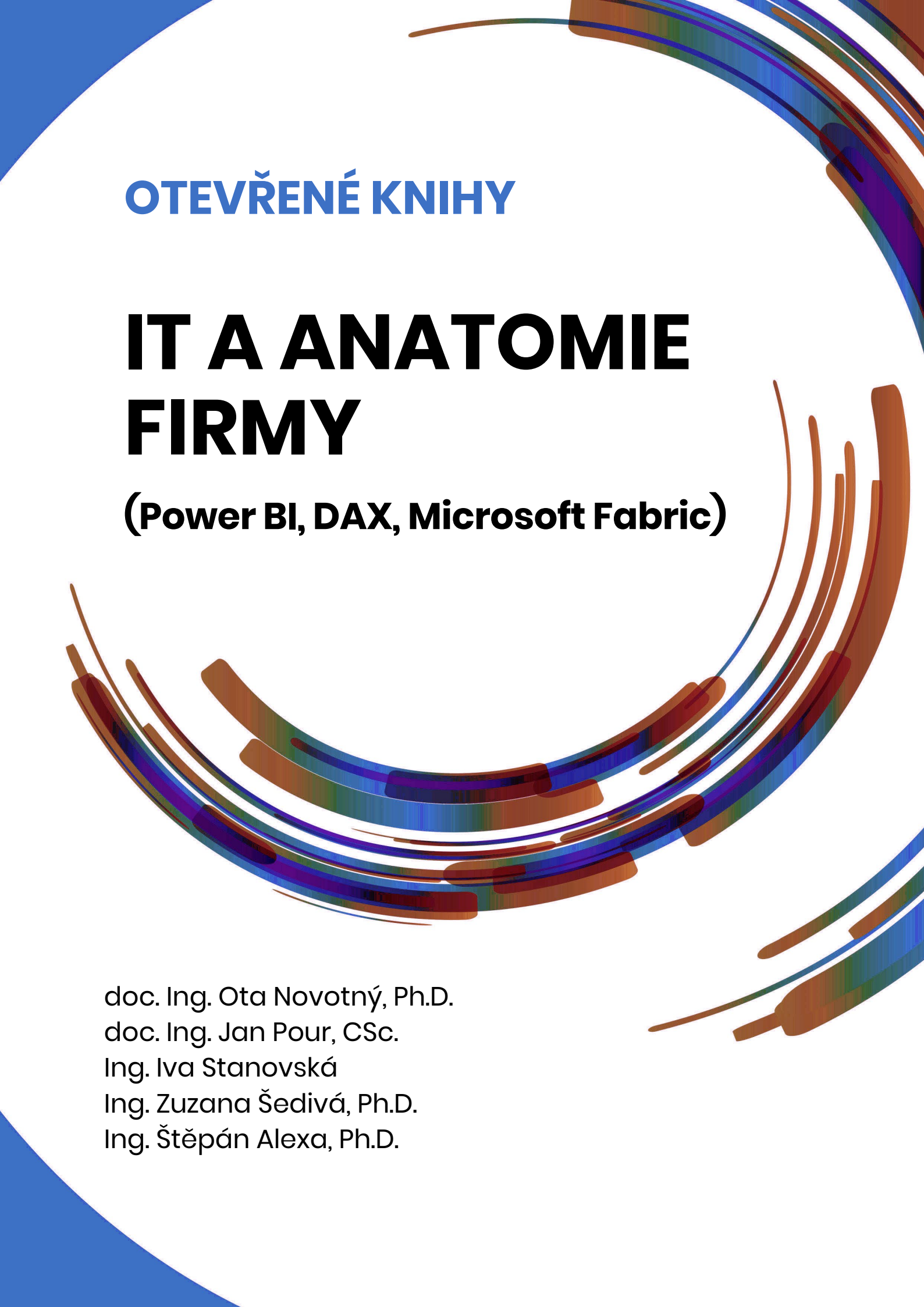




OTEVŘENÉ KNIHY

IT A ANATOMIE FIRMY

(Power BI, DAX, Microsoft Fabric)

doc. Ing. Ota Novotný, Ph.D.

doc. Ing. Jan Pour, CSc.

Ing. Iva Stanovská

Ing. Zuzana Šedivá, Ph.D.

Ing. Štěpán Alexa, Ph.D.

**doc. Ing. Ota Novotný, Ph.D.,
doc. Ing. Jan Pour, CSc., Ing. Iva Stanovská,
Ing. Zuzana Šedivá, Ph.D., Ing. Štěpán Alexa, Ph.D.**

IT A ANATOMIE FIRMY

(Power BI, DAX, Microsoft Fabric)

PROFESSIONAL PUBLISHING

**doc. Ing. Ota Novotný, Ph.D., doc. Ing. Jan Pour, CSc., Ing. Iva Stanovská,
Ing. Zuzana Šedivá, Ph.D., Ing. Štěpán Alexa, Ph.D.**

IT A ANATOMIE FIRMY (Power BI, DAX, Microsoft Fabric)

Kniha byla doporučena k vydání vědeckou radou nakladatelství.

© Autoři

Edition © Professional Publishing, s.r.o.

Obálka: Jan Mottl

První vydání, 2025

ISBN 978-80-88260-84-4



Publikace je zaměřená na **principy a možnosti využití** tří velmi rozšířených a široce využívaných produktů společnosti Microsoft pro **řešení aplikací v oblasti Self Service Business Intelligence**.

Text je rozdělen **do tří oddílů A – C**, každý se samostatným číslováním kapitol:

A. Power BI

B. DAX, Data Analytics Expressions

C. Microsoft Fabric



Poznámka: Pro velkou část příkladů a schémat v této publikaci **byly využity školicí materiály společnosti Seyfor** s laskavým svolením autorů těchto materiálů.

Text se zaměřuje především na **podstatu funkcí** a jejich využití, řešení dílčích a **specifických aspektů** funkcionality ponecháváme s ohledem na rychlý vývoj a časté změny v produktech **na firemních tutoriálech**.

Obsah

Vazby a souvislosti s dalšími publikacemi řady „IT a anatomie firmy“	10
Podniková analytika	10
Oblasti a komponenty řízení	10
A. Power BI.....	12
Úvod	15
1. Self Service Business intelligence, SSBI	16
1.1 Podstatné charakteristiky SSBI	16
1.2 Efekty SSBI.....	16
1.3 Omezení, problémy a předpoklady SSBI.....	17
1.4 Závěry k řešení Self-Service Business Intelligence.....	17
2. Základní principy Power BI.....	18
2.1 Instalace Power BI.....	20
2.2 Postup práce v Power BI	20
2.3 Vizualizace.....	21
2.4 Sestavy (reports)	21
2.5 Řídící panely (dashboardy).....	22
2.6 Dlaždice (tiles)	23
2.7 Pracovní závěry	23
3. Režimy práce s daty	24
3.1 Datový konektor vs. typ připojení.....	25
3.2 Živé připojení (Live connection).....	25
3.3 Přímé dotazování – DirectQuery	25
3.4 Import dat.....	26
3.5 Kompozitní režim (Composite Mode)	26
3.6 Režim Direct lake.....	26
3.7 Pracovní závěry	27
4. Řešení sémantických modelů	28
4.1 Sémantický model	28
4.2 Analytická příprava sémantického modelu	29
4.2.1 Určení podnikových procesů pro řešení	29
4.2.2 Určení granularity, resp. podrobnosti dat	29
4.2.3 Identifikace dimenzí	30
4.2.4 Definování faktů, resp. ukazatelů.....	30
4.3 Postup vytvoření sémantického modelu a aplikace v Power BI desktop.....	31

4.3.1	Zdroje dat pro vytvoření sémantického modelu v Power BI desktop	32
4.3.2	Import dat z datových zdrojů a jejich transformace.....	32
4.3.3	Definování relací mezi tabulkami	35
4.3.4	Rychlá míra.....	39
4.4	Pracovní závěry	40
5.	Reporty a vizuály v Power BI.....	41
5.1	Reporty v Power BI Desktop.....	42
5.2	Tvorba vizuálů v sestavách	42
5.2.1	Individuální hodnoty. Karta (Card)	44
5.2.2	KPI, Klíčový ukazatel výkonu.....	45
5.2.3	Měřidlo (Gauge).....	46
5.2.4	Tabulky, matice	46
5.2.5	Podmíněné formátování, datové pruhy.....	47
5.2.6	Nastavení ikon k datům	48
5.3	Filtrování, průřezy a drilování	49
5.3.1	Filtry (filters) v sestavách	49
5.3.2	Průřezy (slicers) v sestavách.....	50
5.3.3	Nastavení numerického průřezu	52
5.4	Podrobná analýza (Drill through).....	52
5.5	Analytické funkce v grafech Power BI	54
5.5.1	Interakce vizuálů, zvýrazňování prvků napříč vizuály na stránce sestavy.....	55
5.5.2	Procházení hierarchie – drill down.....	55
5.6	Pracovní závěry	56
6.	Power BI Service a mobilní aplikace	57
6.1	Publikace připravených reportů do služby Power BI	57
6.2	Vybrané funkce Power BI Service	58
6.3	Vytvoření a sdílení dashboardů	59
6.4	Použití datové brány pro obnovu dat ve službě Power BI	59
6.5	Mobilní aplikace Power BI	61
6.6	Pracovní závěry	62
7.	Příprava a plánování projektů v Power BI.....	63
7.1	Způsoby nasazení (deployment) Power BI.....	63
7.2	Role v řešení a užití projektu Power BI.....	65
7.2.1	Návrhář sémantických modelů (Semantic models designer).....	66
7.2.2	Tvůrce reportů (Report Author).....	66
7.2.3	Power BI administrátor.....	66
7.3	Otázky k řešení PBI projektu.	67
7.3.1	Vztah řešení PBI k byznysu	67

7.3.2	Datové zdroje	67
7.3.3	Obsah a distribuce reportů na bázi PBI	67
7.4	Poznámky k licencím	68
7.4.1	Sdílená kapacita (Shared capacity)	68
7.4.2	Vyhrazená, přiřazená kapacita (Dedicated capacity).....	69
7.5	Pracovní závěry	71
8.	Plánování a řešení reportů v Power BI	72
8.1	Proces plánování reportu.....	73
8.2	Doporučení pro vizualizace	74
8.3	Tabulky v reportech.....	75
8.4	Typy grafů a jejich účel	76
8.4.1	Sloupcový graf (Column Graph)	76
8.4.2	Výsečový graf (Pie Chart).....	77
8.5	Faktory ovlivňující srozumitelnost grafu.....	79
8.5.1	Nastavení poměru stran grafu	79
8.5.2	Měřítko osy y.....	79
8.5.3	Rozdílný rozsah hodnot na ose y u více datových řad	80
8.5.4	Popisky osy x	81
8.5.5	Popisky hodnot	81
8.5.6	Doprovodná grafika, značky, vysvětlivky.....	82
8.5.7	Písmo.....	82
8.6	Pracovní závěry	83
9.	Příprava datových zdrojů.....	85
9.1	Konektivita	85
9.2	Změny v dimenzích – SCD (Slowly Changing Dimensions).....	85
9.3	Tabulární model	88
9.3.1	Engine.....	89
9.3.2	Vertipaq engine	89
9.3.3	DirectQuery engine	91
9.3.4	Perspektivy	91
9.3.5	Security handling v tabulárním modelu.....	91
9.3.6	Databázové Role	91
9.3.7	ACL tabulky.....	92
9.4	Nastavení možností Power BI desktop.....	92
9.5	Pracovní závěry	93
10.	Transformace dat, jazyk M	94
10.1	Typy M dotazů	95
10.2	Parametry datového zdroje.....	95

10.3 „Staging“ dotazy.....	96
10.4 Dotazy pro fakta a dimenze.....	96
10.5 Dotazy k parametrům tabulek.....	96
10.6 Dotazy k zajištění bezpečnosti tabulek.....	97
10.7 Dotazy pro customizaci funkcí.....	97
10.8 Pracovní závěry.....	97
11. Závěr	98
12. Zdroje	99
B. DAX, Data Analytics Expressions	102
1. Základní principy a užití DAX.....	104
1.1 Výpočetní předpis v DAX.....	104
1.2 Vytvoření kalkulovaného sloupce	105
1.3 Vytvoření kalkulované míry.....	105
1.4 Vytvoření kalkulované tabulky	105
2. Konstrukce jazyka DAX.....	106
2.1 Datové typy.....	106
2.2 Operátory.....	106
2.3 Sloupce.....	107
2.4 Vazby	107
2.5 Hierarchie	107
2.6 Poznámky	107
2.7 Proměnné (Variables).....	107
3. Kontext vyhodnocování výpočetních předpisů	109
3.1 Filtr kontext	109
3.1.1 Implicitní filtr kontext	109
3.1.2 Explicitní filtr kontext	110
3.2 Řádkový kontext	110
4. Kalkulované sloupce a kalkulované míry	112
4.1 Vytvoření kalkulovaného sloupce a řádkový kontext	112
4.2 Příklady kalkulovaných sloupců.....	112
4.3 Vytváření měř a filtr kontext	113
4.4 Funkce CALCULATE().....	116
4.5 CALCULATE() v řádkovém kontextu	117
4.6 Parametr ALLSELECTED.....	118
4.7 Funkce IF ()	118
4.8 Iterátor	119

5. Řešení vazeb	120
5.1 Řešení ve standardních vazbách	120
5.1.1 Řádkový kontext s vazbami tabulek.....	120
5.1.2 Filtr kontext s vazbami tabulek.....	122
5.2 Parametrické nepropojené tabulky	123
6. Time Intelligence	125
6.1 Základní principy time intelligence	125
6.2 Vygenerování dimenzionální tabulky času	125
6.3 Ukazatele pro pracovní dny	126
6.4 Funkce time intelligence – YTD, QTD, MTD.....	127
6.5 Sledování hodnot za minulé roky.....	129
6.6 Klouzavé ukazatele.....	130
7. Závěr	132
8. Zdroje	133
C. Microsoft Fabric.....	135
1. Úvod	136
2. Microsoft Fabric	137
2.1 Scénáře využití Microsoft Fabric.....	137
2.2 Architektura a komponenty Microsoft Fabric	137
2.2.1 OneLake	137
2.2.2 Data Factory	139
2.2.3 Synapse Data Engineering	139
2.2.4 Synapse Data Science.....	140
2.2.5 Synapse Data Warehousing	140
2.2.6 Synapse Real-Time Analytics	140
2.2.7 Power BI a Direct lake	140
2.2.8 Data Activator.....	140
2.3 Další koncepty Microsoft Fabric.....	140
2.3.1 Shorcuts.....	140
2.3.2 Integrace umělé inteligence.....	140
2.3.3 Data Governance.....	141
2.4 Opodstatnění použití Microsoft Fabric.....	141
3. Zdroje	142

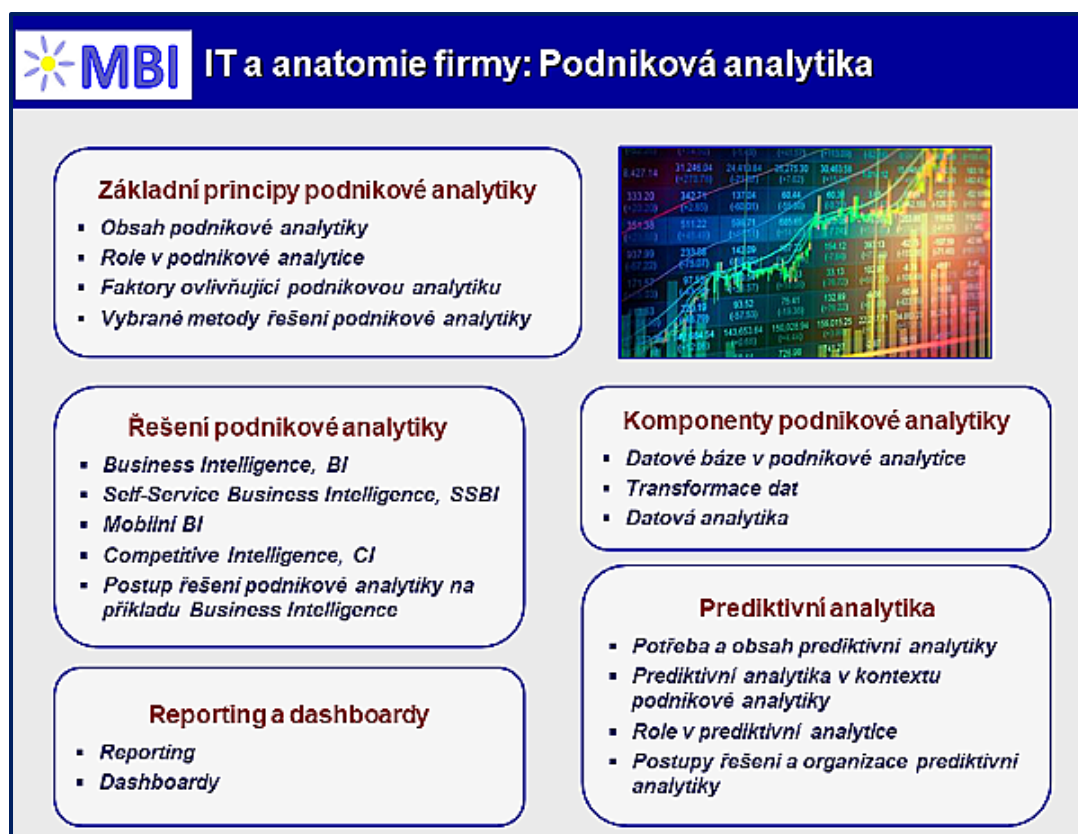
Vazby a souvislosti s dalšími publikacemi řady „IT a anatomie firmy“

S ohledem na značný rozsah funkcionality a dalších momentů, spojených s uvedenými produkty, jsou tyto funkce **obsahem i dalších publikací** uvedené řady, které s nimi souvisejí.

Podniková analytika

Publikace „Podniková analytika“ je **základní publikací** v oblasti analytiky na portálu MBI-AF, který se snaží poskytnout **celkový**, byť relativně stručný **přehled** o principech, postupech, produktech, problémech i řešeních podnikové analytiky v praxi. Zahrnuje otázky jak „**deskriptivní analytiky**“, postavené většinou na nástrojích a přístupech business intelligence, self service business intelligence nebo competitive intelligence, tak **pokročilé, zejména prediktivní analytiky**.

Odkaz: [\[Podniková analytika\]](#)

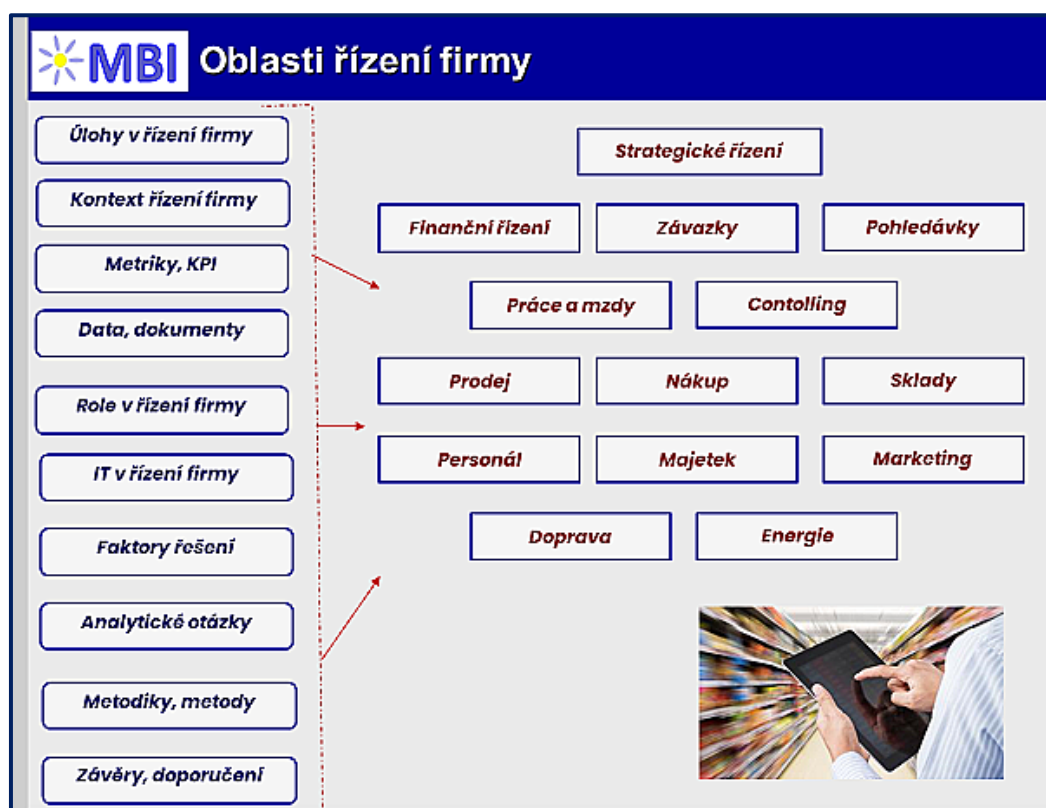


Podniková analytika, struktura publikace

Oblasti a komponenty řízení

Publikace vymezuje obsah řízení firmy z analytického pohledu, pohledu komponent řízení, tj. jednotlivých úloh řízení, datových zdrojů, metrik, IT aplikací a dalších. Jednou ze součástí úloh jsou i analytické úlohy, tedy prostor pro uplatnění Power BI a dalších produktů, jak ukazuje další obrázek.

Odkaz: [\[Oblasti a komponenty\]](#)



Oblasti a komponenty řízení, struktura textu

Další oddíly a kapitoly se již věnují třem výše uvedeným produktům.

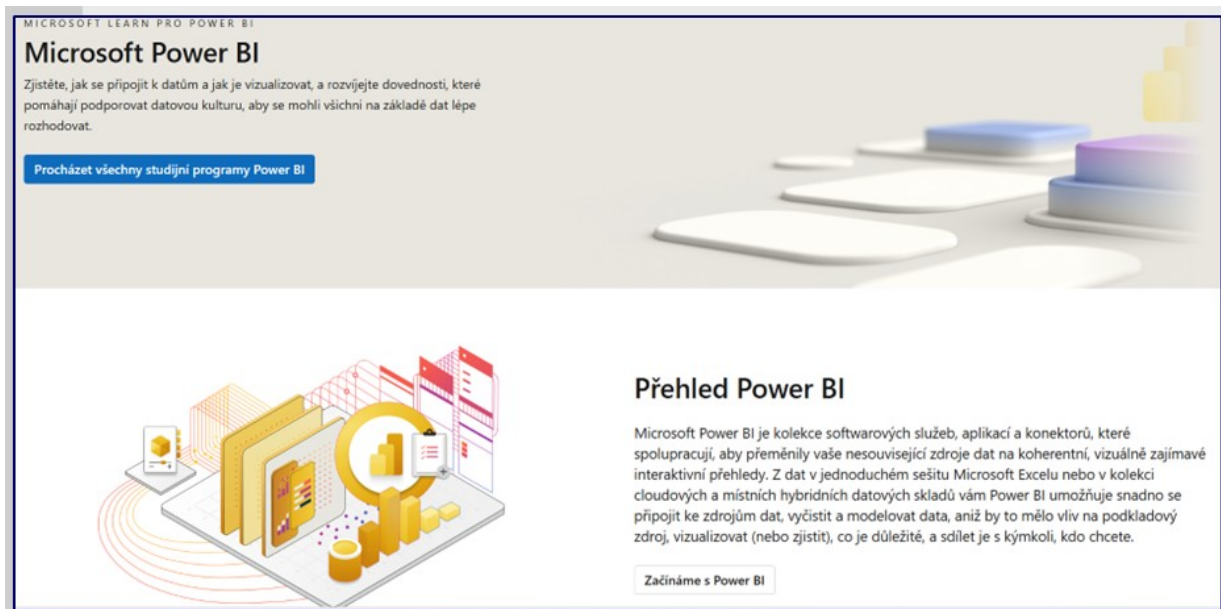
A. Power BI



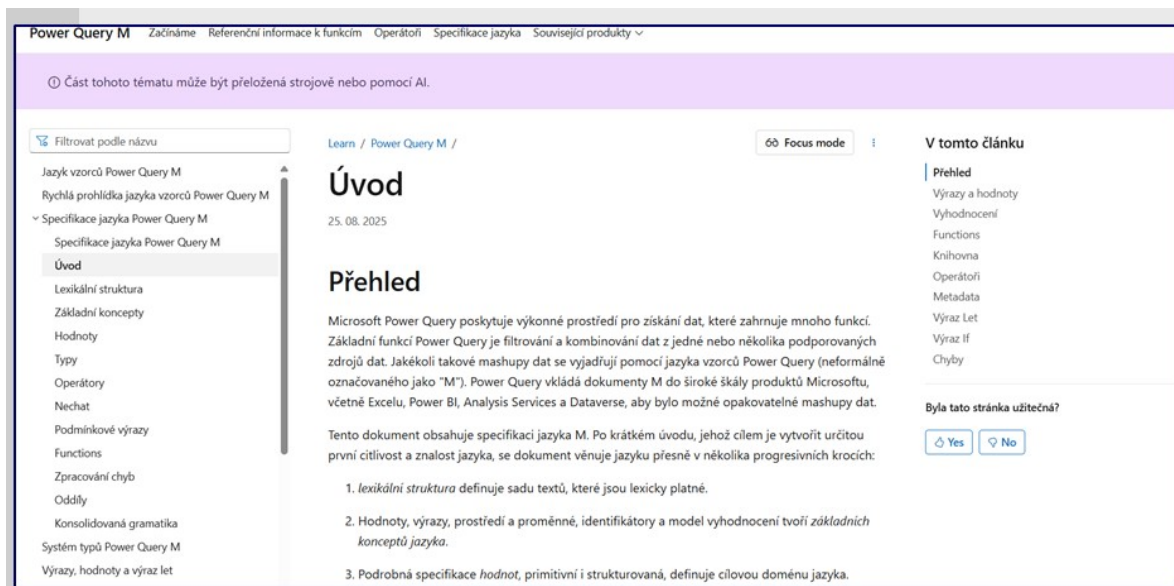
[1] Self Service BI (SSBI) <i>(Základní principy a charakteristiky úloh a aplikací na úrovni SSBI, jejich efekty a možné problémy v řešení.)</i>	
[2] Základní principy Power BI (PBI) <i>(Instalace PBI Desktop, instalace Power BI mobile apps, postup práce v Power BI desktop, vizualizace dat, datové sady, reporty, dashboardy, dlaždice.)</i>	
[3] Režimy práce s daty <i>(Režim založený na fyzickém importu dat do sémantického modelu PBI, režim přímého dotazování – Direct Query, živé připojení (live connection), kompozitní režim, direct lake.)</i>	
[4] Řešení sémantických modelů <i>(Sémantický model, analytická příprava sémantického modelu, postup vytvoření sémantického modelu a aplikace.)</i>	
[5] Reporty a vizuály v Power BI <i>(Vytváření reportů, využití vizuálů, filtrování, průřezy, zvýrazňování a drilování, analytické funkce v grafech.)</i>	
[6] Power BI Service a mobilní aplikace <i>(Publikace reportů, funkce PBI Service, sdílené dashboardy, datové brány, mobilní aplikace.)</i>	
[7] Příprava a plánování projektů v Power BI <i>(Modely řešení a nasazování projektů PBI, role v řešení a užití projektů PBI, podstatné otázky k řešení projektů, licence PBI.)</i>	
[8] Plánování a řešení reportů v PBI <i>(Proces plánování, vizualizace, tabulky a grafy v reportech.)</i>	[9] Příprava datových zdrojů <i>(Data Profiling, tabulární model, SCD.)</i>
[10] Transformace dat a jazyk M <i>(Charakteristika jazyka M, jednotlivé typy M dotazů.)</i>	

Zatímco text tohoto oddílu představuje přehled a principy Power BI, další informace a další detaily nabízejí následující vybrané tutoriály na dále uvedených adresách:

Power BI, přehled funkcí a další témata: [<https://learn.microsoft.com/cs-cz/training/powerplatform/power-bi>]:



Power Query, M: [<https://learn.microsoft.com/cs-cz/powerquery-m/m-spec-introduction>]:



Power Query, jazyk M: [<https://learn.microsoft.com/cs-cz/powerquery-m/>]:

The screenshot shows the Microsoft Learn page for Power Query M. The page has a purple header with the title 'Power Query M' and navigation links: 'Začínáme', 'Referenční informace k funkcím', 'Operátory', 'Specifikace jazyka', and 'Související produkty'. Below the header is a purple banner with the text: 'Část tohoto tématu může být přeložena strojově nebo pomocí AI.' Below the banner is a search bar with the text 'Filtrovat podle názvu' and a 'Learn /' link. The main content area is titled 'Jazyk vzorců Power Query M' and contains a paragraph about Microsoft Power Query. To the left of the main content is a sidebar with a list of topics: 'Jazyk vzorců Power Query M', 'Rychlá prohlídka jazyka vzorců Power Query M', 'Specifikace jazyka Power Query M', 'Systém typů Power Query M', 'Výrazy, hodnoty a výraz let', 'Komentáře', 'Model vyhodnocení', 'Operátory', 'Typy a převod typů', 'Metadatové informace', 'Chyby', 'Datum a čas', 'Formátování textu', 'Functions', 'Výčty', 'Konstanty', and 'Dynamické hodnoty'. To the right of the main content are three boxes: 'Začínáme' (with links to 'RYCHLÉ ZPROVOZNĚNÍ' and 'Rychlá prohlídka jazyka M'), 'Specifikace' (with links to 'ARCHITEKTURA', 'Specifikace jazyka Power Query M', and 'Systém typů Power Query M'), and 'Odkazy' (with a 'REFERENCE' link and several other links).

Power Query M Začínáme Referenční informace k funkcím Operátory Specifikace jazyka Související produkty ▾

🔍 Filtrovat podle názvu Learn /

Jazyk vzorců Power Query M

Microsoft Power Query poskytuje výkonné prostředí pro import dat, které zahrnuje mnoho funkcí. Power Query spolupracuje s desktopovými sešity Analysis Services, Excelu a Power BI a mnoha online službami, jako jsou Prostředky infrastruktury, služba Power BI, Power Apps, Microsoft 365 Customer Insights a další. Základní schopnosti Power Query je filtrovat a kombinovat data z jedné nebo několika podporovaných zdrojů dat. Jakákoli taková datová mashup se vyjadřuje pomocí jazyka vzorců Power Query M. Jazyk M je funkční jazyk s rozlišováním velkých a malých písmen podobný jazyku F#.

Začínáme

- RYCHLÉ ZPROVOZNĚNÍ
- Rychlá prohlídka jazyka M
- KONCEPCE
- Výrazy, hodnoty a výraz let
- Model vyhodnocení
- Typy a převod typů
- Jak kultura ovlivňuje formátování textu

Specifikace

- ARCHITEKTURA
- Specifikace jazyka Power Query M
- Systém typů Power Query M

Odkazy

- REFERENCE
- Referenční informace k funkcím Power Query M
- Referenční informace k výčtu Power Query M
- Referenční dokumentace konstanty Power Query M
- Referenční informace k dynamické hodnotě Power Query M

Úvod

Power BI je **sada softwarových služeb, aplikací a konektorů**, které společně slouží k vytváření, sdílení a využívání aplikací jak v podobě Self Service BI, tak i jako součást podnikových analytických řešení. Umožňuje **spojit a uspořádat rozličné datové zdroje**, od jednoduchých, lokálně uložených tabulek aplikace Excel nebo csv, přes databázové zdroje uložené *on-premise*, po souhrn datových zdrojů provozovaných v cloudu, a vytvářet nad nimi vizuálně přívětivé analytické výstupy. Power BI umožňuje **snadné připojení k datovým zdrojům, vizualizaci dat nebo objevování důležitých skutečností. Celá platforma PBI pak slouží ke sdílení výstupů** jak na webu, tak v mobilních zařízeních.

Power BI za dobu, kdy bylo 24. července 2015 oficiálně uvedeno na trh, mnohonásobně rozšířilo svoji funkcionalitu, i když stále vychází z původních samostatných nástrojů, jako např. *Power Query pro spojování a transformaci dat*, *Power Pivot pro datové modelování*, *Power View pro vizualizaci dat*, doplněné o *Power Map* pro vizualizaci 3D geoprostorových dat a *Power Q&A engine pro otázky a odpovědi v přirozeném jazyce* a o další speciální služby. Navíc PBI zahrnuje konektory na desítky datových zdrojů včetně nástrojů pro *Direct Query* (data nejsou nahrávána do sémantických modelů Power BI, ale jsou uložena v samostatných datových zdrojích). Pro připojení k on-premise datovým zdrojům slouží **Power BI Gateway** (umožňuje rychlý a zabezpečený přenos dat), a to jak osobní i podnikový (*Personal a Enterprise Gateway*). Sdílení pak zajišťuje cloudová služba **Power BI Service a Power BI Mobile apps** (pro přístup k dashboardům a reportům z jakéhokoliv mobilního zařízení se systémy iOS, Android). Vývojáři mohou využít Power BI embeded a prostřednictvím Power BI REST API a Power BI SDK zabudovat dashboardy a reporty přímo do svých vyvíjených aplikací. **Power BI Report Server**, určený pro robustní podniková Business Intelligence řešení, umožňuje sdílení reportů v rámci on-premise řešení v organizacích.

Od června 2023 se stal Power BI součástí platformy Fabric, což je univerzální cloudové řešení v podobě SaaS pro ukládání, správu a analýzu dat. Zahrnuje kromě Power BI řadu dalších služeb, jako je Azure Synapse, Azure Data Explorer. Power BI se soustředí především na vizualizaci dat a vytváření sestav (viz oddíl C).

Oddíl A je věnován prostředí Power BI, přičemž **kapitoly 1 – 5** jsou věnovány základním charakteristikám Power BI a možnostem individuálních využití desktop, **kapitoly 6 – 10** pokrývají obecnější otázky a obecnější kontext pohledu na projekty a řešení úloh v tomto prostředí.

1. Self Service Business intelligence, SSBI



Účelem kapitoly je:

- vymezit hlavní **charakteristiky aplikací a produktů SSBI**, jejich pozitiva a problémy,
- vytvořit základní podklad pro **posouzení jejich možností** při jejich návrzích samotnými uživateli.

1.1 Podstatné charakteristiky SSBI

Smyslem SSBI je na základě nových technologií poskytnout uživatelům prostředí pro realizaci svých analytických úloh bez nutnosti využívání komplexních a obvykle velmi složitých systémů BI.

Samoobslužné BI umožňují např. realizovat **multidimenzionální uložení a zpracování dat**, nabízejí efektivní a **jednoduché přístupy k datům**, poskytují prostředky, např. **jazyk DAX**, pro výpočty a další operace v dimenzionálním prostředí apod. Kromě toho jsou tyto aplikace i **vhodným prostředkem pro pochopení podstaty a způsobu využití větších Business Intelligence systémů**.

Self-Service BI **se snaží vyřešit rozpor mezi dvěma protichůdnými podnikovými silami** – potřebou **flexibility a svobody** při analýze a zkoumání dat koncovými „spotřebiteli“ BI výstupů, které ovšem bohužel nelze docílit s pomocí „dodavatele“ ve formě podnikového IT oddělení. Proti pak stojí **potřeba IT oddělení mít neustálou kontrolu nad daty** a nad vytvářením a distribucí informací uvnitř firmy.

SSBI rozšiřuje klasické tradiční BI prostředí o **možnosti provádění vlastních analýz** nad zpřístupněnými daty a jejich reporting bez nutnosti potřeby zásahu IT oddělení. Dává ne-IT pracovníkům **více možností, větší flexibilitu a větší samostatnost**, kdy je možné získat odpověď na danou otázku ve výrazně kratší době. Přístup SSBI tedy obchází **problém s neustále se měnícími a novými požadavky** a potřebami uživatelů, kdy každý může získat potřebnou informaci, kdykoliv si zamane. Podnikové IT se tak staví do zcela nové role.

Typickým **představitelem** produktů SSBI je **Power BI** společnosti Microsoft, který v současné době patří i k nejrozšířenějším. Schéma jeho architektury představuje Obrázek 2-2:

1.2 Efekty SSBI

Uplatnění SSBI v praxi je spojeno obvykle s těmito efekty:

- Podstatně se zkracuje **doba, potřebná na implementaci** analytických aplikací.
- Analýza dat **samotnými byznys uživateli** a dodání požadovaných výstupů, resp. informací, je rychlejší. Je možné využít i jiná data než z podnikového datového skladu.
- Self-Service BI aplikace nabízejí obdobnou **flexibilitu a výkonnost**, jako je tomu většinou u aplikací, založených na OLAP databázích.
- Základní **příprava aplikací** včetně transformací zdrojových dat je výrazně zjednodušená, takže je dostupná i ne IT pracovníkům, samozřejmě po nezbytném zaškolení.
- Součástí technologií jsou i **programovací prostředky**, jako např. jazyk DAX (Data Analysis Expressions) pro Microsoft Power BI.
- **Finanční, zdrojová i provozní náročnost** Self-Service BI je oproti standardním projektům a provozovaným aplikacím výrazně nižší.
- Self-Service poskytuje i velmi dobrou podporu pro tvorbu **analytických i vizuálně náročnějších aplikací**, jako např. dashboardů, klikovacích map apod.
- **Uživatelé dostávají do rukou data a přístup k nim**, mohou si je kdykoliv podle svých aktuálních potřeb zanalyzovat a vyhodnotit za pomoci poskytnutých nástrojů. Nemusí tak dlouho čekat na výstupy, které jim dodá IT oddělení na základě jejich požadavku.

- Namísto toho, aby **IT oddělení** bylo zahlceno stále novými a měnícími se požadavky na reporty a analýzy, předává podstatný objem práce s daty do rukou samotných uživatelů. Zároveň **se může soustředit na řízení a aktivity, které přinášejí společnosti větší hodnotu**.
- Změna vede **k celkové větší spokojenosti uživatelů s IT oddělením** v organizaci, které se v jejich očích přetváří z podoby úzkého hrdla v rychlosti dodávání BI výstupů, do role flexibilního pomocníka, který lépe dokáže uspokojovat uživatelské byznys požadavky.

