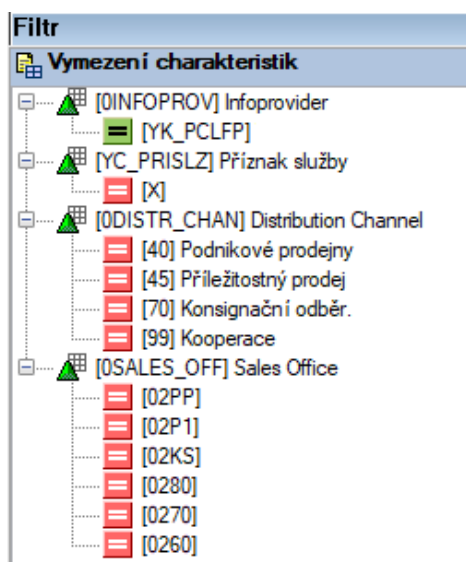
### **3.3.4 Vytvoření queries**

Multiprovider MIS ABC zastřešuje data několika různých infoobjektů. Aby bylo možné z tohoto multiprovideru získat požadovaná data, je třeba vytvořit queries. Při vytváření query se stanoví, zda se výběr omezí jen na některý z infoobjektů v multiprovideru (to je především podstatné, pokud jsou stejné charakteristiky a ukazatele ve více infoobjektech). Dále se mohou vymezit charakteristiky, neboli filtrovat některé hodnoty, které se do koncového reportu vůbec nedostanou. Pro každou tabulku nebo graf ve výsledném reportu není potřeba všech charakteristik a ukazatelů, které infoobjekt obsahuje. V query se přiřadí požadované charakteristiky a ukazatele do řádků, respektive sloupců. Toto uspořádání tvoří výchozí pohled pro danou query. Je možné také zahrnout volné charakteristiky, které uživateli umožní v koncovém reportu tento pohled rozšířit.

Obrázek číslo 22 znázorňuje vymezení charakteristik pro všechny query reportu manažerská tabulka. Z multiprovideru jsou vybírána jen ta data, která pochází z infokostky Faktury ABC. Dále jsou vyloučeny všechny položky, kterým je přiřazen tzv. příznak služby. Tímto příznakem byly v transformaci v BW označeny jednotky s takovým číslem materiálu, který značí, že se jedná o službu nebo o zástupný materiál. Při původním



zpracování manažerské tabulky byly tyto jednotky vyloučeny. V BW systému byly ponechány, protože v jedné query tyto jednotky vyloučeny nebudou. Uživatel si je bude moci zobrazit a exportovat například do excelu pro podrobnější analýzu. Omezení na charakteristikách *cesta odbytu* (distribution channel) a *prodejní kancelář* (sales office) je totožné s filtry, které jsou již nastaveny při nahrávání dat do DSO Faktury ABC. Jedná se o vyloučení jednotek, které pro výstup manažerské tabulky nejsou podstatné. Tím, že je query omezena pouze na infokostku Faktury ABC, nebyly by teoreticky omezení na cestě odbytu a prodejní kanceláři nutné. Je to spíše dvojnásobná pojistka proti tomu, aby v budoucnu po úpravách query do výstupu nepropadla data z jiných infoproviderů, které filtrované jednotky obsahují.



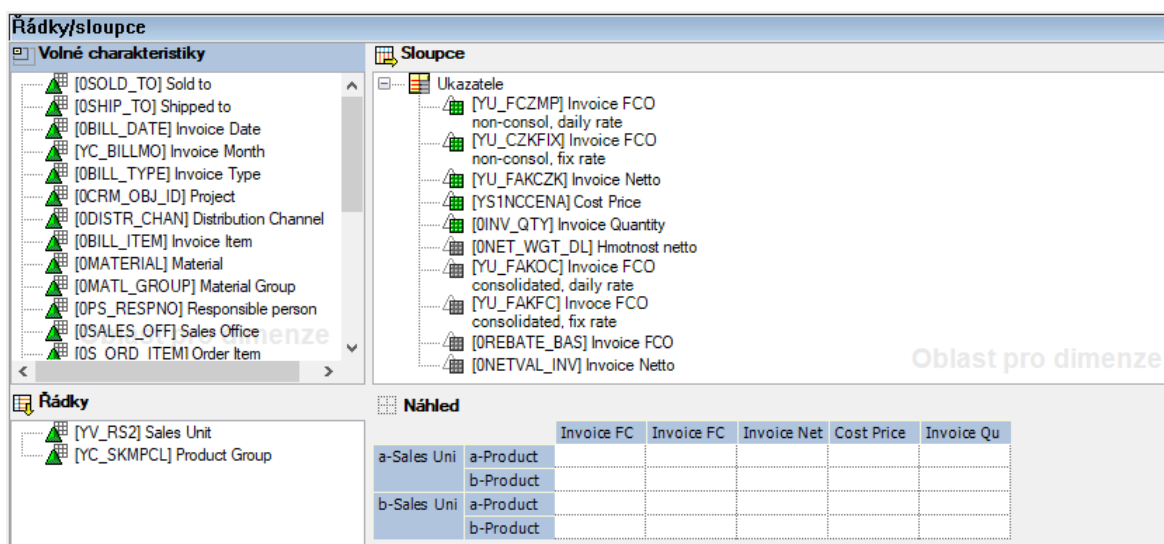
Obrázek 22 – Vymezení charakteristik pro Manažerskou tabulku

Zdroj: interní zdroje společnosti.

Základní dvě query pro report manažerské tabulky jsou ty, které budou odpovídat výstupu z obrázku číslo 15 na straně 46 a dále rozšířené o dodatečné ukazatele. Jak **query\_1** vypadá v programu na jeho tvorbu lze vidět na obrázku číslo 23. V řádcích je zařazena *obchodní jednotka* (sales unit) a pod ní *název skupiny materiálu* (product group). Tato hierarchie stanovuje, že jednotlivé obchodní jednotky budou dále tříděny dle různých názvů skupin materiálu. Ve sloupcích jsou ukazatele třech různých druhů *cen faktur*, *nákladová cena* a *fakturované množství*. Tyto ukazatele jsou defaultně zobrazovány. Dále je do sloupců zařazeno dalších pět ukazatelů (*hmotnost netto* a 4 další druhy *cen faktur*), které jsou potlačeny, ale uživatel si je může nechat zobrazit. Program ukazuje jednoduchý náhled výstupu při současném nastavení, který usnadňuje práci tvorby query. V oblasti volných charakteristik je nutné zahrnout nejen ty charakteristiky, které by si uživatel mohl



chtít do dané tabulky přidat, ale hlavně ty, dle kterých chce výstup filtrovat. V tomto případě to v první řadě jsou charakteristiky časového formátu, jako jsou *datum faktury* (invoice date), *měsíc faktury* (invoice month), *fiskální období* (fiscal period). Také ale například i charakteristiky *zadavatel zakázky* (sold to), *číslo projektu* (project), *cesta odbytu* (distribution channel), *prodejní kancelář* (sales office), *zodpovědná osoba* (responsible person), apod. Mezi volné charakteristiky lze přiřazovat (popřípadě vylučovat) položky libovolně dle požadavků koncového uživatele. Jedinou podmínkou je, aby se charakteristika nacházela ve zdrojovém infoobjektu. **Query\_2** není tak obsáhlá. V řádcích má pouze *název skupiny materiálu* a ve sloupcích tři ukazatele – *FCO cenu (nekonsol., akt. kurz)*, *nákladovou cenu a fakturované množství*. Volné charakteristiky má totožné s první query a víceméně i se všemi ostatními.

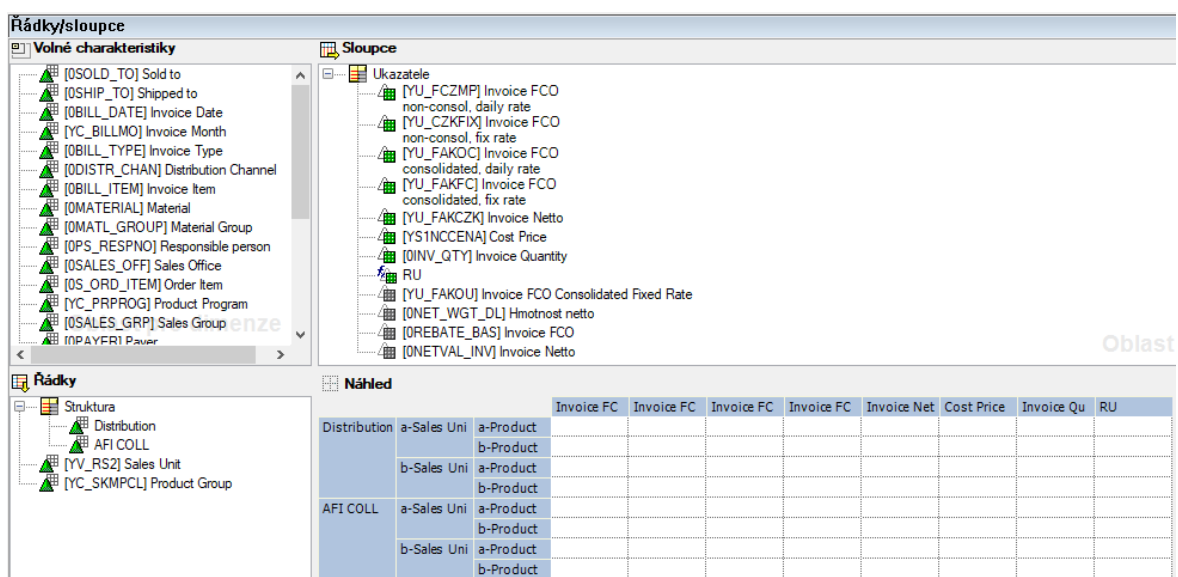


Obrázek 23 – Query\_1 pro Manažerskou tabulku  
Zdroj: interní zdroje společnosti.

Po představení různých možností výstupu z BW zákazníkovi bylo dohodnuto rozšíření Manažerské tabulky o další výstupy. Zákazník měl zájem vidět rozpad fakturací dle jednotlivých měn. Pro to byla vytvořena **query\_3**, které se od query\_2 liší pouze tím, že je zde nahrazen *název skupiny materiálu* v řádcích za charakteristiku *měna* (currency). **Query\_4** v koncovém reportu utváří kompletní tabulku. Většinu volných charakteristik má nyní v řádcích. Tato query slouží uživatelům k podrobnějšímu pohledu na jednotlivé položky.