1.3 Omezení, problémy a předpoklady SSBI

Řešení, provoz a užití aplikací SSBI vyžaduje v některých případech nastavení určitých předpokladů, případně vyřešení následujících omezení a problémů:

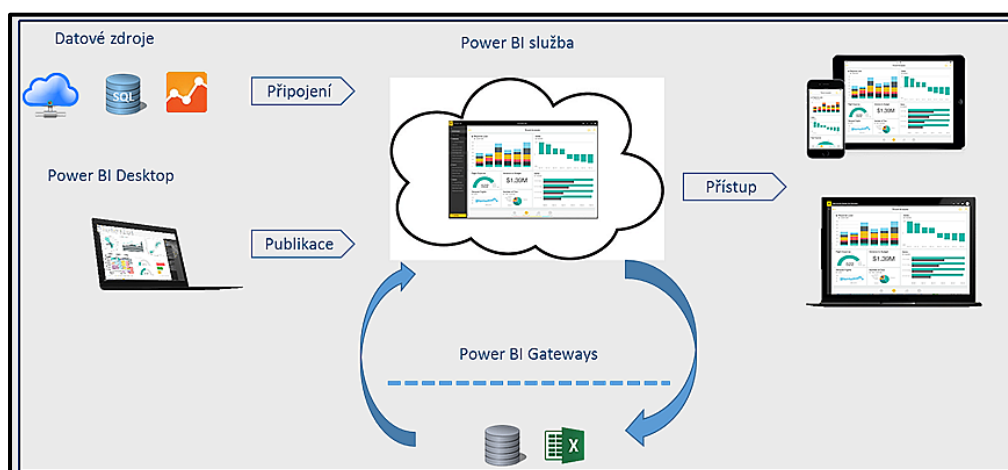
- Self-Service BI řešení se mohou realizovat **pouze pro úlohy určité kategorie**, tedy relativně jednodušší a izolovanější úlohy.
- Self-Service BI má částečně **omezené možnosti čištění a konsolidace dat**, které u standardních BI aplikací představují jednu z pracovně nejnáročnějších, ale i finálně nejefektivnějších částí.
- Problémem je dosažení **integrace dat** v rámci podniku, které jinak nabízejí celopodnikové datové sklady.
- S předchozí poznámkou souvisejí i podstatně omezené **možnosti celopodnikového reportingu**, např. centrálního reportingu v rámci nadnárodních společností apod.
- I když je příprava analytických aplikací na bázi Self-Service BI uživatelsky výrazně dostupnější, přesto je nutné, aby pro kvalitnější aplikace **uživatelé ovládali alespoň základní principy analytických metod**, tj. především dimenzionálního a datového modelování.
- Je zde zvýšené riziko **špatných rozhodnutí** na základě „špatných dat“ od byznys uživatelů.
- Mění se **pozice útvarů IT** z producenta výstupů do pozice tzv. „supervizora“. Je na něj tedy kladena tíha řízení a kontroly celého prostředí podnikové analytiky a produkovaných výstupů.
- Pouhou instalací SSBI nástrojů se požadované přínosy nezískají. Je potřeba, **aby uživatelé pochopili přínos** a možnosti daných nástrojů a začali je využívat.

1.4 Závěry k řešení Self-Service Business Intelligence



- Vznik Self-Service Business Intelligence, SSBI znamenal významný mezník v rozvoji podnikové analytiky. Vytvořil **předpoklady pro samostatnou práci uživatelů** v návrhu a využití analytických a plánovacích aplikací.
- Návrhy SSBI **na straně uživatelů** vytvářejí nároky na **zvyšování znalostí**, nejen pokud jde o provozované produkty, ale i **metody**, uplatňované při analýze a návrzích, zejména multidimenzionálního modelování a datového modelování.
- Nasazení aplikací SSBI se **liší obvykle podle velikosti a potřeb firem**. U menších firem může pokrývat jejich potřeby v celém rozsahu a nahradit komplexní BI systémy, u větších firem tomu tak často není a aplikace SSBI se stávají doplňkem komplexních BI aplikací, poskytujícím uživatelům větší operativnost při jejich práci.
- Na trhu je již adekvátní **nabídka produktů SSBI i souvisejících služeb**. To vytváří i odpovídající nároky na kvalifikovaný výběr těchto produktů a služeb.
- Současné kvalitní SSBI produkty poskytují **možnosti vstupů z nejrůznějších zdrojů**, tedy databází, internetových zdrojů, typů souborů atd.
- Systémy IT ve firmě, jak zdrojové, tak systémy analytiky, procházejí trvale **vývojem a řadou změn**. To při provázanosti SSBI na zdrojové databáze a aplikace vytváří vysokou **potřebu kvalitního řízení, právě na rozhraní zdrojů a SSBI**.
- I pro řešení úloh SSBI je účelné definovat **sadu analytických otázek**, které by měly být předmětem řešení a diskusí mezi manažery, analytiky a uživateli.

2. Základní principy Power BI



Účelem kapitoly je vymezit principy a vysvětlit jednotlivé součásti Power BI. Kapitola prezentuje i doporučený postup pro vytváření aplikací v Power BI desktop, hlavně při jeho použití jako nástroje pro SSBI. Tato kapitola tak představuje **vstupní informace celé předkládané publikace**.

Power BI (PBI) je **interaktivní a vizuální prostředí pro analýzu dat**, vyvinutý společností Microsoft. Slouží **k extrakci, transformaci a vizualizaci dat** z různých zdrojů, jako jsou databáze, tabulky, API a další. Power BI umožňuje uživatelům vytvářet komplexní vizuální reporty, interaktivní grafy a tabulky, které pomáhají představit a analyzovat data jednotlivým uživatelům. Tím umožňuje lepší **porozumění datům, identifikaci trendů a informované rozhodování** nejen při podnikání. Součástí Power BI jsou také nástroje pro publikování reportů online a sdílení těchto vizualizací s ostatními uživateli, případně zákazníky.



Obrázek 2-1: Principy Power BI (Zdroj: Seyfor, 2024)

Prostředí Power BI se skládá z několika komponent. Zjednodušeně se jedná o Windows aplikaci označovanou jako **Power BI Desktop**, o on-line **softwarovou službu**, běžící v cloudu (typu **SaaS – Software as a Service**), označovanou jako **Power BI service**, a **mobilní aplikaci Power BI apps** pro telefony a tablety se systémy iOS a Android.

Všechny součásti Power BI slouží uživatelům pro **efektivní vytváření, sdílení a využívání přehledů** o jejich byznysu tak, jak je potřebují, nebo jak to vyžaduje jejich role. Manažeři například typicky využívají dashboards prostřednictvím **Power BI mobile apps**, to znamená, že na chytrém telefonu sledují plnění vybraných klíčových ukazatelů výkonnosti (KPI, Key Performance Indicator).

Pro podrobnější pohled na data pak lze využít Power BI aplikaci **služby Power BI** (*Power BI service*), kde jsou k dispozici další důležité dashboards a reporty ze všech podnikových oblastí. Tvůrci reportů používají hlavně **Power BI desktop**, z něhož po odladění transformací dat, datových modelů a potřebných reportů publikují celá BI řešení **do cloudové služby Power BI service**, kde doplní potřebné řídicí panely (dashboards).

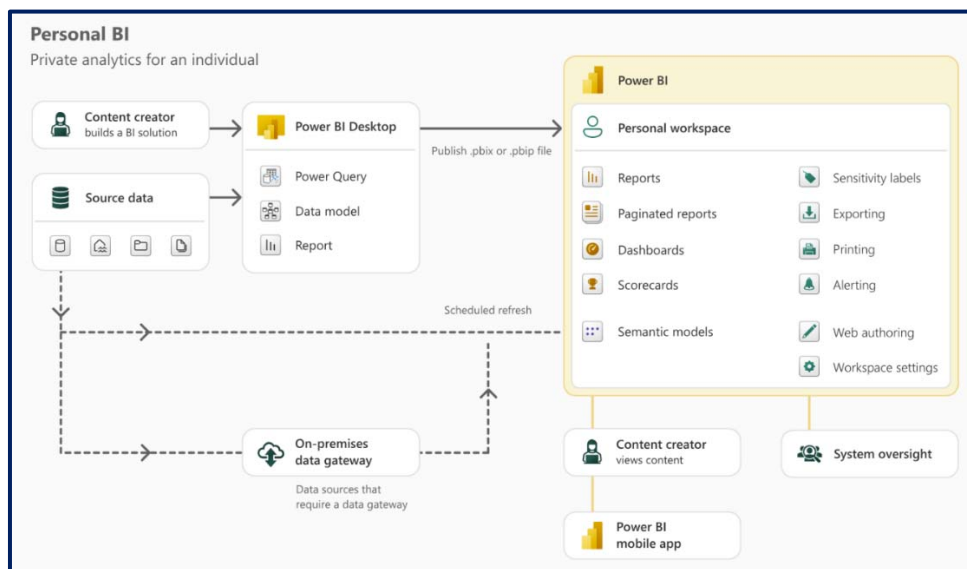
Na druhé straně firmy, které preferují on-premises řešení, využívají speciální verzi Power BI desktop, určenou pro **Power BI Report Server** (*on-premise řešení*) a reporty publikují právě do prostředí Report Serveru. *Power BI Report server je součástí Microsoft SQL Serveru (Enterprise edice) a slouží ke sdílení jak Power BI reportů, tak i tzv. paginated reportů, vyvinutých v prostředí SQL Server Microsoft Reporting services pro zainteresované role a uživatele.* Celkovou architekturu PBI dokumentuje Obrázek 2-2:



Obrázek 2-2: Architektura Power BI

(Zdroj: <https://sqldusty.com/2019/02/15/power-bi-architecture-diagram-v4-is-now-available/>)

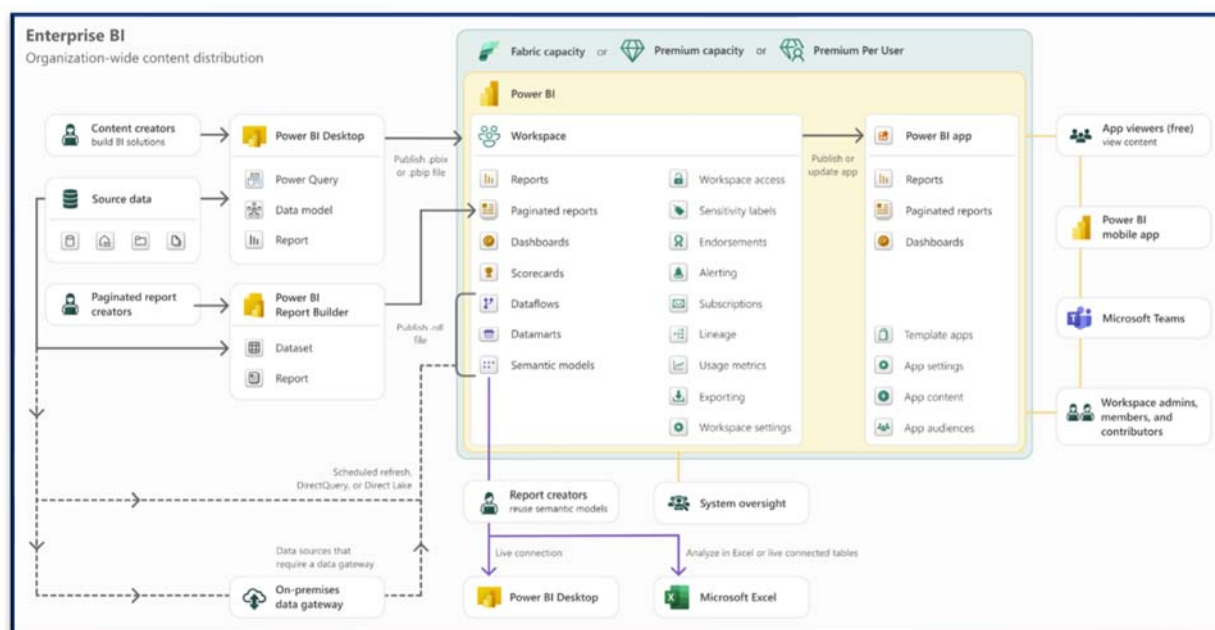
Možnosti využití PBI pro osobní potřeby (SSBI) ukazuje Obrázek 2-3.



Obrázek 2-3: Využití PBI pro individuální potřeby

(Zdroj: <https://learn.microsoft.com/en-us/power-bi/guidance/media/powerbi-implementation-planning-usage-scenario-personal-bi/usage-scenario-personal-bi-no-legend.svg>)

Způsob využití PBI ve firmách dokumentuje Obrázek 2-4.



Obrázek 2-4: Využití PBI ve firmách (Zdroj: Seyfor, <https://learn.microsoft.com/en-us/power-bi/guidance/media/powerbi-implementation-planning-usage-scenario-enterprise-bi/usage-scenario-enterprise-bi-no-legend.svg>)

2.1 Instalace Power BI

Power BI Desktop aplikace je volně dostupná aplikace **ke stažení** například přímo na stránkách Microsoft Power BI <https://powerbi.microsoft.com/en-us/desktop/>.

Mobilní aplikace Power BI jsou zdarma k dispozici na *Windows Store*, *App Store* či *Google Play*.

Power BI service je cloudová služba, dostupná po přihlášení na webu Microsoft Power BI <https://app.powerbi.com/>.

Power BI Report Server je součástí SQL Server Reporting Services 2016 Enterprise a vyšší.

2.2 Postup práce v Power BI

Příprava SSBI řešení zpravidla **začíná v Power BI Desktop**, kam je třeba buď načíst, transformovat a uspořádat vstupní data do sémantického dimenzionálního modelu, tedy do vhodné struktury pro reportování, nebo se napojit na data, umístěná mimo Power BI s využitím služby *Direct Query* a vytvořit sémantický datový model, aniž by data opustila svůj zdroj, případně se napojit živým připojením na analytický model a data ležící v některé analytické databázi, ať již v cloudu nebo on-premise, a nad těmito daty **vytvořit potřebné reporty**. Následně jsou **reporty** buď využívány přímo daným uživatelem v desktopové aplikaci v self-service režimu, nebo jsou **publikovány do Power BI service**, kde jsou vytvořeny **dashboards**, případně další **nové vizualizace** nad publikovaným sémantickým modelem. Zde jsou pak sdíleny prostřednictvím PowerBI app uživatelům, kteří tak mohou informace využívat prostřednictvím této cloudové služby ve webovém prohlížeči nebo na svých mobilních zařízeních s využitím mobilních Power BI apps.

Power BI je dále k dispozici též jako tzv. **Power BI Embedded**, určené k přímému **zabudování pomocí Power BI REST APIs** či **Power BI SDK** do vyvíjených **zákaznických aplikací**. Od verze MS SQL Server 2016 Enterprise lze Power BI reporty **publikovat** a sdílet prostřednictvím **MS SQL Power BI Report Server** (pro přípravu sémantických modelů a reportů je určena speciální verze desktopové aplikace, optimalizovaná pro Power BI Report Server).



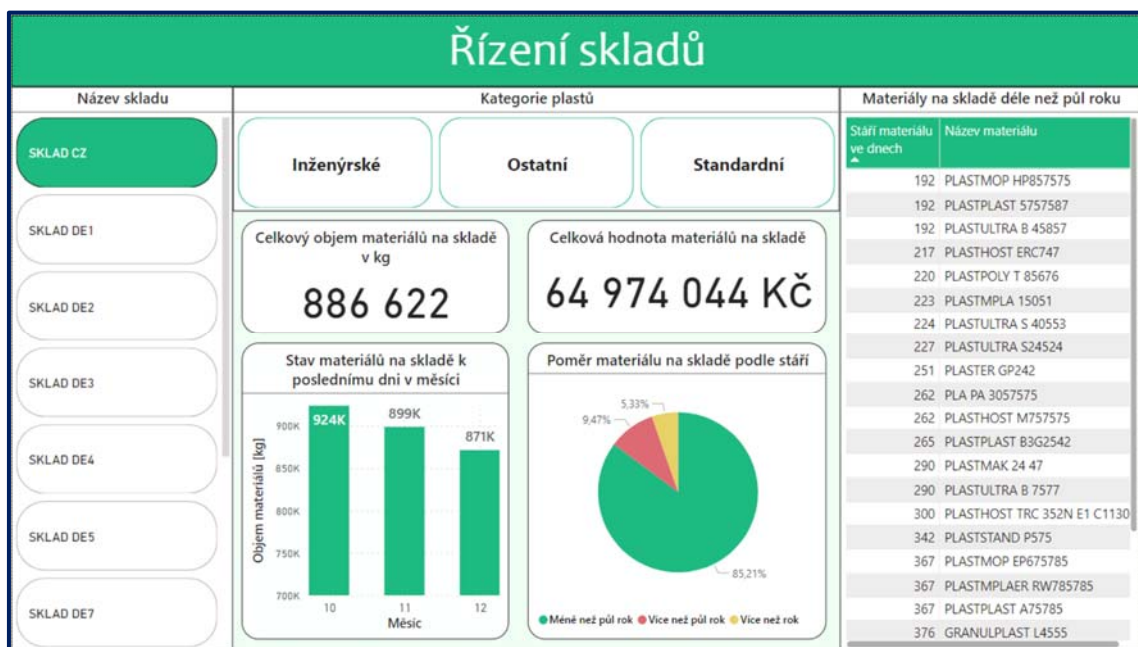
Stavební bloky Power BI

Vytvářené řešení v PBI je realizovatelné prostřednictvím **několika základních stavebních bloků**:

- Vizualizace – *Visualizations*.
- Sémantické modely (původně označované jako datasets) – *Semantic models*.
- Sestavy (reporty) – *Reports*.
- Řídicí panely (dashboards) – *Dashboards*.
- Dlaždice – *Tiles*.

2.3 Vizualizace

Vizualizace (často **označovaná též jako vizuály**) jsou vizuální reprezentací dat. Jedná se například o jednoduché vyjádření jedné hodnoty sledované metriky v podobě jednoho čísla nebo hodnot metrik podle zvolených dimenzí **pomocí různých typů grafů** – spojnicových, sloupcových, bublinových, prostřednictvím různě složitých map s barevnými rozlišeními oblastí podle velikosti zobrazovaných hodnot a podle teritorií. Vizualizace zahrnuje také **interaktivní přehledové tabulky** a řadu dalších zajímavých prvků pro grafické vyjádření dat v takové podobě, která uživateli rychle zpřístupní kontext a význam dat. Microsoft průběžně sadu dostupných vizuálů zlepšuje a doplňuje a zároveň **umožňuje využívat vizuály vlastní**. Příklad skupiny vizuálů ukazuje Obrázek 2-5.

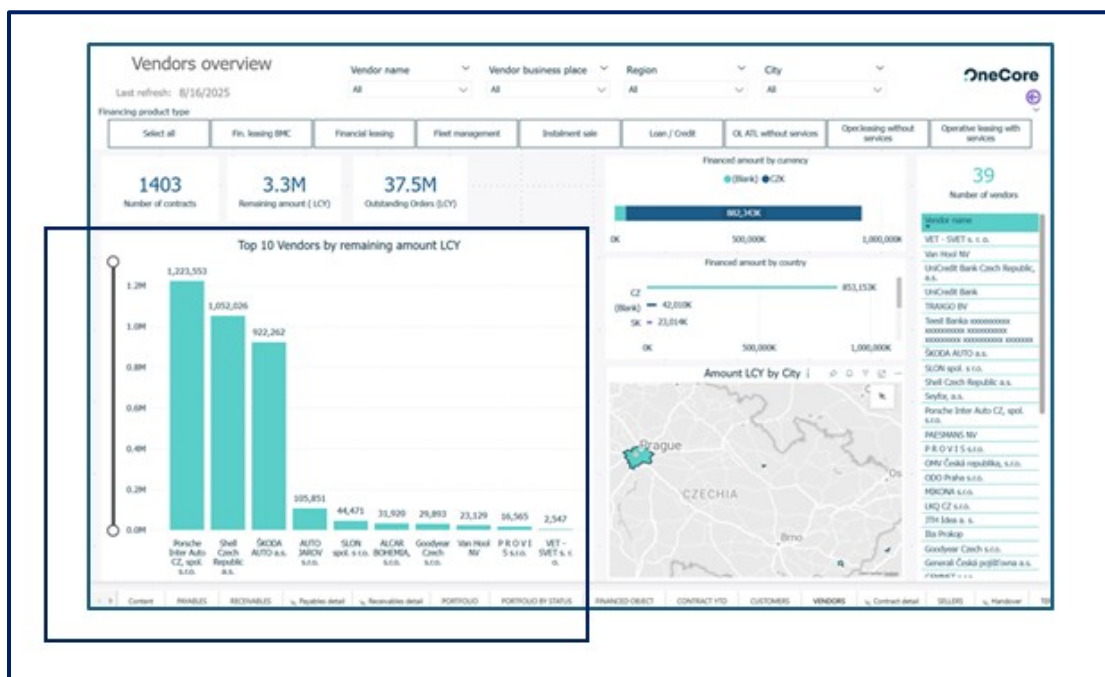


Obrázek 2-5: Příklady vizuálů v Power BI (Zdroj: Čondlová, 2024)

Jednotlivé vizuály lze na stránkách sestavy vzájemně uspořádat, překrývat, měnit jejich vlastnosti, přidávat jim nápovědu apod. Uživatelé, kteří nejsou vlastníky sestavy, mohou s dlaždicemi interagovat, nemohou měnit jejich umístění ani jejich velikost.

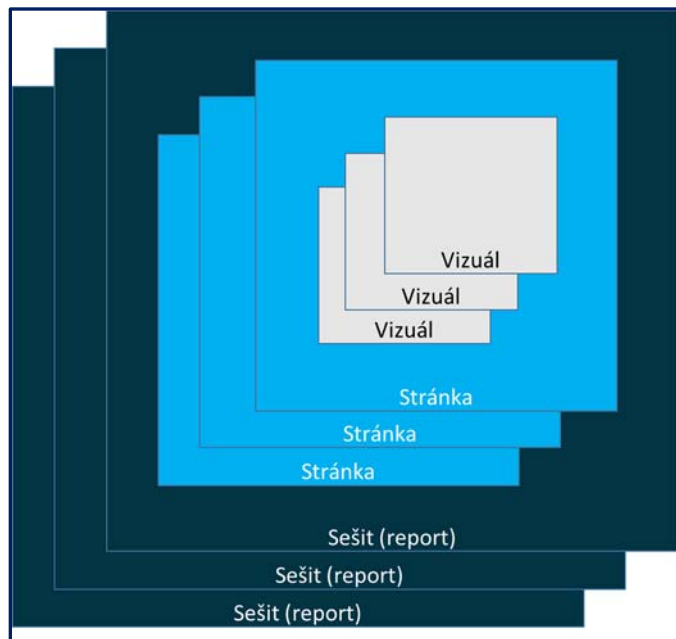
2.4 Sestavy (reports)

Sestavou se v PBI rozumí **souhrn vizuálů, umístěných na jedné či více stránkách (listech sestavy)**. **Jednotlivé prvky sestavy jsou vzájemně provázané**. Obrázek 2-6 je příkladem vícestránkové sestavy v Power BI service. Každá stránka je umístěna **na jedné záložce / listu**, stejně jako jsou listy uspořádány v excelovém sešitě. Všechny stránky sestavy jsou **vytvořeny nad jednou datovou sadou**.



Obrázek 2-6: Příklad sestavy s více listy v Power BI a vyznačením jednoho vizuálu (Zdroj: Seyfor OneCore reporting demo)

Struktura řešení, vytvořeného v Power BI desktop, je znázorněna na následujícím schématu (Obrázek 2-7).



Obrázek 2-7: Struktura sestavy v Power BI desktop

2.5 Řídící panely (dashboards)

Power BI dashboard je **seskupením vybraných vizuálů z jedné či více stránek jednoho či více reportů**, sdružených na jednom místě. Lze je vytvářet až ve službě Power BI (ne tedy v Power BI desktop). Slouží pro získání rychlého přehledu o prezentovaných skutečnostech. Řídící panel lze vytvářet

pouze v Power BI service, odkud je možné jej sdílet s dalšími uživateli. Ti jej mají přístupný buď též prostřednictvím aplikací v *Power BI service*, nebo prostřednictvím mobilní aplikace. Ve všech případech musí uživatelé **mít platnou licenci Power BI PRO, nebo přidělenou „free“ licenci v rámci kapacity Fabric**, vyšší než 64 CU.

Poznámka: Na rozdíl od jiných nástrojů pro datovou analýzu a vizualizace jako je Tableau nebo Qlik, Microsoft chápe sestavu (report) a dashboard mírně jinak. Obecně se dashboardem rozumějí grafická vyjádření, resp. prezentace sledovaných metrik, které jsou důležité z pohledu byznysu. Metriky jsou prezentovány formou grafů, tabulek, budíků a jiných grafických ukazatelů. V Power BI stránky reportů obsahují i grafické zobrazení sledovaných údajů. (Zdroj: *IT a anatomie Firmy. Tvorba Dashboardů*. https://databusiness.cz/wp-content/uploads/page/3284/AF_II_05_04_Dashboardy.pdf).

2.6 Dlaždice (tiles)

Dlaždicí se rozumí jeden vizuál, umístěný na řídicím panelu. Obrázek 2-6 obsahuje zarámovanou část sestavy, která může být umístěna jako samostatný prvek na řídicí panel, dashboard. Při vytváření dashboardů lze následně dlaždice ve službě Power BI různě uspořádat nebo měnit jejich velikost.

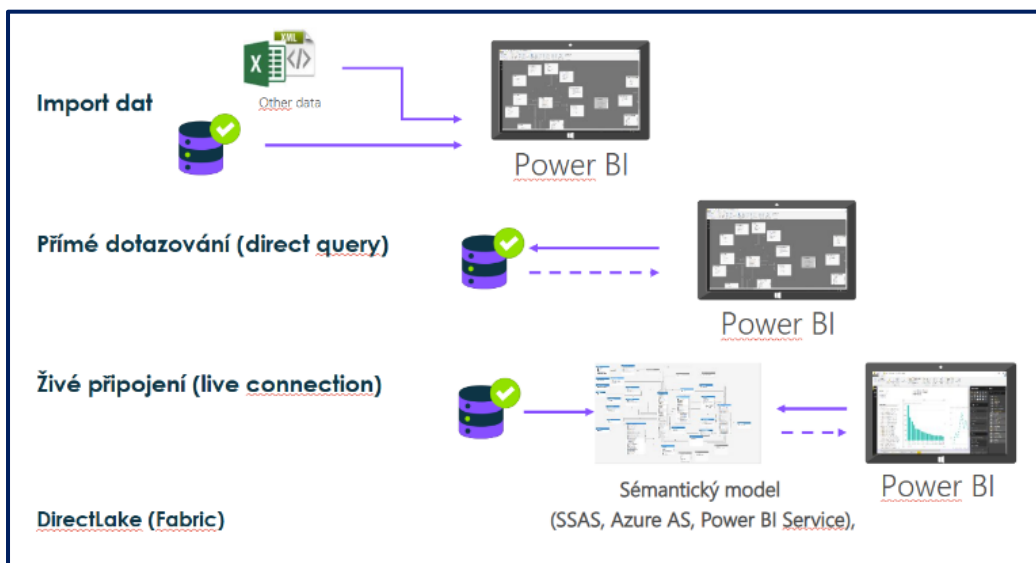
2.7 Pracovní závěry



Z kapitoly vyplývají následující závěry:

- Power BI (PBI) je interaktivní a vizuální prostředí pro analýzu dat.
- PBI slouží **k extrakci, transformaci a vizualizaci dat** z různých zdrojů, jako jsou databáze, tabulky, API a další.
- PBI zahrnuje několik komponent:
 - Windows aplikace **Power BI Desktop**,
 - službu v cloudu typu **SaaS – Software as a Service**, označovanou jako **Power BI service**,
 - mobilní aplikace **Power BI apps** pro telefony a tablety.
- **Postup PBI řešení:**
 - zpravidla začíná **v Power BI Desktop**, kam je třeba buď načíst, transformovat a uspořádat zde vstupní data do sémantického dimenzionálního modelu,
 - jinou možností je využití **služby Direct Query**, kde se vytvoří **sémantický datový model**, aniž by data opustila svůj zdroj, případně se napojí živým připojením na analytický model a data, ležící v některé analytické databázi,
 - nad daty se vytvářejí potřebné reporty, které jsou buď využívány přímo daným uživatelem **v desktopové aplikaci** v self-service režimu, nebo jsou **publikovány do Power BI service**, kde jsou vytvářeny **dashboardy**.
- Řešení v PBI využívá **několik základních stavebních bloků**:
 - Vizualizace – *Visualizations*.
 - Sémantické modely (datasets) – *Semantic models*.
 - Sestavy (reporty) – *Reports*.
 - Řídicí panely (dashboardy) – *Dashboard*.
 - Dlaždice na řídicích panelech – *Tiles*.

3. Režimy práce s daty



Power BI podporuje tyto **základní režimy práce s daty**, a to:

- Živé připojení (*Live connection*).
- Režim přímého dotazování – *DirectQuery*.
- Režim založený na fyzickém importu dat do sémantického modelu PBI.
- Kompozitní režim (*Composite Mode*).
- Režim *Direct Lake*.

Účelem kapitoly je stručně tyto režimy objasnit a určit jejich obvyklé využití.

Principy uvedených režimů dokumentuje

Obrázek 3-1.

Typy připojení Power BI k datovým zdrojům				
	Živé připojení (live)	Přímé dotazování (direct query)	Import	Kompozitní
	- Vhodné pro enterprise řešení - Žádná data nejsou v paměti - Úplný reporting - Žádné datové transformace - Žádné datové modelování - Vysoký výkon - Data jsou uložena na zdroji - Dotazy se odesílají ke zdroji - Není vyžadováno obnovení dat	- Žádná data nejsou v paměti - Úplný reporting - Žádné datové transformace - Omezené datové modelování - Nízký výkon - Data jsou uložena na zdroji - Dotazy se odesílají ke zdroji - Není vyžadováno obnovení dat	- Žádná data nejsou v paměti - Úplný reporting - Plná datové transformace - Plné datové modelování - Vysoký výkon - Data jsou uložena v Power BI - Dotazy se neodesílají ke zdroji - Je vyžadováno obnovení dat	- Kombinace Direct Query a Importu - Větší objemy dat - Nevyužívá Analysis Services
Zdroj dat	Analysis Services + Power BI sémantický model	SQL databáze	SQL databáze	SQL databáze (import) + SQL databáze (direct query)
Zdroj dat	Data + Model	Data		Data
Kde jsou uloženy komponenty Power BI	Vizuály	Model + Vizuály	Data + Model + Vizuály	Data + Model + Vizuály

Obrázek 3-1: Režimy připojení PBI ke zdrojovým datům a uložení sémantického modelu
Zdroj: <https://www.linkedin.com/pulse/power-bi-connectivity-types-lazaros-viastikopoulos/>

3.1 Datový konektor versus typ připojení

Před vysvětlením režimů práce s daty je třeba stručně vysvětlit rozdíl mezi typy připojení (*Connectivity Types*) a datovými konektory (*Data Connectors*), jelikož se tyto dva pojmy často zaměňují.

Typ připojení (režim práce s daty) pak určuje, jakým způsobem Power BI s daty pracuje – zda se data importují přímo do Power BI (*Import*), nebo zda Power BI odesílá dotazy přímo do zdrojové databáze bez ukládání dat (*DirectQuery*, *Live Connection*). S příchodem SaaS Fabric se pak jedná o čtvrtý typ tzv. *Direct Lake* režim.

Datový konektor v Power BI Desktop je rozhraní, které umožňuje připojit se k různým typům zdrojů dat, například k databázím (SQL Server, Oracle, SAP HANA), souborům (Excel, CSV) nebo webovým službám. Každý typ konektoru zajišťuje správné propojení a komunikaci se zdrojem dat včetně ověřování a získávání potřebných informací o datech.

Rozdíl tedy spočívá v tom, že konektor definuje, k jakému zdroji se připojujete, a typ připojení určuje, jak se s daty v rámci tohoto spojení nakládá.

3.2 Živé připojení (Live connection).

Prvním typem připojení, používaným primárně v podnikových řešeních, je přímé napojení Power BI ke službě **MS SQL Analysis Service** či **Azure Analysis Services**, nebo k **sémantickému modelu ve službě Power BI** (*Power BI Service*). Tento režim je označován jako **živé připojení (Live connection)**. Je velmi podobné *DirectQuery* s tím hlavním rozdílem, že zdrojem pro vizualizace je sémantický model, připravený a uložený buď v analytických službách (on-premise *SQL Server Analysis Services* nebo *Azure Analysis Service*), nebo ve službě *Power BI* a při připojení naživo není definován žádný dotaz a v paměti Power BI nejsou uložena žádná data.

Při práci s vizuály se dotazy odesílají na externí zdroj SSAS/AAS/Power BI Service. Na rozdíl od *DirectQuery* **není** přímo v Power BI k dispozici **prostředí pro tvorbu a úpravu sémantického modelu**. Není tedy možné definovat nové vypočítané sloupce, ukazatele, hierarchie, vztahy atd. Vše probíhá a je připravené v externím sémantickém modelu. V závislosti na umístění zdroje dat může být nutné nakonfigurovat datovou bránu (*Data Gateway*). Aktualizace dat v modelu probíhá na zdroji. V Power BI desktop probíhá pouze příprava vizualizací.

3.3 Přímé dotazování – DirectQuery

Práce s daty v režimu **přímého dotazování do zdrojové databáze** znamená, že tato databáze leží mimo Power BI (*DirectQuery*), buď **přímo, on-premise v infrastruktuře vlastníka dat, nebo v cloudu**. Tento režim je vhodné využít, pokud se jedná o velký objem zdrojových dat, pokud podléhají zdrojová data častým změnám, pokud to vyžaduje bezpečnost dat, nebo je-li zdrojem dat multidimenzionální databáze obsahující kalkulované ukazatele, což platí pro SAP BW a SAP HANA (Zdroj: <https://powerbi.microsoft.com/en-us/documentation/powerbi-desktop-directquery-about/>).

Režim přímého dotazování se **nastavuje ve funkci Get Data** pro připojení k datovému zdroji v Power BI Desktop **volbou DirectQuery**. Chování tohoto připojení je následující:

- Nejdříve je třeba **vybrat zdroj dat**. Pro relační databáze se dále **vybírá sada tabulek** a pro každou z nich je možné ještě **definovat dotaz**, který vrací potřebnou množinu dat. U multidimenzionálních datových zdrojů, jako je SAP BW a SAP HANA, se vybírá pouze datový zdroj.
- Po připojení **neprobíhá import** žádných dat.
- Model a definice výpočtů metrik se vytvářejí v Power BI.
- **Při vytváření vizuálů** v prostředí Power BI Desktop jsou **dotazy zasílány do definovaného podkladového zdroje dat**, který vrátí potřebné údaje. Doba nutná k obnovení vizuální hodnoty závisí na výkonu příslušného zdroje dat.
- **Pro promítnutí změn** ve zdrojových datech do vizuálů na sestavách je třeba použít **funkci Refresh**, která zajistí spuštění potřebných dotazů do zdrojové databáze.
- Po publikování sestav do služby Power BI bude **výsledkem datová sada služby Power BI**, stejně jako při importu do datové sady Power BI. **Součástí** této datové sady však **nejsou žádná data, ale pouze model s definovanými výpočty**.
- Při **otevírání existujícího reportu** nebo **při tvorbě nového** reportu ve službě Power BI se **služba znovu dotazuje zdroje dat**, aby získala potřebná data. V závislosti na umístění zdroje

dat může být nutné **nakonfigurovat datovou bránu (Data Gateway)** pro přístup k on-premise datovému zdroji (stejně jako je to třeba pro režim importu).

- Pokud jsou vizuály z reportů nebo celé stránky reportů připojeny jako dlaždice na **dashboard** a má-li být otevření dashboardu rychlé, je vhodné **nastavit plán automatické aktualizace hodnot v dashboardu** (například každou hodinu). Četnost aktualizace by měla odrážet četnost změn podkladových dat a potřebu uživatelů vidět nejnovější data. Při otevření dashboardu jeho dlaždice odrážejí hodnoty dat platné v době posledního obnovení, což nutně nemusí být poslední změny, provedené v podkladovém datovém zdroji. Pro **obnovení hodnot v dlaždicích** dashboardu je v tomto případě potřeba zvolit ruční aktualizaci pomocí funkce **Obnovit dlaždice řídicího panelu (Refresh)**.

3.4 Import dat

Import je **nejběžnějším typem připojení** v Power BI při použití jako SSBI, a tedy i výchozí volbou při připojení k většině zdrojů dat. Při použití importu jsou všechna načtená data uložena v modelu Power BI (soubor .pbix) a držena v jeho paměti. Proto je **přístup k datům velmi rychlý**, protože čtení z paměti je nejrychlejší ze všech typů připojení. Součástí Power BI je výpočetní engine „VertiPaq (xVelocity)“, in-memory sloupcová databáze, která data komprimuje a ukládá do paměti.

Protože jsou data uložena v modelu Power BI, všechny dotazy při interakci s vizuály směřují do tohoto modelu, nikoli zpět do zdrojového systému. **Do zdroje dat** se Power BI dotazuje pouze **při obnově dat (refresh)**. V Power BI je **refresh dat nutný provést vždy manuálně, ve službě je nutné naplánovat** obnovu sémantického modelu, aby reporty obsahovaly aktuální informace.

Import nabízí plné využití možností Power BI Desktop – transformace a úpravy dat, modelování, vytváření relací mezi daty, DAX funkce i tvorbu reportů.

Je však třeba **sledovat velikost modelu**, protože existují limity. Pro licenci Power BI Pro může být velikost modelu maximálně 1 GB po kompresi, s Power BI Premium Per user až 400 GB, v licencování Fabric pak podle zvolené kapacity.

Pro tento režim je tedy **charakteristické**:

- Data i sémantický model v PWBI jsou uloženy přímo v Power BI.
- Je možné využívat kombinace různých zdrojů dat.
- Transformace dat a výpočty se realizují přímo v Power BI.
- Aktualizace dat se provádí v Power BI desktop na vyžádání, v okamžiku potřeby, ve službě podle nastavení frekvence obnovy dat.
- V závislosti na umístění zdroje dat může být nutné nakonfigurovat datovou bránu (*Data Gateway*) pro přístup k on-premise datovému zdroji.

3.5 Kompozitní režim (Composite Mode)

Kompozitní režim je založený na **kombinaci využití režimu importu dat a režimu Direct Query**. Využití je spojeno především s velkými systémy, kde rozsah zejména faktových tabulek by mohl narážet na omezení u režimu, založeného na fyzickém transferu. V tomto případě se **pro přístup k rozsáhlým tabulkám využívá Direct Query** a fyzický transfer je použit většinou pro agregované tabulky menšího rozsahu.

3.6 Režim Direct lake

Režim **Direct lake** v Power BI reprezentuje způsob práce s daty v Power BI, který je **součástí platformy Microsoft Fabric**. Umožňuje okamžitý **přístup k datům, uloženým v datalake** (datovém jezeru), bez nutnosti jejich fyzického importu do Power BI. Díky tomu lze pracovat s velmi rozsáhlými datovými sadami, zajistit tím vyšší výkon, nižší latenci při zachování flexibility a rychlosti zpracování.

Tento režim **kombinuje výhody Direct Query (přímé dotazování) a režimu importu**, avšak data zůstávají uložena přímo v lakehouse, nebo v delta tabulkách. Power BI přistupuje k datům přímo a optimalizovaně. Výsledkem je analýza dat v reálném čase při zachování škálovatelnosti, bezpečnosti a správy dat v prostředí Microsoft Fabric.

Direct lake režim je vhodný zejména pro organizace, které potřebují pracovat **s velkými objemy dat**, požadují rychlé odezvy a chtějí využít centralizované datové úložiště bez omezení kapacit klasických

importních režimů, i když režim importu v řadě případů, hlavně v případě menších datových sad do 1 GB (zatím v roce 2025), má stále ještě rychlejší odezvu, a zároveň umožňuje využít plnou funkcionalitu Power BI nástroje.

Režim Direct Lake lze využít pouze v těch organizacích, které využívají Fabric kapacitu a licencované prostředí SaaS. Tato publikace se proto tímto režimem zabývá pouze okrajově, jelikož je určena spíše pro specifické účely využití.

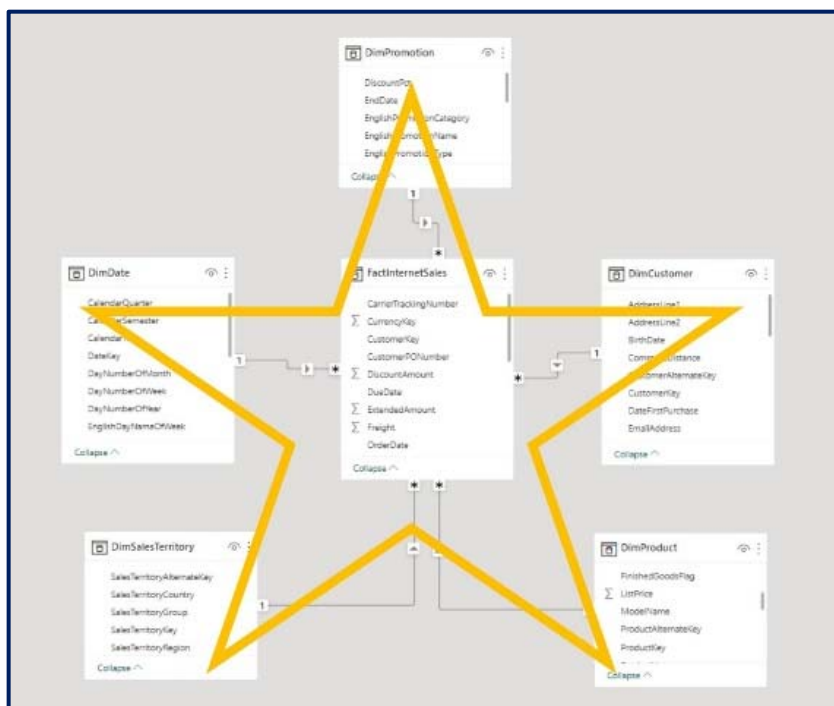
3.7 Pracovní závěry



Z kapitoly vyplývají následující závěry:

- Power BI nabízí tyto **základní režimy** práce s daty:
 - Živé připojení (Live connection).
 - Režim přímého dotazování – DirectQuery.
 - Import dat.
 - Kompozitní režim (Composite Mode)
 - Režim Direct Lake
- **U menších** (většinou individuálních) **úloh** a s relativně menšími objemy dat při osobním využití nebo SSBI se obvykle využívá import dat, **pro velké objemy dat** v korporátních systémech se užívá režim živého připojení k sémantickým modelům, přímého napojení na sémantické modely, nebo kompozitní režim.
- V moderních řešeních, využívajících platformu **Microsoft Fabric**, se v podnikové praxi pro velké objemy dat využívá režim **Direct lake**.

4. Řešení sémantických modelů



Sémantický model je kolekce dat, která obsahuje tabulky, sloupce a vztahy mezi nimi, kterou uživatelé Power BI používají **k vytváření vizualizací**. Zjednodušeně řečeno, jedná se o databázi, určenou pro reportování. Sémantický model vzniká transformací dat **z datových zdrojů**.

Účelem kapitoly je vymezit podstatu sémantických modelů a následně určit základní kroky k jejich vytvoření a využití.

4.1 Sémantický model

Sémantický model může vzniknout jak využitím jednoho jednoduchého sešitu, vytvořeného v MS Excel, tak kombinací mnoha různých datových zdrojů, jejich filtrací a uspořádáním. Do jednoho sémantického modelu lze například **zkombinovat data získaná z veřejných zdrojů**, umístěných někde na webových stránkách (například kursovní lístek ČNB), **data z excelového sešitu** (například roční plán prodeje) a **data z podnikové SQL databáze ERP** systému o prodeji zboží v různých měnách a o souvisejících číselnících jako zdrojích dimenzí.

Příprava kvalitního sémantického modelu na základě potřeb a požadavků byznysu tvoří **většinu práce** realizace BI aplikací.

K přípravě sémantického modelu (ETL) lze využít mnoho nástrojů. *Pro účely této publikace je popsána speciální komponenta PBI, **Power Query**. Transformační kroky v Power Query* zajistí, aby se do sémantického modelu dostala jen ta data, která jsou pro další práci a reporting podstatná. Například z kursového lístku pomůže vybrat pouze ty měny, ve kterých jsou vedeny ceny zboží ve zmíněné aplikaci.

Důležitou částí Power BI je **množství zabudovaných datových konektorů**. Zajišťují jednoduché připojení k řadě datových zdrojů, potřebné filtrování vstupních dat a jejich načtení do sémantického modelu. Konektory existují pro již zmíněné excelové sešity, csv soubory, řadu MS SQL databází, a to vlastní podnikové, nebo v Azure, Oracle DB, služby jako Salesforce, Microsoft D365 aplikace, Google Analytics, Facebook a pro desítky dalších typů datových zdrojů.

Sémantický model **může ležet i mimo Power BI**, podle toho se rozlišují režimy modelů (viz kapitola 3), které zahrnují:

- **Import** – data i datový model jsou „uvnitř“ PBI, lze kombinovat různé zdroje, v PBI probíhají transformace a výpočty, aktualizace dat na vyžádání.
- **Živé připojení** – sémantický model leží mimo Power BI.
- **Direct Query** – datový model je v PBI, data zůstávají ve zdrojové databázi, dotaz do zdroje probíhá v okamžiku zobrazení dat, dotazy je možné upravit, lze doplnit výpočty, nastavit propojení tabulek.
- **Kombinovaný režim** – část dat je importována do Power BI modelu, část zůstává na zdroji. Tento režim je označován jako „Composite“.
- **Direct lake** – model je vytvářen ve Fabric SaaS a data leží v datovém jezeru.

4.2 Analytická příprava sémantického modelu

Analytická příprava **sémantického modelu je založená na obdobných principech, jako je tomu u ostatních úloh podnikové analytiky**, které jsou náplní dokumentu na portálu MBI-AF „Principy a postupy řešení SSBI a BI úloh“ [Postupy]. Na tomto místě je doplněno několik poznámek vzhledem k PBI, které zahrnují:

- Určení podnikových procesů a oblastí řízení pro řešení.
- Určení granularity, resp. podrobnosti dat.
- Identifikace dimenzí.
- Definování faktů.
- Definování a výpočet metrik (ukazatelů).

4.2.1 Určení podnikových procesů pro řešení

Určení podnikových procesů, které mají být předmětem řešení PBI projektu, **vychází ze stanovených priorit v podniku**. Prioritní jsou obvykle ty byznys procesy, které pokrývají **hlavní problematiku** nebo potřeby řízení firmy a na druhé straně **jsou k dispozici potřebná data a kapacity** pro jejich řešení. Stanovení priorit by mělo být **založeno na konsensu** v rámci BI týmů, a to na řešitelské, ale zejména uživatelské straně.

4.2.2 Určení granularity, resp. podrobnosti dat

Granularita dat znamená úroveň detailu, na které jsou data v tabulce zaznamenána. Jinými slovy, jak „jemné“ nebo „hrubé“ jsou jednotlivé řádky dat. **Úroveň detailu** dat v datasetu, resp. sémantickém modelu, **ovlivňuje detail analytických úloh** a současně objem dat, uložených v sémantickém modelu.

Vysoká granularita:

- Každý řádek představuje jednotlivou událost nebo transakci. Např. každá položka na účtence v obchodě, každé kliknutí na webu, každá minuta sledování videa.
- Výhody: umožňuje detailní analýzu, sledování trendů, chování uživatelů.
- Nevýhody: velké množství dat, náročnější na výkon a údržbu.

Nízká granularita:

- Každý řádek představuje souhrn více událostí. Např. denní tržby za obchod, měsíční návštěvnost webu, průměrná doba sledování videa za týden.
- Výhody: menší objem dat, rychlejší zpracování.
- Nevýhody: méně detailů, omezené možnosti analýzy.

S uvedenými poznámkami souvisí i **tyto charakteristiky řešení**:

- Tabulky jsou rozděleny na **faktové** tabulky a **dimenzionální** tabulky.

- Každý **záznam (řádek) faktové tabulky** reprezentuje **byznys aktivitu nebo určitý sledovaný fakt o nějakém objektu** (např. cena celé objednávky nebo ceny jednotlivých zboží na řádcích objednávky, klik na banner na webu).
- Každý **záznam dimenzionální tabulky** reprezentuje **určitou sledovanou vlastnost o nějakém objektu** (například popisné údaje o právě jednom produktu z katalogu zboží).
- Každý řádek faktové tabulky musí vždy **mít stejný význam** včetně odkazů na dimenze.
- Je nutné **určit, co každý řádek** ve faktové tabulce **bude reprezentovat**, toto odsouhlasit s byznysem.
- Je možné, že jednotlivé řádky ve zdrojových datech obsahují **více aktivit nebo objektů** (např. objednávka a dodávka). V tom případě se doporučuje tabulky **rozdělit**.
- **Vysoká úroveň** detailu umožňuje vyšší možnosti analytiky, obvykle i s více dimenzemi, při **nízké úrovni** detailu je nutné ověřit s byznysem, že takové řešení vyhovuje potřebám.

4.2.3 Identifikace dimenzí

Dimenze popisují **kontext byznysových událostí**. Například ve firmě zabývající se prodejem zboží dimenze pomáhají odpovědět na otázky typu „*kdy k prodeji došlo*“, „*kdo zboží koupil*“, „*co přesně bylo prodáno*“, „*kde se prodej uskutečnil*“, „*kdo byl prodejce*“? Dimenze umožňují **filtrovat, seskupovat a porovnávat** fakta podle různých pohledů.

Příklady dimenzí:

- Produkt: katalogové číslo produktu, název produktu, značka, kategorie, výrobce.
- Zákazník: IČ, název zákazníka, typ zákazníka (B2B / B2C), region.
- Datum: den, měsíc, kvartál, rok.
- Prodejna: pobočka, prodejce, kanál prodeje (e-shop, kamenná prodejna).

Při identifikaci dimenzí je nutné **brát v úvahu:**

- Navrhované dimenze je účelné posuzovat **z pohledu možností filtrování dat v reportech**.
- Je nutné určovat, **které dimenze** jsou pro řešení **zásadní** a kontrolovat, zda jsou pro ně data v datových zdrojích.
- Je třeba posuzovat, které dimenze lze **pro filtrování dat využívat současně**.
- Při **nízké úrovni granularity** faktové tabulky se obvykle pracuje **s menším počtem dimenzí**.
- Je nutné **udržovat řízení verzí** u dimenzionálních tabulek a jejich atributů (pokud se údaje v dimenzích mohou v čase měnit).
- Pro dimenzionální tabulky se **doporučují denormalizovaná star schémata**, která zejména zvyšují výkon při dotazech do sémantických modelů.
- Pro kvalitní návrhy, řešení a úplnost systému dimenzí je účelné **aplikovat ve firmách procesy a zavádět řešení MDM (Master Data Management)**.

4.2.4 Definování faktů, resp. ukazatelů

Fakta představují **konkrétní události**, které chceme sledovat a vyhodnocovat. Jsou to **číselné hodnoty**, které odpovídají na otázky typu „*kolik zboží bylo prodáno*“, „*jaký byl obrát*“, „*jaký byl zisk*“, „*kolik objednávek bylo vyřízeno*“?

Příklady faktů ve firmě prodávající zboží:

- Počet prodaných kusů.
- Tržby v Kč.
- Sleva poskytnutá zákazníkovi.
- Náklady na dopravu.
- Marže.

Tyto údaje se ukládají do faktových tabulek. **Faktová tabulka obsahuje čísla** (např. tržby) a **odkazy na dimenze**. Díky tomu může uživatel snadno analyzovat například tržby podle produktů, počet objednávek podle prodejen, marži podle prodejců, vývoj prodeje v čase.

Fakta reprezentují **numerické sloupce** v rámci faktových tabulek. Nad hodnotami sloupců faktů se pak, kromě prostého zobrazení, **realizují agregace, průměrování hodnot a další** výpočetní operace v podobě definice metrik.

Pojmy **sloupec (column)** a **metrika (measure)** často obsahují čísla, vypadají tedy podobně, ale **jejich význam a použití jsou v byznysu zásadně odlišné**.

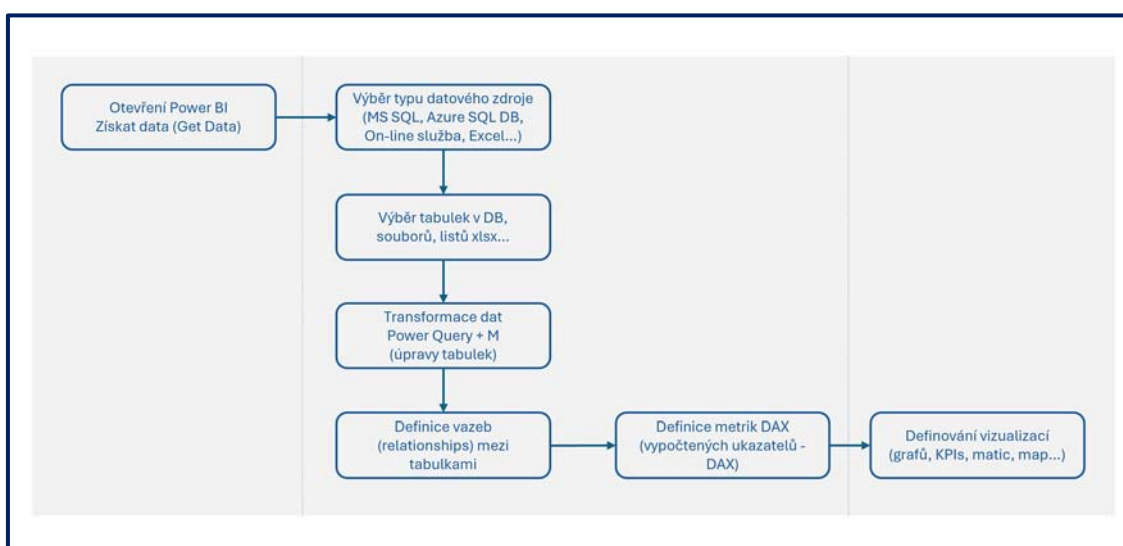
- **Sloupec je uložená hodnota** v tabulce. Je to konkrétní údaj, který je zaznamenán pro každou transakci nebo řádek. Představuje surová data, která popisují jednotlivé byznys události (viz výše, např. „kolik kusů bylo prodáno, za jakou cenu, s jakou slevou“).
- **Metrika je vypočítaná hodnota**, která vzniká na základě sloupců. Metriky se vypočítávají podle jejich **vzorců a v kontextu použitých filtrů** až při jejich použití (např. v reportu nebo vizualizaci). Hodnoty metrik nejsou tedy nikde ve faktových tabulkách uloženy, v sématickém modelu jsou pouze jejich definice.

V souvislosti s faktovými **tabulkami je tedy nutné řešit**:

- Jak jasné **definovat byznys logiku (obsah, význam)** pro jednotlivá fakta, resp. **kalkulace pro metriky** (ukazatele). Je účelné **k tomu udržovat kvalitní dokumentaci** a metadata.
- Je nezbytné detailně **verifikovat byznys logiku** fakt a faktových tabulek **s byznysem**.
- Je nutné **analyzovat dostupnost a kvalitu datových zdrojů** vzhledem k definovaným faktovým tabulkám.

4.3 Postup vytvoření sémantického modelu a aplikace v Power BI desktop

Postup vytvoření datové sady a aplikace v prostředí Power BI Desktop v případě zvoleného režimu práce „import“ dokumentuje Obrázek 4-1.



Obrázek 4-1: Postup vytvoření sémantického modelu a reportu v prostředí Power BI Desktop

Operace **výběr typu datového zdroje a jeho import** do tabulek v Power BI se opakuje postupně pro všechny datové zdroje, které je potřeba využít. Následně probíhá **úprava načtených dat** (pro účely transformace jsou načteny vždy jen vzorky řádků každé tabulky) **do požadované formy dimenzí a faktů v Power Query**, tj. úprava názvů tabulek, sloupců, filtrování – odstranění nepotřebných sloupců či řádků, rozdělení nebo naopak sloučení sloupců, úprava formátů, doplnění vypočtených sloupců, odstranění duplicitních řádků v dimenzionálních tabulkách, nahrazení vazebních atributů (business klíčů) umělými (*surrogate*) klíči typu integer, které budou použity pro vazby mezi dimenzemi a fakty a řada dalších operací. Následně **se definují relace mezi tabulkami dimenzí a faktů (1:M)**. Po vytvoření sémantického modelu je pak již možné vytvářet jednotlivé stránky reportu, obsahující kombinace vizuálů v podobě různých typů grafů, tabulek, matic, KPIs, map a dalších.

4.3.1 Zdroje dat pro vytvoření sémantického modelu v Power BI desktop

Power BI **podporuje širokou škálu datových zdrojů**, jak již bylo zmíněno v úvodu. Data lze importovat ze souborů, relačních nebo dimenzionálních databází vlastních podnikových nebo cloudových jako je Microsoft Azure, ale i přímo z množství on-line služeb. Microsoft neustále přidává další a další typy datových zdrojů, jejich výčet lze nalézt na stránkách Microsoftu (například <https://learn.microsoft.com/en-us/power-bi/connect-data/desktop-data-sources>).

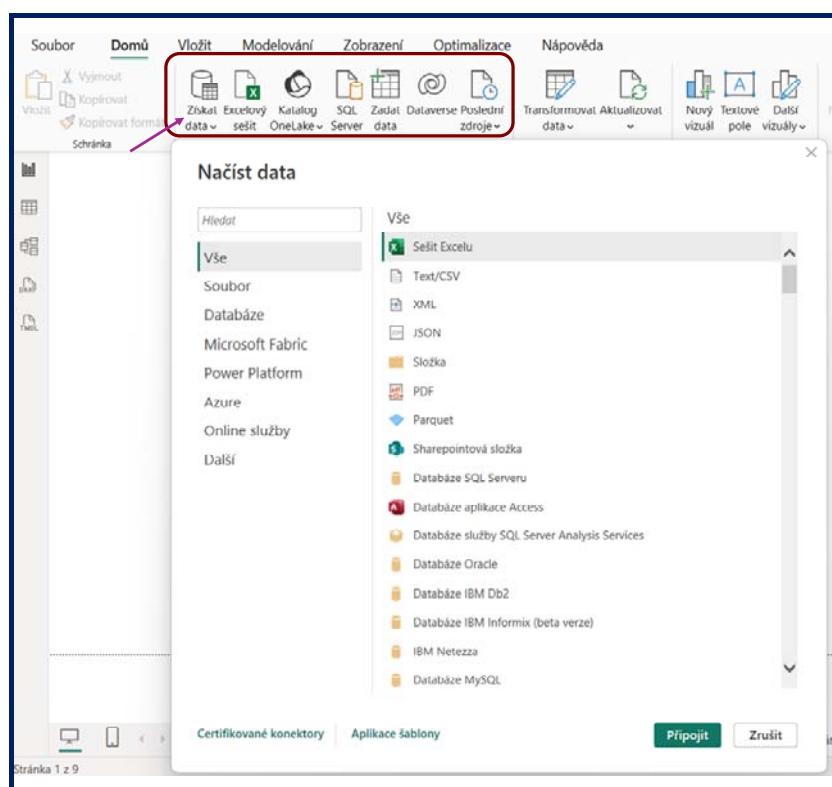
Hlavní skupiny konektorů do datových zdrojů jsou:

- Soubory (např. Excel, Text/CSV, XML, JSON, PDF, Parquet, folder).
- Databáze (konektory na více než 40 DB systémů).
- Fabric (např. PowerBI sémantické modely, datové služby, zahrnuté do Microsoft Fabric Synapse Data Warehouse Data Marts, nebo Data lake storage).
- Power Platform (např. datawarehouse nebo dataflows).
- Azure (konektory na řadu Azure služeb).
- Online služby (konektory na desítky online služeb včetně SaaS aplikací nebo sociálních sítí).
- Ostatní (desítky konektorů například na web, SharePoint seznamy, API OData Feed, Active Directory, Microsoft Exchange, Hadoop File (HDFS), ODBC a mnoho dalších).

Aktuální seznam služeb je nabízen v Power BI Desktop ve volbě *Get Data*.

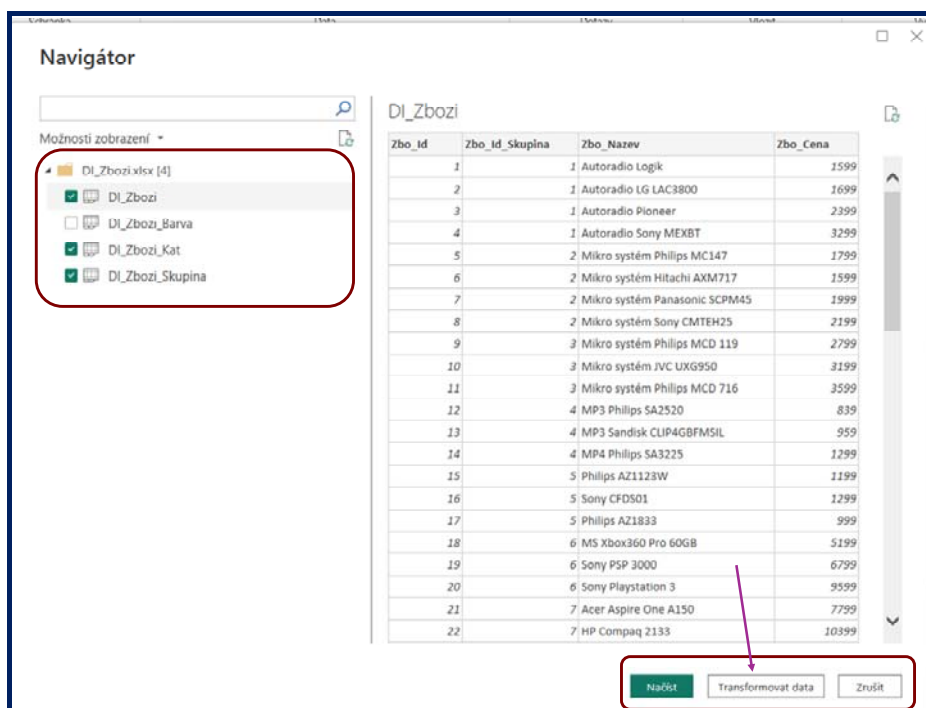
4.3.2 Import dat z datových zdrojů a jejich transformace

Pro **načtení dat z datových zdrojů** do Power BI Desktop slouží prvotně **funkce** pro vybrané vstupy dat, jak ukazuje Obrázek 4-2, ikona **Získat data** (*Get Data*) s nabídkou všech dostupných konektorů na datové zdroje a služby. Pro urychlení jsou samostatně v menu vyjmenované **často využívané zdroje**, tj. Excelový sešit, One Lake katalog, SQL Server, Dataverse a případně možnost Zadat vlastní data manuálně (Enter data).



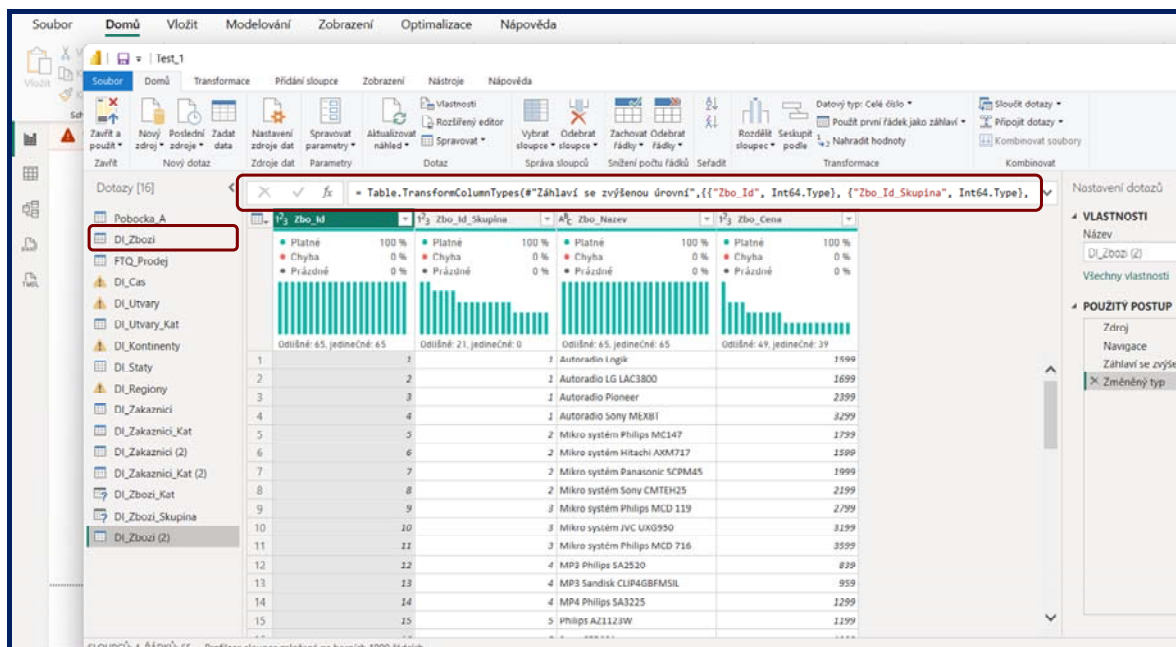
Obrázek 4-2: Přidání dat do Power BI

V následujícím příkladu se data importují ze souborů Excel (Obrázek 4-3.)



Obrázek 4-3: Výběr zdroje dat pro import do Power BI Desktop s přechodem na transformaci

Pokud je určen zdroj dat, v dalším kroku se určuje **konkrétní soubor v konkrétním adresáři**, nebo se definuje **připojení** k databázi či on-line zdroji.



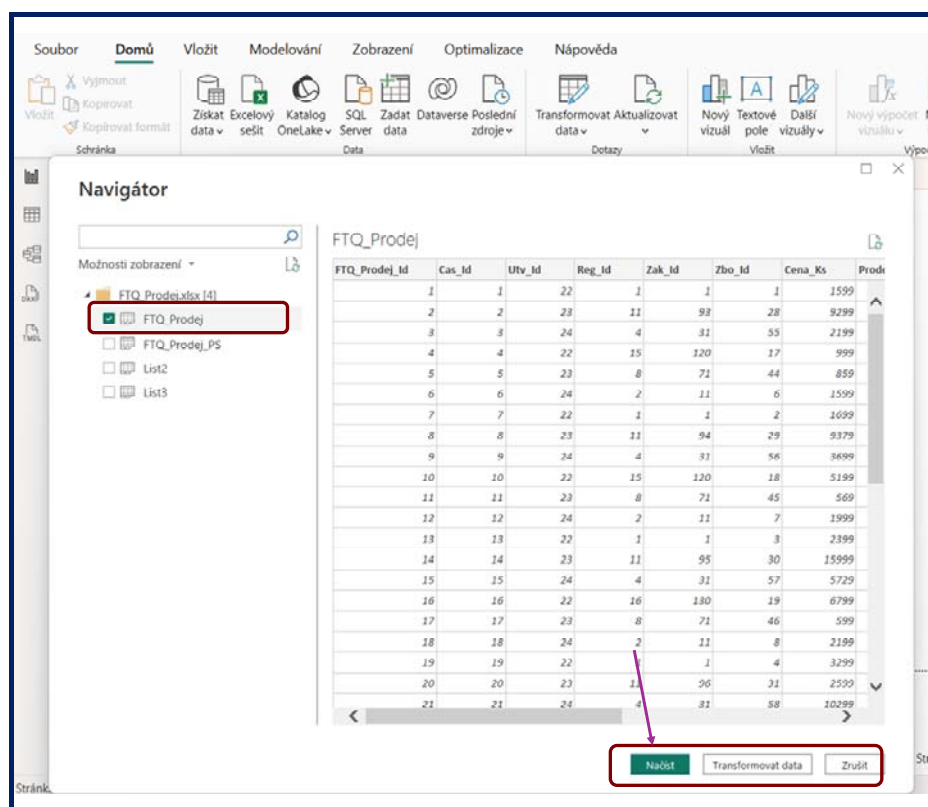
Obrázek 4-4: Power Query editor Power BI Desktop

V případě zdrojových dat v **Excelu** je dále třeba **vybrat i příslušné listy**, kde jsou potřebná data umístěna, u **databázových zdrojů** jde o **výběr datových tabulek**. Pro databázi se po připojení k databázi a zadání příslušných oprávnění vybírají databázové tabulky. Pro každý typ zdroje je **postup závislý na typu zdroje**, postup je řízen aplikací PBI.

V příkladu uživatel vybere potřebné soubory a zvolí **Transformovat data (Transform Data)**. Tím jsou načteny vybrané entity (datové tabulky) ve shodné struktuře, jako tomu bylo ve vstupních souborech a je otevřen nástroj **Power Query** pro transformaci dat. Volbu **Load** nedoporučujeme, vždy je vhodné zkontrolovat, jak data vypadají a upravit je do dimenzionálního modelu.

Spuštěním funkce **Transformovat data** otevře Power BI **Power Query editor**. Zde je možné data **upravovat, formátovat a vybírat jen potřebné sloupce a řádky** do podoby, která je potřeba pro tvorbu sestav. Jedná se v podstatě o **funkcionalitu ETL nástroje**, ale velmi uživatelsky příjemného a intuitivního, kdy většinu transformací lze jednoduše naklikat z široké nabídky transformačních funkcí. Uživatel má i možnost složitější transformace naprogramovat **v jazyce M**, který je nedílnou součástí Power Query.

Vznik a obsah faktové tabulky, jejímž zdrojem je rovněž list v Excelu, ukazuje (Obrázek 4-5).



Obrázek 4-5: Načtení faktové tabulky

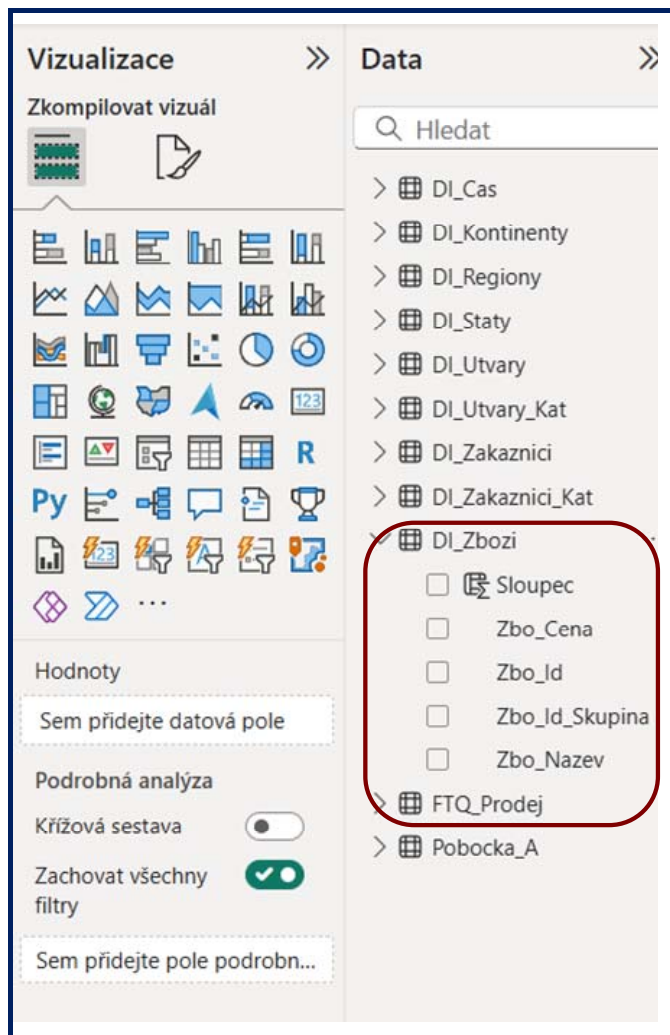
V případě vstupu z databáze se po výběru příslušné databáze **zobrazí seznam všech DB tabulek** a je možné si prohlédnout jejich strukturu (sloupce) a obsah několika prvních řádků.

Před definováním vztahů mezi tabulkami je vhodné tabulky upravit. **Power Query** nabízí širokou **škálu formátovacích funkcí**, jak je vidět z nabídky na obrázku (Obrázek 4-4). Další úpravy lze provádět s využitím jazyka DAX. Všechny **úpravy tabulek** lze provádět v **Power Query editoru**¹. Jedním

¹ Pokud není Power BI používán v režimu Direct Query (přímého dotazování). V tomto případě je možné veškeré změny provádět jen na úrovni datového modelu v příslušném prostředí, které Direct Query umožňuje. Stejně tak i v případě živého připojení (live connection) k jinému sémantickému modelu například tabulárnímu modelu MS Analysis Services, nebo AAS nebo Power BI sémantickému modelu je nutné upravovat přímo tento model, například v Microsoft Visual studiu, nebo Tabular Editoru (a tedy v Power úpravy provádět nelze).

z prvních kroků je **pojmenování sloupců**, i když jejich názvy je možné měnit i kdykoliv později, typicky při tvorbě reportů. Ve většině případů se **názvy sloupců automaticky přebírají z pojmenování v datových zdrojích**. Je však vhodné využít možnosti **přejmenování sloupců tabulek** tak, aby ve výsledných reportech lépe odpovídaly potřebám uživatelů.

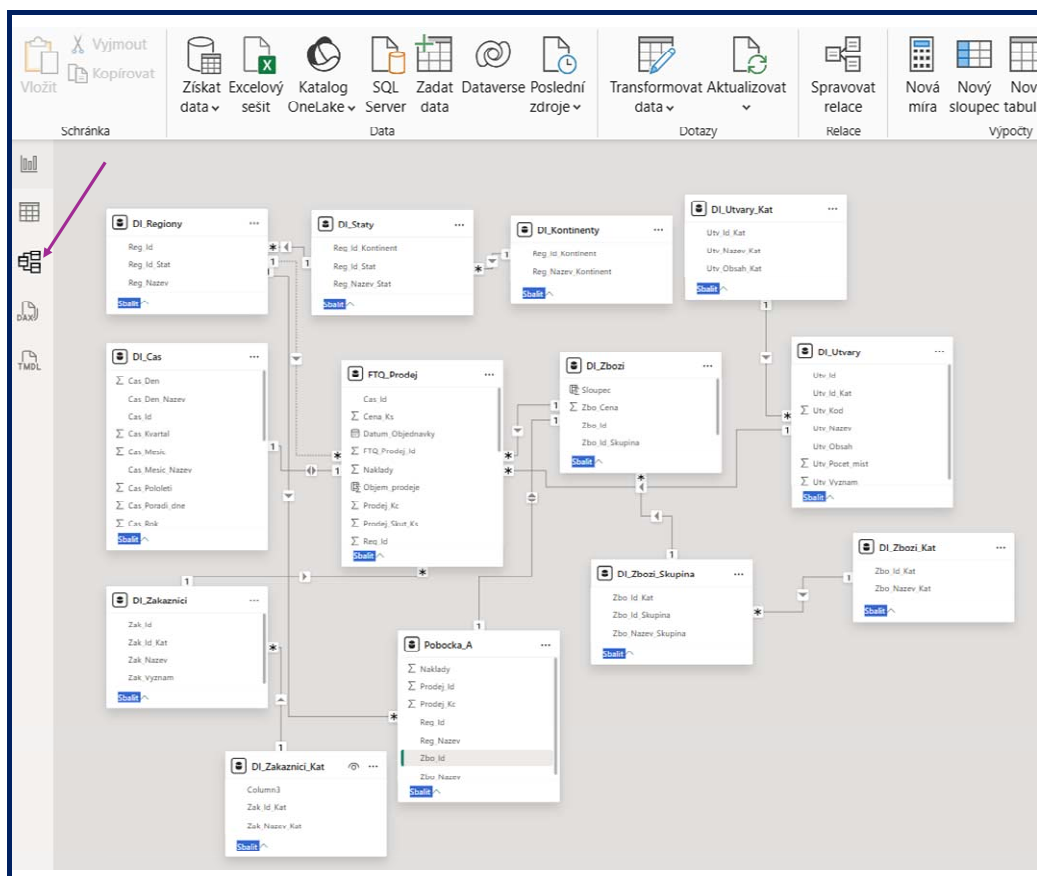
Obrázek 4-6 pak dokumentuje zobrazení přehledu všech načtených tabulek.



Obrázek 4-6: Zobrazení načtených položek

4.3.3 Definování relací mezi tabulkami

Tabulky se stejně jako ve všech aplikacích BI rozlišují na faktové a dimenzionální. **Vztahy** mezi tabulkami **se vytvářejí automaticky při importu dat** z datových zdrojů, pokud Power BI relace rozpozná, tj. tabulky obsahují **stejně názvy a datové typy** atributů. Prvotní návrh po načtení z datových zdrojů uváděného příkladu zobrazuje Obrázek 4-7. Je důležité vždy **zkontrolovat**, zda vztahy vytvořené automaticky dávají smysl, zda nevznikly vztahy M:N nebo vztahy mezi faktovými tabulkami. Model je vždy nutné upravit do dimenzionální podoby a doplnit potřebné náležitosti **manuálně**.



Obrázek 4-7: Sémantický datový model po úpravě v Power Query a načtení datových zdrojů do Power BI

Přehled veškerých vytvořených relací poskytuje funkce **Spravovat relace** (Manage relationships), jak dokumentuje Obrázek 4-8.

Z: tabulka (sloupec)	Relace	Do: tabulka (sloupec)	Stav
DI_Regiony (Reg_Id_Stat)	1 to many	DI_Staty (Reg_Id_Stat)	Aktivní
DI_Staty (Reg_Id_Kontinent)	1 to many	DI_Kontinenty (Reg_Id_Kontine...)	Aktivní
DI_Utvary (Utv_Id_Kat)	1 to many	DI_Utvary_Kat (Utv_Id_Kat)	Aktivní
DI_Zakaznici (Zak_Id_Kat)	1 to many	DI_Zakaznici_Kat (Zak_Id_Kat)	Aktivní
FTQ_Prodej (Cas_Id)	1 to many	DI_Cas (Cas_Id)	Aktivní
Σ FTQ_Prodej (Reg_Id)	1 to many	DI_Regiony (Reg_Id)	Neaktivní
FTQ_Prodej (Utv_Id)	1 to many	DI_Utvary (Utv_Id)	Aktivní
FTQ_Prodej (Zak_Id)	1 to many	DI_Zakaznici (Zak_Id)	Aktivní
FTQ_Prodej (Zbo_Id)	1 to many	DI_Zbozi (Zbo_Id)	Aktivní
Pobocka_A (Reg_Id)	1 to many	DI_Regiony (Reg_Id)	Aktivní
Pobocka_A (Zbo_Id)	1 to many	DI_Zbozi (Zbo_Id)	Aktivní

Obrázek 4-8: Správa relací

V tomto případě lze **relace aktualizovat a upravovat** (druhá možnost je v grafickém zobrazení, Obrázek 4-7.), a také vytvářet manuálně další relace s využitím funkce „**Nový vztah**“ (*New relationship*) na předchozím obrázku. Možnosti vytvoření takové relace dokumentuje následující obrázek (Obrázek 4-9).

← Nový vztah

Vyberte tabulky a sloupce, které spolu souvisejí.