Dalším požadavkem na přidání výstup do reportu bylo vytvoření přehledu prodeje dle kanálů. **Query\_5** (znázorněna na obrázku číslo 24) zavádí dva kanály – distribution a AFI

COLL. Ty dále dělí dle *obchodní jednotky* a *názvu skupiny materiálu*. Ve sloupcích této query jsou všechny druhy ceny faktur v CZK, *nákladová cena*, *fakturované množství* a ukazatel *RU* – podíl konsolidované ceny FCO v aktuálním kurzu a nákladové ceny. Do kanálu s názvem Distribution spadají *prodejní kanceláře 0291 až 0294* a *zadavatel zakázky 334700* (jedná se tedy o *obchodní jednotky COL* a *WWRD*). Do kanálu AFI COLL patří pouze jednotky s *názvem skupiny materiálu collections, old catalogue, legacy, výprodej* a *zákaznický ceník*. Tato omezení jsou nastavena v oblasti řádků query. Poslední **query\_6** obsahuje pouze jeden řádek – *Bespoke*. Je to agregační tabulka, která zahrnuje pouze položky s *názvem skupiny materiálu VA Collections*.

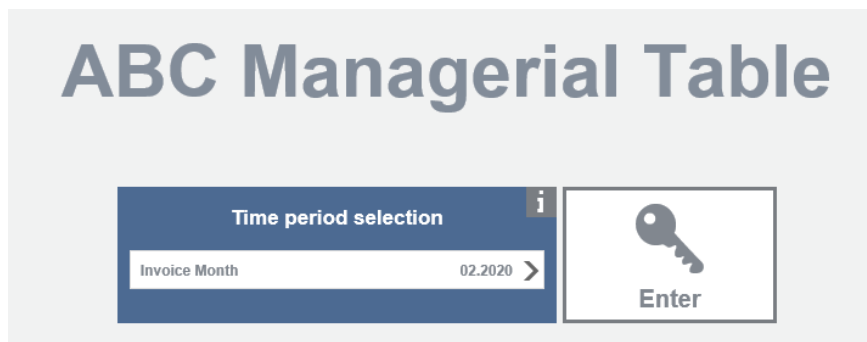


Obrázek 24 – Query\_5 pro Manažerskou tabulku  
Zdroj: interní zdroje společnosti.

### 3.3.5 Vytvoření koncového reportu

Po přípravě všech požadovaných queries, bylo potřeba vytvořit jednotné místo, kde se mají jejich výstupy zobrazovat v přehledné grafické podobě pro koncového zákazníka. Stejně jako u předchozích skupin reportů (Mastertable, Forecasting, Pipeline) se jednalo o interaktivní stránku v prostředí webového prohlížeče. Při vstupu do reportu se nejprve zobrazí úvodní obrazovka (viz obrázek číslo 25). Zde se volí časové období, za které chce uživatel výstup sledovat. Ve filtrovací řádce na obrázku je charakteristika *měsíc fakturace* (Invoice month), která je navázaná na všech šest queries, které ji mají zahrnutou ve volných charakteristikách. Jakmile se potvrdí výběr, omezí se výstup jen na vybrané měsíce. Defaultně je uživateli navrhován předchozí měsíc, jelikož tak je zpravidla manažerská tabulka sledována. Hlavní výhodou omezení časového období před načtením

koncových reportů je výrazné zkrácení času nahrávání všech queries. Postupem času, jak objem dat v infokostce Faktury ABC narůstá, dochází k nárůstu náročnosti zobrazení všech (tedy nevyfiltrovaných) dat. Pokud je výstup vždy omezen pouze na jediný měsíc, je tento problém redukován.



Obrázek 25 – Manažerská tabulka, úvodní obrazovka  
Zdroj: interní zdroje společnosti.

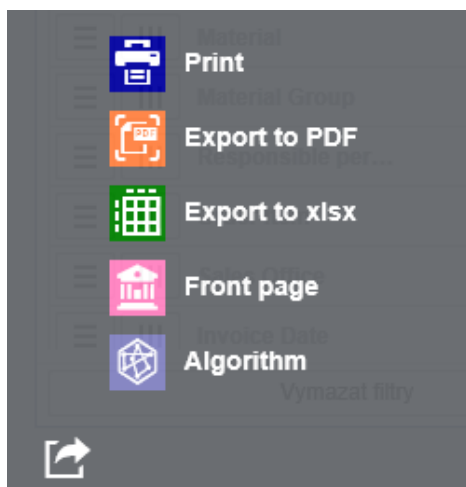
Po kliknutí na tlačítko Enter na úvodní obrazovce je uživateli představen výstup, který znázorňuje obrázek číslo 26. Větší, pravá část reportu je oblastí pro zobrazení dat manažerské tabulky. V jeho horní části se nachází záložky – Final realization (finální realizace) a Channels (kanály). Finální realizace odpovídá jednoduchému výstupu, který vznikl při ručním zpracování manažerské tabulky v excelu. Navíc je k tomu přidána tabulka s rozdělením prodejů dle měny. Každá tabulka je napojena na svou query a v hlavičce má název výstupu a měsíc, za který je realizována. Záložka finální realizace obsahuje tři tabulky – Final realization (**query\_1**), Product group breakdown (**query\_2**) a Invoice value in original currency (**query\_3**). Druhá záložka, kanály, zahrnuje dvě tabulky napojené na **query\_5** a **query\_6** – Aggregation by channels.

Levý panel reportu slouží koncovému uživateli k tomu, aby si mohl modifikovat výstup. V jeho horní části je znovu možnost měnit výběr měsíce fakturace. Pod tím jsou dvě dlaždice – Aggregation a Complete table. První je zapnuta defaultně. Při výběru druhé se na celé oblasti pro data zobrazí velká tabulka neagregovaných dat, která je napojena na **query\_4**. Pod dlaždicemi se nachází filtrovací panel. Ten funguje nezávisle pro každou dlaždici a záložku. Pomocí filtru je možné omezit hodnoty některé z charakteristik, což ovlivní všechny tabulky v dané záložce. Navíc pro **query\_1** a **query\_4** lze pomocí tohoto panelu měnit rozložení řádků a sloupců (a přidávat charakteristiky a skryté ukazatele) odpovídajících tabulek.



Obrázek 26 – Manažerská tabulka, základní pohled  
Zdroj: interní zdroje společnosti.

V levé dolní části celého reportu je tlačítko, které zapíná zobrazení menu panelu (viz obrázek číslo 27). V něm má uživatel na výběr vytisknout nebo převést do formátu pdf současný report. Dále také může exportovat všechny tabulky při zachování stávajícího filtrování do formátu xlsx. Menu nabízí možnost vrátit se na úvodní obrazovku nebo také zobrazit algoritmus přiřazování *obchodní jednotky a názvu skupiny materiálů* (obdobný tomu, jaký je popsán v kapitole Algoritmus plnění charakteristik).



Obrázek 27 – Manažerská tabulka, menu  
Zdroj: interní zdroje společnosti.

### 3.3.6 Přístup uživatelů do reportu Manažerská tabulka

Uživatelé mohou do reportu přistoupit pomocí odkazu, který byl umístěn do vstupní obrazovky MIS ABC. Ta původně obsahovala skupiny reportů Pipeline, Mastertable a Forecasting. Verze doplněná i o Manažerskou tabulku se nachází na obrázku číslo 28.



Obrázek 28 – Nová vstupní obrazovka MIS ABC

Zdroj: interní zdroje společnosti.

Aby se uživatel dostal do této vstupní obrazovky, musí k tomu mít přidělen přístup. Všechny výstupy z BW jsou umístěny na portál na firemním intranetu. Přidělování odkazů uživatelům na jednotlivé reporty je v oprávnění správců těchto reportů. Ne však každý, kdo měl přístup do předchozích třech reportů, automaticky uvidí data v Manažerské tabulce. Aby se uživateli zobrazila data v tabulkách reportu, musí k tomu mít přidělené oprávnění na dané queries. Tato oprávnění se přidělují přímo v systému SAP BW a jsou odlišná od přidělování přístupů k odkazům na report. Tímto postupem je lze např. zajistit, aby zaměstnanci controllingu měli přístupnou pouze skupinu reportů ManTab, zaměstnanci obchodního oddělení pouze skupinu reportů Pipeline a vrcholoví manažeři společnosti ABC mohli přistupovat ke všem čtyřem skupinám reportů.

### 3.3.7 Produktivní využívání reportu

Po dokončení automatické Manažerské tabulky v BW byla stále ještě několik měsíců současně tvořena i původním způsobem. Během této doby byly podrobně sledovány odchylky ve výstupech mezi těmito dvěma způsoby. Jedná se o poměrně časově náročnou část, která vyžaduje mnoho kontrolních tabulek a analýz, které pátrají po příčinách, proč se některé údaje liší. Ve většině případů se jedná o nesrovnalosti při zadání od zákazníka,

kteře byly následně v algoritmu v BW upraveny. Chybovost vzniká také při zadávání údajů do zdrojového systému obchodními zástupci. Při ručním zpracování Manažerské tabulky si těchto chyb zodpovědná osoba všimne a opraví. Nicméně při automatickém zpracování tyto opravy možné nejsou a z toho důvodu mohou být některá data v koncovém výstupu přiřazena nesprávně. Automatizovaná manažerská tabulka naopak vede k tomu, aby oddělení controllingu více dbalo na to, aby byla data do zdrojového systému zadávána správně.

I přes výše zmíněné překážky dosahoval report uspokojivých výsledků o pár měsíců dříve, než bylo očekáváno a nyní je již 8 měsíců produktivně využíván. Největší výhodou oproti původnímu zpracování je automatizace procesu a s tím spojené ušetření časových kapacit odborných zaměstnanců společnosti ABC. Další nespornou výhodou automatizované Manažerské tabulky je její aktuálnost. Původně byl výstup tvořen pouze jedenkrát měsíčně. Do reportu z BW jsou nahrávána aktuální data každý den a je možné výstup sledovat nejen za předchozí měsíc, ale například jen za poslední týden nebo za poslední půl rok, apod. Lze také jednoduše filtrovat hodnoty výstupů dle jednotlivých charakteristik, což tvoří vynikající podklad uživatelům pro dodatečné analýzy. Ty pak slouží jako podklad pro tvorbu strategií celé společnosti. Vše je také dostupné na jediném místě na portálu firemního intranetu, kde k němu mají všechny oprávněné osoby jednoduchý přístup.

## 4. Aplikace statistických metod v MIS ABC

Tato kapitola je věnována aplikaci vybraných statistických metod pro predikci úspěšnosti projektů. Tyto predikce jsou založeny na datech, které se nacházejí v MIS ABC. První podkapitola představuje využití statistické metody. Následně je podrobněji představeno hodnocení projektů a jsou vytvořeny různé modely na základě těchto hodnocení. Tyto modely jsou mezi sebou porovnány pomocí relevantních ukazatelů. Poslední část této kapitoly vysvětluje možnost aplikace výsledků do datového toku pro MIS ABC.

### 4.1 Využití statistické metody

V této kapitole jsou popsány vybrané statistické metody, které jsou následně využity při samotné analýze dat obsažených v MIS ABC.

#### $\chi^2$ test nezávislosti v kontingenční tabulce

Kontingenční tabulka zpravidla znázorňuje vztah mezi dvěma kategoriálními znaky. Jedná se o speciální typ dvourozměrné tabulky (viz tabulka 2). Obsahuje sdružené rozdělení četností pro dvě proměnné, kde pozorované absolutní četnosti jsou značeny  $n_{ij}$ . Ty představují kombinaci kategorií daných dvou znaků. Okrajové (marginální) četnosti se značí  $n_{i\cdot}$  a  $n_{\cdot j}$  a představují součty jednotlivých řádků a sloupců. Z marginálních četností jsou odvozovány četnosti odpovídající nezávislosti sledovaných znaků. Označují se jako očekávané četnosti a pro jejich výpočet je využíván vzorec (1) (Wegner, 2013).

**Tabulka 2:** Kontingenční tabulka

$x_i$ ( $i = 1, 2, \dots, r$ )	$y_j$ ( $j = 1, 2, \dots, s$ )	$y_1$	$y_2$	...	$y_s$	Součty četností $n_{i\cdot}$
$x_1$		$n_{11}$	$n_{12}$	...	$n_{1s}$	$n_{1\cdot}$
$x_2$		$n_{21}$	$n_{22}$	...	$n_{2s}$	$n_{2\cdot}$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_r$		$n_{r1}$	$n_{r2}$	...	$n_{rs}$	$n_{r\cdot}$
Součty četností $n_{\cdot j}$		$n_{\cdot 1}$	$n_{\cdot 2}$	...	$n_{\cdot s}$	$n$

Zdroj: vlastní zpracování.

$$n'_{ij} = \frac{n_{i.} \times n_{.j}}{n} \quad (1)$$

Zda jsou znaky v kontingenční tabulce nezávislé, je testováno tzv.  $\chi^2$  (chí-kvadrát) testem nezávislosti. Tento test je založený na porovnávání očekávaných (teoretických) četností s pozorovanými (empirickými). Předpokladem pro test je, že všechny očekávané četnosti nabývají alespoň hodnoty 5. Výhodiskem pro test je nulová a alternativní hypotéza:

$H_0$ : proměnné  $x$  a  $y$  jsou nezávislé;

$H_1$ : non  $H_0$ .

Hypotéza  $H_0$  je testována pomocí testové statistiky  $G$ . Ta má při platnosti nulové hypotézy chí-kvadrát rozdělení ( $\chi^2$ ). Počet stupňů volnosti rozdělení se řídí počtem řádků  $r$  a počtem sloupců  $s$  v kontingenční tabulce. Její výpočet zobrazuje vzorec (2). Vyšší podobnost očekávaných a pozorovaných četností svědčí ve prospěch testované hypotézy  $H_0$ . Vypočtená hodnota  $G$  je následně porovnána s kritickým oborem  $W$  dle vzorce (3). Pokud spadá hodnota testové statistiky  $G$  do kritického oboru, je  $H_0$  zamítnuto a  $H_1$  přijato. V tom případě se prokáže hypotéza o závislosti proměnných  $x$  a  $y$  (Hendl, 2012).

$$G = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}} \sim \chi^2[(r-1)(s-1)] \quad (2)$$

$$W \equiv \{G: G \geq \chi^2_{1-\alpha}[(r-1)(s-1)]\} \quad (3)$$

## Test lineárního trendu v kontingenční tabulce

Pokud je proměnná  $x$  nebo  $y$  ordinální, chí-kvadrát test nezávislosti tyto informace o pořadí ignoruje. V tomto případě je vhodnější využití testových statistik, které tyto proměnné posuzují jako kvantitativní, nikoliv kvalitativní. Pokud jsou proměnné ordinální, je běžné přiřazení trendu. Jak se zvyšuje úroveň  $x$ , tak má  $y$  tendenci se buď snižovat, nebo zvyšovat (Agresti, 2007).

Během testu lineárního trendu v kontingenční tabulce nejprve jednoduchá analýza přiřadí skóre jednotlivým kategoriím a změří stupeň lineárního trendu. Testová statistika, která je citlivá na pozitivní a negativní lineární trendy, využívá informace o korelacích v datech.



Korelace mezi  $y$  a  $x$  se vypočítá dle vzorce (4). Přičemž  $u_i$  označuje skóre pro jednotlivé řádky a  $v_j$  pro jednotlivé sloupce.

$$r = \frac{\sum_{i,j}(u_i - \bar{u})(v_j - \bar{v})p_{ij}}{\sqrt{[\sum_i(u_i - \bar{u})^2 p_{i+}][\sum_j(v_j - \bar{v})^2 p_{+j}]}} \quad (4)$$

Korelace nabývá hodnot od -1 do 1.  $H_0$  a  $H_1$  zůstávají stejné jako u předchozího testu. Zde se však testuje pomocí testové statistiky  $M^2$  – vzorec (5). Tato testová statistika se zvyšuje s rostoucím  $r$  a se zvyšujícím se rozsahem vzorku  $n$ . Pro velké  $n$  má  $M^2$  přibližně chí-kvadrát rozdělení s jedním stupněm volnosti. Vysoké hodnoty tedy svědčí proti nezávislosti proměnných.  $M^2$  (stejně jako  $G^2$ ) nerozlišuje vysvětlované a vysvětlující proměnné. Prohozením řádků a sloupců test dojde ke stejným výsledkům (Agresti, 2007).

$$M^2 = (n - 1)r^2 \quad (5)$$

## Model logistické regrese

Závislá proměnná často nabývá jen dvou hodnot (ano, ne). Výsledkem binární proměnné  $Y$  může v tom případě být pouze 1 (úspěch), nebo 0 (neúspěch). Distribuční rozdělení  $Y$  je specifikováno pravděpodobností  $P(Y = 1) = \pi$  (úspěchu) a  $P(Y = 0) = (1 - \pi)$  (neúspěchu). Hodnota  $\pi$  se mění tím, jak se mění hodnota vysvětlujících proměnných  $x$ . Pro popsání závislosti na dané proměnné  $x$  se  $\pi$  nahrazuje za  $\pi(x)$ .

Vztah mezi  $\pi(x)$  a  $x$  často nebývá lineární. Fixní změna  $x$  může mít menší vliv když je  $\pi$  blíže 0 nebo 1, než když je uprostřed svého rozsahu. Právě pro modelování pravděpodobnosti poměru pozitivních výsledků v závislosti na hodnotě proměnné se využívá logistická regrese. Logistická regresní funkce znázorňuje vzorec (6), po úpravě vzorec (7). Výraz  $\log(\pi(x) / (1 - \pi(x)))$  se nazývá logit. Hodnoty  $\alpha$  a  $\beta$  jsou regresní koeficienty a k jejich odhadu  $a, b$  je použita iterativní metoda nejmenších čtverců.

$$\pi(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} \quad (6)$$

$$\log \frac{\pi(x)}{1 - \pi(x)} = \alpha + \beta x \quad (7)$$

Tento model je možné využít pro predikci pravděpodobnosti při nastavených hodnotách  $x$ . Vysvětlující proměnná  $x$  může být kvantitativní, ale i kvalitativní. V případě, že je i vícerozměrná, má odpovídající model pak tvar dle vzorce (8). Pomocí daného modelu lze pak předpovědět pravděpodobnost nastání jevu při nových hodnotách nezávisle proměnných.

$$\pi(x) = \frac{\exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m)}{1 + \exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m)} \quad (8)$$

Výraz  $(\pi(x) / (1 - \pi(x)))$  se označuje jako odds ratio (podíl šancí) udává skutečně podíl šancí výskytu pro dvě různé hodnoty nezávisle proměnných. Vzorec (7) je možná dále upravit na vzorec (9). Tento exponenciální vztah poskytuje interpretaci pro  $\beta$ . Šance se znásobí  $e^\beta$  krát za každý nárůst  $x$  o jednotku. Čím vyšší hodnota odds ratio, tím větší vliv má změna dané vysvětlující proměnné na celkový model.

$$\frac{\pi(x)}{1 - \pi(x)} = \exp(\alpha + \beta x) = e^\alpha (e^\beta)^x \quad (9)$$

Pro zvážení vhodnosti zařazení jednotlivých vysvětlujících proměnných do modelu je využíván test poměrem věrohodností (likelihood-ratio test). Test porovnává věrohodnost kompletního modelu ( $L_1$ ) s modelem, ze kterého byla daná vysvětlující proměnná vyřazena ( $L_0, \beta_i = 0$ ). Testová statistika  $-2(L_0 - L_1)$  má také u velkého počtu  $n$  chí-kvadrát rozdělení s jedním stupněm volnosti. Malé hodnoty p-value naznačují, že model byl významně vylepšen zařazením odpovídající proměnné. Vysoké hodnoty naopak naznačují, že by z modelu daná proměnná mohla být vypuštěna.

## 4.2 Predikce úspěšnosti projektů

V kapitole 2.2.1 byla představena skupina reportů Pipeline. Ta zkoumá příležitosti a projekty z různých úhlů pohledu. Poskytuje zákazníkovi přehled o obchodních příležitostech a běžících projektech z pohledu množství, peněz, času a úspěšnosti získávání. Každý rok se společnosti naskytne množství příležitostí v řádu tisíců. Pro management firmy je podstatné znát odhad množství projektů, které firma získá. Díky tomu mohou sledovat například plnění plánu pro daný fiskální rok.

Prodejní cyklus je určen datem zahájení a předpokládaným datem ukončení příležitosti. Prodejní cyklus produktu nebo služby začíná rozpoznáním příležitosti, tzn. stanovením

možné šance prodeje. Proces končí zakázkou nebo odmítnutím ze strany zákazníka. Projekty a příležitosti jsou rozlišeny dle fáze, v které se nachází. Informaci o tom poskytuje charakteristika s názvem *Status projektu ABC*. Status projektů může nabývat následujících hodnot:

- 21 – Open (otevřené příležitosti),
- 22 – InProcess (právě zpracovávané příležitosti),
- 23 – Won (vyhrané projekty),
- 24 – Lost (ztracené příležitosti/projekty),
- 25 – Completed (dokončené projekty),
- 26 – Shipped Out (dodané projekty),
- 27 – Confirmed (potvrzené projekty),
- 28 – Installed (instalované projekty).