Z tabulky

FTQ_Prodej

Skut_Ks	Reg_Id	Rok	Utv_Id	Zak_Id	Zbo_Id
	1	2015	22	1	1
	11	2015	23	93	28
	4	2015	24	31	55

Do tabulky

DI_Zbozi

Sloupec	Zbo_Cena	Zbo_Id	Zbo_Id_Skupina	Zbo_Nazev
79	1599	1	1	Autoradio Lo...
79	1699	2	1	Autoradio LG ...
79	2399	3	1	Autoradio Pio...

Kardinalita

N:1

Směr křížového filtru

Jednoduché

☒ Aktivovat tuto relaci

☐ Použít filtr zabezpečení v obou směrech

☐ Předpokládat referenční integritu

Uložit **Zrušit**

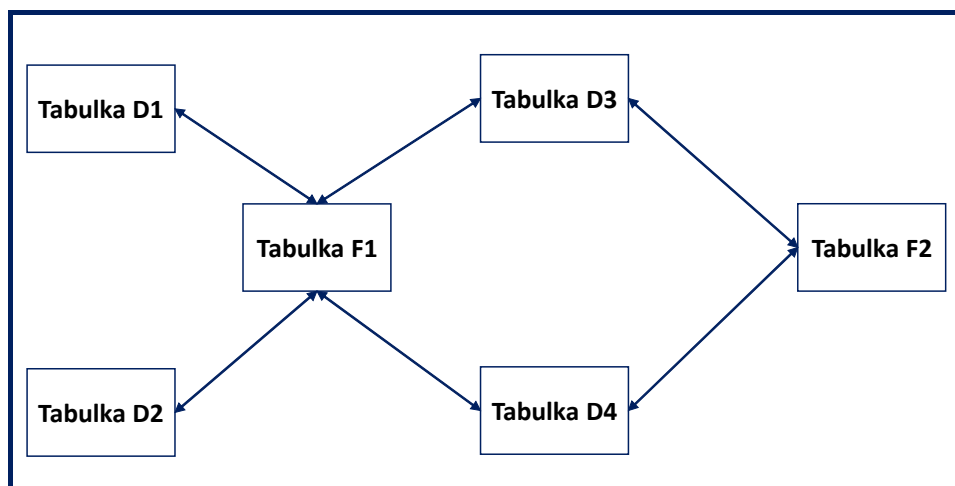
Obrázek 4-9: Vytvoření vztahu mezi entitami

Power BI navrhne automaticky i **kardinalitu a způsob filtrování záznamů** v provázaných tabulkách (entitách) **prostřednictvím dané relace**. Jsou možné **relace Many to One** (* : 1), **One to One** (1:1) a **One to Many** (1:*) a **filtrování Jednoduché** (Single), nebo **Obousměrné** (Both).

Po určení kardinality a navigace pro filtrování je ještě možné **určit, zda je vztah aktivní**. Pokud je volba zaškrtnuta, znamená to, že aktivní vztah je výchozím vztahem pro reportování a cross filtrování mezi vybranými tabulkami. V případech, kdy existuje více než jeden vztah mezi danými dvěma tabulkami, Power BI Desktop aktivní vztah **použije automaticky při vytváření vizualizací**, v nichž jsou použity obě tabulky.

Pokud chce uživatel **využít jiný (neaktivní) vztah**, je nutné ho pro příslušnou akci (například výpočet ukazatele) **aktivovat** (buď nejprve původní aktivní vztah deaktivovat a poté zvolit jako aktivní vybraný jiný vztah anebo využít DAX funkci „USERELATIONSHIP“). Mezi dvěma tabulkami může být **aktivní** v jednu chvíli **pouze jeden vztah**.

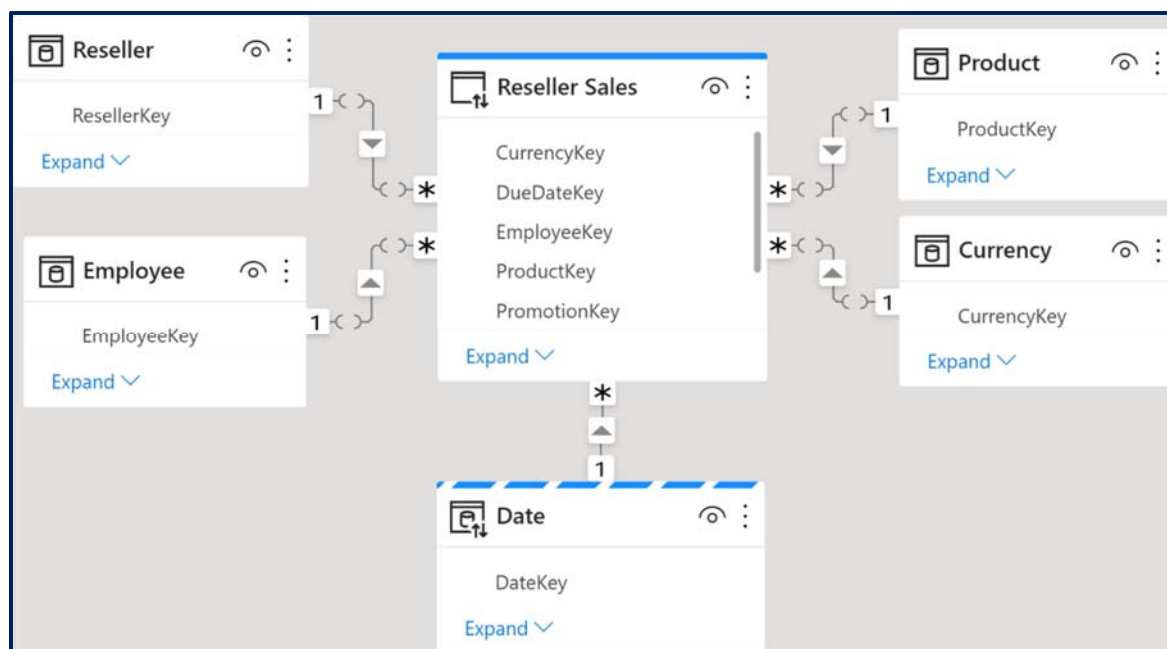
Nastavení **obousměrného filtrování (Both)** umožňuje Power BI Desktop **pracovat s provázanými tabulkami, jako kdyby byly jedinou tabulkou**. Křížové filtrování *Both* funguje dobře **při využití STAR schéma**. Existují však situace, **kdy obousměrné nastavení nelze použít**, protože by vedlo k nejednoznačným výsledkům při reportování. Například v případě modelu, uvedeného na následujícím obrázku (Obrázek 4-10), což je struktura běžné relační databáze. Proto obecně **nastavení obousměrného filtrování** v sémantických modelech **není doporučeno**. Vždy je nutné být velmi obezřetný při jeho výjimečném použití.



Obrázek 4-10: Nevhodně navržené schéma databáze v Power BI pro použití „Both“ filtrování

Pokud by byl použit souhrn vybrané položky (sloupce) z faktové tabulky F2 a pak by bylo vybráno filtrování pro vybranou položku faktové tabulky F1, **nebylo by zřejmé, kterou cestou má filtr zafungovat**, zda přes horní dimenzi či spodní dimenzi (tabulku D3 nebo D4). Typickým příkladem pro tento vzor je situace, kdy tabulka F1 je např. tabulkou prodeje a tabulka F2 je tabulkou plánovaných výnosů, dimenze D3 představuje například region a D4 zboží. Pokud by uživatele zajímalo, jak je plněn plán výnosů prodejem, je takový model nepoužitelný. Řešením je buď jeden ze vztahů na dimenzi zrušit či zneaktivnit, nebo dimenzionální tabulky duplikovat (s jiným názvem). Takovým dimenzím se říká „**role playing dimenze**“.

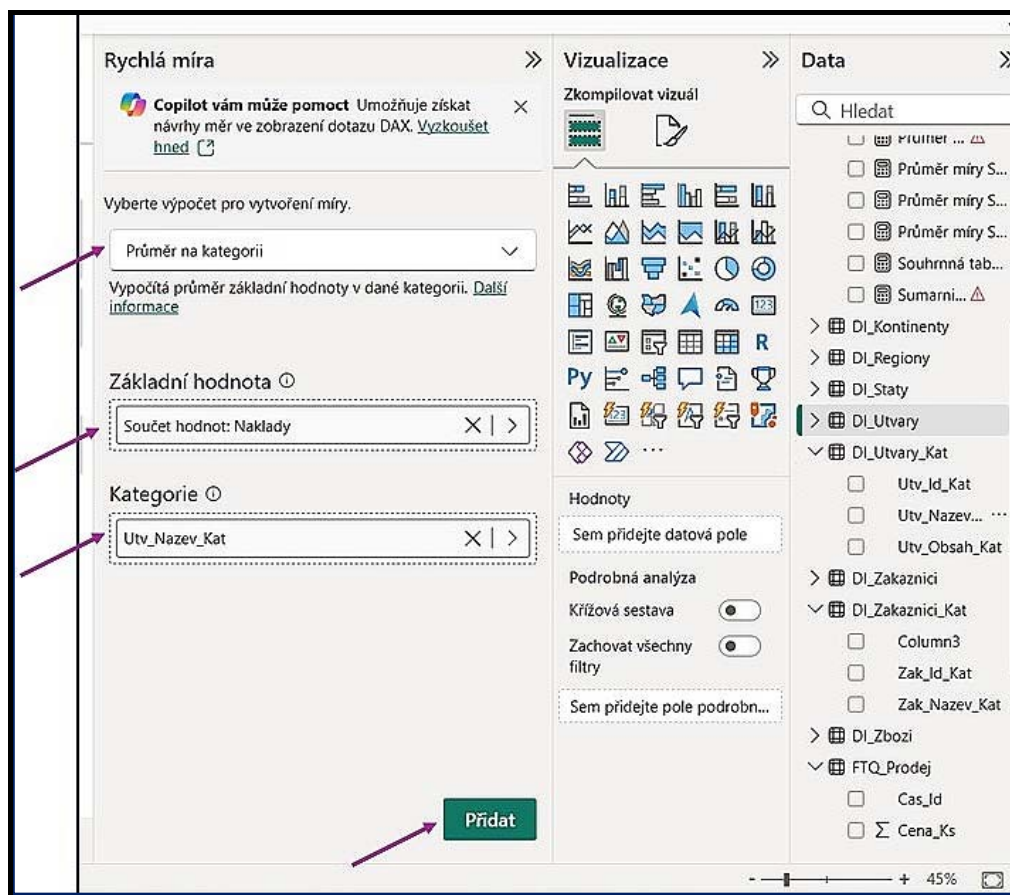
V případě použití kompozitního modelu jsou tabulky na bázi režimu Direct Query označeny v záhlaví plnou silnou čarou (*Reseller Sales*, viz další příklad), duální tabulky podporující import dat i Direct Query čárkovanou čarou, tabulky získané importem dat jsou standardní, bez označení. To dokumentuje Obrázek 4-11.



Obrázek 4-11: Řešení kompozitního modelu (Zdroj: Decler, Powell, 2022)

4.3.4 Rychlá míra

Pokud jsou vytvořeny všechny potřebné relace, je možné **doplnit** potřebné **vypočítané ukazatele (measures)**, a to buď s využitím možností nabídky Power BI desktop, a to **funkcí „Rychlá míra“**, nebo je definovat **prostřednictvím jazyka DAX**.

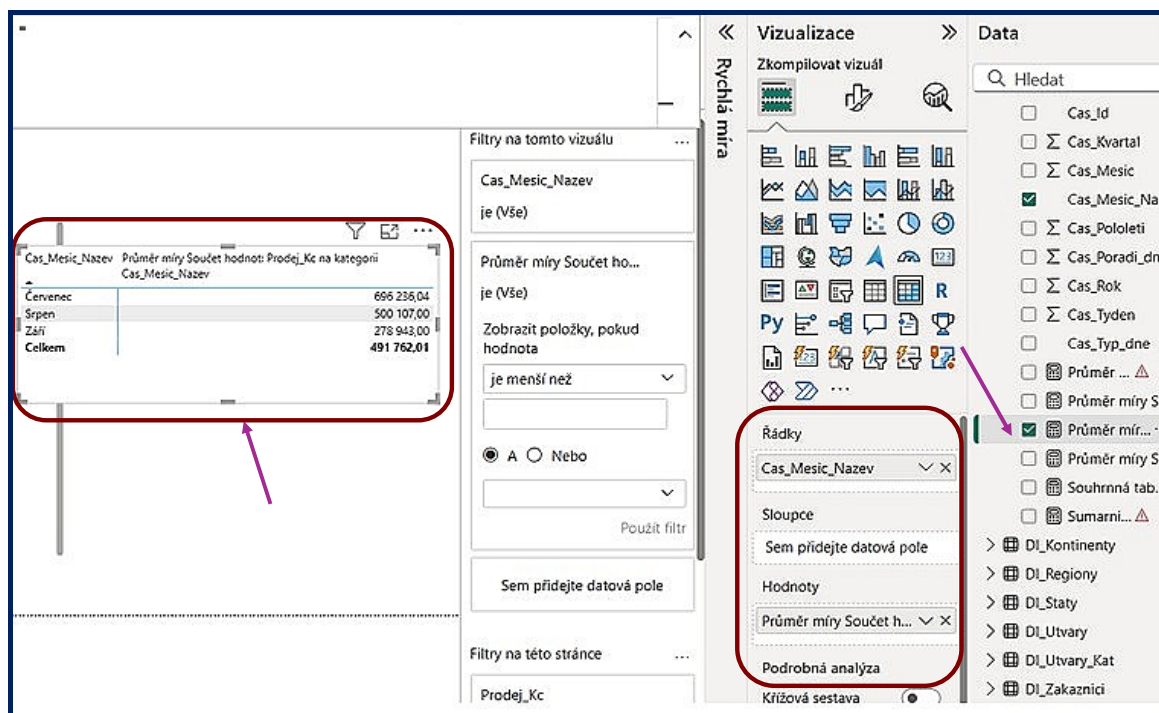


Obrázek 4-12: Příklad pro vytvoření rychlé míry

Prostřednictvím funkce „**Rychlá míra**“ lze realizovat **ukazatele podle kategorií** (průměry, minima, maxima, vážené průměry), vytvářet **filtrované ukazatele** (hodnoty podle filtru, rozdíly hodnot k vybrané hodnotě podle filtru a procentní rozdíly), vytvářet **ukazatele s využitím Time Intelligence** (výpočty hodnot od začátku roku – YTD, od začátku čtvrtletí – QTD, od začátku měsíce – MTD apod.), dále lze využít **souhrny** (součty, rozdíly, součiny, podíly, procentní rozdíly), různé **matematické funkce**, nebo vybrané **textové funkce** a další. Power BI Desktop automaticky **generuje výpočty v DAXu**.

Více informací, vysvětlení a způsob využití uvedených skupin funkcí pro rychlý výpočet ukazatelů je k dispozici přímo na stránkách powerbi.com (<https://powerbi.microsoft.com/en-us/documentation/powerbi-desktop-quick-measures/>). Vypočtené ukazatele **lze přidávat kdykoliv** i v průběhu práce na reportech, stejně jako lze přidávat další vypočtené sloupce do databázových tabulek.

Následně je již možné nad takto definovanou datovou sadou (databází) **realizovat výsledné aplikace**, tj. jednotlivé sestavy, obsahující vhodné vizuály (výběry hodnot, grafy, tabulky, KPI, mapy a další) z nabídky Power BI či vizuály vlastní (*custom*). Kromě postupu, obsaženého v dalším textu této publikace, je více než vhodné **dodržovat doporučení pro vizualizaci a prezentaci dat**.



Obrázek 4-13: Příklad vytvořené rychlé míry

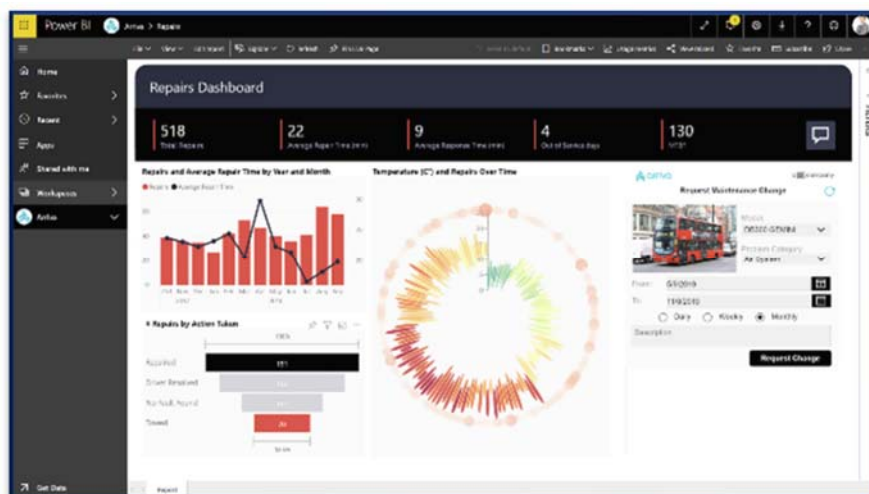
4.4 Pracovní závěry



Kapitola představuje **shrnutí nejpodstatnějších principů a obsahu sémantických modelů** Power BI. K tomu patří:

- Uvedený **doporučený postup řešení aplikací** je v daném případě **velmi zjednodušený** a bude dále doplněn o další aspekty, spojené s řešením projektů Power BI.
- Jednou z podstatných charakteristik Power BI jsou široké **možnosti vizualizace dat** pro nejrůznější účely prezentací a vytváření různých druhů reportů, dashboardů, převážně pro manažery a specialisty firmy.
- **Báze dat**, na které se uskutečňují prezentační a další operace Power BI, se označují jako **sémantické modely**, představované faktovými a dimenzionálními tabulkami a jejich relacemi. Fyzická realizace modelu, obsahující i data, se obvykle označuje termínem **dataset**, který budeme i nadále v dokumentu využívat společně s pojmem sémantický model.
- **Vytvoření sémantického modelu** zahrnuje identifikaci potřebných datových zdrojů, jejich nahrání do sémantického modelu a úpravy, nebo nastavení potřebných relací.

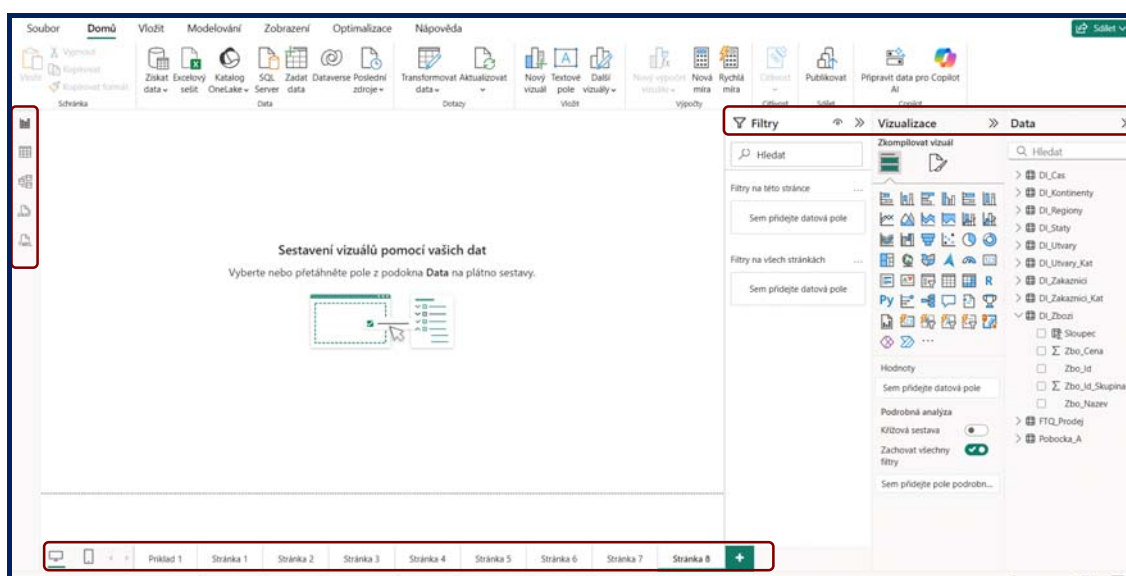
5. Reporty a vizuály v Power BI



Účelem kapitoly je v návaznosti na předchozí kapitolu doplnit další detaily k vytváření reportů a aplikací v Power BI. Text kapitoly využívá **obecné principy a termíny spojené s řešením BI, resp. SSBI** aplikací, které se zde neopakují, ale **odkazuje na speciální dokument** na portálu MBI-AF, věnovaný postupům řešení těchto úloh [Postupy].

Základem uživatelského rozhraní jsou **vizualizace dat (či vizuály)**, které umožňují interakci s daty, načtenými a uloženými v datasech aplikace Power BI, nebo s daty, uloženými v databázích, k nimž Power BI přistupuje v režimu přímého dotazování (*Direct Query*), viz dále. Vizualizace dat **se realizuje v sestavách (reportech)**. Pro publikování reportů do služby *Power BI Service* lze vybrané vizuály z reportů umístit na řídicí panely (dashboards) a ty sdílet vybraným skupinám uživatelů anebo jednotlivým uživatelům, kteří mají příslušnou licenci *Power BI Pro*, viz dále.

V Power BI může report (sestava) **obsahovat jednu nebo více stránek**. Všechny stránky dohromady jsou společně označovány jako **report**. Nová stránka se přidá klikem na „+“ vedle poslední záložky s názvem stránky. Stránky je možné **přejmenovat** stejně jako v excelovém sešitu, tj. pravým tlačítkem myši na záložce sestavy **funkci „Přejmenovat“**.



Obrázek 5-1: Pracovní plocha v Power BI desktop

Význam základních ikon pro účel a formát plochy (Obrázek 5-1, vlevo nahoře) je v následující tabulce:

	Zobrazení sestavy (Report view) slouží pro prezentaci uživatelského rozhraní, tedy reportů, jejichž základními prvky jsou již zmíněné vizuály , viz dále.
	Zobrazení tabulky (Data view) zajišťuje zobrazení importovaných řádků zvolené tabulky datasetu a současně odpovídající metadata jako počet řádků tabulky nebo počet unikátních hodnot ve sloupci. Metadata jsou zobrazena na spodní straně okna Data view, Počet unikátních hodnot je podstatný z hlediska vyšších nároků na datový prostor nebo času při realizaci dotazů.
	Zobrazení modelu (Model view) nabízí ekvivalent databázového diagramu odpovídajícího zobrazení tabulek a jejich vazeb datasetu, jak dokumentuje Obrázek 4-7.
	Zobrazení dotazu DAX slouží pro zobrazení výpočetních předpisů a dotazů v jazyce DAX.
	TMDL je zkratka pro Tabular Model Definition Language , což je skriptovací jazyk, který slouží k definování objektů v tabulárním datovém modelu. Power BI je postavené na tabulárním modelu, TMDL umožňuje pracovat s tímto modelem pomocí kódu místo grafického rozhraní.

5.1 Reporty v Power BI Desktop

Využití vizuálů pro tvorbu reportů se volí v pravém menu pracovní plochy Power BI, jak ukazuje Obrázek 5-1.

Základními **prvky reportu jsou vizuály** (či vizualizace), které se vkládají do sestavy výběrem z menu na pravé straně pracovní plochy (Obrázek 5-1, vpravo, **část Vizuály – Visualisation**). Kromě řady vizuálů, které nabízí přímo Power BI, lze přidávat i uživatelské vizuály, které jsou ostatními uživateli zpřístupněné, nebo lze vytvářet zcela **vlastní vizuály s využitím jazyka TypeScript**, uložit je a následně načíst do Power BI. Do sestavy se dají též **vkládat textová pole, obrázky a různé tvary** (viz příslušné funkce v menu).

Pro každou část reportu (vizuál, stránku nebo celou sestavu) je možné **nastavit filtry**, které se uplatní právě jen v kontextu vizuálu, stránky, nebo sestavy, pro který je definován. K definici filtrů slouží v pravé části pracovní plochy Power BI menu **Filtry – Filters** (Obrázek 5-1 nahoře).

Vše, co se na sestavě vyskytuje – tedy jednotlivé datové prvky, každý vizuál i každou stránku sestavy – lze **formátovat prostřednictvím řady funkcí** a voleb, které Power BI nabízí.

5.2 Tvorba vizuálů v sestavách

Vizuály jsou různé typy grafického zobrazení dat. Jejich vytváření je uživatelsky velmi jednoduché. Existuje **několik postupů, jak vizuály uplatňovat**. V následujícím textu jsou uvedeny **dva nejběžnější**.

První, **nejjednodušší způsob** je následující:

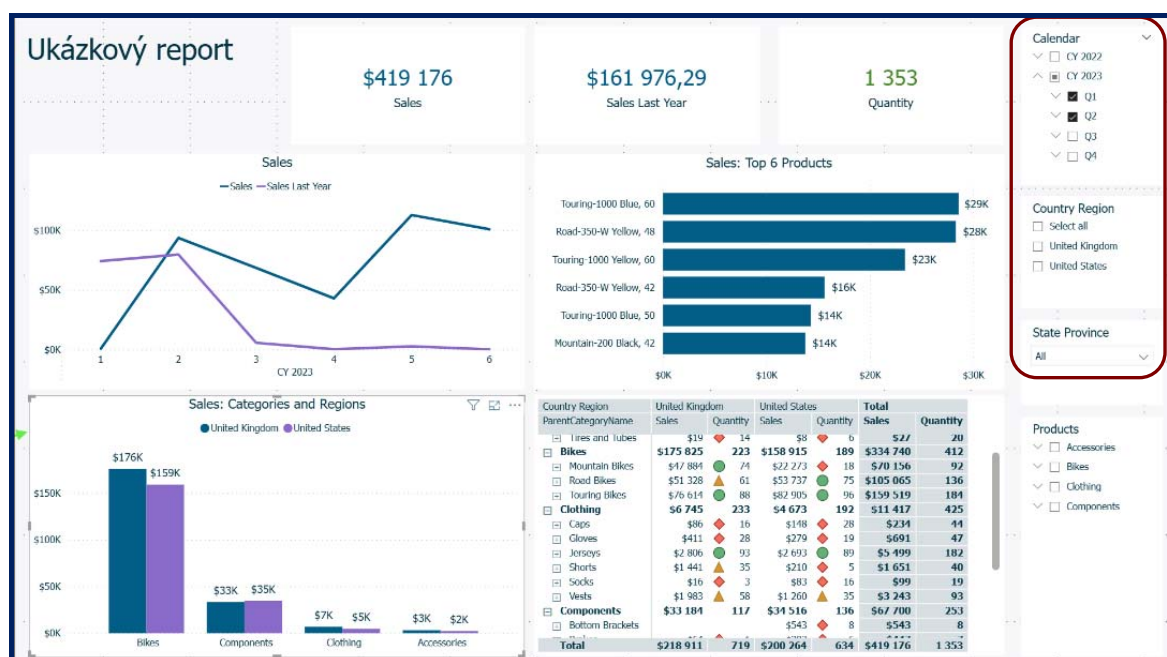
1. Uživatel vybere levým tlačítkem myši **pole z některé datové tabulky** (Obrázek 5-1, vpravo nahoře, sekce **Data**) a tažením myši pole umístí na plochu listu reportu.
2. **Power BI podle typu pole sama vybere nejvhodnější vizuální zobrazení** a umístí ho na dané pozici pracovní plochy listu reportu.

3. Dále lze do daného vizuálu **přidávat další pole, upravovat způsob zobrazení polí a agregovaných polí**, s použitím pravého tlačítka myši lze například vybrat, zda se má hodnota pole zobrazovat jako průměr, maximum, minimum apod., přidávat legendu, popisky (*tooltips*) dat a řadu dalších.
4. Pokud automaticky vybraný **vizuál nevyhovuje**, lze z **palety nabízených vizuálů vybrat jiný** (pouhým stiskem levého tlačítka myši) a Power BI vybraná pole zobrazí v tomto grafickém vyjádření (pokud vybraný vizuál pro zvolená pole dává smysl).

Druhým obvyklejším postupem je tento:

1. **Uživatel vybere** z palety dostupných **vizuálů (Vizualizace – Visualisation) jeden** a umístí prázdný vizuál na plochu listu sestavy.
2. Následně uživatel tažením myši **umístí do tohoto vizuálu potřebná pole** z datových tabulek (**Data**), která chce vybranou vizualizací prezentovat.
3. **Další parametry**, jako například výběr dimenzí pro osu grafu, podle nichž se mají hodnoty zobrazovat, výběr polí pro legendu, výběr polí, která se mají zobrazovat v *tooltipech*, výběr polí, která určují sytost zobrazovaných barev apod., se nastavují **přetažením příslušných polí do sekce parametrů pro pole (Hodnoty)**, umístěné pod výběrem vizuálů.

Formátování každého vizuálu je možné **na ikonkách pod názvem „Vizualizace– Visualisation“** a pro každý typ vizuálu je různě bohaté. Microsoft postupně s každou novou verzí Power BI přidává další a další možnosti úprav zobrazení (Obrázek 5-2).



Obrázek 5-2: Formátování vizuálů

S prezentací dat souvisejí i známé *tooltips*, což je nápověda, která se zobrazí, když uživatel najede kurzorem myši nebo klepne na určitý prvek na stránce či v aplikaci. Jde o **kontextovou nápovědu**, která slouží uživatelům k pochopení určité funkce daného prvku. Power BI je nabízí standardně a použití dokumentuje Obrázek 5-3:

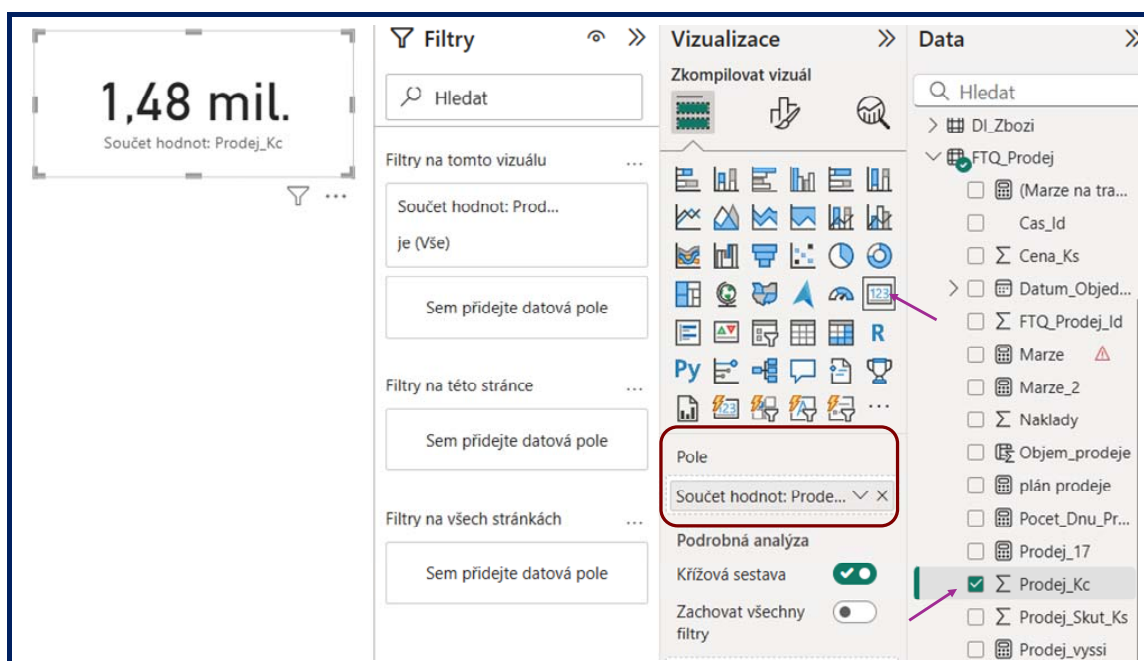


Obrázek 5-3: Použití tooltipů u vizuálu

5.2.1 Individuální hodnoty. Karta (Card)

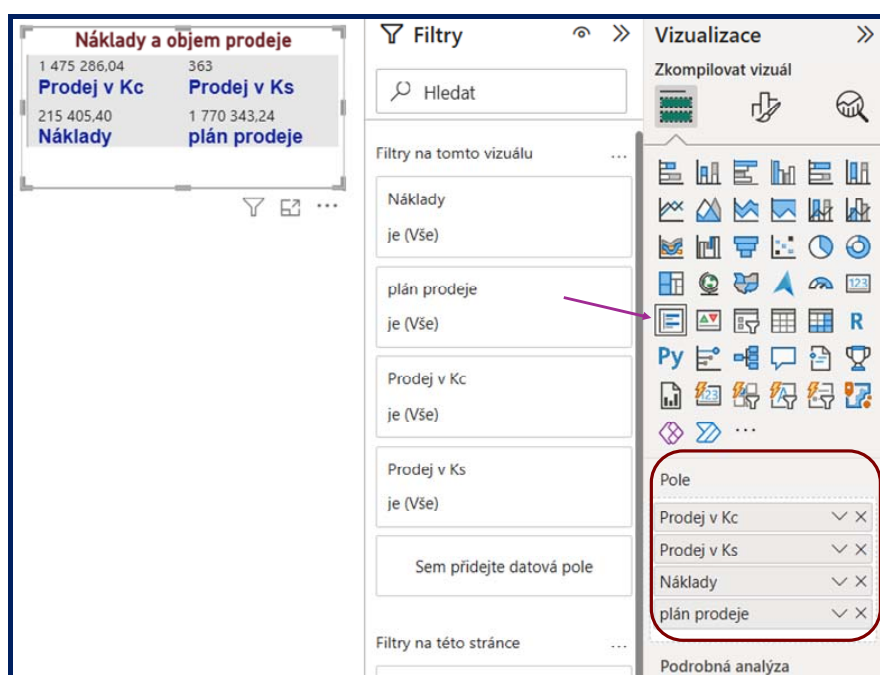
Karta zachycuje **nejvýznamnější ukazatele a jejich hodnoty**, jako např. objem prodeje v měně. Jsou obvykle umístěny **na horní nebo levé straně stránky reportu** a jsou i součástí dashboardu v rámci PBI služby. Jde o hodnotu z jednoho sloupce tabulky nebo míru v DAXu. Jsou často první informací, na kterou uživatel zaměří pozornost a podle ní se rozhoduje, zda bude sledovat i další zobrazené vizuály. Proto je účelné výběru těchto ukazatelů na kartách věnovat při analýze značnou pozornost. Využití karty ukazuje Obrázek 5-4.

Barvu karty lze podmíněně formátovat podle skutečné hodnoty. Může obsahovat i doprovodný text.



Obrázek 5-4: Individuální hodnota, Karta (Card)

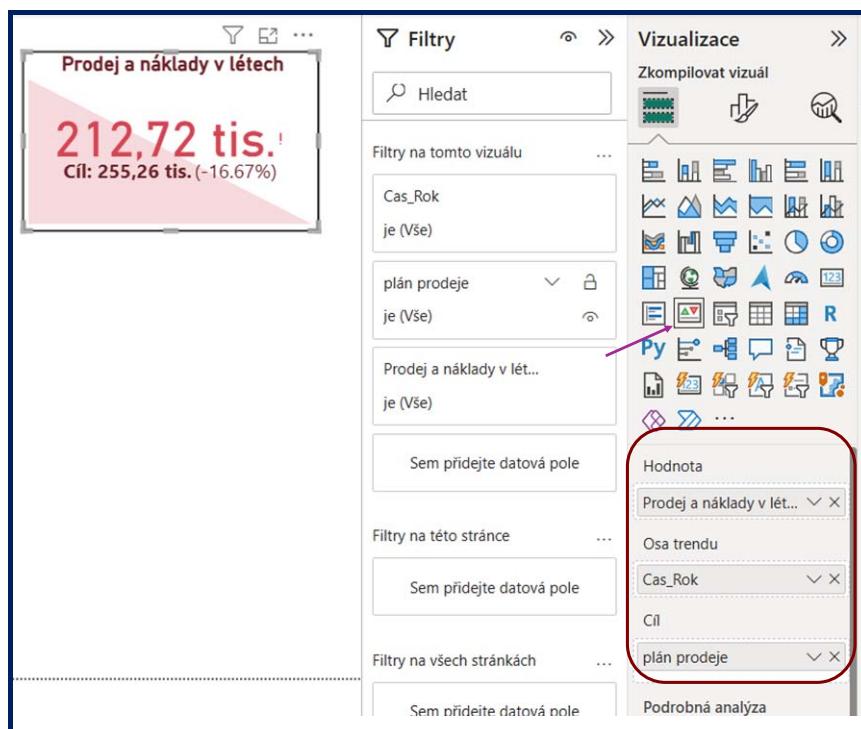
Další možností je využít kartu pro zobrazení více klíčových ukazatelů, jak demonstruje Obrázek 5-5:



Obrázek 5-5: Karta s více ukazateli

5.2.2 KPI, Klíčový ukazatel výkonu

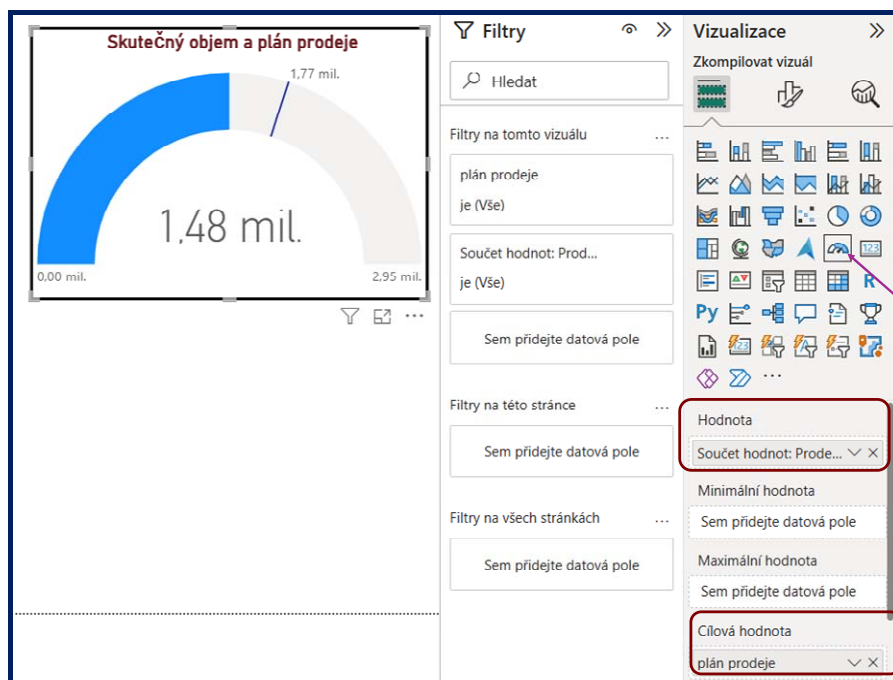
Vizuál „Klíčový ukazatel výkonu“ (*KPI, Key Performance Indicator*) je velmi často využíváný, neboť demonstruje **hlavní ukazatele pro byznys**, a to jejich **hodnotu**, **vývojový trend** v čase (v daném případě v letech) a vztah ke specifikované **cílové, resp. plánované hodnotě**. To dokumentuje Obrázek 5-6. Jak je z obrázku patrné, KPI v tomto případě není pozitivní.