V MIS ABC je toto dělení zjednodušeno tak, že projekty 25, 26 a 28 tvoří společně kategorii Realized (realizované). Pro potřeby statistické analýzy v této diplomové práci je potřeba pouze vědět, zda projekt byl úspěšný, nebo ne. Mezi úspěšné se řadí všechny Won (23) a Realized (25, 26, 28). Neúspěšné jsou pouze ty se statusem projektu ABC Lost (24). U projektů Confirmed (27) stále není jisté, zda se stanou úspěšnými. Proto jsou tyto projekty spolu s příležitostmi Open a InProcess (21, 22) z analýzy vyloučeny.

Výběr pro statistickou analýzu je omezen datem založení příležitosti. Jsou zahrnuty pouze projekty a příležitosti založené ve fiskálním roce 2019 a 2020 (období od 1. 4. 2018 do 31. 3. 2020). Dohromady to je 2901 příležitostí a projektů, z toho 577 úspěšných, 1448 neúspěšných a 876 nezařazených. Výběrový soubor pro statistickou analýzu obsahuje tedy 2025 jednotek.

#### **4.2.1 Rating příležitostí**

Skupina reportů Pipeline dělí projekty a příležitosti pomocí charakteristiky *Rating příležitostí* do čtyř skupin – A, B, C, D (kde A je skupinou s nejslibnějšími příležitostmi a projekty). Do těchto skupin probíhá řazení na základě ukazatele *Overall project rating*, kterému bude věnována následující kapitola.

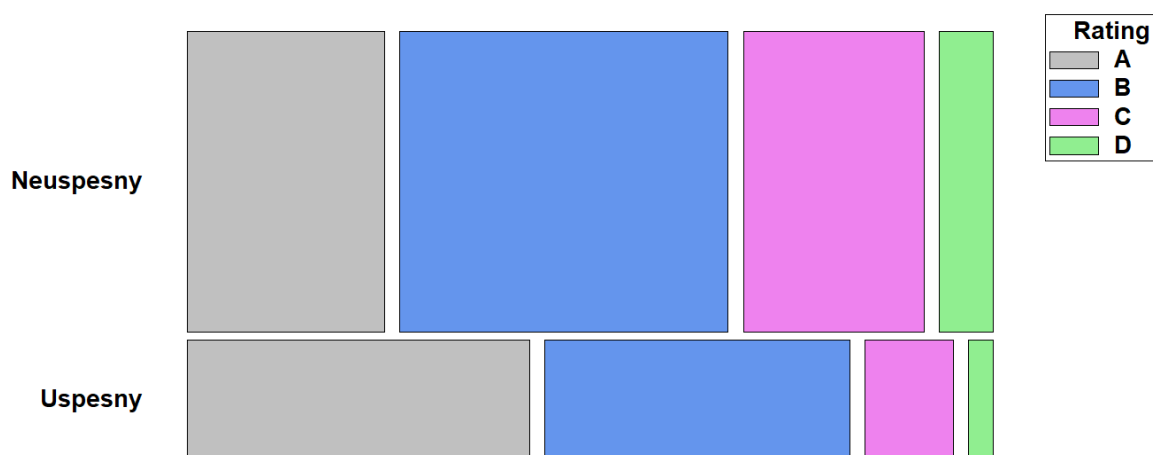
Tabulka 3 zobrazuje rozdělení zkoumaných 2025 jednotek do jednotlivých skupin, kde první řádek obsahuje pozorované četnosti a druhý procentuální podíl ze součtu řádku. Lze

vidět, že nejvíce je projektů s ratingem B (42 % z celkového počtu), následně pak projekty s ratingem A a C. Jen 6 % příležitostí má rating D. Již na první pohled je patrná závislost ratingu na tom, zda bude projekt úspěšný. Například rating A tvoří 45 % úspěšných projektů a jen 26 % neúspěšných. Naopak rating D jen 3 % úspěšných, ale 7 % neúspěšných. Jen u ratingu B není patrné, že by zde byl rozdíl v úspěšnosti. Graficky tuto skutečnost znázorňuje mozaikový graf na obrázku číslo 29.

**Tabulka 3: Kontingenční tabulka (úspěšnost a rating příležitosti)**

	A	B	C	D	Celkem
<b>Neúspěšný</b>	376	625	344	103	1448
	25,97%	43,16%	23,76%	7,11%	71,51%
<b>Úspěšný</b>	260	231	67	19	577
	45,06%	40,03%	11,61%	3,29%	28,49%
<b>Celkem</b>	636	856	411	122	2025
	31,41%	42,27%	20,30%	6,02%	100,00%

Zdroj: vlastní zpracování dle interních zdrojů společnosti.



*Obrázek 29 – Úspěšnost příležitostí dle ratingu*

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

Závislost potvrzuje test lineárního trendu (jeho podrobné výsledky se nachází v tabulce A1 přílohy A). Hodnota P-Value značí, že na více než 99% hladině významnosti může být tento test nezávislosti zamítnut. Pozorovaná hodnota úspěšnosti je tedy závislá na hodnotě ratingu. Korelační koeficient nabývá hodnoty -0,208. To značí, že čím nižší rating příležitost má, tím je větší pravděpodobnosti, že bude neúspěšná. Nicméně tato hodnota poukazuje na spíše slabší korelaci.

## 4.2.2 Overall project rating

*Overall project rating* je ukazatel, který má každá příležitost a projekt. Nabývá hodnoty od 25 do 100 (v %). Tento ukazatel se nachází v reportingu pro společnost ABC, nicméně defaultně je zobrazována pouze charakteristika *Rating příležitosti*, která je odvozena dle následujícího klíče:

- Overall project rating od 25 do 44 – Rating D,
- Overall project rating od 45 do 58 – Rating C,
- Overall project rating od 59 do 70 – Rating B,
- Overall project rating od 71 do 100 – Rating A.

Tabulka 4 obsahuje vybrané statistické ukazatele Overall project ratingu při rozděleném základním souboru na neúspěšné a úspěšné projekty. Průměr je vyšší u úspěšných projektů o přibližně 6 bodů. Směrodatná odchylka je obdobná pro obě kategorie – v průměru se od průměru hodnoty liší o 13 bodů. Graficky tuto situaci zobrazuje krabicový graf na obrázku číslo 30. Je na něm vidět, že jak průměr, tak medián Overall project ratingu je vyšší u úspěšných projektů. Přibližně v místě, kde mají neúspěšné projekty průměr, se nachází první kvartil úspěšných projektů.

**Tabulka 4:** *Overall project rating (charakteristika)*

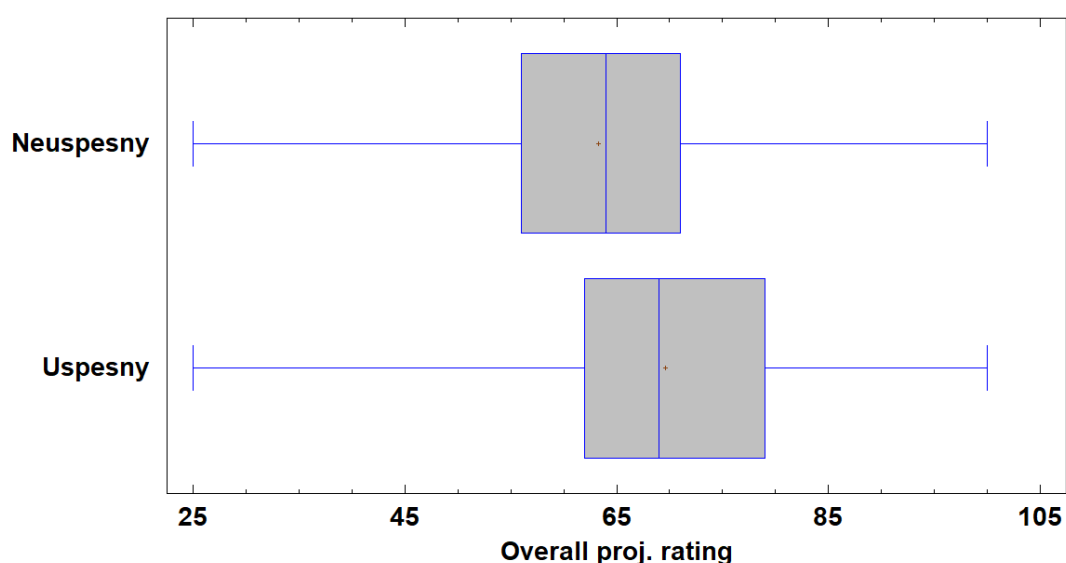
	Četnost	Průměr	Směrodatná odchylka	Variační koeficient
<b>Neúspěšný</b>	1448	63,27	12,96	20,49%
<b>Úspěšný</b>	577	69,6	12,72	18,27%
<b>Celkem</b>	2025	65,08	13,2	20,29%

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

Samy o sobě tyto srovnání důkaz o závislosti neposkytují. Na data byla z toho důvodu aplikována logistická regrese, kde úspěšnost projektu je závislou proměnnou (0 neúspěšný, 1 úspěšný) a Overall project rating kvantitativní nezávisle proměnnou. Podrobné údaje o tomto logistickém modelu se nachází v příloze A (tabulka A2).

Na základě P-Value analýzy rozptylu tohoto modelu lze tvrdit, že existuje statisticky významný vztah mezi proměnnými. Procento rozptylu vysvětlené modelem je v hodnotě 4 %. Zároveň test dobré shody potvrzuje vhodnost modelu na základě pozorovaných a očekávaných hodnot. Odhadovaný Odds ratio (podíl šancí) je roven 1,04. Znamená to, že odhadovaná pravděpodobnost úspěšnosti projektu se násobí 1,04 krát za každý nárůst

hodnoty Overall project rating o jeden bod (to je 4% nárůst). Tabulka 5 ukazuje inverzní předpovědi získané modelem proloženým daty. Tyto inverzní předpovědi indikují hodnotu Overall project ratingu při které model dosáhne určité procentuální pravděpodobnosti. Na hladině významnosti 95 % má příležitost, jejíž hodnota Overall proj. rating je v rozmezí 85,5 až 96,3, pouze 50% šanci úspěchu. Je to dáno především tím, že primární účel tohoto ukazatele je identifikovat pro společnost ABC nejzajímavější příležitosti, nikoliv ty s největší šancí na úspěch. Na Overall proj. rating se tedy nelze dívat jako na procentuální šanci na úspěch daného projektu. Pro vytvoření modelu, který by tuto pravděpodobnost lépe postihoval, je vhodné se blíže podívat na jednotlivé faktory ovlivňující rating.