Obrázek 5-6: KPI, Klíčový ukazatel výkonu

5.2.3 Měřidlo (Gauge)

„Měřidlo“ (*Gauge*) je obdobné vizuálu KPI. V daném případě je skutečná hodnota porovnávána s cílovou (např. plánovanou) hodnotou a tento vztah dokumentuje graficky, jak ukazuje Obrázek 5-7:



Obrázek 5-7: Měřidlo (Gauge)

5.2.4 Tabulky, matice

Tabulky a matice jsou vizuály, nabízející detailní přehledy hodnot zvolených ukazatelů podle vybraných dimenzí. To dokumentuje Obrázek 5-8, a to hodnoty tržeb podle jednotlivých zákazníků (v řádcích) a útvarů (ve sloupcích). Zatím co vizuál „Tabulka“ (*Table*) uvádí základní přehledy hodnot, „Matice“ (*Matrix*) nabízí přehled adekvátní kontingenčním tabulkám, tj. s využitím více dimenzí nebo struktur v řádcích nebo ve sloupcích a jejich vzájemné kombinace.

The matrix visualization displays the following data:

Zak_Nazev	Prodejní oddělení 11	Prodejní oddělení 12	Prodejní oddělení 13	Celkem
Akademik			155 966,89	155 966,89
Amondon	115 996,38			115 996,38
Bárta inženýring		13 293,00		13 293,00
Fiber log		15 594,00		15 594,00
Háček			168 762,00	168 762,00
ICM sklady CZ		56 893,01		56 893,01
Knihy všem	410 408,00			410 408,00
Kuhnhakel		161 749,89		161 749,89
Luxor			20 402,00	20 402,00
Pavelka - obuv		2 509,00		2 509,00
Pepo Tour		59 621,46		59 621,46
Síma Josef		44 780,00		44 780,00
Velká hypoteční	25 193,00		114 175,16	139 368,16
Celkem	587 174,47	428 805,52	459 306,05	1 475 286,04

Obrázek 5-8: Zobrazení matice hodnot objemů prodeje podle zákazníků a prodejních oddělení

Formu tabulky nebo matice lze upravovat, a to její jednotlivé části (hodnoty, záhlaví sloupců i řádků atd.) na základě volby ikony vpravo nahoře, jak ukazuje Obrázek 5-9.

Název zboží	Tržby	Marže_2 (Marže na transakci)	Pocet_Dnu_Prodeje
Acer Aspire 5730Z	83 645,00	76 953,40	76 953,40
Acer Aspire One A150	23 397,00	17 157,80	17 157,80
ASUS VW193B	11 277,00	8 269,80	8 269,80
Asus X51L	13 199,00	2 639,80	2 639,80
Autoradio LG LAC3800	10 306,13	7 909,63	3 954,82
Autoradio Logik	6 466,36	4 210,96	2 105,48
Autoradio Pioneer	24 253,89	20 870,09	10 435,05
Autoradio Sony MEXBT	19 794,00	17 154,80	17 154,80
BC LEXZ1380	6 495,00	5 975,40	5 975,40
Bosch 250	1 299,00	259,80	259,80
DLDAP1160	7 794,00	6 754,80	6 754,80
DWLGS10	10 354,66	9 552,06	4 776,03
Espresso Delonghi E2100	78 393,00	73 913,40	73 913,40
Espresso Delonghi ESAM2600	51 495,00	47 375,40	47 375,40
Fujitsu Siemens AMILO Pa 3515	45 597,00	39 517,40	39 517,40
HP Compaq 2133	72 793,00	64 473,80	64 473,80
HP Pavilion A6611CS	56 893,01	43 603,91	21 831,96
HP Photosmart D5460	5 697,00	4 177,80	4 177,80
Celkem	1 475 286,04	1 259 880,64	15 947,86

Obrázek 5-9. Formální úpravy tabulky / matice

Rychlou orientaci v datech umožňují **funkce podmíněného formátování** podle hodnot vybraného ukazatele jako jsou barvy hodnot, datové pruhy, ikony k hodnotám a další.

5.2.5 Podmíněné formátování, datové pruhy

Jednou z běžných možností podmíněného formátování je nastavení datových pruhů podle hodnot vybraného ukazatele. Zobrazení takto formátované matice dokumentuje Obrázek 5-10. Nastavení způsobu podmíněného formátování je při použití ikony fx (Obrázek 5-10, vlevo dole). V tom případě se otevře okno pro nastavení formátování, Obrázek 5-11.

Zbo_Nazev	Prodejní oddělení 11	Prodejní oddělení 12
Acer Aspire 5730Z	83 645,00	
Acer Aspire One A150	23 397,00	
ASUS VW193B		11 277,00
Asus X51L	13 199,00	
Autoradio LG LAC3800	10 306,13	
Autoradio Logik	6 466,36	
Autoradio Pioneer	24 253,89	
Autoradio Sony MEXBT	19 794,00	
BC LEXZ1380		6 495,00
Bosch 250		
DLDAP1160		7 794,00
DWLGS10		10 354,66
Espresso Delonghi E2100		
Espresso Delonghi ESAM2600		
Fujitsu Siemens AMILO Pa 3515	45 597,00	
HP Compaq 2133	72 793,00	
HP Pavilion A6611CS		56 093,01
HP Photosmart D5460		5 697,00
HPD12560		12 995,00
IP4600		19 593,00
Kodak BC		
Celkem	587 174,47	428 805,52

Obrázek 5-10: Podmíněné formátování s využitím pruhů podle hodnot

Datové pruhy - Datové pruhy

Styl formátování: Datové pruhy
Platí pro: Jen hodnoty

Které pole máme vzít jako základ?: Tržby
Směr pruhu: Zleva doprava

Minimum: Nejnižší hodnota
Maximum: Nejvyšší hodnota

Kladný pruh: ☒
Záporný pruh: ☐
Osa: ☒

☐ Zobrazit jenom pruh

[Další informace o podmíněném formátování](#)

OK Zrušit

Obrázek 5-11: Nastavení datových pruhů

5.2.6 Nastavení ikon k datům

Pro rychlou orientaci v jednotlivých datech v tabulce nebo matici lze také využívat **ikony nastavované na základě pravidel** ve vazbě k jejich hodnotám, resp. intervalům hodnot. Volbu a zobrazení ikon u jednotlivých hodnot v tabulce nebo matici dokumentuje Obrázek 5-12.

Nastavením ikon podle intervalů hodnot na základě **volby fx** (Obrázek 5-12, dole) se zobrazí okno s možnostmi nastavení konkrétních ikon. To demonstruje Obrázek 5-13,

Hodnoty prodeje podle zboží

Zbo_Nazev	Součet hodnot: Prodej_Kc
Acer Aspire 5730Z	83 645,00
Acer Aspire One A150	23 397,00
ASUS VW193B	11 277,00
Asus X51L	13 199,00
Autoradio LG LAC3800	10 306,13
Autoradio Logik	6 466,36
Autoradio Pioneer	24 253,89
Autoradio Sony MEXBT	19 794,00
BC LEXZ1380	6 495,00
Bosch 250	1 299,00
DLDAP1160	7 794,00
DWLG510	10 354,66
Espresso DeLonghi E2100	78 393,00
Espresso DeLonghi ESAM2600	51 495,00
Fujitsu Siemens AMILO Pa 3515	45 597,00
HP Compaq 2133	72 793,00
HP Pavilion A6611CS	56 893,01
HP Photosmart D5460	5 697,00
Celkem	1 475 286,04

Filtry

Hledat

Filtry na tomto vizuálu

Součet hodnot: Pro...
je (Vše)

Zbo_Nazev
je (Vše)

Sem přidejte datová pole

Filtry na této stránce

Sem přidejte datová pole

Filtry na všech stránkách

Sem přidejte datová pole

Vizualizace

Formátovat vizuál

Udělte prvkům buňky odlišné formátování z podmínek. Nastavte podmínky výběrem fx.

Vizuál Obecné

Elementy buňky

Použít nastavení pro

Datová řada
Součet hodnot: Prod...

Barva pozadí ☒

Barva písma ☒

Datové pruhy ☒

Ikonky ☒

Webová URL ☐

Obrázek 5-12: Zobrazení ikon k datům

Ikony - Ikony

Styl formátování: Pravidla

Platí pro: Jen hodnoty

Které pole máme vzít jako základ?: Součet hodnot: Naklady

Souhrn: Součet

Rozložení ikon: Nalevo od dat

Zarovnání ikon: Nahoře

Styl: [Color-coded icons]

Pravidla

Pokud hodnota	Operátor	Hodnota	Procenta	do	Operátor	Hodnota	Procenta	pak	Ikona	Akce
>=	0	Procenta	do	<	33	Procenta	pak	[Red diamond]		↑ ↓ ×
>=	33	Procenta	do	<	67	Procenta	pak	[Yellow triangle]		↑ ↓ ×
>=	67	Procenta	do	<=	100	Procenta	pak	[Green circle]		↑ ↓ ×

[Další informace o podmíněném formátování](#)

OK Zrušit

Obrázek 5-13: Nastavení ikon

5.3 Filtrování, průřezy a drilování

Pro zkoumání hodnot dat v reportech **podle různých podmínek** nabízí Power BI řadu možností. Patří k nim nastavení a práce s **filtry (Filters)** a využití **průřezů (Slicers)**.

Velmi silnou vlastností Power BI je **křížové zvýrazňování prvků ve vizuálech** podle výběru prvků v některém z vizuálů (včetně řádků tabulek). Neposlední možností pro analýzu hodnot dat ve vizuálech je **hierarchický rozpad a drilování (drillthrough filters)** podle hierarchických dimenzí.

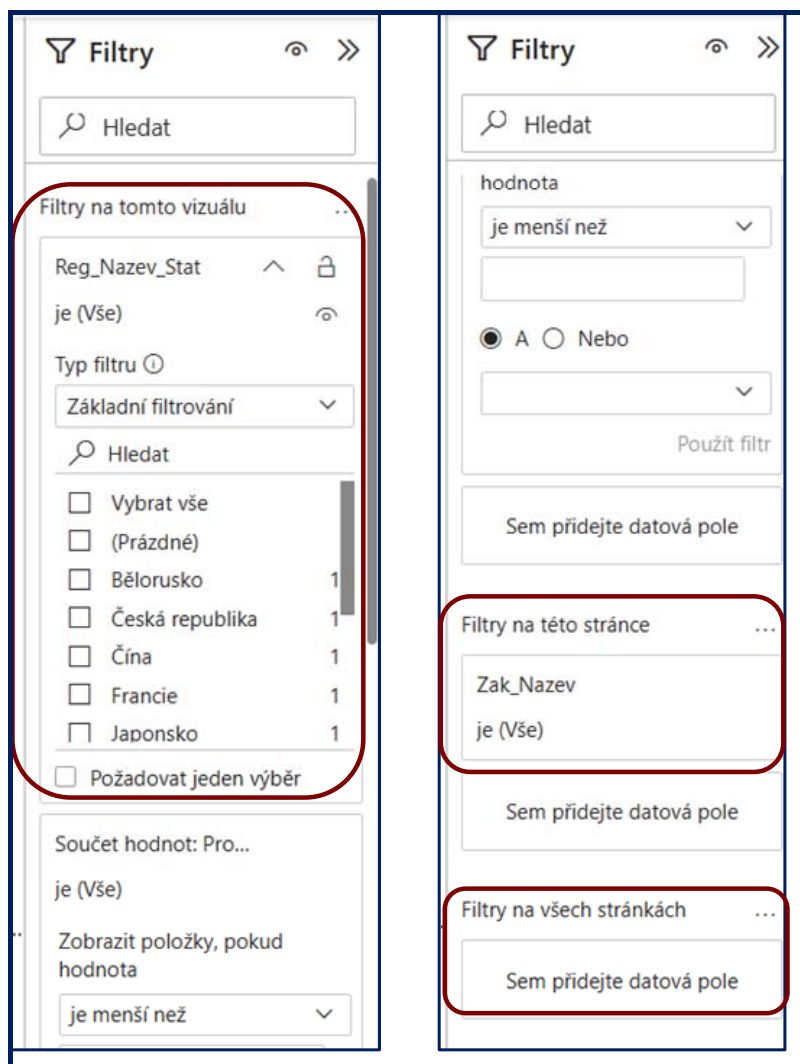
5.3.1 Filtry (filters) v sestavách

Filtry mají v Power BI **různý dosah**, podle toho, **pro jakou část reportu jsou nastaveny**. Lze nastavit **následující filtry**:

1. **Filtr pro celou sestavu (Filters on all pages)**: Nastavení hodnot filtru se uplatňuje na všech stránkách (listech) sestavy bez ohledu na to, na které stránce byl filtr nastaven. Všechny hodnoty ve všech vizuálech na všech stránkách jsou profiltrovány (zobrazují pouze data, která vyhovují nastavení stránkovým filtrům). Změna nastavení filtru se projeví na všech stránkách sestavy.
2. **Filtr pro jednu stránku sestavy (Page level filters)**: Jak název napovídá, jsou nastavením filtru profiltrovány všechny hodnoty ve všech vizuálech pouze na té stránce, na které je filtr definován. Na ostatních stránkách se jeho změny neprojeví.
3. **Filtr pro jeden vizuál (Filters on this visual)**: Nastavení filtru se uplatňuje jen a pouze pro vybraný vizuál. Pokud chce uživatel vidět, jaká jsou aktuální nastavení těchto filtrů, je třeba, aby na stránce reportu myší vybral příslušný vizuál a v sekci **Filtry** se dané nastavení zobrazí. Automaticky se do filtru nastaví všechna datová pole a ukazatele, která jsou použita ve vybraném vizuálu. Do filtru lze přidat z nabídky **Data** jakoukoliv další položku (metriku nebo datové pole).
4. **Filtry pro drilování hodnot (Drillthrough filters) napříč stránkami sestavy**: Pokud je některé z polí, uvedených v tomto filtru, použito v jakémkoliv vizuálu na jakékoliv stránce reportu, lze z tohoto vizuálu přejít na možnost **Drillthrough** a vybrat přechod na příslušnou stranu reportu, kde je tento filtr nastaven.

Definice omezení pro filtr je velmi jednoduchá. Možnosti vycházejí z typu vybraného pole. Pro všechny typy polí lze použít **základní filtrování**, kdy uživatel zaškrtně vybrané hodnoty ze všech dostupných hodnot příslušného pole. **Další typy filtrů se liší podle datového typu pole**. Volba **pokročileho**

filtrování nabízí běžné řetězcové funkce, jako „obsahuje / neobsahuje / začíná / přesně obsahuje/ je prázdné“ apod.). Užitečný je i **filtr pro nejlepší či nejhorší zvolený počet výskytů** datových záznamů podle hodnoty pole, vybrané pro daný vizuál.

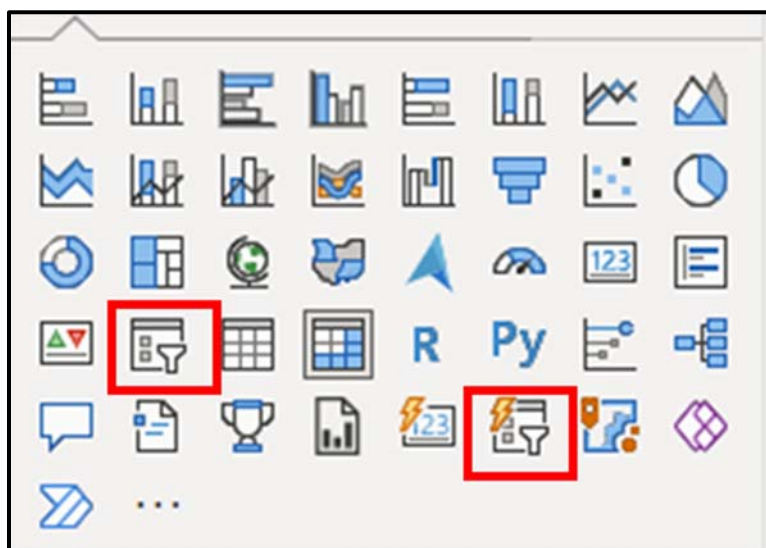


Obrázek 5-14: Příklad nastavení filtru vizuálu

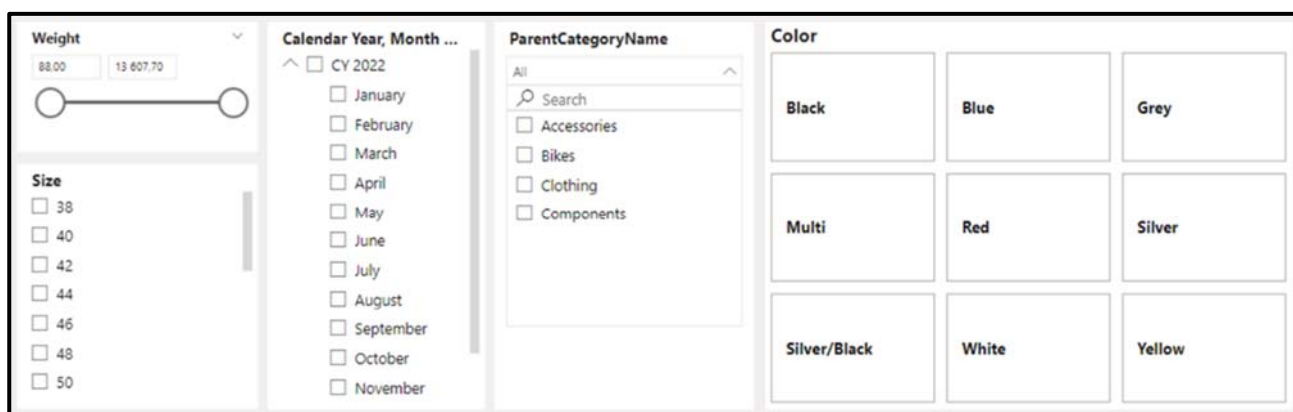
Zrušení nastavení každého filtru lze pomocí ikonky **guma**. Zrušení celého filtru pak křížkem (x) vpravo vedle názvu pole ve filtru.

5.3.2 Průřezy (slicers) v sestavách

Použití filtrů popsaných výše slouží pro specifické omezení výstupů, jak to zamýšlí tvůrce sestavy. Pro koncové uživatele reportů ve službě Power BI je vhodné tyto filtry skrýt (přeškrtnutá ikonka oka). Pro rychlé výběry či omezování hodnot je **doporučeno použít průřezy, resp. slicers**. Jedná se o typ vizuálních prvků, podobných filtrům. Jsou ale **umístěné přímo na stránkách sestavy**, takže vybrané hodnoty jsou pro uživatele přímo viditelné a také použití je přímočařejší. Například pro dimenzi **Datum** lze nastavit výběr **od data do data**, všechna *data po*, všechna *data před*, výběr konkrétního data nebo relativní zadání vybraného počtu dnů, týdnů, měsíců či roků. Příklady představuje Obrázek 5-15 a Obrázek 5-16.

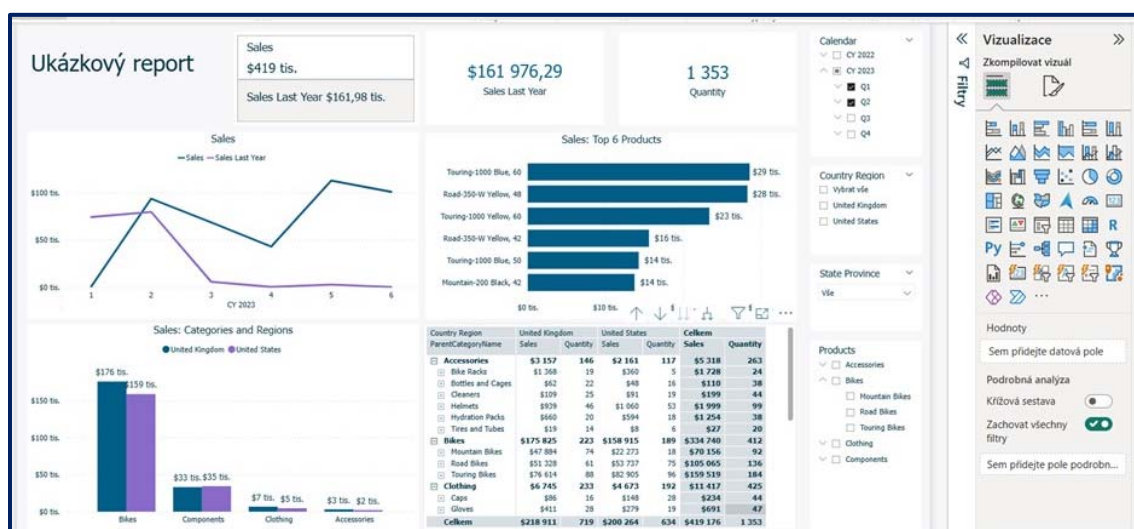


Obrázek 5-15: Výběr vizuálu typu průřez (slicer)



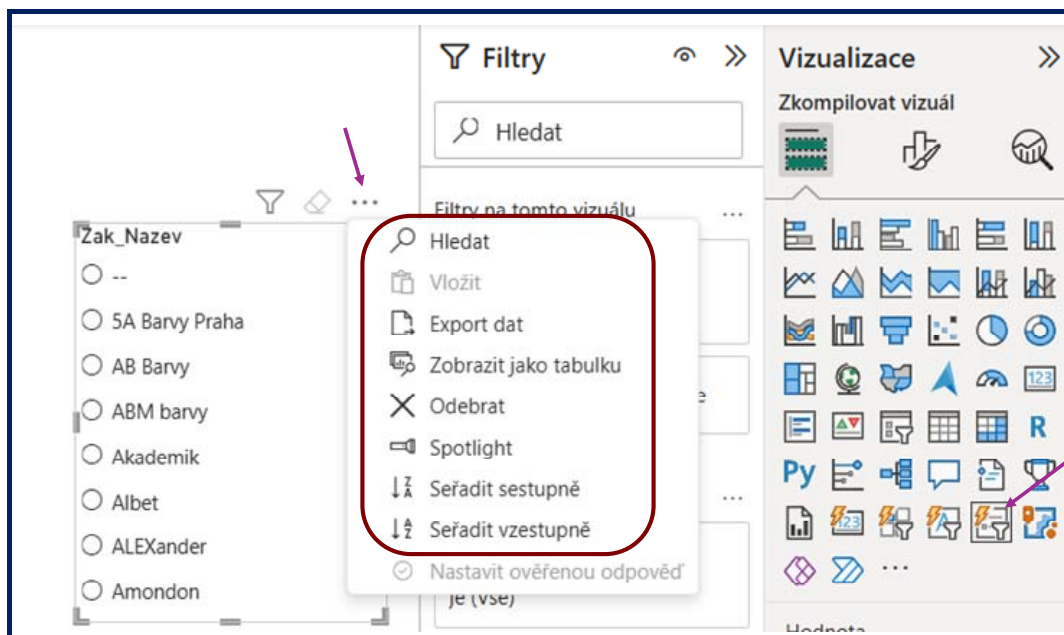
Obrázek 5-16: Příklady formátování průřezů

Report s filtry a průřezy dokumentuje Obrázek 5-17:



Obrázek 5-17: Report s filtry a průřezy

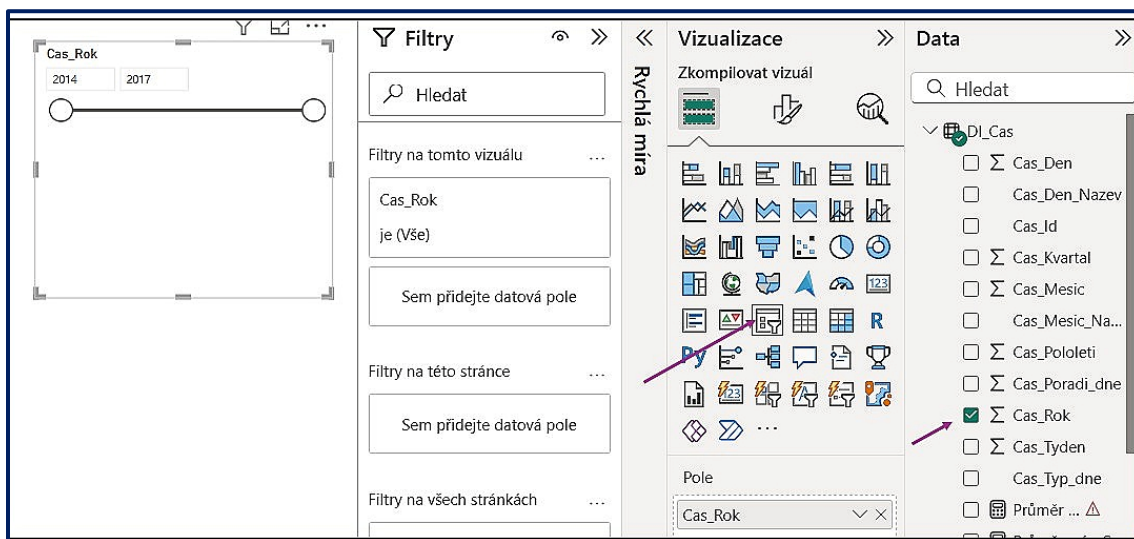
Ve spojitosti s průřezy je vhodné zmínit ještě možnost **přidání výběrového pole do průřezů typu seznam**. Nastaví se výběrem z kontextového menu (tři tečky v pravém horním rohu vizuálu). Zadaním hodnoty do vzniklého výběrového pole (za ikonkou lupy) se omezí seznam hodnot ve výběru jen na ty hodnoty, které obsahují zadaný řetězec (Obrázek 5-18). Hodnoty ale nejsou vybrané, je třeba je aktivně vybrat (zatrhnout).



Obrázek 5-18: Omezení seznamu hodnot ve výběru

5.3.3 Nastavení numerického průřezu

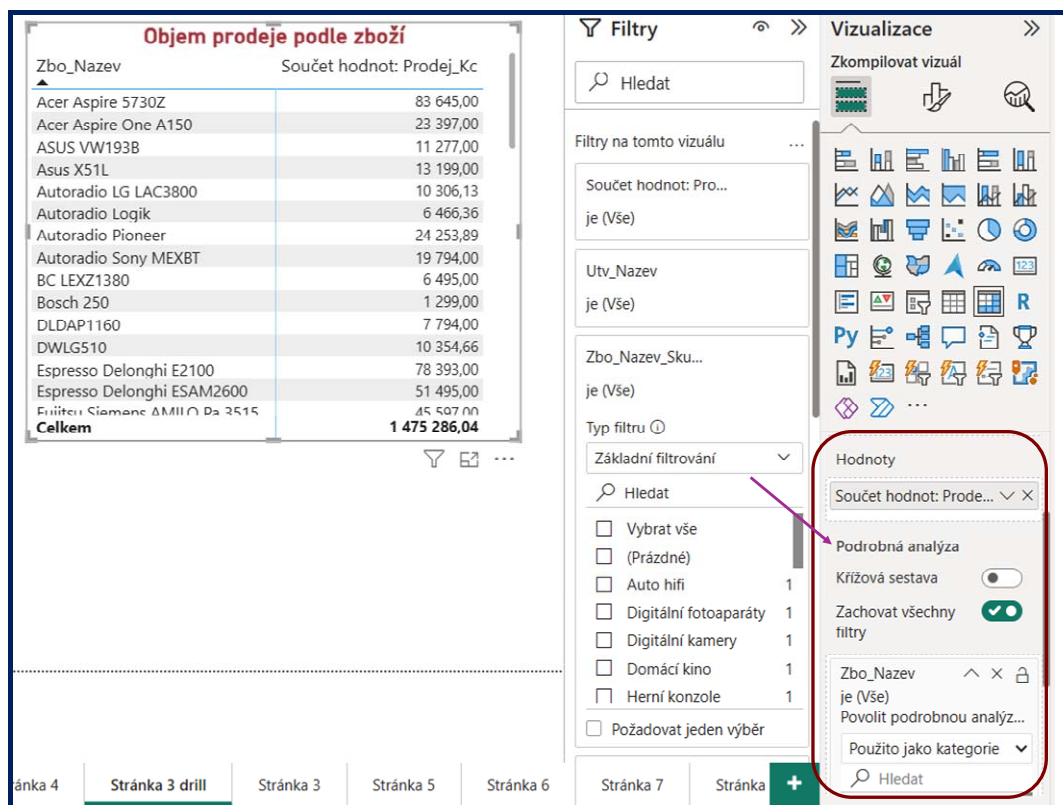
Nastavení průřezu, založeného na numerických datech, dokumentuje Obrázek 5-19:



Obrázek 5-19: Nastavení numerického fitru

5.4 Podrobná analýza (Drill through)

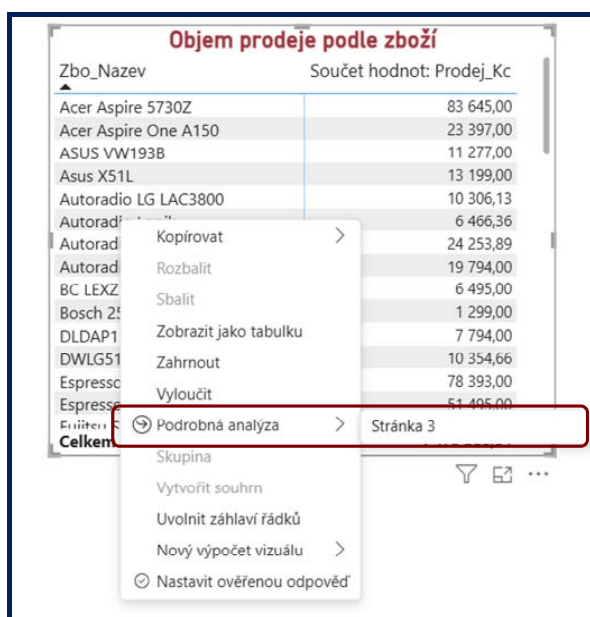
„Podrobná analýza“ (*Drill through*) v Power BI umožňuje přechod ze souhrnného pohledu na detailní stránku sestavy, filtrovanou podle konkrétní hodnoty. Nastavuje se pomocí speciální stránky a pole pro Podrobnou analýzu *drill through*. To dokumentuje **Obrázek 5-20**, vpravo dole, kde pro podrobnou analýzu byla zvolena položka dimenze název zboží (*Zbo_Nazev*). Předmětem analýzy bude objem prodeje zboží v Kč.



Obrázek 5-20: Podrobná analýza, zadání, Drill Through

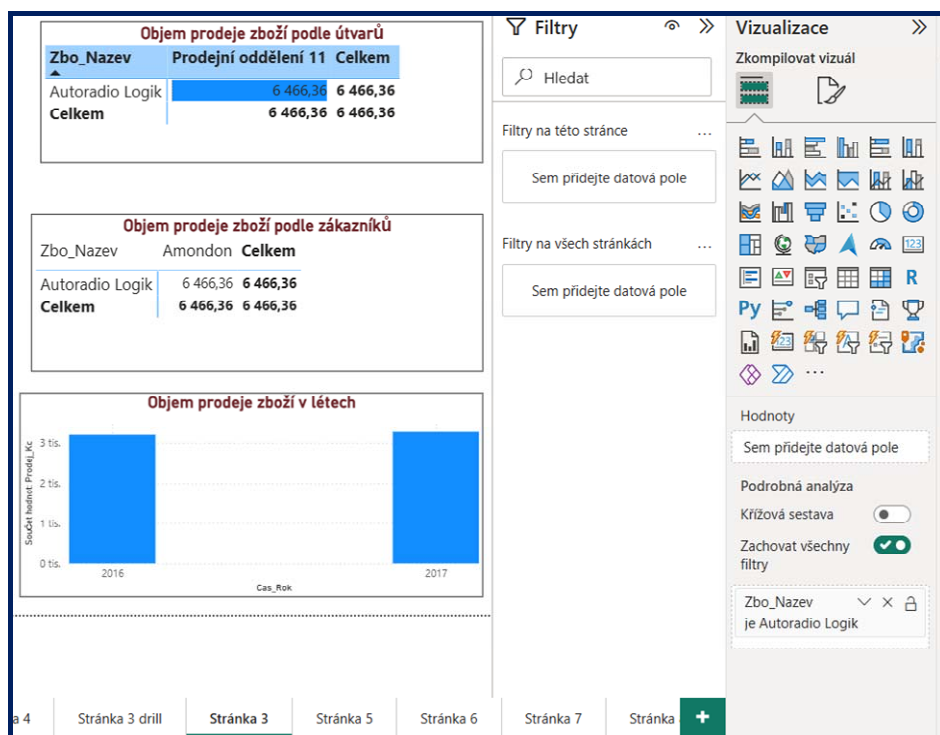
Je třeba dále vytvořit detailní stránku s vizuály (tabulky, grafy), které budou zobrazovat podrobnosti o vybrané hodnotě. V našem případě je to „Stránka 3“.

Na základní, souhrnné stránce („Stránka 3 drill“), při použití pravého tlačítka na vizuál, obsahující pole, použité pro podrobnou analýzu (*Zbo_Nazev*), se objeví kontextové menu s možností volby detailní stránky („Stránka 3“). Power BI ji pak přesměruje na detailní stránku s filtrem podle vybrané hodnoty. To dokumentuje Obrázek 5-21.



Obrázek 5-21: Výběr položky pro podrobnou analýzu (Drill Through)

Detailní stránka pak nabízí vizuály pro podrobnou analýzu k vybranému prvku dimenze, v daném případě „Autoradio Logik“, jak ukazuje Obrázek 5-22.



Obrázek 5-22: Zobrazení podrobné analýzy vybrané položky (Drill Through)

5.5 Analytické funkce v grafech Power BI

Pro vybrané typy vizuálů nabízí Power BI desktop i Power BI Service (v režimu úprav sestav) vybrané analytické funkce. Ve sloupcových a plošných grafech (*Bar chart* a *Area chart*) jsou dostupné **funkce pro zobrazení křivek trendu, konstantní hodnoty** (zadáva uživatel), **maxima, minima, průměru, mediánu a percentilu**. U skládaných grafů (*Stacked bar chart* a *Stacked area chart*) je možné využít pouze přímkou konstantní hodnoty.

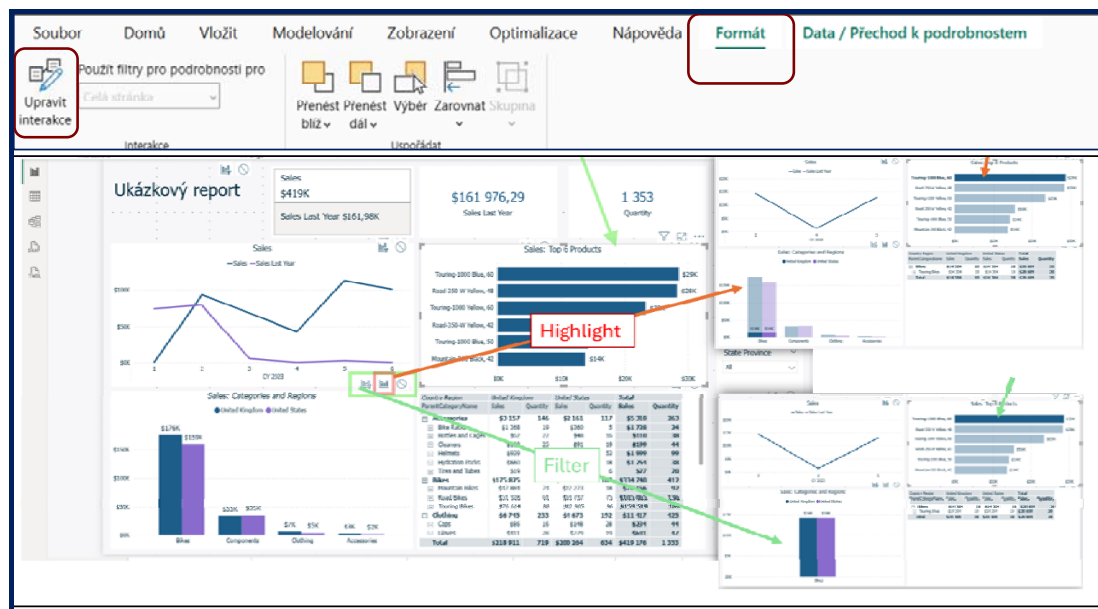
V čárovém grafu (*Line chart*) lze využít i další zabudované analytické funkce, jako je přidání lines pro min, max, *medián a další*, jak ukazuje Obrázek 5-23:



Obrázek 5-23: Rozšíření o další analýzy v čárovém grafu.

5.5.1 Interakce vizuálů, zvýrazňování prvků napříč vizuály na stránce sestavy

Velmi účelnou vlastností Power BI je **interaktivita většiny vizuálů a jejich vzájemné provázání**. Ve většině vizualizací lze zvýraznit vybranou hodnotu klikem na konkrétní prvek, resp. sloupec ve sloupčovém grafu, výšek v koláčovém grafu, bublinu v bublinovém grafu, místo v mapě, ale i řádek v tabulce. **Podle nastavení způsobu vzájemných interakcí pak zareagují všechny vizuály na dané stránce reportu**, jak znázorňuje Obrázek 5-24. Standardně dojde ke **zvýraznění hodnot podle výběru ve vizuálech typu graf a k profiltrování hodnot v negrafických a ostatních grafických vizuálech** (tabulkách, kartách, KPI, mapách, vodopádovém grafu, grafu typu měřidlo apod.).



Obrázek 5-24: Nastavení vzájemných interakcí vizuálů.

5.5.2 Procházení hierarchie – drill down

Práce s hierarchiemi, drillování a rozpad na podrobnější úroveň dat podle zvolených dimenzí je další charakteristikou Power BI. Lze využít jak **hierarchických dimenzí**, tak **hierarchii vytvořit pro konkrétní vizuál** zařazením více dimenzí jako prvků os v grafech, viz Obrázek 5-25. Podobný charakter měla i funkce „Podrobné analýzy“, kterou jsme zmínili již výše.



Obrázek 5-25: Práce s hierarchiemi, drillování

5.6 Pracovní závěry



Z kapitoly vyplývají následující **závěry**:

- Reporty jsou založeny **na skladbě grafických objektů různého typu**, které zde mají označení **vizuál**.
- Power BI nabízí celou **škálu vizuálů (tabulky grafy, mapy)** včetně možnosti vytváření vlastních vizuálů uživatelem.
- Podstatnou funkcí, jako i u jiných aplikací, je **filtrování dat** a s tím spojenéi dynamické úpravy reportů.
- **Průřezy, resp. slicery**, slouží pro lepší prezentaci dat pro filtrování s tím, že se v tabulkách nebo schématech přímo na reportu uvádějí konkrétní hodnoty pro výběr do filtrů.
- Vzhledem k tomu, že dimenze v BI mají obvykle **hierarchickou strukturu prvků**, PBI poskytuje funkcionalitu **k procházení těchto prvků po jednotlivých úrovních, tzv. drillování**.

6. Power BI Service a mobilní aplikace



Služba Power BI (Power BI Service) je aplikace, vyvíjená společností Microsoft pro publikování, vytváření a sdílení BI aplikací. Je poskytována **jako služba (IaaS)** a je dostupná uživatelům po přihlášení na adrese <https://app.powerbi.com>. **Účelem** kapitoly je detailněji charakterizovat jednotlivé charakteristiky a možnosti **Power BI Service** při použití datové brány pro aktualizaci dat. Poslední část kapitoly je věnovaná možnostem mobilních aplikací na bázi Power BI.

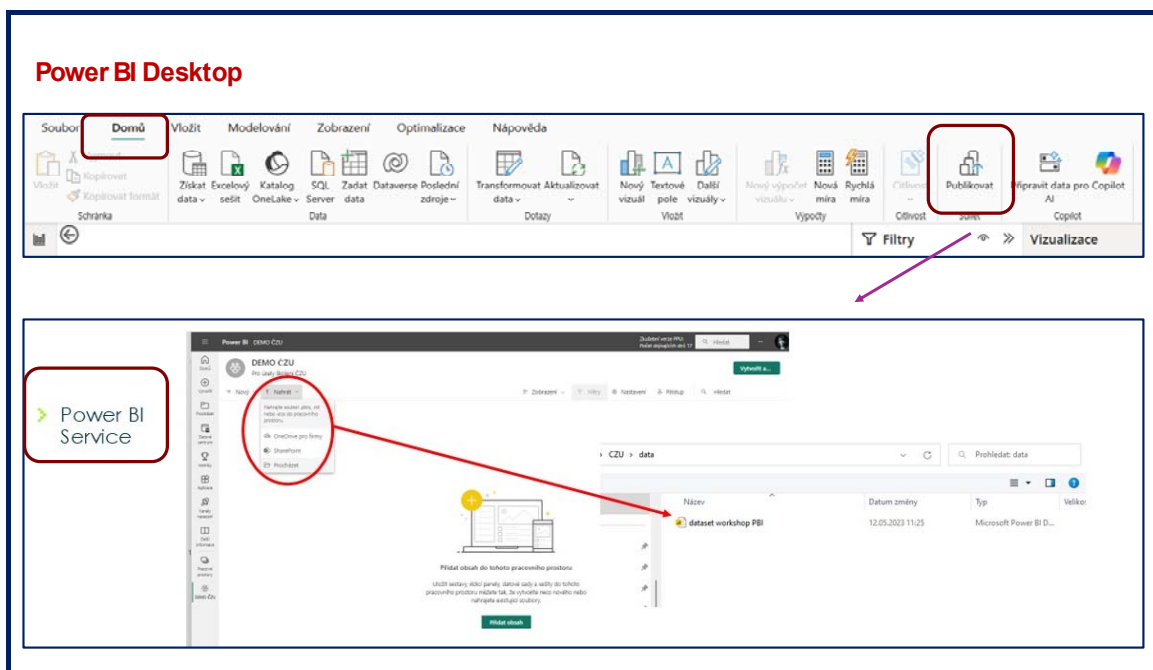
Využití služby je závislé na tom, **zda se jedná o uživatele s předplatným Power BI Pro, nebo uživatele bez předplatného (Free user)**. Pouze uživatelé s předplatným **Power BI Pro** nebo **Power BI Premium per User** mohou publikovat obsah do pracovních prostorů aplikací, konzumovat aplikace bez kapacity Fabric, sdílet dashboards a přihlásit se k odběru změn v reportech a dashboardech. **Free uživatelé** se mohou připojit ke všem datovým zdrojům prostřednictvím všech možností připojení, jako je DirectQuery, realizovat živé připojení a použití datové brány, nemohou ale konzumovat publikovaný obsah, pokud tento obsah není licencován předplatným Microsoft Fabric F64 a vyšším. Bližší informace o licencování jsou uvedeny přímo na stránkách Power BI (<https://powerbi.microsoft.com/en-us/pricing/>).

6.1 Publikace připravených reportů do služby Power BI

Sdílet reporty a datové sady vytvořené v Power BI Desktop lze prostřednictvím služby **Power BI Service**. **Uživatel**, který sdílí data, **musí být přihlášen k službě Power BI Service (Sign in na <https://app.powerbi.com/>)**. Následující text se věnuje publikování pro účely práce více uživatelů. **Free uživatelé** mohou publikovat obsah pouze do svého pracovního prostoru, bez možnosti sdílet ho s jinými uživateli.

Nejprve je potřeba **vytvořit pracovní prostor přímo v Power BI Service** a poté **z Power BI Desktop vytvořenou aplikaci** do pracovního prostoru služby Power BI **publikovat**.

Power BI Service obsahuje datovou sadu včetně dat, která byla importována z Power BI Desktop, a zahrnuje rovněž všechny reporty ve stavu, v jakém byly ve chvíli publikace. Pokud **v Power BI Desktop dojde následně k jakékoliv změně**, ať už se jedná o datový model, strukturu tabulek, vypočtené ukazatele nebo o změny v reportu, **je nutné publikaci opakovat**, změny se do Power BI Service nepromítají automaticky. Výjimkou je situace, pokud je Power BI Desktop aplikace umístěna na sdíleném úložišti OneDrive for Business a do Power BI Service je publikována z tohoto úložiště. Pak jsou změny promítány vždy po aktualizaci sešitu Power BI Desktop na OneDrive. Způsob publikace do Power BI Service dokumentuje Obrázek 6-1.



Obrázek 6-1: Publikace reportu do PBI Services

6.2 Vybrané funkce Power BI Service

Uživatelé, kteří mají **licenci Power BI Pro** a **správce pracovního prostoru** jim udělil přístupová práva alespoň na čtení, uvidí po přihlášení do služby **powerbi.com** **pracovní prostor a v něm publikovanou sestavu**. Mohou přecházet mezi stránkami sestavy, měnit dostupné filtry, zobrazovat a analyzovat vizualizovaná data a využívat dostupnou funkcionalitu vizualizací, stejně, jak je to popsáno v předchozích kapitolách, věnovaných Power BI Desktop.

Uživatelé s právy editovat mohou navíc sestavy vytvářet a upravovat. Do editačního módu sestavy se přepnou funkcí **Upravit sestavu**. Vpravo se zobrazí panel s nabídkou vizuálů, sekcí pro datová pole, použitá ve vizuálech, formátovací sekcí a sekcí pro nastavení filtrů. Úplně vpravo je pak zobrazen panel se seznamem datových tabulek, tj. s dimenzemi a faktovými tabulkami. Funkcionalita a ovládání jsou **obdobné jako v Power BI Desktop**.

Pokud však uživatelé provedou jakoukoliv **úpravu a uloží ji, tato změna se projeví jen v Power BI Service, do aplikace Power BI Desktop se nepromítá!** Nově lze využít funkci **Stáhnout pbix** a původní aplikaci **pbix** nahradit. Přesto je více než nutné **stanovit jasné postupy**, kdo a na kterém místě bude vytvořené Power BI aplikace spravovat a rozvíjet.

Uživatelé s editačními právy mohou také **vytvářet dashboards** a **umísťovat na ně jednotlivé vizuály z listů sestavy** či celé listy sestavy.

Zcela unikátní funkcionalitou, i díky níž je Microsoft hodnocen jako leader v oblasti **Business Analytics**, je funkce **Získat rychlý přehled (Quick Insights – ikonka žárovky)**. Pro všechna data v datové sadě Power BI Service **vygeneruje grafické zobrazení všech možných závislostí v datech** včetně jejich stručného popisu, aniž by uživatel musel vytvářet sestavy a definovat vlastní vizuální zobrazení. Každý takto vygenerovaný vizuál lze následně připnout na dashboard a mít ho tak stále k dispozici.

Neméně užitečnou funkcionalitou je **možnost analyzovat data**, umístěná v Power BI Servis **v aplikaci MS Excel** (je nutné nainstalovat doplňky, které Power BI při spuštění funkce doporučí). Systém vytvoří automatické propojení do databáze Azure, kde je datová sada uložena (je třeba v Excelu povolit datová připojení), a kontingenční tabulku. Nyní je již možné analyzovat data, vytvářet kontingenční grafy tak, jak je uživatel zvyklý z běžné práce s kontingenčními tabulkami.

Podrobné tutoriály, videa a doporučení k ovládání a pro práci s Power BI Service jsou k dispozici přímo na stránkách <https://powerbi.microsoft.com>.

6.3 Vytvoření a sdílení dashboardů

Řídicí panely neboli dashboardy lze vytvářet pouze v Power BI Service. Po vytvoření je možné je sdílet vybraným skupinám uživatelů či konkrétním uživatelům a také je zobrazit v mobilních zařízeních prostřednictvím Power BI Apps. Power BI Desktop, ani Power BI Report Server funkcionalitu tvorby a sdílení dashboardů neposkytují.

Dashboard je **tvořen dlaždicemi**, jejichž obsahem je buď **jeden vizuální prvek** (graf, mapa, karta s hodnotou, KPI apod.), **nebo celá stránka sestavy**.

Vytvoření dashboardu je velmi jednoduché. Pro připnutí živé stránky sestavy nebo jednoho vizuálu slouží **ikonka špendlíku**. U každého vizuálu (kromě slicerů) se špendlík objeví při najetí myši do prostoru tohoto vizuálu. Klepnutí myši na ikonku špendlíku je vyvolán dialog, který nabídne umístění daného prvku na dashboard, včetně možnosti vytvořit dashboard nový.

Pokud je na řídicí panel **připnuta živá stránka**, pak je **zachována její funkcionalita** a lze s ní pracovat přímo na dashboardu (využívat zvýrazňování napříč vizuály, drilovat na nižší úroveň v grafech apod.). Pokud je připnut **pouze jeden vizuál**, kliknutí na něj způsobí **přechod na stránku sestavy**, z níž byl vizuál připnut.

Je třeba upozornit na **několik specifíků v chování dashboardů v aplikaci Power BI Service**. Pokud je v sestavě nastaven nějaký filtr, nebo jsou ve vizuálu na sestavě zvýrazněné některé prvky díky vzájemné interakci s jinými vizuály a daný vizuál je připnut na dashboard při tomto nastavení, **následná změna nastavení v reportu se na dashboardu neprojeví**. I když by byl vizuál v sestavě zrušen nebo nahrazen jiným, dlaždice v dashboardu se tato změna nijak nedotkne. **Pokud jsou však aktualizována data v sémantickém modelu** (nebo v externí databázi v případě *DirectQuery*), pak se **na dashboardu projeví změněné hodnoty** vždy po obnovení dlaždic. Obnovení je možné buď spustit **ručně** pomocí funkce **Obnovit dlaždice řídicího panelu (Refresh)**, **nebo probíhá automaticky**, pokud je naplánována pravidelná obnova dat příslušné datové sady.

Nastavení automatické obnovy dat předpokládá, že musí být **aktivní datová brána Data Gateway**. Nastavení dále vyžaduje **autorizaci přístupu ke zdrojům dat** zadáním přihlašovacích údajů ke každému zdroji. Zapnutí funkce **Zajištění aktuálnosti dat** zpřístupní volby **nastavení periodicity** aktualizace. Je možné vybrat buď denní aktualizaci anebo týdenní. Je vhodné zaškrtnout **funkci Odeslat email s oznámením selhání aktualizace**. Pokud se tak stane, je email zaslán na adresu uživatele, který danou datovou sadu nakonfiguroval.

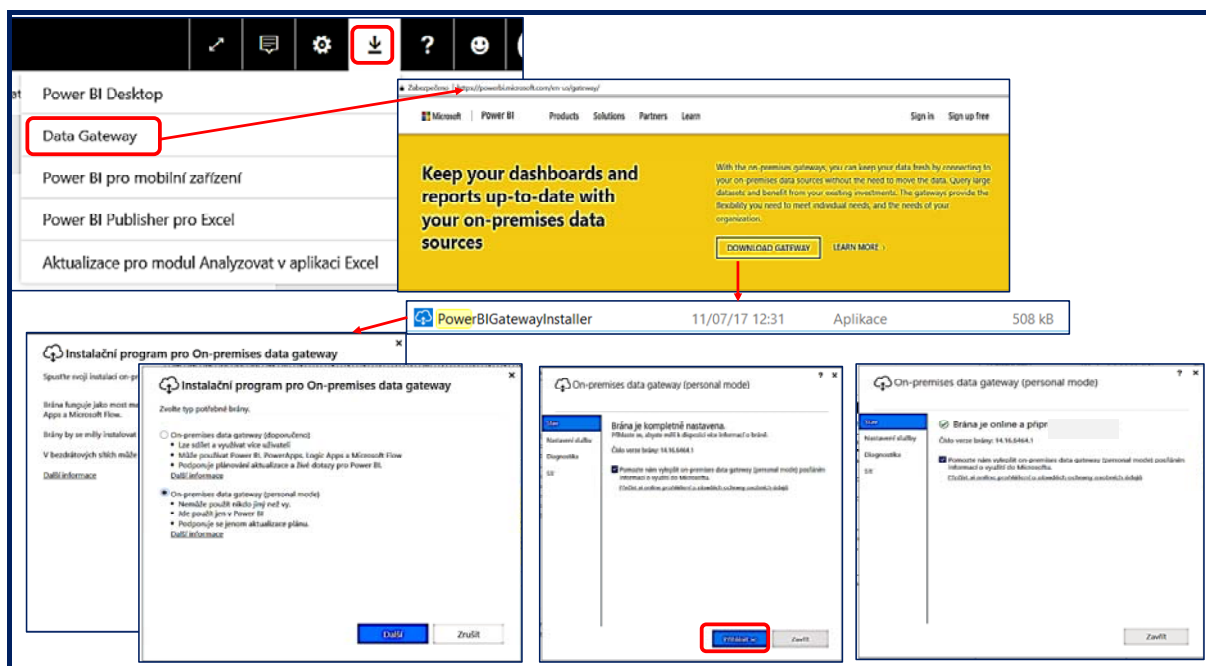
Rozložení dlaždic na dashboardu je možné měnit pouhým přetažením myši. Velikosti dlaždic se upravují z pravého dolního rohu dlaždice pohybem myši. Je možné též upravit **názvy a popisy dlaždic**, nastavit zobrazování času poslední aktualizace hodnot dat, nastavit vlastní hypertextový odkaz a tím zajistit, že klikem na dlaždici bude předáno řízení jinam než na sestavu, odkud pochází vizuál v dlaždici.

Další nastavení a chování dashboardů si uživatelé postupně osvojí používáním Power BI Service, případně s využitím bohaté dokumentace na stránkách <https://powerbi.microsoft.com>.

Zpřístupnění dashboardu vybraným skupinám uživatelů či jednotlivým uživatelům **se nastavuje pro každý dashboard**. Uživatelé musí mít licenci *Power BI Pro* a musí být zařazeni do stejného účtu organizace (*Tenantu*). Uživatelé se zadávají jejich **emailovými adresami**. Součástí nastavení sdílení je i **povolení dalšího sdílení daného reportu** pro zadané uživatele a výběr, zda jim má Power BI zaslat e-mailovou notifikaci, že jim byl dashboard zpřístupněn.

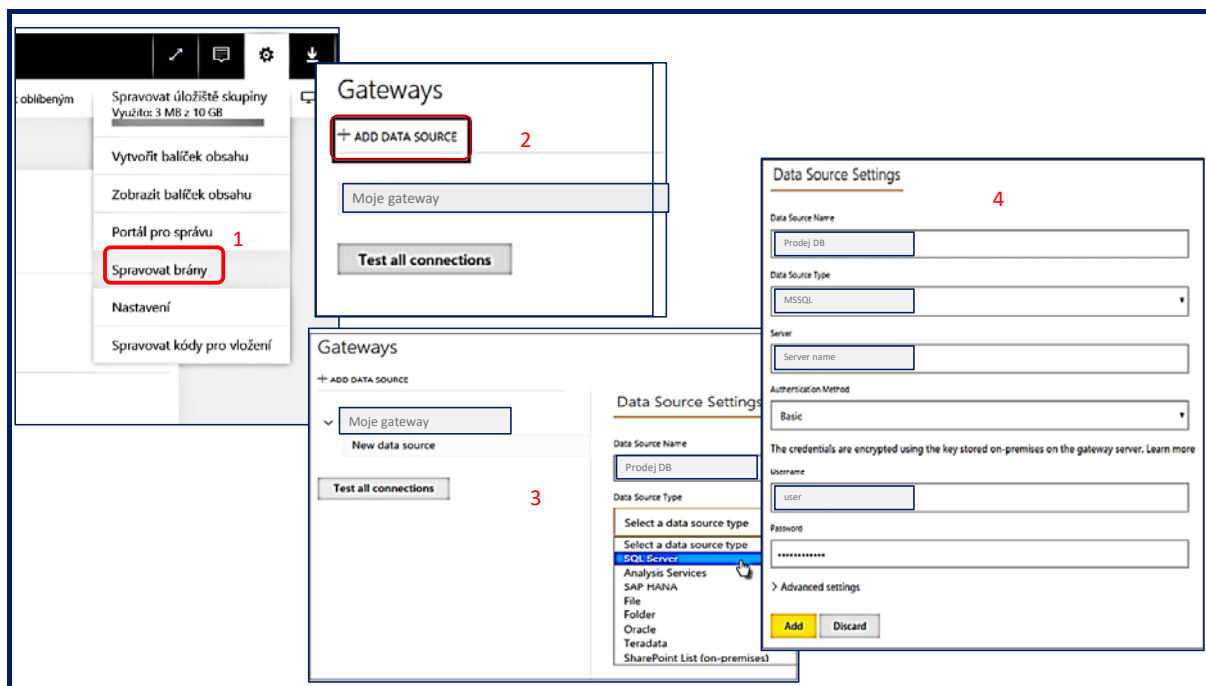
6.4 Použití datové brány pro obnovu dat ve službě Power BI

Pokud má být **Power BI napojen na datový zdroj ležící v infrastruktuře uživatele** a tedy má-li být zajištěna aktualizace hodnot dat, zobrazovaných v reportech a dashboardech, je potřeba mít **nainstalovanou a on-line dostupnou tzv. On-premises datovou bránu (DataGateway)**. Tato brána zajišťuje službě Power BI, resp. službě *Azure Service Bus*, **přístup k on-premise, resp. podnikovým datům**. Jedná se o aplikaci, která je ke stažení na powerbi.microsoft.com, odkaz je též přímo z menu Power BI Service (Obrázek 6-2). Pokud je použit *personal mode*, pak běží pod účtem konkrétního uživatele. Pokud má obnova dat proběhnout v pořádku, je třeba, aby na počítači, na němž brána běží, byl provoz *on-line zabezpečen*.



Obrázek 6-2: Instalace a zpřístupnění DataGateway

Po úspěšném nainstalování brány (brána je on-line a připravena k použití) **je do ní v Power BI Service možné přidat datový zdroj (Data source)**, jak znázorňuje Obrázek 6-3, na základě funkce **Spravovat brány (Manage gateways)**. Následně je třeba v nabídce zvolit bránu a kliknout na funkci **Přidat datový zdroj (Add data source)** a zvolit typ datového zdroje (v našem ilustrativním příkladu je vybrán SQL Server), vyplnit název datového zdroje, **vybrat požadovaný server, databázi** a její přístupové údaje. Pokud je nastavení správné, zobrazí se hlášení o úspěšném propojení. Nakonec je ještě potřeba k danému datovému zdroji **nastavit uživatelským přístupová práva**.



Obrázek 6-3: Přidání a nastavení datového zdroje v DataGateway

Pracovat je možné jak s řídicími panely, tak se sestavami. Klikem na konkrétní dlaždici dashboardu se na mobilním telefonu **zobrazí příslušný vizuál**, dotekem prstu se zobrazí podrobnosti. **Z dlaždice je možné přejít přímo na sestavu (ikonka sloupcového grafu)**. Sestavy se zobrazují při otočení telefonu na šířku. Oblíbenou funkcionalitou manažerů je možnost vizuál opatřit poznámkami, kreslením či ikonkami označit co je na grafech zaujalo a takto doplněný obrázek odeslat například emailem kolegům s komentářem, jakou akci mají na základě prezentovaných údajů přijmout (Obrázek 6-5).

Díky využití mobilních aplikací Power BI mohou být manažeři neustále informováni o stavu a vývoji své firmy.

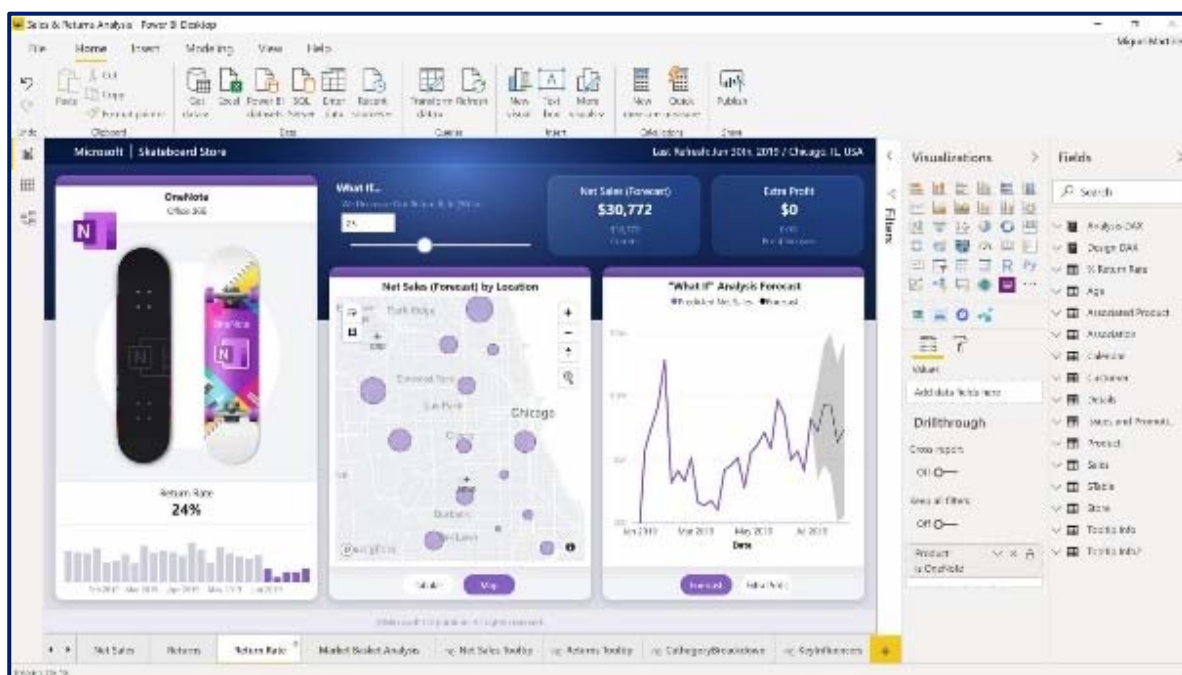
6.6 Pracovní závěry



Z kapitoly vyplývají následující **závěry**:

- **Reporty a sémantické modely**, vytvořené v Power BI Desktop, lze sdílet prostřednictvím služby Power BI Service, uživatel **musí být přihlášen k službě Power BI Service (Sign in)**.
- Uživatelé s **licencí Power BI Pro**, a pokud **správce pracovního prostoru** jim udělil přístupová práva, uvidí po přihlášení do služby **powerbi.com pracovní prostor a v něm publikovanou sestavu**.
- Uživatelé s editačními právy mohou také **vytvářet dashboardy a umísťovat na ně jednotlivé vizuály z listů sestavy** či celé listy sestavy.
- Funkce **Získat rychlý přehled (Quick Insights – ikonka žárovky)** znamená, že pro všechna data v datové sadě Power BI Service **vygeneruje grafické zobrazení všech možných závislostí v datech** včetně jejich stručného popisu, aniž by uživatel musel vytvářet sestavy a definovat vlastní vizuální zobrazení.
- Další funkcionalitou je **možnost analyzovat data**, umístěná v Power BI Service, **v aplikaci MS Excel**.
- Řídicí panely neboli dashboardy **lze vytvářet pouze v Power BI Service**. Po vytvoření je možné je sdílet vybranými skupinami uživatelů či konkrétními uživateli a také je zobrazit v mobilních zařízeních.
- Dashboard je **tvořen dlaždicemi**, jejichž obsahem je buď **jeden vizuální prvek** (graf, mapa, karta s hodnotou, KPI apod.), **nebo celá stránka sestavy**.
- Pokud má být zajištěna aktualizace hodnot dat, zobrazovaných v reportech a dashboardech, je potřeba mít **nainstalovánu a on-line dostupnou tzv. On-premises datovou bránu (DataGateway)**. Tato brána zajišťuje službě Power BI, resp. službě Azure Service Bus, **přístup k on-premise, resp. k podnikovým datům**.
- Mobilní aplikace Power BI apps zajišťují přístup k dashboardům a reportům, publikovaným v Power BI Service, nebo k reportům, umístěným na Power BI Report Serveru z mobilních zařízení.

7. Příprava a plánování projektů v Power BI



Řešení projektů a aplikací na bázi Power BI můžeme sledovat **v užším kontextu**, zaměřeném na základní **vytvoření aplikace v Power BI desktop**, což byla náplň předchozích kapitol, nebo **v širším kontextu**, respektujícím i různé **projektové aspekty a přístupy**. Právě tento **širší kontext** pohledu na Power BI **je účelem** této kapitoly (podle Deckler, G., Powell, B., 2022).

Kapitola se zaměřuje především **na tyto body**:

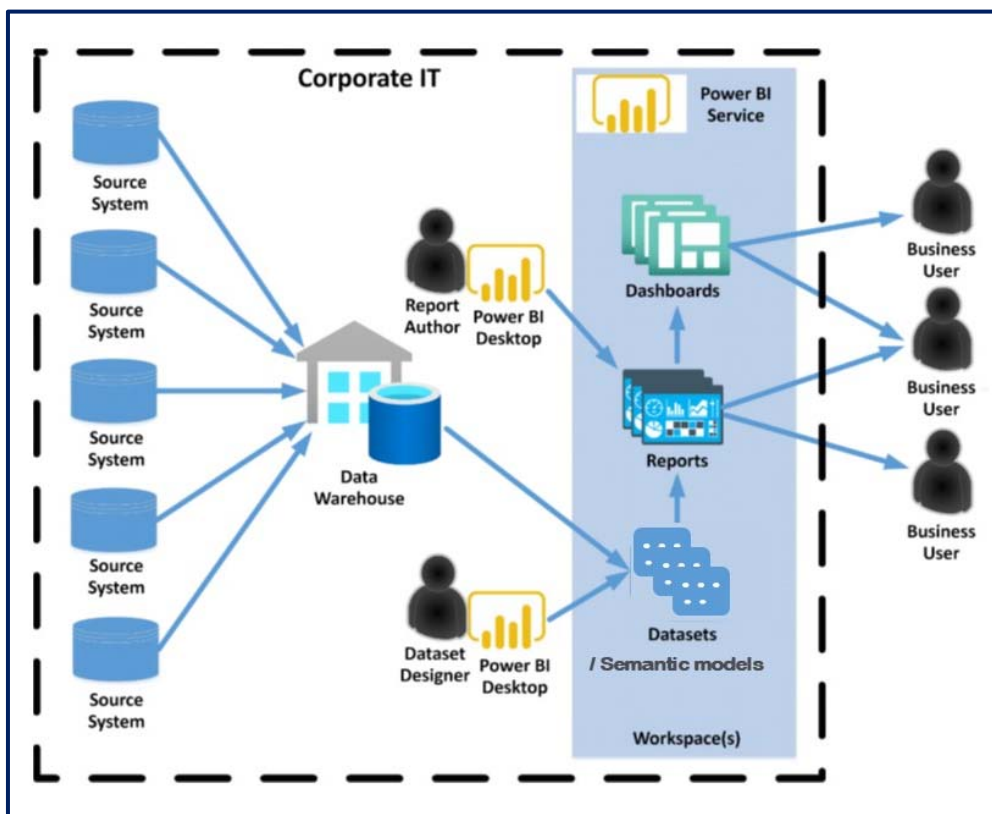
- Způsoby nasazení (deployment) Power BI.
- Role v řešení a užití projektu.
- Hlavní otázky řešení projektu.
- Poznámky k licencím.

7.1 Způsoby nasazení (deployment) Power BI

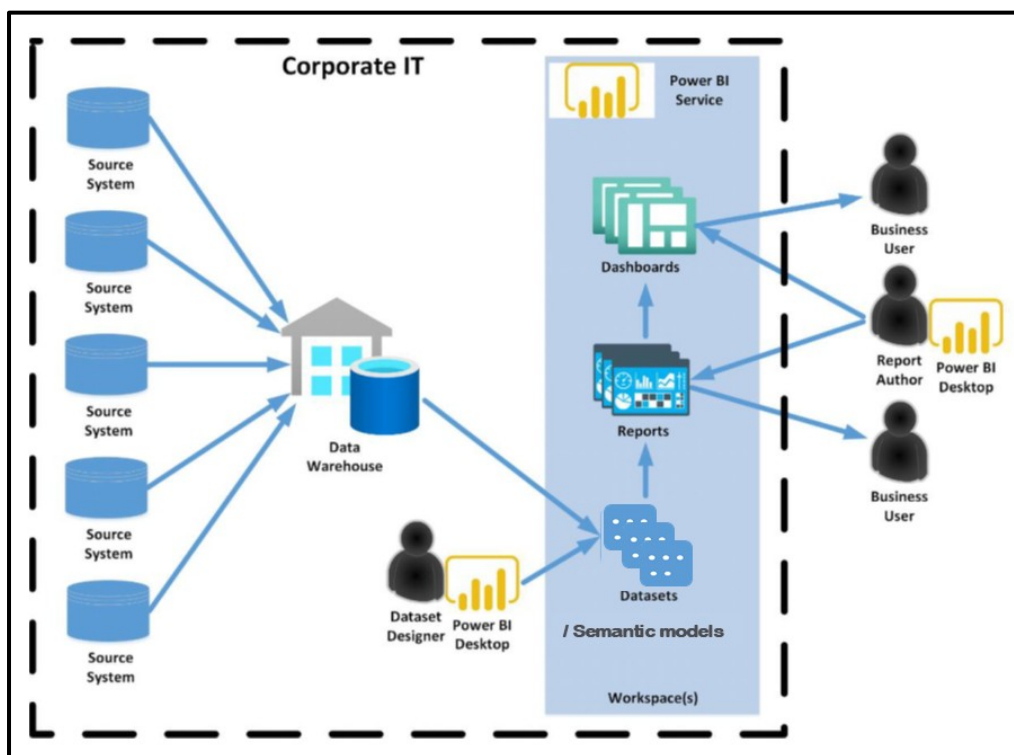
Podle (Deckler, G., Powell, B., 2022) se rozlišují tyto módy nasazení Power BI s následujícími charakteristikami:

- **Podnikové BI (Corporate BI):**
 - Reporty a dashboardy jsou vlastněné IT útvary.
 - Sémantické modely jsou rovněž vlastněné IT útvary.
- **Self-Service Vizualizace:**
 - Reporty a dashboardy jsou vlastněné byznysem.
 - Sémantické modely jsou vlastněné IT útvary.
- **Self-Service BI:**
 - Reporty a dashboardy jsou vlastněné byznysem.
 - Sémantické modely jsou rovněž vlastněné byznysem.

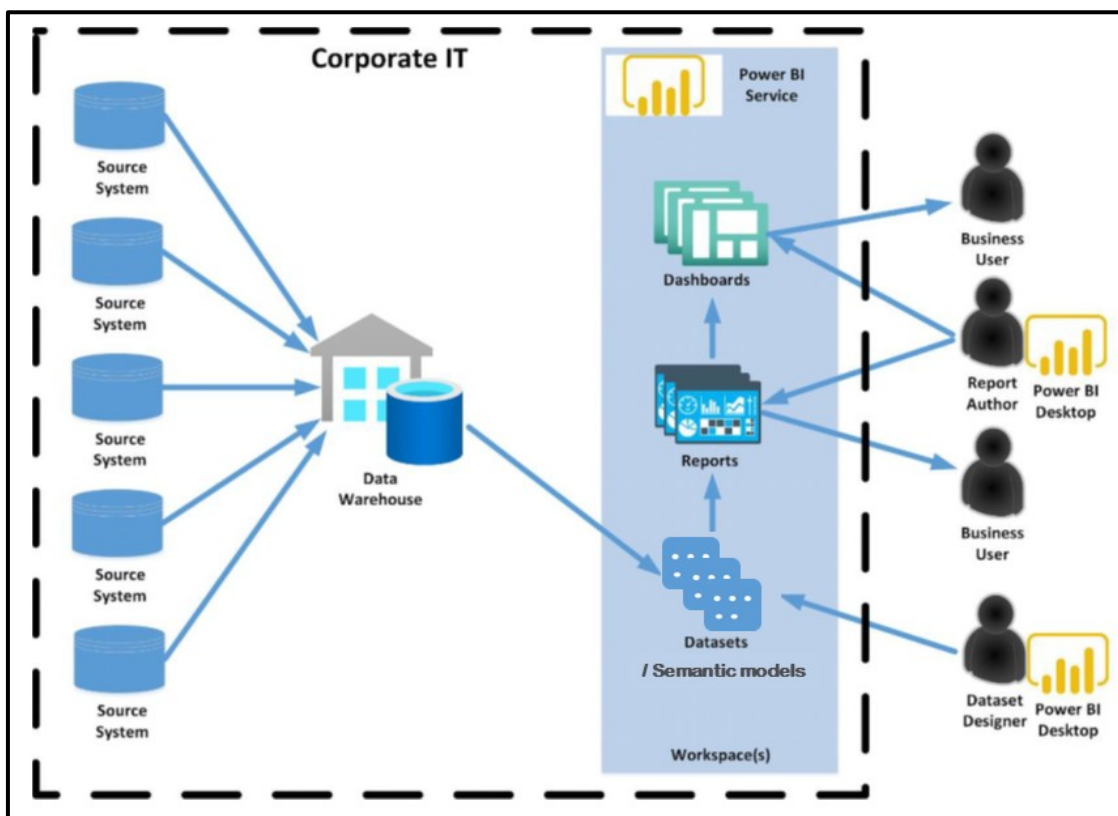
Uvedené módy jsou dokumentované následujícími schémata:



Obrázek 7-1: Podnikové BI, Corporate BI (Zdroj: Deckler, G., Powell, B., 2022 – upraveno)



Obrázek 7-2: Self-Service Vizualizace (Zdroj: Deckler, G., Powell, B., 2022 – upraveno)



Obrázek 7-3: *Self-Service BI* (Zdroj: Deckler, G., Powell, B., 2022 – upraveno)

Doplňující **poznámky**:

- V rámci projektu **je možné migrovat** z jednoho způsobu nasazení do jiného. Řada firem využívá často **mix uvedených způsobů** – podle konkrétních podmínek a potřeb.
- Ukazuje se, že **přílišné sdílení vlastnictví jednotlivých součástí** řešení (bez jasně vymezených zodpovědností) **obvykle vede k problémům** v komunikaci, problémům v řešení verzí, kvality a konzistence v rámci projektu.
- **Corporate BI** se užívá především ve velkých společnostech a poskytuje obvykle nejlepší možnosti řízení projektu v rámci celého řízení BI a kvalitního řízení verzí aplikací, jejich přenositelnosti a jejich optimálního výkonu. V tomto případě jsou uživatelé „pouze“ konzumenty reportů a dashboardů.
- V případě **Self-Service Vizualizace** je vytváření a údržba sémantických modelů na straně IT nebo specializovaných BI týmů, zatímco reporty jsou řešeny na straně uživatelů s příslušným oprávněním, a to i pro ostatní uživatele. Výhodu lze zde předpokládat ve vyšší úrovni byznys řešení reportů.
- Pokud jde o **Self-Service BI**, pak IT / BI týmy zajišťují pouze potřebou infrastrukturu a monitorování provozu, byznys uživatelům se nabízí maximální volnost a flexibilita při řešení aplikací. To může být výhodou, ale často je to spíše rizikem z pohledu řízení celého BI systému. Vzhledem k tomu, že je nutné zajišťovat potřebnou úroveň bezpečnosti a konzistence dat, jsou obvykle takto vytvořené aplikace po odsouhlasení převáděny pod řízení IT/BI útvarů.

7.2 Role v řešení a užití projektu Power BI

V rámci celé škály rolí v rámci BI a jejich podílu na řešení se v souvislosti s projekty Power BI zaměříme na tyto role:

- Návrhář sémantických modelů (*Semantic models designer*).
- Tvůrce reportů (*Report Author*).
- Power BI administrátor.

7.2.1 Návrhář sémantických modelů (*Semantic models designer*)

Návrhář sémantických modelů zajišťuje tyto funkce:

- analyzuje a **navrhuje struktury a obsah sémantických modelů** (ideálně star schéma), průběžně **komunikuje s vlastníky** datových zdrojů,
- zjišťuje, dokumentuje, upravuje modely a **informuje o změnách** v datových zdrojích,
- posuzuje **kvalitu zdrojových databází**,
- **monitoruje provoz** sémantických modelů a řeší případné problémy,
- řeší transformace dat prostřednictvím nástrojů k tomu vhodných, například v **Power Query** s **využitím jazyka M, nebo v Dataflow Gen 2, nebo v noteboocích** pro definování tabulek datového modelu,
- nastavuje **autentifikaci a autorizaci uživatelů** vzhledem k přístupům k modelům,
- podle potřeby programuje **metriky v jazyce DAX** pro užití v reportech,
- definuje a realizuje případný **mix přístupů k sémantickému modelu** na bázi fyzického transferu dat či s využitím funkcí DirectQuery s vyhodnocení výkonů v obou přístupech,
- podle potřeby spolupracuje na řešení a monitoruje **provoz sémantických modelů, vytvářených uživatelskými BI týmy** v rámci módu Self-Service BI,
- **navrhuje kontrolní a opravné procedury** v souvislosti s čištěním, resp. zajištěním kvality dat,
- zajišťuje funkce **bezpečnosti modelů**, zejména **RLS (row-level security)** nebo **OLS (object-level security)**, určující, kteří uživatelé nebo jejich skupiny mají přístup k datům, resp. reportům na nich vytvořených.

7.2.2 Tvůrce reportů (*Report Author*)

Tvůrce reportů zajišťuje tyto funkce:

- **konzultuje s uživateli** jejich problémy a požadavky na reporty, kooperuje na **analýzách požadavků** uživatelů a jejich konsolidaci,
- navrhuje a projednává s uživateli **hrubý dimenzionální model** řešení,
- specifikuje **obsah a strukturu hlavních reportů**, ukazatelů a jim odpovídajících dimenzí,
- definuje a řeší potřebné **kalkulace pro reporty** na bázi jazyka DAX,
- uplatňuje efektivní principy a možnosti **vizualizace dat** podle charakteru reportů a potřeb jednotlivých typů uživatelů,
- respektuje a uplatňuje **obsahové i formální standardy** firmy při navrhování reportů,
- podle potřeby **realizuje reporty s uplatňováním** podmíněného formátování dat, drill down, drill through a další specifických funkcí pro efektivní interaktivní využití reportů,
- navrhuje a řeší **aktualizací reportů při změnách dimenzí**, resp. faktových tabulek,
- kooperuje na **přípravě datové základny** pro zahájení provozu aplikací,
- podílí se na **řešení analytických problémů nebo chyb**, vyplývajících z přípravy provozu.