Obrázek 30 – Krabicový graf (Overall proj. rating a úspěšnost)

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

Tabulka 5: Inverzní předpověď pro Overall project rating

Percent	Overall proj. rating	Dolní mez	Horní mez
5,0	14,93	0,9392	24,25
10,0	33,97	24,75	40,18
15,0	45,75	39,38	50,17
20,0	54,62	50,21	57,87
25,0	61,95	58,83	64,56
30,0	68,35	65,8	70,95
35,0	74,17	71,53	77,36
40,0	79,61	76,51	83,74
45,0	84,82	81,1	90,05
50,0	89,93	85,5	96,33
55,0	95,05	89,85	102,7

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

### 4.2.3 Faktory Overall project ratingu

Hodnota ukazatele Overall project rating vzniká výpočtem z hodnot sedmi různých faktorů, které nabývají hodnot od 1 do 4. Jsou to faktory, které blíže specifikují danou příležitost. Při zadávání příležitosti do systému SAP zodpovědná osoba vyplní také hodnoty ke všem 7 faktorům. Z nich je automaticky vypočítán Overall project rating. Pokud má například všech 7 faktorů hodnotu 4, je Overall project rating 100. Naopak pokud mají hodnotu 1, je rating 25. Jednotlivé faktory a hodnoty, kterých mohou nabývat, jsou podrobněji představeny níže.

#### Relationship with decision maker (vztah k rozhodovateli)

- 1 – Žádný vztah s rozhodovatelem nebo není známo, kdo to je, nebo neexistuje způsob, jak ho kontaktovat,
- 2 – Rozhodovatel zná značku společnosti ABC,
- 3 – Rozhodovatel zná značku společnosti ABC a osobně člena obchodního týmu,
- 4 – Máme skvělý vztah s rozhodovatelem (vždy to je osoba, nikoliv společnost).

#### Design specs (konstrukční specifikace)

- 1 – Název společnosti ABC se na specifikacích vůbec neobjevuje,
- 2 – Na specifikacích je uveden pouze název společnosti ABC,
- 3 – Výkresy společnosti ABC jsou součástí specifikace,
- 4 – Společnost ABC je specifikována spolu s návrhy vyvinutými ve firmě ABC.

#### Support of the designer (podpora projektanta)

- 1 – Projektant podporuje konkurenty společnosti ABC,
- 2 – V nabídce není známa nebo deklarovaná podpora designu žádnému dodavateli,
- 3 – Projektant zná společnost ABC, líbí se mu její značka a předpokládá se, že by ji mohl upřednostnit.
- 4 – Projektant otevřeně prohlašuje, že v nabídce podpoří společnost ABC.

#### Commission (provize)

- 1 – Projektant podporuje konkurenty společnosti ABC,
- 4 – Jakákoliv zúčastněná strana v průběhu vývoje projektu dostává provizi od prodejní kanceláře společnosti ABC.

### **Budget fit (rozpočet)**

- 1 – Žádné informace o rozpočtu zákazníka,
- 2 – Umístění areálu je prestižní,
- 3 – Za projektem stojí silný investor nebo se projekt váže k luxusnímu hotelu,
- 4 – Rozpočet na projekt je vysoký, klient utratí peníze za danou oblast / místo.

### **Success rate with the client (úspěšnost s klientem)**

- 1 – Společnost ABC neměla v minulosti s klientem žádné projekty,
- 2 – Ve vzdálené minulosti proběhlo pár realizovaných projektů,
- 3 – Klient dobře zná a oceňuje výkon společnosti ABC.
- 4 – Bohatá historie s klientem, společnost ABC je preferovaným dodavatelem.

### **Competition (konkurence)**

- 1 – Žádné informace o konkurenci,
- 2 – Konkuruje jen čínské společnosti,
- 3 – Konkuruje společnosti značek nižší kvality a druhořadé české společnosti.
- 4 – Buď na projektu pracuje jen společnost ABC, nebo jakákoliv společnost ze specifického seznamu je také zapojena.

Každý z faktorů je ordinální proměnnou a očekává se, že čím vyšší je jeho hodnota, tím zajímavější by příležitost pro společnost měla být. Některé z faktorů zároveň mohou naznačovat, zda příležitost bude úspěšná. Pro každý faktor zvlášť je sestavena tabulka četností, která jednotlivé hodnoty faktoru rozděluje na úspěšné a neúspěšné pro zkoumaný výběrový soubor 2025 příležitostí a projektů. Všechny tyto tabulky spolu s testy lineárního trendu s vypočítanými skóry pro každý faktor se nachází v příloze A (tabulky A3 až A9).

U všech faktorů je na 95% hladině významnosti pomocí testu lineárního trendu zamítnuta nezávislost úspěšnosti na daném faktoru. Nicméně u faktoru **Competition** to je pouze na 95,3% hladině významnosti. Koeficient korelace s úspěšností k tomu má hodnotu pouze 0,044. Nízkou korelací s úspěšností disponuje také faktor **Commision**. Navíc k tomu je záporný (-0,0579), což značí, že když není provize od společnosti ABC v projektu zahrnuta, má vyšší šanci na úspěšnost. U ostatních faktorů se korelační koeficient pohybuje v rozmezí od 0,10 do 0,17.



Následně byl aplikován model logistické regrese, nejprve za využití všech faktorů. Výběrovým souborem v tomto případě bylo pouze 75 % (1525 příležitostí) náhodně vybraných položek z původního zkoumaného souboru. Zbýlých 25 % (500 příležitostí) bude sloužit k porovnání úspěšnosti jednotlivých modelů na datech, které nesloužili k sestavení těchto modelů.

Odhad parametrů (provedený metodou maximální věrohodnosti) tohoto modelu se nachází v tabulce A10 v příloze A. U parametrů **Success rate with client**, **Commission**, **Competition** a **Support of the designer** se odhady pohybují na hranici nebo pod hodnotou směrodatné odchylky pro danou hodnotu parametru. Je to jeden z indikátorů, že model by mohl být přesnější vypuštěním některé z těchto faktorů. Upravené procento rozptylu vysvětlené tímto modelem je 5 %. Ačkoliv test nezávislosti potvrzuje vhodnost tohoto modelu testy hypotéz o parametrech (na základě testu poměrem věrohodností) naznačují, že stejné čtyři parametry by mohly být z modelu vyloučeny (viz tabulka 6).

**Tabulka 6:** Testy poměrem věrohodností pro faktory

Faktor	Chí-kvadrát	Df	P-Value
Succ. rate with clnt	2,879	3	0,4107
Budget fit	21,48	3	0,0001
Commission	0,5331	1	0,4653
Competition	4,112	3	0,2496
Design specs	24,67	3	0,0000
Supp. of the design	2,525	3	0,4708
Rel. with decision	21,07	3	0,0001

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

Vzhledem k testům lineárních trendů jednotlivých faktorů, zmíněných výše, byly z modelu nejprve vyloučeny faktory **Commission** a **Competition**. Zároveň bylo zjištěno, že korelační koeficient naznačuje vzájemnou korelaci mezi hodnotami 1, 2 a 3 pro **Success rate with client** a mezi hodnotami 2 a 3 pro **Supp. of the design** (viz příloha A, tabulka A11). Z těchto 3, respektive 2 hodnot, byla vytvořena pouze 1 skupina. U faktoru **Supp. of the design** je model vhodnější při použití pouze dvou skupin. Naopak u **Succ. rate with clnt** tato změna model zhoršuje. Bylo tedy uvažováno, zda tento faktor ponechat, nebo také vyloučit. Bez tohoto faktoru model dosahuje horších výsledků u testu nezávislosti. Při porovnání upraveného procenta rozptylu vysvětlené modelem jsou výsledky srovnatelné (5,28 % oproti 5,41 %). Odhadnutý regresní model metodou maximální věrohodnosti se nachází v tabulce 7. Tento model bude využitý pro porovnání s původním modelem založeným pouze na ukazateli Overall project rating.

**Tabulka 7: Regresní model (5 faktorů)**

<b>Parametr</b>	<b>Odhad</b>	<b>Směrodatná odchylka</b>	<b>Odhadnuté Odds Ratio</b>
Konstanta	0,493	0,1783	
Design specs=1	-0,6853	0,1757	0,504
Design specs=2	-0,5809	0,1829	0,5594
Design specs=3	-0,04316	0,1628	0,9578
Rel. with decision=1	-0,9159	0,3087	0,4002
Rel. with decision=2	-0,7975	0,1958	0,4505
Rel. with decision=3	-0,4483	0,1506	0,6387
Supp. of the design=1	0,1688	0,3341	1,184
Supp. of the design=2&3	-0,1687	0,1492	0,8447
Budget fit=1	-0,916	0,1948	0,4001
Budget fit=2	-0,5444	0,2017	0,5802
Budget fit=3	-0,3197	0,1974	0,7263
Succ. rate with clnt=1	-0,2841	0,2151	0,7527
Succ. rate with clnt=2	-0,05709	0,2028	0,9445
Succ. rate with clnt=3	0,07083	0,1649	1,073

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

#### 4.2.4 Porovnání modelů

Aby bylo modely možné porovnat, i model logistické regrese pouze s faktorem Overall project rating musí být vytvořen pouze na omezeném základním souboru 75 % příležitostí. Podrobné informace o tomto modelu se také nacházejí v příloze A (tabulka A12, A13). Upravené procento rozptylu vysvětlené modelem je přibližně o dva procentní body nižší u modelu jen s Overall project ratingem. Jak již bylo zmíněno v předchozí kapitole, u modelu založeného na jednotlivých faktorech, P-Value u testu poměrem věrohodností u dvou faktorů přesahují hranici 0,05, což upozorňuje na jejich možnou nevhodnost zařazení. U druhého modelu tento test pro jediný kvantitativní faktor (Overall project rating) potvrzuje jeho vhodnost.

Oba modely jsou považovány za vhodné na základě chí-kvadrát testu nezávislosti očekávaných a pozorovaných hodnot. Bližší pohled na úspěšnost odhadu modelů pro omezený základní soubor nabízí tabulka 8. Pro vybrané mezní hodnoty vyčísluje správnost predikce úspěšných a neúspěšných projektů pro model A (založený na 5 faktorech) a model B (založený jen na Overall project ratingu). Mezní hodnota (cutoff) představuje hranici predikované procentuální úspěšnosti příležitosti, při které bude daná příležitost považována za úspěšnou. Například pokud by byla u modelu A zvolena mezní hodnota 0,1, tak všechny příležitosti, kterým model predikoval 10% a vyšší pravděpodobnost úspěchu, by byly považovány za úspěšné. V tom případě by model

správně predikoval 98,68 % skutečně úspěšných projektů a pouze 3,36 % neúspěšných. Celkem by správně predikoval jen 31,67 % projektů. Výběr mezní hodnoty je subjektivní, pro porovnání se však nabízí prostřední hodnota 0,5. Model A je při této mezní hodnotě přesnější u predikce úspěšných projektů než model B (18,32 % oproti 10,82 %). U neúspěšných projektů jsou výsledky predikce srovnatelné (94,87 % oproti 96,27 %). Nejen pro mezní hranici 0,5, ale pro jakoukoliv zvolenou, je model A přesnější (nebo stejně přesný) jako model B co se celkového procenta korektně predikovaných výsledků projektů týče.