7.2.3 Power BI administrátor

Power BI administrátor je obvykle nastaven v Azure Active Directory a zajišťuje tyto funkce:

- kooperuje na **přípravě sémantických modelů** pro zahájení provozu aplikací,
- **konfiguruje celé PBI řešení**, aktivuje nebo deaktivuje jednotlivé funkce, např. nahrávání dat, přístupy k reportům, nastavuje přístupy externích uživatelů apod.,
- podílí se na **řešení problémů nebo chyb**, vyplývajících z přípravy provozu aplikací,
- zajišťuje **analýzu datových struktur zdrojových systémů**, jejich kvalitu a dostupnost,
- **monitoruje a řeší problémy provozu**, problémy aktualizace sémantických modelů, problémy výkonu sémantických modelů, přístupy uživatelů a další,
- **definuje pravidla provozu** v rámci Power BI Admin Portal,
- spolupracuje s datovým architektem na návrhu **datové architektury pro Power BI**,

- spolupracuje na **vyhodnocení stavu a kvality dat**, na formulaci principů řízení jejich kvality,
- konzultuje a posuzuje možné **varianty přístupů k řešení** projektů z pohledu datových zdrojů,
- spolupracuje na **specifikaci nástrojů pro transformace dat**.

7.3 Otázky k řešení PBI projektů

Analytické otázky k řešení PBI projektů vycházejí ze společného základu k řešení analytických aplikací konkretizovaných na prostředí Power BI. Otázky jsou rozdělené do několika skupin.

7.3.1 Vztah řešení PBI k byznysu

- Jak zvýšit **úspěšnost a výkonnost** byznysu díky vysoké kvalitě PBI aplikací?
- Provádí se **vyhodnocení ekonomiky** dané oblasti řízení pravidelně, promítá se do aplikací PBI?
- Dochází k hlubší **analýze dosahovaných parametrů** produktů, zboží a služeb, které se promítají do aplikací PBI?
- Jak zajistit potřebnou požadovanou **komplexnost a kvalitu** analytických aplikací PBI?
- Jak dosáhnout požadované **flexibility aplikací** vzhledem k momentálním potřebám manažerů a specialistů firmy?
- Jak dosáhnout potřebné **kvalifikace a motivace** uživatelů na řešení aplikací v prostředí PBI?
- Jak racionálně vymezit **očekávané efekty** PBI aplikací?

7.3.2 Datové zdroje

- **Jaké datové zdroje**, s jakými charakteristikami budou třeba pro plánované řešení PBI?
- Jsou potřebná **data dostupná** v rámci datového skladu?
- Které **další datové zdroje** jsou pro potřebná data k dispozici?
- Jak často mají být data v sémantickém modelu **aktualizována**?
- Je potřeba vidět data a jejich **změny v reálném čase**?
- Po kolik let budou třeba **historická data**?
- Je třeba sledovat **historii v datech podle určitých dimenzí** (např. produktů, zákazníků)?
- Jak mají být **nastavena práva na přístup** k datům pro určité uživatele nebo jejich skupiny?
- Obsahují data i některé **citlivé údaje**, např. podle GDPR?
- Bude třeba disponovat s daty **v reálném čase**?

7.3.3 Obsah a distribuce reportů na bázi PBI

- Které **reporty** a s jakým obsahem budou adekvátní jednotlivým úrovním a pozicím řízení?
- Které **metriky** budou pro jednotlivé reporty významné, které budou mít charakter KPI?
- Které **dimenze** ve vztahu k metrikám budou pro reporty relevantní?
- Jak dosáhnout požadovanou **granularitu dat** pro jednotlivé reporty?
- **Kolik uživatelů má mít přístup** k určitým reportům a dashboardům?
- Kolik uživatelů se má **podílet na vytváření** reportů a dashboardů?
- Mají mít uživatelé přístup k reportům a dashboardům i **přes mobilní zařízení**?
- Jak zajistit vysokou **prezentační úroveň** a vizualizaci dat v rámci reportů?

- Jak dosáhnout **konsensu mezi pracovníky** na navrženém obsahu a strukturách reportů?
- Jak posilovat **samostatnost pracovníků firmy** při řešení reportů a využívání analytických nástrojů?
- Jak se mají **řídít verze reportů a dashboardů**, podle jakých pravidel?
- Jak správně stanovit **perspektivu datových zdrojů** z pohledu řešených reportů?
- Jak dosahovat **zkracování doby a časové náročnosti** na přípravu reportů?

7.4 Poznámky k licencím

Existují dvě základní kategorie licencí:

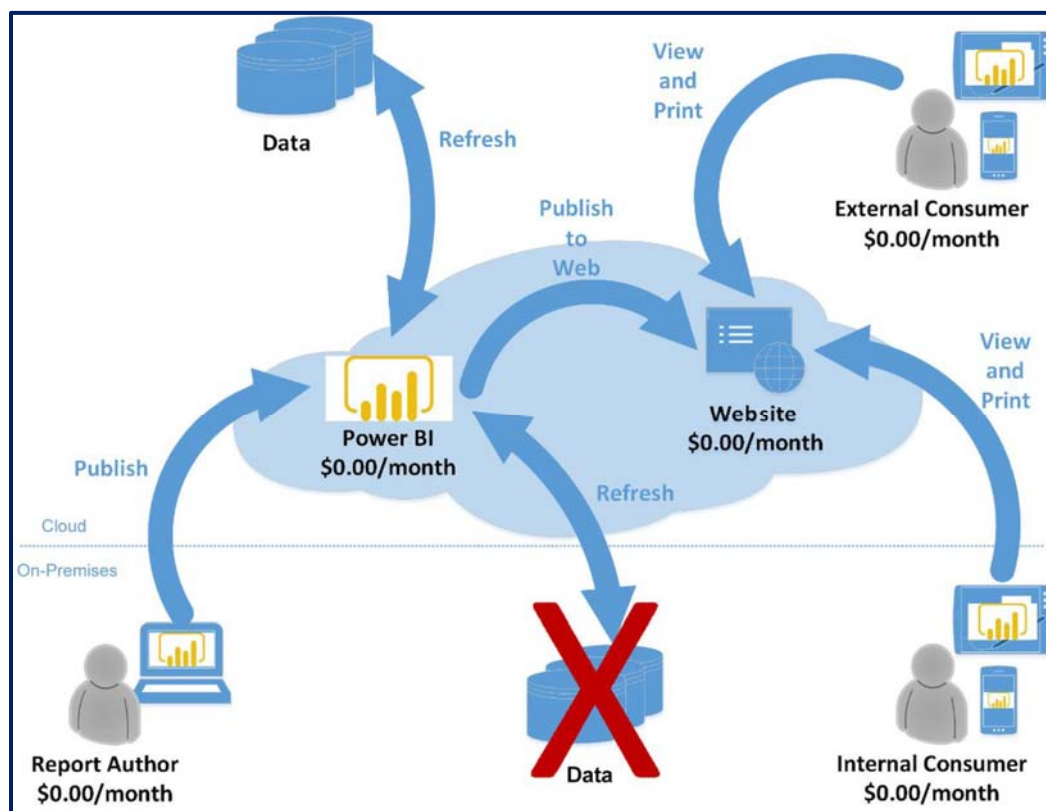
- Sdílená kapacita (Shared capacity).
- Vyhrazená, přiřazená kapacita (Dedicated capacity).

7.4.1 Sdílená kapacita (Shared capacity)

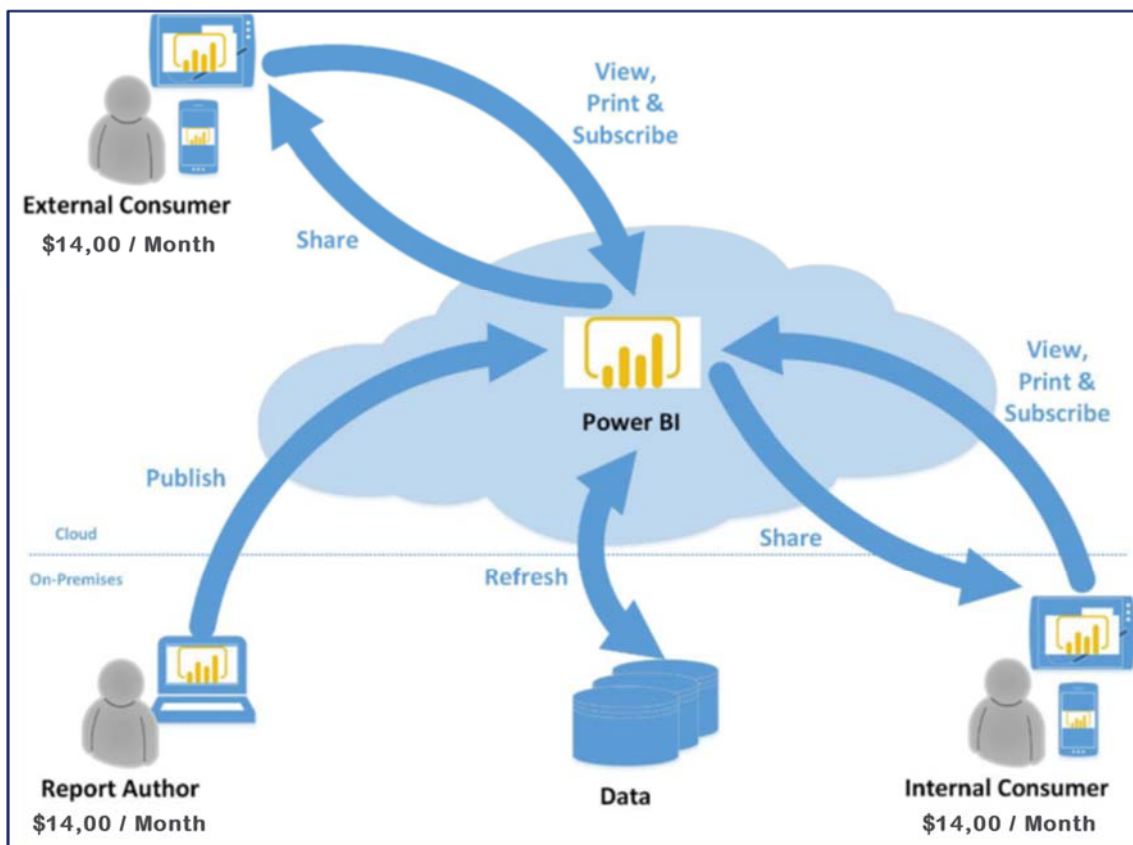
V rámci sdílené kapacity jsou dvě možnosti:

- **Free:**
 - PBI je k dispozici volně k stažení a užití.
 - Autoři reportů mohou používat PBI desktop pro vytváření sémantických modelů a reportů, mohou je publikovat v cloudu PBI services, nemohou zde však aktualizovat data.
 - PBI services v tomto případě neumožňují řadu dalších funkcí (autentifikace apod.).
- **Pro:**
 - Eliminuje omezení pro PBI services, uvedená v předchozím případě.
 - Umožňuje sdílet reporty interními i externími uživateli, kteří ovšem potřebují rovněž „Pro“ licenci.

Obě uvedené možnosti dokumentují následující obrázky:



Obrázek 7-4: Power BI „free“ (Zdroj: Deckler, G., Powell, B., 2022)



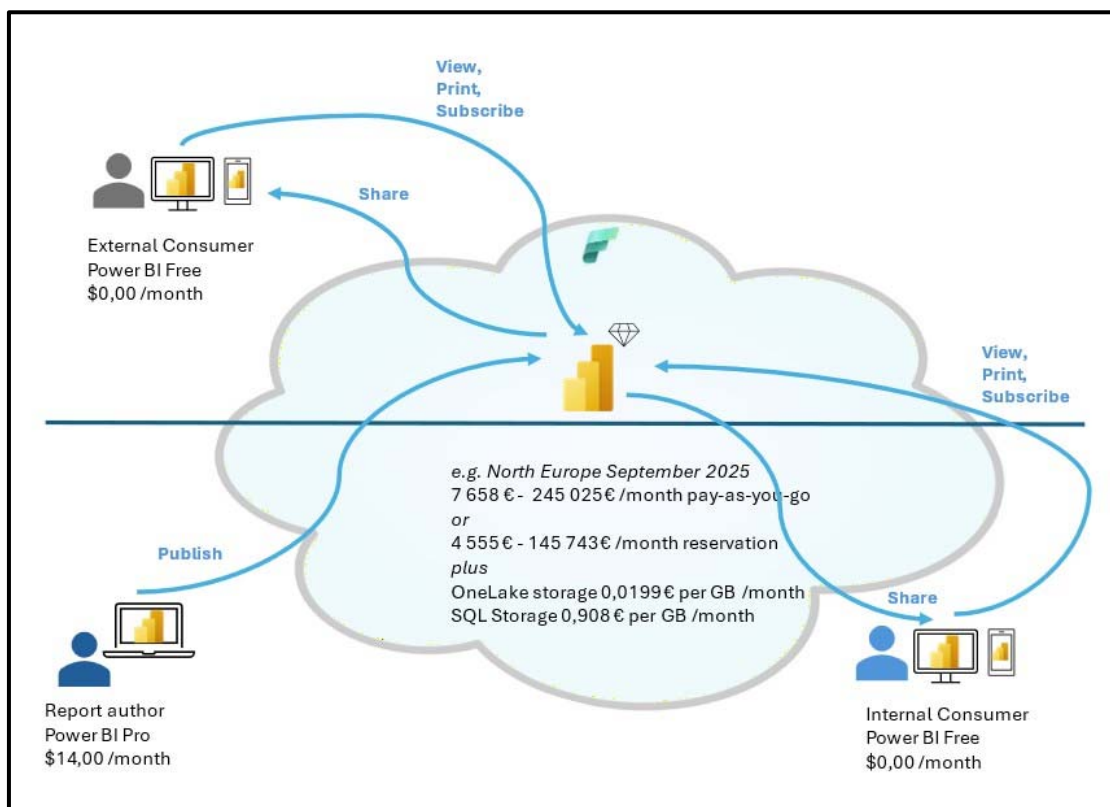
Obrázek 7-5: Power BI „Pro“ (Zdroj: Deckler, G., Powell, B., 2022 – upraveno)

7.4.2 Vyhrazená, přiřazená kapacita (Dedicated capacity)

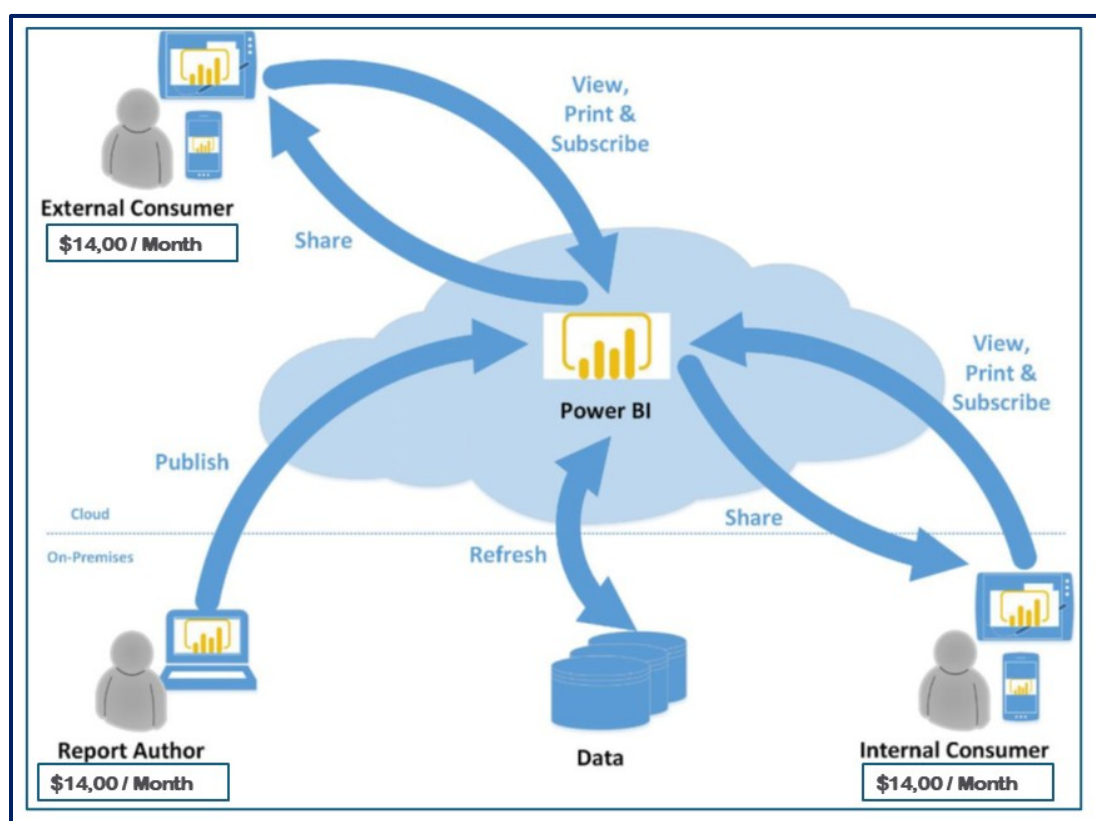
Tato licence rezervuje paměť a capacity pouze pro určitého vlastníka. Umožňuje řadu dodatečných funkcí (větší sémantické modely, častější aktualizace, komplexní reporty, scorecardy atd.). Existují tři varianty:

- **Fabric capacity F64 a vyšší:**
 - Uživatelé s licencí „Free“ mají přístup k reportům a dashboardům, kterým byla přiřazena „Fabric Capacity F64“ a vyšší, tj. s využitím mobilních zařízení a plně interaktivních PBI service funkcí.
 - Uživatelé „PBI Pro“ mohou sdílet dashboardy v cloudu i s „PBI Free“ uživateli.
- **Premium per User:**
 - Funguje téměř identicky s licencí „PBI Pro“, ale s nutností pro cloud vlastnit PPU licenci a umožňuje větší sémantické modely (až 400 GB, pokud je workspace nastaven na velké modely).
- **Embedded:**
 - Je určena pro dodavatele software k použití API pro propojení na PBI vizuály. Reporty a dashboardy do vlastních aplikací.

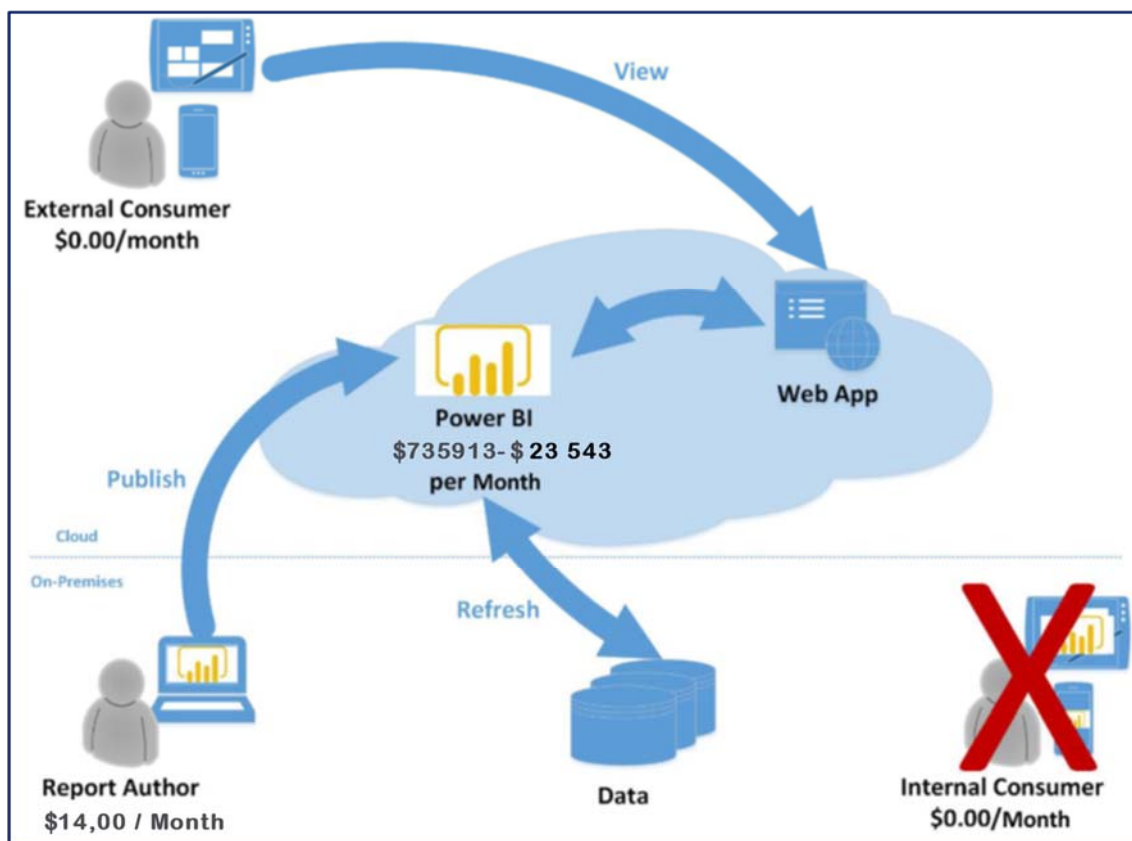
Tyto možnosti dokumentují další tři obrázky:



Obrázek 7-6: Power BI ve Fabric kapacitě > F64 (vlastní zpracování)



Obrázek 7-7: Power BI „Premium Per User“ – upraveno o ceny 2025
(Zdroj: Deckler, G., Powell, B., 2022)



Obrázek 7-8: Power BI „Embedded“ – upraveno o ceny 2025
(Zdroj: Deckler, G., Powell, B., 2022)

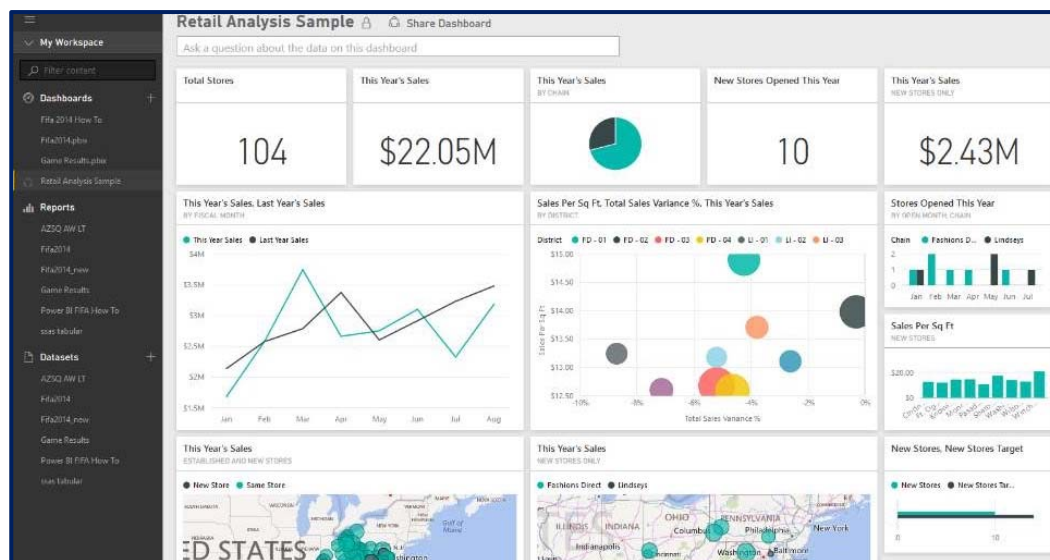
7.5 Pracovní závěry



Z kapitoly vyplývají následující **závěry**:

- **Úspěch PBI projektů**, obdobně i u jiných, úzce souvisí s kvalitou **plánování, organizace a efektivní kooperací** účastníků projektu, a to jak na straně řešitelské, tak i uživatelské.
- **Řízení projektů** je samozřejmě závislé na jeho složitosti a rozsahu vzhledem k firemnímu prostředí, ale i v tomto případě musí být **založeno na odpovídající úrovni nastavených procesů, data governance, řízení kvality dat, datového modelování a způsobů distribuce řešení**.

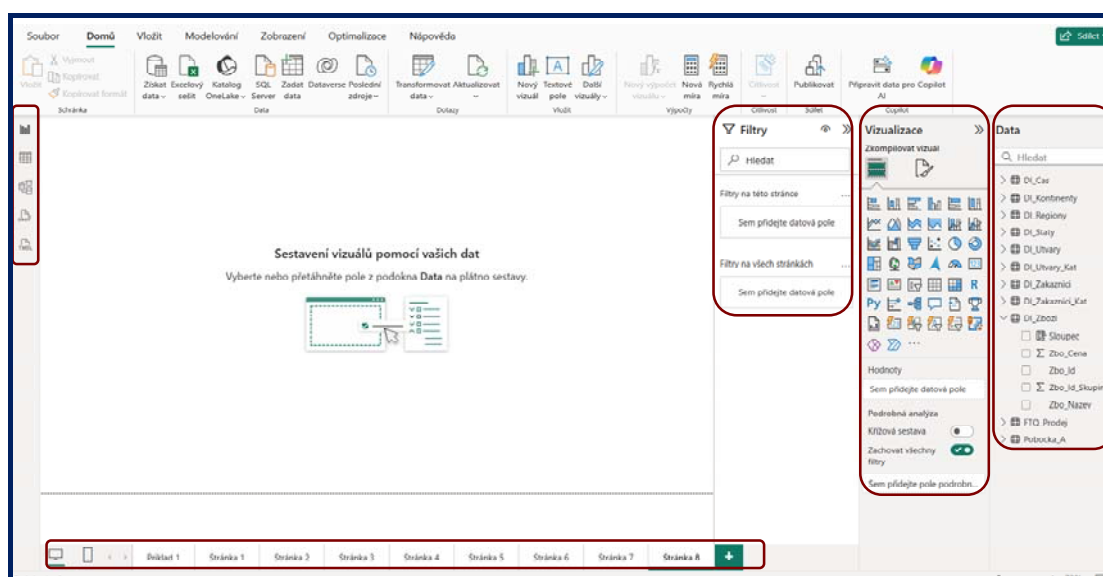
8. Plánování a řešení reportů v Power BI



Specifické nároky na řešení PBI aplikací má obsahová, ale i vizuální stránka reportů, dashboardů a dalších typů výstupů. Součástí jsou i jejich klasifikace a specifické nároky, jako např. pro dashboardy, tj. strategický, taktický, operativní apod.

Účelem kapitoly je specifikovat základní principy a doporučení, spojené s plánováním a řešením reportů v PBI a formulovat hlavní možnosti vizualizace v jejich rámci.

Základem uživatelského rozhraní jsou **vizualizace dat (či vizuály)**, které umožňují interakci s daty, načtenými a uloženými v sémantických modelech aplikace Power BI, nebo s daty, uloženými v databázích, k nimž Power BI přistupuje v režimu přímého dotazování (*Direct Query*), viz dále. Vizualizace dat **se realizuje v sestavách (reportech)**. Po publikování reportů do služby *Power BI Service* lze vybrané vizuály z reportů umístit na řídicí panely (dashboardy) a ty sdílet vybranými skupinami uživatelů nebo jednotlivými uživateli, kteří mají příslušnou licenci *Power BI Pro*, viz dále.



Obrázek 8-1: Pracovní prostor pro tvorbu reportů v Power BI desktop

8.1 Proces plánování reportu

V Power BI může report (sestava) **obsahovat jednu nebo více stránek** (Obrázek 5-1 vlevo dole). **Všechny stránky** dohromady jsou společně označovány jako **report**. Nová stránka se přidá klikem na „+“ vedle poslední záložky s názvem stránky. Stránky je možné **přejmenovat** stejně jako v excelovém sešitu, tj. pravým tlačítkem myši na záložce sestavy **funkcí „Přejmenovat“**.

Dříve, než se přistoupí k vytváření vizuálů v rámci reportů, je účelné reporty dobře naplánovat a zajistit tak jejich odpovídající hodnotu pro byznys. To znamená uskutečnit proces naplánování reportu(ů). Proces plánování reportu má (podle Deckler, G., Powell, B., 2022) šest základních kroků, viz další obrázek.



Obrázek 8-2: Proces plánování reportu

K jednotlivým krokům doplníme **několik poznámek**:

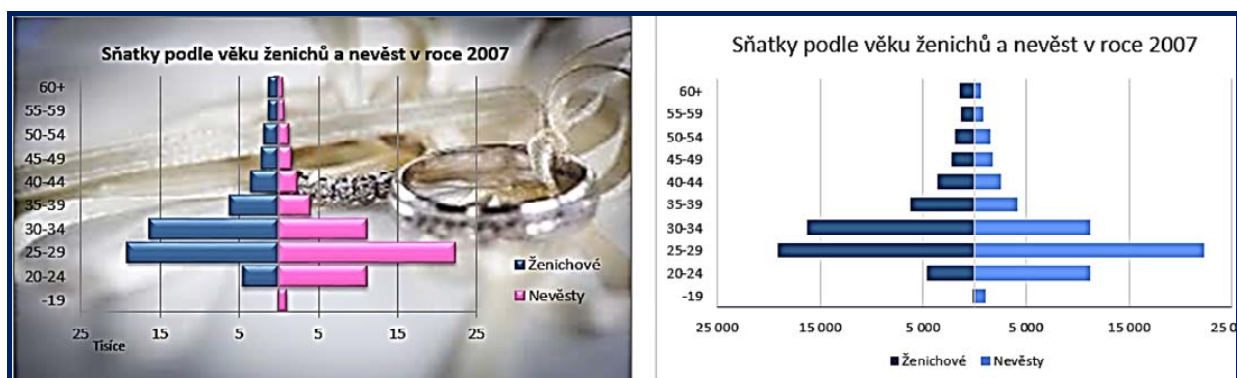
- **Identifikace uživatelů, konzumentů reportu:**
 - identifikace uživatelů reportu a jejich potřeb a priorit,
 - „storytelling“, to je zaměření na nejdůležitější KPI a vysvětlení, proč jsou vyšší nebo nižší, než se očekávalo, vysvětlení odchylek od normálu a významných trendů,
 - podnikoví analytici budou vyžadovat větší detail, spíše tabulky, více slicerů,
 - manažeři budou naopak preferovat nižší interaktivitu a intuitivní vizuály a schopnost zaměření na nejpodstatnější informace komplexního charakteru.
- **Definování byznys otázek pro report:**
 - specifikace otázek z byznysu, na které má report odpovídat,
 - nejdůležitější byznys otázka by měla být umístěna vlevo nahoře na reportu.
- **Posouzení, zda sémantický model odpovídá otázkám:**
 - prověření, že sémantický model obsahuje všechny potřebné metriky a dimenze vzhledem k byznys otázkám,
 - posouzení, zda kalkulované metriky v DAX mají odpovídající logiku k otázkám, zda vazby mezi faktovými a dimenzionálními tabulkami pokrývají podstatu otázek,
 - specifikace existujících omezení v sémantickém modelu.

- **Určení interaktivity reportu:**
 - určení, zda u konkrétního reportu půjde o vícestránkový report, realizovaný na bázi Power BI Builderu,
 - specifikace možností interaktivity v rámci reportu,
 - určení kombinace různých typů reportů.
- **Definování přístupů a distribuce reportu:**
 - specifikace možných forem přístupů k reportu a přístupů vzhledem k uživatelům,
 - určení, zda bude report využíván i na mobilních zařízeních, nebo bude zasílán emailem.
- **Náčrt formy a vzhledu reportu:**
 - vytvoření základního schématu reportu a umístění jednotlivých vizuálů,
 - doporučuje se nejpodstatnější informace umístit do středu,
 - na rozdíl od vytváření sémantického modelu, vyžadujícího větší časový objem, lze řešení formy reportu realizovat obvykle agilním přístupem,
 - doporučuje se využívat i diagram reportů při komunikaci s uživateli i vedením projektu,
 - diagram reportů specifikuje report, který je z pohledu byznysu hlavní, souhrnný, a reporty, které na něj navazují a jsou podstatně detailnější,
 - diagram specifikuje i možnosti funkce „drill through“, umožňující přecházet na velmi detailní i zdrojová data.

8.2 Doporučení pro vizualizace

Další přehled obsahuje několik hlavních doporučení podle (Deckler, G., Powell, B., 2022, upraveno) pro řešení vizualizace reportů:

- **Minimalizace** nepodstatných, resp. méně podstatných detailů:
 - každý vizuál by měl být jasně spojen s účelem reportu a byznys otázkami, které má report řešit,
 - vizualizace by neměla odpovídat především zájmům návrháře a jeho názorům,
 - je účelné eliminovat prvky, které se přímo neváží k tématu nebo k pochopení obsahu reportu,
 - Obrázek 8-3 ukazuje rozdíl mezi řešením vizualizace statistických dat. Zatímco pruhový graf vlevo obsahuje fotografii v pozadí grafu a sloupce jsou zobrazeny ve 3D grafice se zaoblením hran sloupců, působí graf vpravo velmi jednoduše a stroze. Řešení vlevo je příkladem přípravy vizualizace pro komerční prezentaci s cílem upoutat na téma, samotná statistika dat zde hraje vedlejší roli. Naopak graf vpravo na Obrázek 8-3 je příkladem jednoduchého a přehledného řešení v analytické úloze,



Obrázek 8-3: Rozdíl v řešení vizualizace dat

- data v analytických úlohách informačního systému by měla být prezentována tak, aby žádné nepodstatné prvky (texty, doprovodná grafika apod.) zbytečně neodváděly pozornost od základních informací,
- pro správný výběr grafického vyjádření je třeba si uvědomit, že pracujeme s různými typy dat, které mají rozdílný účel i způsob zpracování.
- Užití **jednoduchých a jasných titulků** stránek reportů a vizuálů:
 - textová pole pro jména reportů a stránek reportů musí být jednoduchá a současně srozumitelná,
 - je dobré, když reporty obsahují datum poslední aktualizace, resp. obnovy.
- **Umístění vizuálů** a jejich skupin na ploše má usnadnit efektivní a logickou navigaci v reportu:
 - nejdůležitější vizuály se umísťují do levého horního rohu plochy,
 - vizuály, které spolu logicky souvisejí, mají být umístěny do stejného prostoru na ploše.
- Použití **méně výrazných a přirozených barev** pro vizuály:
 - je dobré minimalizovat použití velmi jasných nebo naopak tmavých barev,
 - výrazné barvy je účelné použít pro významné hodnoty i s využitím podmíněného formátování,
 - je třeba respektovat firemní standardy pro formální stránku vizuálů.
- Použití **rozdílných barev v grafech** pro rozlišení různých prvků a hodnot:
 - např. rozdílné barvy se používají pro rozdílné kategorie produktů.
- **Umístění vizuálů vertikálně a horizontálně** má být na ploše tak, aby mezi nimi byl adekvátní prostor:
 - rozsah mezer musí jasně rozlišovat vizuály jako separátní entity.
- **Pozadí plochy** má být přirozeně kontrastní k vizuálům:
 - obvykle se doporučuje bílá barva.
- U sloupcových grafů se **doporučuje seřadit sloupce** podle vybrané míry:
 - toto seřazení je aplikovatelné u nominálních proměnných, jako jsou produktové kategorie. Není vhodné u ordinálních kategorií (typicky se jedná například o roky nebo měsíce v roce).
- **Minimalizace velkých prázdných prostor** na ploše.
- Je účelné **doplňovat vizuály o cílové nebo doplňující hodnoty** (maxima, průměry apod.):
 - analytické reporty je dobré efektivně doplňovat informacemi trendového charakteru, prediktivní analytikou, manažerskými odhady.
- Všechny reporty mají mít **stejnou nebo obdobnou formu**:
 - reporty v rámci projektu se mají řídit stejnými pravidly, obvykle založenými na firemních standardech.

Podrobněji se faktorům, ovlivňujícím využitelnost a srozumitelnost grafů, vrátíme v podkapitole 8.5.



Poznámky v dalších podkapitolách vycházejí zejména z textu (Šedivá, Z. a další, 2018). **Účelem je charakterizovat možnosti** vizualizace **na obecnější úrovni** s tím, že některé z nich mají odraz ve vizuálech Power BI, v jiných případech jde o obecnější doporučení s případnou realizací v uživatelsky vytvářených vizuálech.

8.3 Tabulky v reportech

Tabulky dat představují jednu z forem prezentace dat a jsou **vhodné** tam, kde uživatel vyžaduje sledovat **detailní informace o sledovaných veličinách**. K tabulkám patří tyto charakteristiky:

- Tabulka představuje **jednoduchou strukturu** uspořádání dat. Jednoduchost uspořádání dat je dobrým **předpokladem snadné orientace** v datech tabulky. Vzhledem k velkému množství zobrazovaných dat je třeba zvážit, **zda se pro cíl prezentace forma tabulky hodí**, zda neprezentuje příliš detailní data, která nejsou ze strany uživatele vyžadována.

- Tabulky jsou **součástí reportů především na operativní a taktické úrovni**. Pro lepší přehlednost a snazší orientaci v datech je účelné **doplnit údaje i dalšími grafickými prvky**, které zdůrazní důležité odchylky, limity atd. Proto je vhodné **doplnit do dat určitý kontrast**, buď barvou, změnou umístění, nebo například přidáním piktogramu k datu.
- Při zpracování tabulek se vizualizace často řeší **prostřednictvím podmíněného formátování** buněk, tj. formátu, který se uplatní při splnění podmínky. Hodnoty jsou zvýrazněny barvou písma, barvou buňky, ikonami, které často upozorňují na trendy, odchylky apod.
- Zde je třeba upozornit na možné **riziko nepřehlednosti dat v tabulce**, pokud je použito příliš mnoho vizualizačních prvků, nebo příliš mnoho barev. Také v případě, že není uvedena legenda pro dané vizualizační prvky, nemusí být na první pohled jasné, co vlastně představují a na základě jakých pravidel jsou použity (např. nejsou uvedeny limity hodnot pro použitá barevná schémata).

8.4 Typy grafů a jejich účel

Grafy jsou velmi častou vizualizační komponentou pro zobrazení ekonomických (statistických) dat. Ne vždy je zvolen správný výběr grafu pro daná data a cíl jejich prezentace. Abychom mohli určit, který typ grafu je nejvhodnější pro prezentaci dat, **je třeba si uvědomit, s jakým cílem, resp. myšlenkou, chceme data prezentovat**. Těmta data můžeme prezentovat z různých úhlů pohledu a zdůraznit tím jiný aspekt výstupu. Je důležité předem stanovit, pro koho prezentaci vytváříme, z jakého pohledu na ni bude uživatel nahlížet a jakou informaci potřebuje z dat co nejdříve získat.

Určujícím aspektem výběru grafu by měl být vztah mezi daty, podle kterého chceme data prezentovat. Tento vztah dat lze rozdělit do několika následujících typů (Šedivá, Z. a další, 2018):

- **nominální porovnání** – grafické znázornění kvantitativních hodnot v různých kategoriích, např. počet obyvatel v zemích EU, počet prodaných automobilů apod.,
- **rozdílný (odchylka)** – znázorňuje, jak se daná hodnota odchyluje od průměru nebo od očekávané hodnoty, např. plánované náklady vs. skutečné náklady,
- **časová řada** – prezentuje změny hodnot v časovém období, např. výše tržeb v jednotlivých měsících,
- **distribuce** – prezentuje rozložení dat kolem dané centrální hodnoty, např. sledujeme výši důchodů ve skupině seniorů, výšku hráčů v basketbalovém týmu atd.,
- **korelace** – představuje vztah dvou a více hodnot, např. výše tržeb v závislosti na denní době prodeje, výše mzdy vzhledem k dosaženému vzdělání,
- **vztah částí vůči celku** – porovnává podskupinu (podmnožinu) dat vůči celku, např. podíl fixních nákladů na celkových nákladech v rozpočtu projektu,
- **pořadí** – vzájemně porovnává dvě a více hodnot v pořadí podle sledované veličiny, např. žebříček studentů podle studijních výsledků.