**Tabulka 8:** Porovnání úspěšnosti odhadů modely A a B

Mezní hodnota	Úspěšné		Neúspěšné		Celkem	
	model A	model B	model A	model B	model A	model B
0	100	100	0	0	29,7	29,7
0,1	98,68	99,12	3,36	2,43	31,67	31,15
0,2	85,65	90,07	29,76	21,92	46,36	42,16
0,3	62,03	58,5	65,49	62,59	64,46	61,38
0,4	36,42	30,24	86,57	85,35	71,67	68,98
0,45	26,05	17,66	92,26	92,54	72,59	70,3
<b>0,5</b>	<b>18,32</b>	<b>10,82</b>	<b>94,87</b>	<b>96,27</b>	<b>72,13</b>	<b>70,89</b>
0,55	11,04	2,21	97,01	99,16	71,48	70,36
0,6	7,95	0,66	97,95	99,81	71,21	70,36
0,7	0	0	100	100	70,3	70,3
0,8	0	0	100	100	70,3	70,3
0,9	0	0	100	100	70,3	70,3
1	0	0	100	100	70,3	70,3

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

Další porovnání modelů je provedeno jejich aplikací na zbylých 25 % příležitostech, které nebyly pro sestavení modelů využity. Z těchto 500 příležitostí jich ve skutečnosti bylo 124 úspěšných a 376 neúspěšných, což je 24,8% úspěšnost. Průměr pravděpodobnosti úspěšnosti je při aplikaci modelu A 29,43 % a modelu B 28,98 %. Na základě tohoto indikátoru nelze potvrdit vhodnost jednoho modelu nad druhým.

Klasifikační tabulka (viz tabulka 9) vyčísluje správně odhadnuté úspěšné a neúspěšné projekty pro oba modely. Zvolený mezní bod je zde 0,5. Model A správně klasifikoval 76,6 % příležitostí (383 z 500) a model B pouze 74,2 % (371 z 500). Model B také dokázal správně identifikovat jen 8 úspěšných projektů. Oproti tomu druhý model jich identifikoval 28. Lépe toto vystihuje citlivost, která měří podíl skutečných pozitiv (úspěšných projektů), které jsou správně označeny jako takové. Pro model A tato hodnota činí 22,5 %, pro model

B pouze 6,9 %. Naopak specifičnost, která měří podíl skutečných negativ, které jsou správně označeny jako takové, lépe vychází pro model B (96,5 % oproti 94,4 % u modelu A). Pro praktické využití je vedle celkové úspěšnosti predikce modelu lepší, aby model uměl spíše identifikovat příležitosti, které budou úspěšné, než ty neúspěšné. Z těchto důvodů je model A vhodnější. Nejen že pro něj lépe vychází předchozí výsledky, ale také je jednodušší v tom smyslu, že využívá jen 5 faktorů, kdežto model B využívá Overall project rating, který se skládá ze 7 faktorů.

**Tabulka 9:** Klasifikační tabulka (predikce modely A a B, mezní hodnota 0,5)

Skutečný status	Predikce modelem A		Predikce modelem B		Celkem
	Úspěšný	Neúspěšný	Úspěšný	Neúspěšný	
Úspěšný	28	96	8	116	<b>124</b>
Neúspěšný	21	355	13	363	<b>376</b>

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

#### 4.2.5 Modely s přidávanými proměnnými

Pro sestavení dalších modelů byly uvažovány i další faktory, u které by na úspěšnost projektu mohly mít vliv. Jedná se o charakteristiky a ukazatele, které se již v datovém toku MIS ABC nacházejí. Mezi charakteristiky, u kterých jednotlivé testy lineárního trendu potvrzují vzájemnou závislost s úspěšností projektu, patří *Prodejní jednotka*, *Produktový program* (8 kategorií, do kterých se příležitosti člení) a *Kalendářní měsíc očekávané doby dokončení příležitosti*. Dále v úvahu připadá začlenění ukazatele *Očekávaný obrat*. Jedná se o částku, kterou obchodní zástupce očekává při úspěšném dokončení projektu.

Z těchto nových i všech předchozích proměnných se hledá co nejvhodnější kombinace, která co nejlépe vysvětluje pravděpodobnost úspěšnosti projektu. Nejprve je využita metoda postupné regrese, která do modelu postupně přidává a odebírá proměnné na základě jednotlivých testů poměrem věrohodnosti. Uvažovaná hladina významnosti pro začlenění a vyčlenění proměnné z modelu je stanovena na 99 %.

Výsledkem je model C (viz tabulka A14 v příloze A). Tento model obsahuje charakteristiky *Design specs*, *Relationship with decision maker*, *Budget fit*, *Prodejní jednotka*, *Kalendářní měsíc* a ukazatel *Očekávaný obrat*. Oproti modelu A postrádá charakteristiky *Support of the designer* a *Success rate with client*. Upravené procento rozptylu vysvětlené modelem C je 9,23 %, což je zlepšení o 4 procentní body oproti modelu A. Při aplikaci modelu na zkoumaných datech (75 % základního souboru)

a zvolené mezní hodnotě 0,5 klasifikuje správně 74,56 % příležitostí. Jak si model C vede při predikci úspěšnosti u zbylých 25 % příležitostí, zobrazuje tabulka 10. Model C správně dokázal klasifikovat 75,2 % příležitostí, což je horší výsledek než u modelu předchozího. Citlivost je však na úrovni 27,42 % (o 5 procentních bodů lepší výsledek). Specifičnost má hodnotu 90,96 %.

**Tabulka 10:** Klasifikační tabulka (predikce modelem C, mezní hodnota 0,5)

Skutečný status	Predikce modelem C		Celkem
	Úspěšný	Neúspěšný	
Úspěšný	34	90	<b>124</b>
Neúspěšný	34	342	<b>376</b>

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

Ručně byly také sestavovány modely s různými kombinacemi výše uvedených proměnných. Většina z nich nedosahovala kvalit modelu C ani modelu A. Veškeré vhodné modely však obsahují alespoň tři z původních sedmi faktorů, a to konkrétně *Design specs*, *Relationship with decision maker* a *Budget fit*. Tyto charakteristiky jsou tím pádem dobrým indikátorem toho, zda příležitost bude úspěšná.

Za zmínku stojí ještě model, který k modelu D přidává charakteristiky *Support of the designer* a *Produktový program*. Má srovnatelnou hodnotu upraveného procenta rozptylu vysvětlené modelem (9,05 %). Při aplikaci na zbylých 25 % příležitostí klasifikuje tento model správně 76 % příležitostí. Zároveň jeho citlivost je rovna 29 %. Nicméně tyto menší zlepšení jsou na úkor toho, že je model s přidávanými charakteristikami složitější. Pro praktické využití není tento model tolik vhodný.

### 4.3 Aplikace výsledků do datového toku

Vybraný model lze následně aplikovat do datového toku v SAP BW. Nejprve je zapotřebí vytvořit nový ukazatel – *Pravděpodobnost úspěchu projektu*, který bude uchovávat procentuální hodnoty úspěšnosti pro jednotlivé projekty a příležitosti. Tento ukazatel by byl zaveden do Infokosky Projekty ABC a Multiprovideru MIS ABC.

K plnění ukazatele by docházelo v transformaci mezi DSO Příležitosti ABC a již zmiňovanou Infokostkou. Vstupními hodnotami by byly jednotlivé kategoriální a kvantitativní proměnné, které model obsahuje. Například u modelu A by to bylo 5 charakteristik: *Design specs*, *Relationship with decision maker*, *Budget fit*, *Support*

*of the designer* a *Success rate with client*. Využito by bylo vzorce (8), kde jednotlivé  $x_i$  představují řádky z tabulky 7. Hodnotami  $\beta_i$  jsou odhady jednotlivých parametrů modelu. Výsledkem je hodnota pravděpodobnosti, která indikuje, zda příležitost bude úspěšná, či nikoliv.

Hodnoty nově vzniklé charakteristiky by byly zobrazovány ve skupině reportů Pipeline. Zde by manažer, který do reportu nahlíží, mohl vidět nejen atraktivitu projektu (na základě hodnoty *Overall project ratingu*), ale také jeho šanci na úspěch (na základě hodnoty *Pravděpodobnosti úspěchu projektu*). Kombinace těchto dvou ukazatelů napomůže rozhodování o tom, do jaké příležitosti a projektu věnovat omezené zdroje společnosti a kterým příležitostem naopak nedávat tak vysokou prioritu. Zároveň by report také umožnil porovnání průměrných hodnot těchto ukazatelů pro skupiny projektů (filtrovaných dle různých charakteristik).

Na závěr je nutné dodat, že všechny modely byly sestaveny na základě dat z posledních dvou fiskálních let. Pro co nejpřesnější výsledky a aktuálnost nového ukazatele je vhodné model stejným postupem sestavovat opakovaně v pravidelných intervalech (například jednou za rok) na základě aktuálních dat.

## Závěr

Cílem této diplomové práce bylo v první řadě zpracování manažerského informačního systému společnosti ABC. Nejprve je zhodnocen současný stav tohoto MIS. Práce podává informace o třech skupinách reportů, které MIS obsahoval. Na základě požadavku zákazníka byl rozšířen o další skupinu reportů – tzv. Manažerskou tabulku. Jedná se o nástroj sledování a hodnocení plnění plánu tržeb v konsolidaci. Práce podrobně popisuje postup, dle kterého oddělení controllingu společnosti ABC tuto tabulku manuálně každý měsíc sestavovalo. V návaznosti na to práce řeší problematiku automatizaci tohoto postupu a realizaci Manažerské tabulky v rámci systému BW. Snahou bylo přiblížit algoritmy plnění vybraných charakteristik a ukazatelů. Dále také přehledně znázornit datový tok pro Manažerskou tabulku v systému BW, a to od zdrojových systémů až po konečný infoobjekt. V posledním kroku byla zpracována podoba koncového reportu. Ten data ze systému BW získává pomocí předem vytvořených queries.

Největší výhodou oproti původnímu zpracování je automatizace procesu a s tím spojené ušetření časových kapacit odborných zaměstnanců společnosti ABC. Další nespornou výhodou automatizované Manažerské tabulky je její aktuálnost. Původně byl výstup tvořen pouze jedenkrát měsíčně. Do reportu z BW jsou nahrávána aktuální data každý den a je možné výstup sledovat nejen za předchozí měsíc, ale například jen za poslední týden nebo za poslední půl rok. Lze také jednoduše filtrovat hodnoty výstupů dle jednotlivých charakteristik, což tvoří vynikající podklad uživatelům pro dodatečné analýzy. Ty pak slouží jako podklad pro tvorbu strategií celé společnosti.

Dalším cílem bylo vytvoření nástroje, který predikuje procentuální šanci na úspěch projektů a příležitostí společnosti ABC. Toho bylo dosaženo prostřednictvím statistické analýzy dat obsažených v MIS ABC. Ten původně obsahoval ukazatel Overall project rating, jehož primární účel je identifikovat pro společnost ABC nejzajímavější příležitosti, nikoliv ty s největší šancí na úspěch. Pomocí chí-kvadrát testů nezávislosti a testů lineárního trendu v kontingenčních tabulkách byly analyzovány jednotlivé faktory tohoto ukazatele. Na základě těchto faktorů a dalších charakteristik bylo vytvořeno několik modelů logistické regrese. Při porovnání s modelem založeným pouze na původním ukazateli, mají tyto modely mnohem lepší hodnotu citlivosti. Například u modelu C je to nárůst ze 7 % na 27 %. To značí, že oproti modelu, který byl založen pouze na ukazateli Overall project rating, dokáží tyto další modely přesněji identifikovat úspěšné

projekty. Zlepšení je možné také pozorovat u hodnoty upraveného procenta rozptylu vysvětleného modelem, který poukazuje na jeho kvalitu. Zde se jedná o nárůst až o 4 procentní body. Poslední část práce nastiňuje možnosti aplikace těchto výsledků do datového toku v BW. Začlenění modelů do manažerského informačního systému ABC by napomohlo manažerům v rozhodování o tom, do jaké příležitosti a projektu vložit omezené zdroje společnosti a kterým příležitostem naopak nedávat tak vysokou prioritu.



# Seznam použité literatury

## Citace

AGRESTI, Alan. 2007. *An Introduction to Categorical Data Analysis*. 2nd ed. New Jersey: JohnWiley & Sons, Inc. ISBN 978-0-471-22618-5.

HENDL, Jan. 2012. *Přehled statistických metod: analýza a metaanalýza dat*. 4. vyd. Praha: Portál. ISBN 978-80-262-0200-4.

INMON, W. H. 2002. *Building the Data Warehouse*. 3rd ed. Chichester: Wiley. ISBN 978-0-764-59944-6.

KIMBALL, Ralph a Margy ROSS. 2013. *The data warehouse toolkit: The definitive guide to dimensional modeling*. 3rd ed. Indianapolis: John Wiley & Sons. ISBN: 978-1-118-53080-1.

LUHN, H. P. 1984. A Business Intelligence System. *IBM Journal of Research and Development* [online]. Issue 4 [cit. 2020-07-21]. DOI: 10.1147/rd.24.0314. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5392644>.

NOVOTNÝ, Ota, Jan POUR a David SLÁNSKÝ. 2005. *Business Intelligence: jak využít bohatství ve vašich datech*. 1. vyd. Praha: Grada. ISBN 978-80-2476-685-0.

POUR, Jan, Miloš MARYŠKA a Ota NOVOTNÝ. 2012. *Business intelligence v podnikové praxi*. 1. vyd. Praha: Professional Publishing. ISBN 978-80-7431-065-2.

SKLENÁK, Vilém, 2001. *Data, informace, znalosti a Internet*. 1. vyd. Praha: C. H. Beck. ISBN 80-7179-409-0.

VERCELLIS, Carlo. 2009. *Business intelligence: data mining and optimization for decision making*. 1st ed. Chichester: Wiley. ISBN 978-0-470-51139-8.

WEGNER, Trevor. 2013. *Applied Business Statistics: Methods and Excel-based Applications*. 3rd ed. South Africa: Juta and Company. ISBN 978-0-7021-7774-3.

## Bibliografie

PROQUEST. 2019. Databáze článků ProQuest [online]. Ann Arbor, MI, USA: ProQuest. [cit. 2019-10-12]. Dostupné z <http://knihovna.tul.cz/>.

SYNEK, Miloslav. 2011. *Manažerská ekonomika*. 5. vydání. Praha: Grada. ISBN 978-80-247-3494-1.

# Seznam příloh

**Příloha A:** Statistické výpočty (7 stran)

## Příloha A: Statistické výpočty

**Tabulka A11: Test lineárního trendu (Rating příležitosti)**

Test	Korelace	P-Value	
Test lineárního trendu	-0,2083	0,0000	
Řádek	Skóre	Sloupec	Skóre
Neúspěšný	724,5	A	318,5
Úspěšný	1737,0	B	1064,5
		C	1698,0
		D	1964,5

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A12: Model logistické regrese (Overall project rating, 100 % dat)**

Odhadnutý regresní model (Maximální věrohodnost)						
Parametr	Odhad	Směrodatná odchylka	Odhadnuté Odds Ratio			
Konstanta	-3,531	0,2841				
Overall proj. rating	0,03926	0,004124	1,04			
Analýza rozptylu						
Zdroj	Rozptyl	Df	P-Value			
Model	99,06	1	0,0000			
Rezidua	2321,	2023	0,0000			
Celkem (kor.)	2420,	2024				
Procento rozptylu vysvětlené modelem = 4,093						
Upravené procento = 3,928						
Test poměrem věrohodností						
Faktor	Chi-kvadrát	Df	P-Value			
Overall proj. rating	99,06	1	0,0000			
Chi-kvadrát test dobré shody						
Třída	Interval logit	n	Úspěšný Pozorováno	Úspěšný Očekáváno	Neúspěšný Pozorováno	Neúspěšný Očekáváno
1	méně než -1,371	406	63,0	63,12	343,0	342,9
2	-1,371 to -1,057	486	113,0	114,7	373,0	371,3
3	-1,057 to -0,8611	366	92,0	102,8	274,0	263,2
4	-0,8611 to -0,547	376	134,0	124,7	242,0	251,3
5	-0,547 a více	391	175,0	171,6	216,0	219,4
Celkem		2025	577,0		1448,	
Chi-kvadrát = 2,78 s 3 stupni volnosti P-value = 0,4268						

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A13: Četnosti a test lineárního trendu (Rel. with decision)**

	1	2	3	4	Celkem
Neúspěšný	129	354	563	402	1448
	8,91%	24,45%	38,88%	27,76%	71,51%
Úspěšný	24	86	212	255	577
	4,16%	14,90%	36,74%	44,19%	28,49%
Celkem	153	440	775	657	2025

	7,56%	21,73%	38,27%	32,44%	100,00%
<b>Test nezávislosti</b>					
<i>Test</i>	<i>Korelace</i>		<i>P-Value</i>		
Test lineárního trendu	0,1780		0,0000		
<i>Řádek</i>	<i>Skóre</i>		<i>Sloupec</i>		<i>Skóre</i>
Neúspěšný	724,5		1		77,0
Úspěšný	1737,0		2		373,5
			3		981,0
			4		1697,0

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A14:** Četnosti a test lineárního trendu (Design specs)

	1	2	3	4	Celkem
Neúspěšný	403	368	304	373	1448
	27,83%	25,41%	20,99%	25,76%	71,51%
Úspěšný	117	89	147	224	577
	20,28%	15,42%	25,48%	38,82%	28,49%
Celkem	520	457	451	597	2025
	25,68%	22,57%	22,27%	29,48%	100,00%
<b>Test nezávislosti</b>					
<i>Test</i>	<i>Korelace</i>		<i>P-Value</i>		
Test lineárního trendu	0,1477		0,0000		
<i>Řádek</i>	<i>Skóre</i>		<i>Sloupec</i>		<i>Skóre</i>
Neúspěšný	724,5		1		260,5
Úspěšný	1737,0		2		749,0
			3		1203,0
			4		1727,0

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A15:** Četnosti a test lineárního trendu (Supp. of the design)

	1	2	3	4	Celkem
Neúspěšný	64	279	759	346	1448
	4,42%	19,27%	52,42%	23,90%	71,51%
Úspěšný	25	86	255	211	577
	4,33%	14,90%	44,19%	36,57%	28,49%
Celkem	89	365	1014	557	2025
	4,40%	18,02%	50,07%	27,51%	100,00%
<b>Test nezávislosti</b>					
<i>Test</i>	<i>Korelace</i>		<i>P-Value</i>		
Test lineárního trendu	0,1094		0,0000		
<i>Řádek</i>	<i>Skóre</i>		<i>Sloupec</i>		<i>Skóre</i>
Neúspěšný	724,5		1		45,0
Úspěšný	1737,0		2		272,0
			3		961,5
			4		1747,0

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A16: Četnosti a test lineárního trendu (Commission)**

	1	4	Celkem
Neúspěšný	1208	240	1448
	83,43%	16,57%	71,51%
Úspěšný	508	69	577
	88,04%	11,96%	28,49%
Celkem	1716	309	2025
	84,74%	15,26%	100,00%
Test nezávislosti			
Test	Korelace	P-Value	
Test lineárního trendu	-0,0579	0,0091	
Řádek	Skóre	Sloupec	Skóre
Neúspěšný	724,5	1	858,5
Úspěšný	1737,0	4	1871,0

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A17: Četnosti a test lineárního trendu (Budget fit)**

	1	2	3	4	Celkem
Neúspěšný	560	398	357	133	1448
	38,67%	27,49%	24,65%	9,19%	71,51%
Úspěšný	150	138	177	112	577
	26,00%	23,92%	30,68%	19,41%	28,49%
Celkem	710	536	534	245	2025
	35,06%	26,47%	26,37%	12,10%	100,00%
Test nezávislosti					
Test	Korelace		P-Value		
Test lineárního trendu	0,1658		0,0000		
Řádek	Skóre	Sloupec		Skóre	
Neúspěšný	724,5	1		355,5	
Úspěšný	1737,0	2		978,5	
		3		1513,5	
		4		1903,0	

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A18: Četnosti a test lineárního trendu (Succ. rate with clnt)**

	1	2	3	4	Celkem
Neúspěšný	375	313	469	291	1448
	25,90%	21,62%	32,39%	20,10%	71,51%
Úspěšný	87	95	209	186	577
	15,08%	16,46%	36,22%	32,24%	28,49%
Celkem	462	408	678	477	2025
	22,81%	20,15%	33,48%	23,56%	100,00%
Test nezávislosti					
Test	Korelace		P-Value		
Test lineárního trendu	0,1633		0,0000		
Řádek	Skóre	Sloupec		Skóre	
Neúspěšný	724,5	1		231,5	
Úspěšný	1737,0	2		666,5	

		3	1209,5
		4	1787,0

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A19: Četnosti a test lineárního trendu (Competition)**

	1	2	3	4	Celkem
Neúspěšný	606	150	362	330	1448
	41,85%	10,36%	25,00%	22,79%	71,51%
Úspěšný	216	59	153	149	577
	37,44%	10,23%	26,52%	25,82%	28,49%
Celkem	822	209	515	479	2025
	40,59%	10,32%	25,43%	23,65%	100,00%
Test nezávislosti					
Test	Korelace		P-Value		
Test lineárního trendu	0,0441		0,0470		
Řádek	Skóre		Sloupec	Skóre	
Neúspěšný	724,5		1	411,5	
Úspěšný	1737,0		2	927,0	
			3	1289,0	
			4	1786,0	

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A20: Model logistické regrese (7 Faktorů, 75 % dat)**

Odhadnutý regresní model (Maximální věrohodnost)			
Parametr	Odhad	Směrodatná odchylka	Odhadnuté Odds Ratio
Konstanta	0,2931	0,2441	
Succ. rate with clnt=1	-0,2427	0,2178	0,7845
Succ. rate with clnt=2	-0,01378	0,2046	0,9863
Succ. rate with clnt=3	0,07818	0,1656	1,081
Budget fit=1	-0,9583	0,2141	0,3835
Budget fit=2	-0,5827	0,2073	0,5584
Budget fit=3	-0,4175	0,2048	0,6587
Commission=1	0,1306	0,1798	1,14
Competition=1	0,1748	0,1741	1,191
Competition=2	0,1599	0,2325	1,173
Competition=3	0,3573	0,1776	1,429
Design specs=1	-0,7355	0,1794	0,4793
Design specs=2	-0,6305	0,186	0,5323
Design specs=3	-0,08761	0,1654	0,9161
Supp. of the design=1	0,1698	0,3382	1,185
Supp. of the design=2	-0,1017	0,2041	0,9033
Supp. of the design=3	-0,201	0,1545	0,8179
Rel. with decision=1	-0,937	0,3107	0,3918
Rel. with decision=2	-0,8384	0,2033	0,4324
Rel. with decision=3	-0,48	0,1553	0,6188

<b>Analýza rozptylu</b>			
<i>Zdroj</i>	<i>Rozptyl</i>	<i>Df</i>	<i>P-Value</i>
Model	133,2	19	0,0000
Rezidua	1722,	1505	0,0001
Celkem (kor.)	1855,	1524	
Procento rozptylu vysvětlené modelem = 7,18			
Upravené procento = 5,024			
<b>Test poměrem věrohodností</b>			
<i>Faktor</i>	<i>Chi-kvadrát</i>	<i>Df</i>	<i>P-Value</i>
Succ. rate with clnt	2,879	3	0,4107
Budget fit	21,48	3	0,0001
Commission	0,5331	1	0,4653
Competition	4,112	3	0,2496
Design specs	24,67	3	0,0000
Supp. of the design	2,525	3	0,4708
Rel. with decision	21,07	3	0,0001
<b>Chi-kvadrát test dobré shody</b>			
Chi-kvadrát = 7,213 s 3 stupni volnosti P-value = 0,06541			

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A21:** Korelační koeficienty mezi vybranými faktory

	Succ. rate with clnt=1	Succ. rate with clnt=2	Succ. rate with clnt=3
Succ. rate with clnt=1	1,0000	0,5410	0,5326
Succ. rate with clnt=2	0,5410	1,0000	0,5622
Succ. rate with clnt=3	0,5326	0,5622	1,0000
	Supp. of the design=1	Supp. of the design=2	Supp. of the design=3
Supp. of the design=1	1,0000	0,2536	0,2874
Supp. of the design=2	0,2536	1,0000	0,5452
Supp. of the design=3	0,2874	0,5452	1,0000

Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A22:** Model logistické regrese (Model B – Overall project rating, 75 % dat)

<b>Odhadnutý regresní model (Maximální věrohodnost)</b>			
<i>Parametr</i>	<i>Odhad</i>	<i>Směrodatná odchylka</i>	<i>Odhadnuté Odds Ratio</i>
Konstanta	-3,636	0,3289	
Overall proj. rating	0,0416	0,004764	1,042
<b>Analýza rozptylu</b>			
<i>Zdroj</i>	<i>Rozptyl</i>	<i>Df</i>	<i>P-Value</i>
Model	83,92	1	0,0000
Rezidua	1772,	1523	0,0000
Celkem (kor.)	1855,	1524	
Procento rozptylu vysvětlené modelem = 4,523			
Upravené procento = 4,307			
<b>Test poměrem věrohodností</b>			
<i>Faktor</i>	<i>Chi-kvadrát</i>	<i>Df</i>	<i>P-Value</i>
Overall proj. rating	83,92	1	0,0000

<b>Chí-kvadrát test dobré shody</b>
Chí-kvadrát = 1,821 s 3 stupni volnosti P-value = 0,6104
Zdroj: vlastní zpracování dle interních zdrojů společnosti.

**Tabulka A23: Model logistické regrese (Model A – 5 faktorů, 75 % dat)**

<b>Odhadnutý regresní model (Maximální věrohodnost)</b>			
<i>Parametr</i>	<i>Odhad</i>	<i>Směrodatná odchylka</i>	<i>Odhadnuté Odds Ratio</i>
Konstanta	0,493	0,493	
Design specs=1	-0,6853	0,1757	0,504
Design specs=2	-0,5809	0,1829	0,5594
Design specs=3	-0,04316	0,1628	0,9578
Rel. with decision=1	-0,9159	0,3087	0,4002
Rel. with decision=2	-0,7975	0,1958	0,4505
Rel. with decision=3	-0,4483	0,1506	0,6387
Succ. rate with clnt=1	-0,2841	0,2151	0,7527
Succ. rate with clnt=2	-0,05709	0,2028	0,9445
Succ. rate with clnt=3	0,07083	0,1649	1,073
Supp. of the design_2=1,0	0,1688	0,3341	1,184
Supp. of the design_2=2&3	-0,1687	0,1492	0,8447
Budget fit=1	-0,916	0,1948	0,4001
Budget fit=2	-0,5444	0,2017	0,5802
Budget fit=3	-0,3197	0,1974	0,7263
<b>Analýza rozptylu</b>			
<i>Zdroj</i>	<i>Rozptyl</i>	<i>Df</i>	<i>P-Value</i>
Model	128,0	14	0,0000
Rezidua	1727,	1510	0,0001
Celkem (kor.)	1855,	1524	
Procento rozptylu vysvětlené modelem = 6,9			
Upravené procento = 5,283			
<b>Test poměrem věrohodností</b>			
<i>Faktor</i>	<i>Chí-kvadrát</i>	<i>Df</i>	<i>P-Value</i>
Budget fit	26,8	3	0,0000
Rel. with decision	20,61	3	0,0001
Design specs	23,37	3	0,0000
Supp. of the design_2	2,067	2	0,3558
Succ. rate with clnt	3,609	3	0,3069
<b>Chí-kvadrát test dobré shody</b>			
Chí-kvadrát = 3,904 s 3 stupni volnosti P-value = 0,272			
Zdroj: vlastní zpracování dle interních zdrojů společnosti.			

**Tabulka A24: Model logistické regrese (Model C, 75 % dat)**

<b>Odhadnutý regresní model (Maximální věrohodnost)</b>			
<i>Parametr</i>	<i>Odhad</i>	<i>Směrodatná odchylka</i>	<i>Odhadnuté Odds Ratio</i>
Konstanta	0,6341	0,4244	
Očekávaný obrat	-9,275E-8	2,638E-8	1,0
Design specs=1	-0,6718	0,1781	0,5108
Design specs=2	-0,58	0,1838	0,5599
Design specs=3	0,005785	0,1695	1,006



Rel. with decision=1	-1,17	0,3021	0,3102
Rel. with decision=2	-0,974	0,1855	0,3776
Rel. with decision=3	-0,5449	0,1466	0,5799
Budget fit=1	-1,058	0,214	0,3471
Budget fit=2	-0,727	0,2187	0,4834
Budget fit=3	-0,4868	0,2138	0,6146
Kalendářní měsíc=1	-0,3978	0,2832	0,6718
Kalendářní měsíc=10	-0,04933	0,2901	0,9519
Kalendářní měsíc=11	0,1413	0,2995	1,152
Kalendářní měsíc=12	0,301	0,3165	1,351
Kalendářní měsíc=2	-0,7268	0,2851	0,4835
Kalendářní měsíc=3	0,5433	0,3212	1,722
Kalendářní měsíc=4	-0,5076	0,3046	0,6019
Kalendářní měsíc=5	-0,4062	0,2815	0,6662
Kalendářní měsíc=6	1,037	0,2977	2,821
Kalendářní měsíc=7	-0,1499	0,2902	0,8608
Kalendářní měsíc=8	0,329	0,2917	1,39
Sales Unit=AFI1	0,9447	0,4616	2,572
Sales Unit=AFI2	-0,02199	0,3882	0,9783
Sales Unit=OTHER	0,1552	0,4897	1,168
Sales Unit=AFI3	0,4022	0,3817	1,495
Sales Unit=AFI4	0,06474	0,3884	1,067
Sales Unit=AFI5	-1,434	0,6082	0,2383
Sales Unit=AFI6	0,2132	0,3598	1,238
Sales Unit=AFI7	-0,308	0,4112	0,7349

#### Analýza rozptylu

<i>Zdroj</i>	<i>Rozptyl</i>	<i>Df</i>	<i>P-Value</i>
Model	232,3	29	0,0000
Rezidua	1623,	1495	0,0110
Celkem (kor.)	1855,	1524	

Procento rozptylu vysvětlené modelem = 12,52

Upravené procento = 9,288

#### Test poměrem věrohodnosti

<i>Faktor</i>	<i>Chi-kvadrát</i>	<i>Df</i>	<i>P-Value</i>
Očekávaný obrat	21,39	1	0,0000
Design specs	24,31	3	0,0000
Rel. with decision	37,55	3	0,0000
Budget fit	27,51	3	0,0000
Kalendářní měsíc	61,06	11	0,0000
Sales Unit	27,19	8	0,0007

#### Chi-kvadrát test dobré shody

Chi-kvadrát = 4,671 s 3 stupni volnosti P-value = 0,1975

Zdroj: vlastní zpracování dle interních zdrojů společnosti.