Výše uvedené **vztahy mezi daty jsou rozhodující pro správné zvolení typu grafu a jeho nastavení**. V následujícím textu budou popsány **pouze vybrané, nejčastěji užívané typy grafů** a budou uvedena i základní pravidla pro jejich tvorbu (Šedivá, Z. a další, 2018).

8.4.1 Sloupcový graf (Column Graph)

Sloupcový graf je **velmi často používaným typem** grafu:

- Představuje nástroj pro **porovnávání kvantitativních a kvalitativních hodnot**, kdy kvalitativními hodnotami jsou často jednotky času.
- Slouží pro **nominální porovnávání kvantitativních hodnot** pomocí sloupců a využívá výšky sloupce pro znázornění hodnoty. Rozdíly v délce sloupce jsou pro člověka dobře rozpoznatelné, takže uživatel rychle získá přehled v porovnávání hodnot. Hodnoty mohou být znázorněny v jedné či ve více datových řadách.
- Sloupcový graf slouží velmi často také pro **vyjádření vztahu hodnoty k celku**, a to formou skládaného (nastavovaného) grafu. Hodnoty jednotlivých kategorií jsou ukládány do jednoho sloupce, takže lze získat přehled o podílu jednotlivých kategorií na celkovém výsledku.

- **Sloupcový skládaný graf** lze vyjádřit buď v absolutních hodnotách, či v procentuálním vyjádření. Sloupce pro % vyjádření podílů vždy mají stejnou délku, která odpovídá 100 %. Je třeba zdůraznit, že tento typ grafu je vhodný jen tehdy, když zobrazované hodnoty jednotlivých kategorií tvoří celek, tedy dohromady 100 %.
- **Obdobou sloupcového grafu je graf pruhový.** Jedná se o horizontální variantu sloupcového grafu. Pruhový graf je vhodný, pokud zobrazujeme více položek, které by při omezené šířce na ose x byly nesnadno zobrazitelné. I tento typ grafu umožňuje skládanou formu.
- Pruhový graf se používá v různých upravených formách **pro zachycení a porovnávání hodnot v časovém období ve formě časových linek (time lines).**
- Pro porovnání dvou datových řad je vhodné použít **pruhový graf se zrcadlovým vynášením dat** na ose x. Tento typ pruhového grafu je **přehledný tehdy, když jsou větší rozdíly hodnot mezi datovými řadami.** Pokud jsou hodnoty obou skupin přibližně stejné, rozdíly jsou málo zřetelné a vizualizace neplní svoji úlohu. Navíc oko čtenáře musí přeskakovat z pravé strany na levou, aby mohlo porovnat hodnoty navzájem. Potom je lépe vynášet data jedním směrem u obou datových řad.

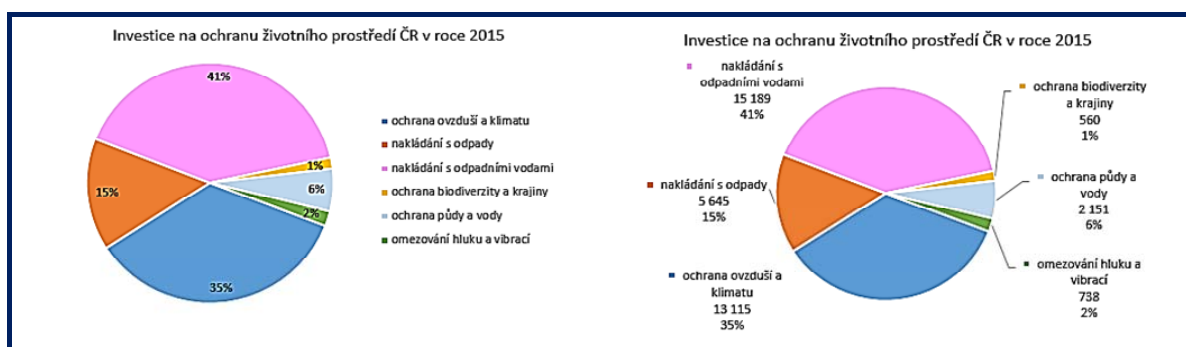
Dvojitý pruhový graf

- Dvojitý pruhový graf je speciální **variantou sloupcového grafu, která je přehledná pro zjištění vztahu mezi dvěma proměnnými.** Typicky se používá pro porovnávání skutečných a plánovaných hodnot (např. tržeb, nákladů, zisku).
- Obě formy sloupcového grafu velmi dobře slouží **k rychlému porovnání rozdílů mezi hodnotami.** Popisky dat lze doplnit ikonami či jinou doprovodnou grafikou.

8.4.2 Výsečový graf (Pie Chart)

Výsečový (koláčový) graf **slouží pro zobrazení vztahu části k celku pro jednu kvantitativní proměnnou.** Graf nemá žádné osy, hodnoty jsou vynášeny ve formě výsečí v kruhu, kdy celý kruh tvoří 100 %. Šířka výseče, tj. její úhel, odpovídá podílu části z celku:

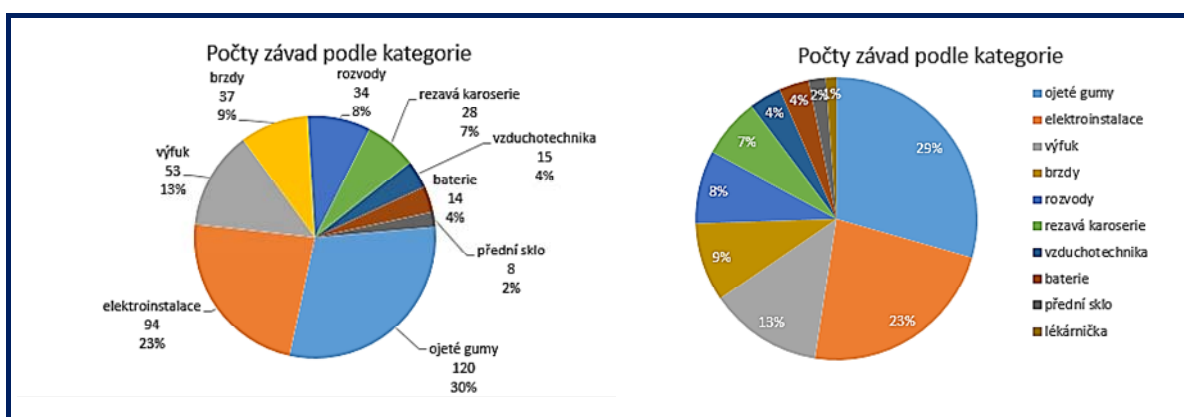
- Ačkoliv je výsečový graf v praxi velmi oblíbeným typem grafu, představuje i určitá **úskalí a chybné interpretace hodnot.** I když je kruh jako forma tvarově uživateli blízká a příjemná, z pohledu práce s daty je podle odborníků v oblasti vizualizace dat (Few, 2004) velmi nepřesný pro vnímání velikosti dat podle tvaru výsečí.
- Lidské oko je schopno dobře **rozpoznat výseče po čtvrtinách kruhu**, tj. 90°, 180°, 270°. Ostatní stupně úhlu výsečí a rozdíly mezi nimi již nejsou dobře rozeznatelné. Přesná čísla oko vizuálně nezvládne odhalit (přeceňuje tupé úhly, podceňuje ostré úhly).
- Na příkladu dat z ČSÚ (CZSO, 2016) na obrázku (Obrázek 8-4), obsahujícím porovnání ročních investic na ochranu životního prostředí v ČR v roce 2015, jsou popsány **nedostatky výsečového grafu.**



Obrázek 8-4: Výsečový graf s popisky

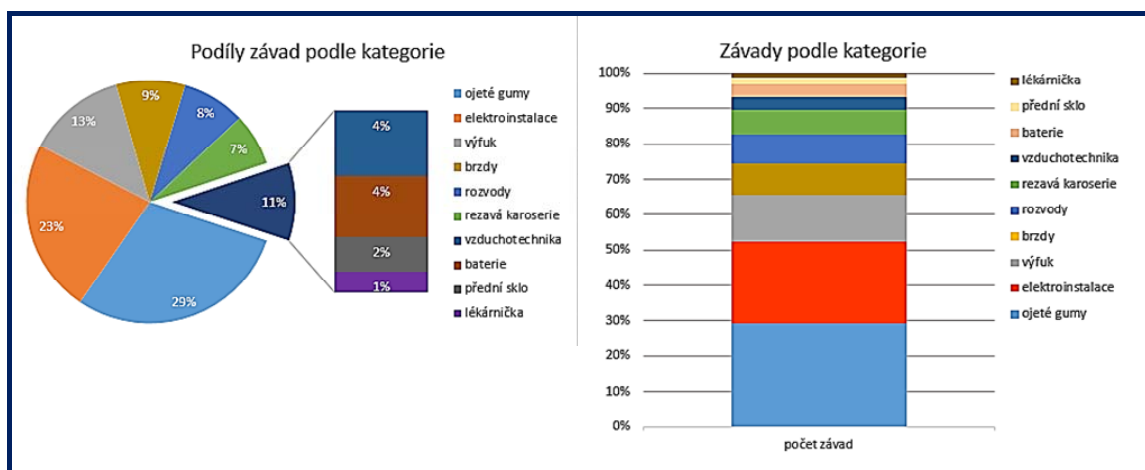
- V koláči grafu je uvedeno pouze **6 výsečí**, ale jejich **porovnání z pohledu velikosti podílů, nelze řešit bez popisků výsečí.** Těžko lze totiž vizuálně odhadnout, zda výseč s 41% podílem je větší a případně o kolik, než výseč s 35%. Lze rozpoznat pouze velké rozdíly mezi velikostí výsečí.

- **Popisky jsou proto nezbytné.** Jejich umístění je však často problematické. Popisky s hodnotami podílů by měly být umístěny **co nejbližše výšečím, a to nejlépe i s názvem kategorie**. Na obrázku (Obrázek 8-4, vlevo) jsou názvy kategorií vloženy mimo koláč grafu v legendě, zatímco v řešení vpravo popisky výšečí obsahují nejen procentní podíl, ale i název kategorie a hodnotu investic, uložené vně koláče.
- **Umístění výšečí** je problematické vzhledem k čitelnosti popisků. Při větším počtu kategorií (výšečí) jsou často popisky velmi špatně čitelné, a proto je žádoucí použít jiný typ grafu.
- Obrázek 8-5 představuje **data o počtu závad automobilů**, ve výšečovém grafu jsou znázorněny hodnoty podle kategorie závady. Je zde více kategorií závad s velmi rozličnými podíly z celku (od velmi nepatrných podílů několika procent až po podíly v desítkách procent).
- Pro lepší přehlednost grafu jsou **obě řešení seříděna sestupně** od největších po nejmenší podíly. Řešení vlevo na obrázku představuje jakéhosi nevzhledného pavouka popisků, u řešení vpravo byla první (největší) výšeč z důvodu lepší přehlednosti umístěna od 0°, ale nejmenší výšeče jsou umístěny vedle sebe velmi blízko, takže není možné umístit úplné informace do popisků a názvy kategorií jsou umístěny v legendě. Procentní podíly u výšečí nejsou dobře čitelné.



Obrázek 8-5: Nevhodné použití výšečového grafu pro více kategorií

- Příklad **s větším počtem kategorií s velkými rozdíly hodnot** lze řešit **použitím grafu s dílčí výšečí**, do které umístíme hodnoty s velmi malými podíly z celku. Na obrázku (Obrázek 8-6) jsou všechny podíly do 5 % kumulovány do jedné výšeče, která je popsána detailně ve vedlejším skládaném sloupcovém grafu.
- Vytržením této výšeče z grafu je **zvýrazněn kontrast k ostatním výšečím, který na ni upozorňuje**. V obrázku vpravo je zvolen také skládaný sloupcový graf, který zobrazuje procentní podíly, jejichž hodnoty snadno odečteme podle mřížky. Mřížku je možné více „zahustit“ – upravit krok na 5 %.



Obrázek 8-6: Alternativní řešení výšečového grafu pro více kategorií

- Při vytváření výšečového grafu je dobré **dodržet následující pravidla**:
 - výšečový graf používat pro malý počet kategorií, optimálně do pěti kategorií, pro větší počet kategorií zvolit jinou alternativu – skládaný sloupcový graf,
 - popisky uvádět přímo u výsečí, legendu používat jen pokud je to nezbytně nutné,
 - pro lepší přehlednost hodnot grafu, je účelné hodnoty seřadit sestupně od největší po nejmenší hodnotu,
 - první výseč umístit od 0° – zajistí se tak lepší čitelnost grafu. Člověk je zvyklý číst hodnoty shora dolů ve směru hodinových ručiček, proto jsou výseče v tomto směru přehlednější.
- **Variantou** výšečového grafu je **prstencový graf**. Také prstencový graf slouží k prezentování **podílů částí z celku**, kdy jednotlivé podíly jsou **představovány délkou oblouku**. Pro umístění oblouků a jejich popis platí stejná pravidla jako u výsečí ve výšečovém grafu.

8.5 Faktory ovlivňující srozumitelnost grafu

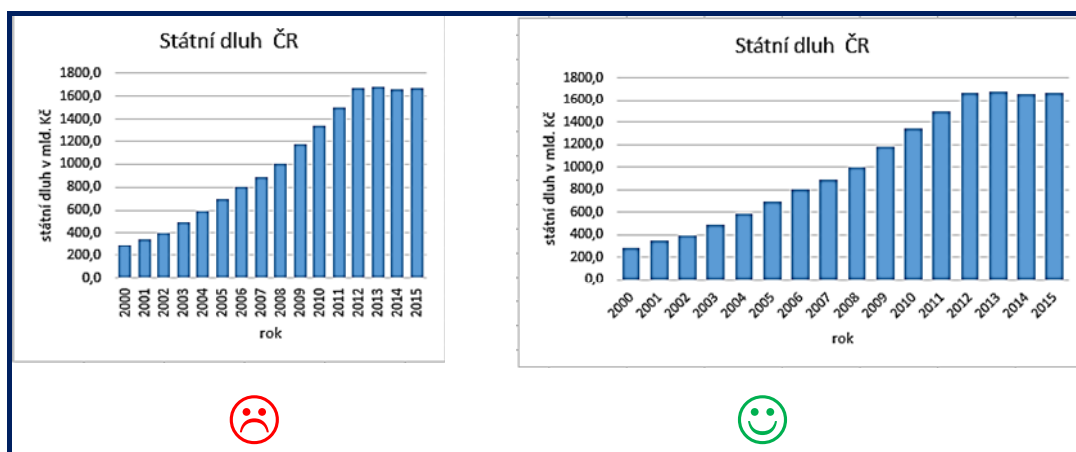
Ani správný výběr typu grafu nezaručí, že budou data správně pochopena a interpretována. **Nedostatečný popis, nebo naopak příliš mnoho popisků či nevhodné nastavení** parametrů grafu může způsobit špatnou interpretaci informace ze strany uživatele. V předchozím textu jsme uvedli rady pro návrh u konkrétních typů grafů. Nyní si uveďme ještě **několik obecných pravidel** pro správně graficky navržený graf.

Základním cílem návrhu je čistý, přehledný a srozumitelný design. Prezentovat data tak, **aby se uživatel v grafu dobře orientoval a získal informace rychle**. Aby nemusel dlouho přemýšlet, co vlastně data představují a nemusel ještě informace dlouho dohledávat, nebo dokonce sám dopočítávat.

8.5.1 Nastavení poměru stran grafu

Tento faktor může velmi **nepříjemně ovlivnit vnímání hodnot a jejich vztahů**:

- Na příkladu dat z ČSÚ (CZSO, 2017a), zobrazovaných v grafu (Obrázek 8-7), je ukázáno **vlevo nevhodné nastavení** poměru stran grafu, který vede k mylné interpretaci strmějšího růstu státního dluhu ČR, než je tomu v grafu vpravo.
- Vpravo vývoj dluhu není opticky tak prudký. Pro určení poměru stran os grafu je vhodné využít pravidlo třetin, tj. poměr šířka: výška = 3:2.



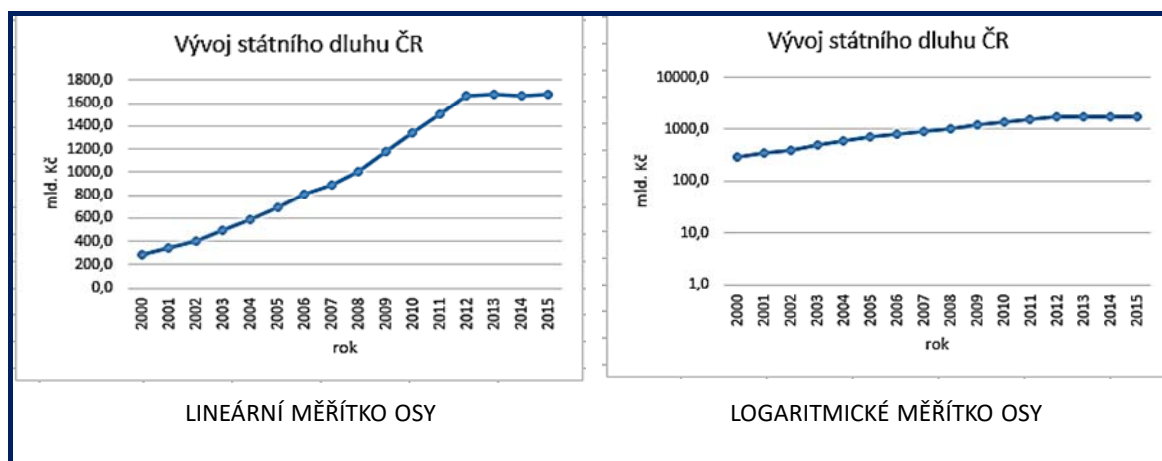
Obrázek 8-7: Správné nastavení poměru stran grafu

8.5.2 Měřítko osy y

Především u sloupcových a spojnicových **grafů je třeba data velmi dobře popsat**, aby uživatel neinterpretovat data chybně:

- Například **osa y** může být nastavena **kromě lineárního měřítka také v logaritmickém měřítku**.

- Obrázek 8-8 demonstruje na datech z (CZSO, 2017a) **vlevo použití lineárního měřítka** osy y, kde je zobrazen vývoj státního dluhu v jednotlivých letech období let 2000 až 2015.
- Graf **vpravo** má nastavenou osu **y v logaritmickém měřítku**, a proto data v grafu představují rychlost růstu státního dluhu.
- Pokud bude graf vpravo takto obecně pojmenován („Vývoj státního dluhu v ČR“), potom může **uživatel snadno toto nastavení měřítka osy y přehlédnout** a získat chybnou informaci. Proto je **nutné u grafu vpravo změnit název grafu a upřesnit i název osy y**.

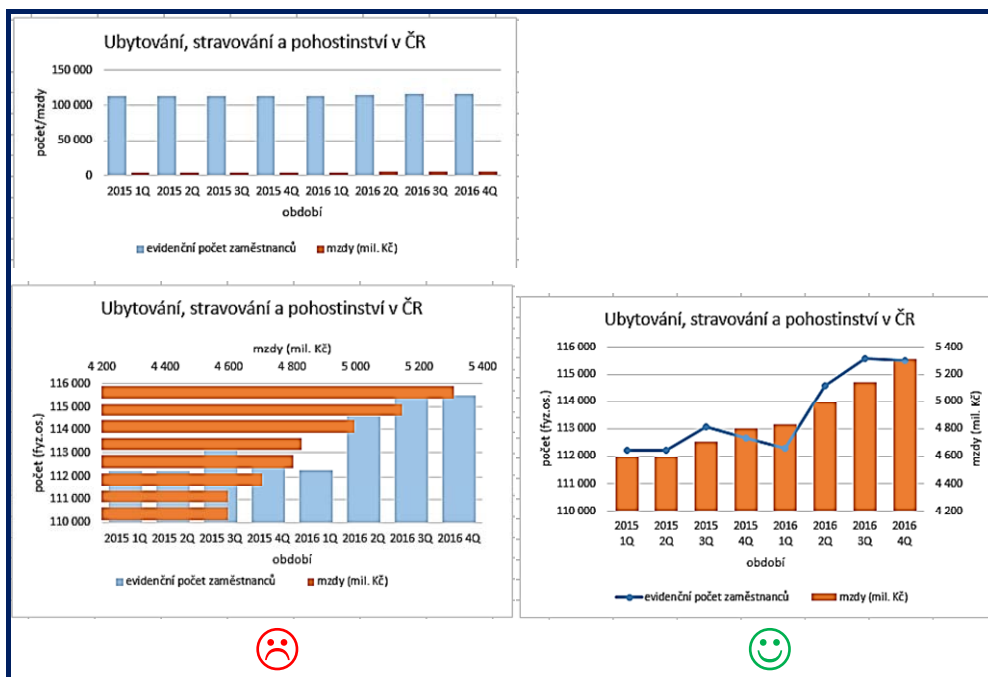


Obrázek 8-8: Nastavení měřítka osy y

8.5.3 Rozdílný rozsah hodnot na ose y u více datových řad

Velmi častým faktorem, způsobujícím špatnou čitelnost, je **nevhodné nastavení typu grafu a jeho osy y**.

- Na následujícím příkladu dat, převzatých z (CZSO, 2017c), jsou na Obrázek 8-9 zobrazeny **počty zaměstnanců a výše mezd v mil. Kč** v jednotlivých čtvrtletích let 2015 a 2016 v sektoru ubytování a stravování v ČR.
- Pokud bychom prezentovali tyto údaje v jednom grafu, vynášené na společné ose y, budou **rozdíly výšky sloupců mezi počtem zaměstnanců a výší mzdy tak velké**, že mzdy jsou téměř **nečitelné** (řešení vlevo nahoře).
- **Alternativou** řešení je **použití kombinovaného grafu** s hlavní a vedlejší osou y.



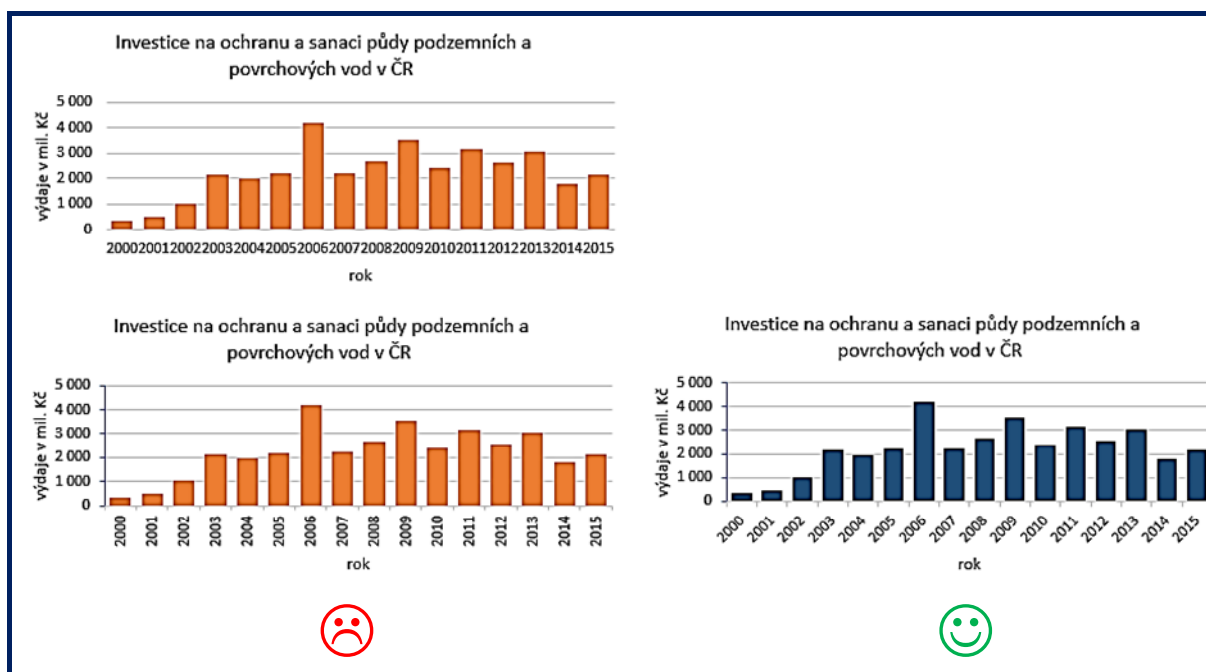
Obrázek 8-9: Použití vedlejší osy y

- Na grafu **vlevo dole** však byla zvolena zcela **nevhodná varianta kombinace horizontálního a vertikálního směru osy y** u sloupcového grafu. Graf je tak **velmi nesrozumitelný**.
- **Rozdílný charakter dat** (mzdy v řádech tisíců, počet zaměstnanců v řádech statisíců) lze vhodně zvýraznit nejen **rozdílným typem grafu (sloupcový vs. spojnicový)**, **ale především hlavní a vedlejší osou y** pro každou datovou proměnnou tak, jak je to použito v řešení grafu vpravo dole.

8.5.4 Popisky osy x

Popisky os x a y, používané u převážné většiny grafů, **by měly být vždy dobře čitelné a měly by vhodně doplňovat prezentované hodnoty**:

- **Problematickým** často bývá řešení, kdy je **na ose x v grafu zobrazeno mnoho popisků**.
- **Správně** by měly být **popisky v horizontálním směru** pro snazší čtení, ale to vždy není možné. Obrázek 8-10 demonstruje problém nastavení směru popisků u více hodnot na ose x.
- **Horizontální ani vertikální směr** zdola nahoru u popisků vlevo na obrázku **není přehledným uspořádáním**. V tomto případě lze použít **kompromisní řešení** – nastavit úhel cca 45° pro snadné čtení popisků osy x (viz řešení vpravo).



Obrázek 8-10: Směr popisků určuje jejich čitelnost

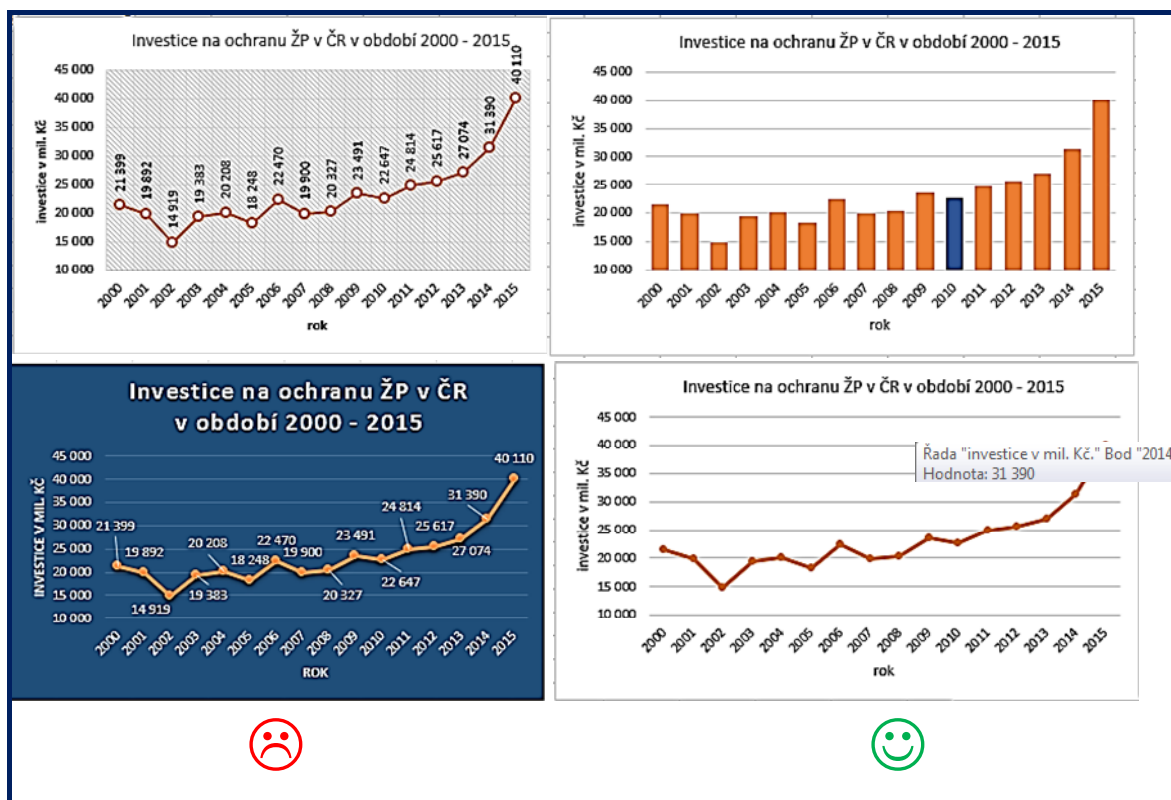
8.5.5 Popisky hodnot

I zde platí **prvořadý cíl: čitelnost**, kterou lze dodržet, když:

- umístíme popisky v horizontálním směru,
- nastavíme kontrastní barvu a velikost písma popisků vzhledem k barvě pozadí,
- popisky umístíme tam, kde jsou nejlépe čitelné – mimo sloupec, mimo výseč.

Pokud je v grafu málo místa pro popisky hodnot, potom je **lépe popisky hodnot vůbec nepoužít**:

- Lze **porovnat přehlednost grafů** na Obrázek 8-11.
- Příklady **vlevo** obsahují popisky, které však **nejsou příliš čitelné** (především u grafu vlevo dole).
- **Alternativním řešením** může být **sloupcový graf**, u kterého lze hodnoty lépe přečíst podle mřížky, nebo na příkladu **vpravo dole využít bublin** s popiskem pro zvolený rok v grafu.



Obrázek 8-11: Umístění popisků hodnot

8.5.6 Doprovodná grafika, značky, vysvětlivky,

Grafický návrh grafu v analytických aplikacích by měl být **čistý, jednoduchý, oproštěný od nadbytečných dat**, která čtenáře zatěžují zbytečnými informacemi a odvádějí pozornost:

- Toto pravidlo se týká i **používání jakékoliv doprovodné grafiky, značek, obrázků**, vysvětlivek a také písma.
- Doprovodná grafika by na sebe neměla zbytečně upozorňovat** s výjimkou situace, kdy je to žádoucí, např. upozornit na odchylku dat, výjimečnou hodnotu apod. Z toho vyplývá, že omezíme používání doprovodných obrázků, textur a ozdobných linek jen na místa, kde mají plnit svůj účel.
- Např. pokud používáme **ohraničení sloupců** či ploch v plošném grafu, budou tyto linky nevýrazné, tenké, s neutrální barvou. Nebudou tak rušit tvary grafu.
- Velmi nevhodné** je **používání různých textur a vzorků**. Šrafování se využívá především jako způsob řešení kontrastu u většího množství barevných ploch, např. u kartografické mapy, tepelní mapy apod.
- Plocha vyplněná vzorkem** je výrazným kontrastním prvkem vzhledem k ostatním plochám, vyplněným jednou barvou.

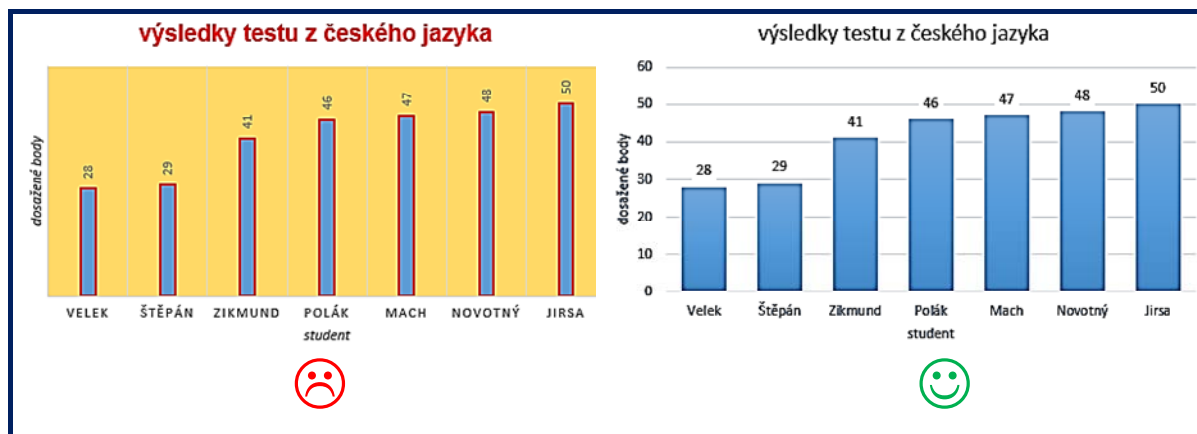
8.5.7 Písmo

Texty se v grafu používají v hlavním **nadpisu grafu**, v **popiscích os**, u **popisků hodnot** či v **legendě**. Základní **pravidla pro výběr písma** jsou následující:

- používat **fonty, které jsou velmi dobře čitelné i v malých velikostech písma** (např. v popiscích hodnot na ose x, kde jsou údaje často hustě blízko sebe),
- proto se využívají **především bezpatkové (sans-serif) fonty** Arial, Calibri, Tahoma. Jsou objemově úsporné s jemným tahem písma,
- nehodí se fonty používané v tiskových médiích** jako Times New Roman, Book Antiqua, Courier New, apod. Jedná se o patkové (serif) fonty, které jsou v malých velikostech hůře čitelné,

- **používat jeden font pro všechny texty v grafu**, nemíchat více fontů písma v témže grafu, měnit pouze velikost písma,
- **vyvarovat se používání různých řezů písma**, tzn. nekombinovat písmo tučné, kurzívu nebo písmo podtržené v jednom grafu.

Můžeme porovnat **správné a chybné řešení grafu** na Obrázek 8-12. Kolik chyb, které porušují výše uvedená pravidla, v grafickém návrhu sloupcového grafu lze najít?



Obrázek 8-12: Chybný a správný návrh grafického řešení sloupcového grafu

Chybný návrh vlevo (Obrázek 8-12) obsahuje souhrnně **tyto chyby**:

- **různé typy (fonty) písma** – v nadpisu grafu je použit Arial, ostatní texty jsou v Calibri,
- jsou **kombinovány různé řezy písma**: kurzíva u nadpisů osy x, tučné písmo v nadpisu grafu, kapitálky u popisků hodnot na ose x,
- příliš **úzké sloupce** a velký prostor mezi sloupci,
- **popisky hodnot** u sloupců jsou uvedeny **vertikálně** – špatně se čtou,
- **barva pozadí grafu** je v syté žluté, popisky hodnot ve světle šedé barvě – malý kontrast písma vůči pozadí vede ke špatné čitelnosti hodnot,
- **sloupce jsou orámovány neúměrně** tlustou čarou vůči šířce sloupce – odvádí to pozornost od hodnot a porovnání rozdílů mezi sloupci,
- **barva hlavního nadpisu** grafu a orámování sloupců je **červená** – tento kontrast odvádí pozornost od samotných hodnot,
- **v pozadí byla použita vertikální mřížka**, která zde plní jen funkci vzorku, odděluje vizuálně prostor mezi sloupci. Zcela je zde potlačena podpora rychlého a snadného čtení hodnot na ose y, k čemuž je tento typ grafu primárně určen,
- příklad **vpravo** na obrázku prezentuje velmi **jednoduchý a přehledný návrh**.

8.6 Pracovní závěry

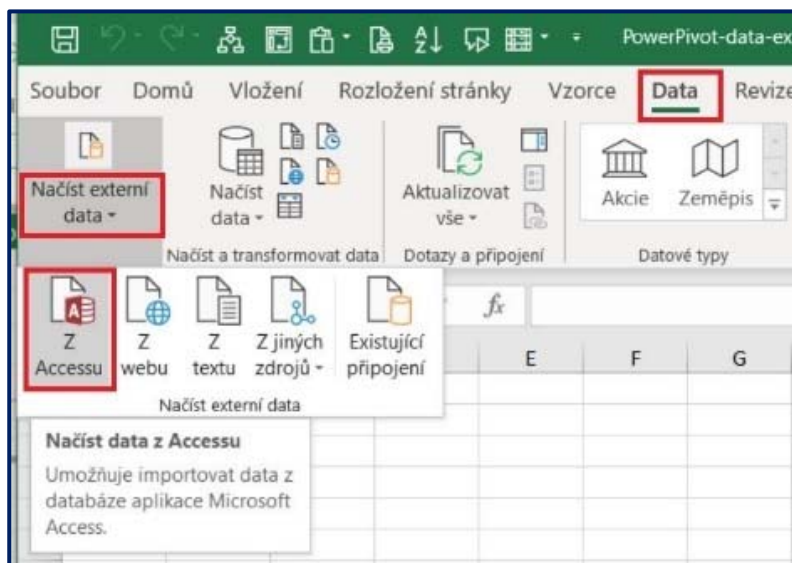


Z kapitoly vyplývají následující **závěry**:

- Jednotlivé reporty **je třeba dobře naplánovat** a zajistit tak jejich **odpovídající hodnotu pro byznys**. Plánování má zahrnovat:
 - Identifikaci uživatelů, konzumentů reportu.
 - Definování byznys otázek pro report.
 - Posouzení, zda dataset odpovídá otázkám.
 - Určení interaktivity reportu.
 - Definování přístupů a distribuce reportu.
 - Náčrt formy a vzhledu reportu.

- Pro reporty je dobré **dodržovat řadu doporučení** jak obsahových, tak formálních, zejména:
 - Minimalizace nepodstatných, resp. méně podstatných detailů.
 - Užití jednoduchých a jasných titulků stránek reportů a vizuálů.
 - Umístění vizuálů a jejich skupin na ploše má usnadnit efektivní a logickou navigaci v reportu.
 - Použití méně výrazných a přirozených barev pro vizuály.
 - Použití rozdílných barev v grafech pro rozlišení různých prvků a hodnot.
 - Umístění vizuálů vertikálně a horizontálně na ploše tak, aby mezi nimi byl adekvátní prostor.
 - Pozadí plochy má být přirozené kontrastní k vizuálům.
 - U sloupcových grafů se doporučuje seřadit sloupce podle vybrané míry.
 - Minimalizace velkých prázdných prostor na ploše.
 - Je účelné doplňovat vizuály o cílové nebo doplňující hodnoty (maxima, průměry apod.).
 - Všechny reporty by měly mít stejnou nebo obdobnou formu.
- Při užití grafů v reportech je podstatné určit, který typ grafu je nejvhodnější pro prezentaci dat, **je třeba si uvědomit, s jakým cílem, resp. myšlenkou, chceme data prezentovat**. Tátáž data můžeme prezentovat z různých úhlů pohledu a zdůraznit tím jiný aspekt výstupu.
- Je důležité předem **stanovit, pro koho prezentaci vytváříme**, z jakého pohledu na ni bude uživatel nahlížet a jakou informaci potřebuje z dat co nejdříve získat.

9. Příprava datových zdrojů



Účelem kapitoly je specifikovat některé další aspekty spojené s přípravou datových zdrojů, a to v širším a rovněž obecnějším kontextu řešení. Kapitola zahrnuje tyto aspekty, spojené s datovými zdroji:

- základní možnosti konektivity vzhledem k datovým zdrojům,
- Data Profiling a řešení kvality dat,
- řešení změn v dimenzích, SCD,
- principy tabulárního modelu, využívaného pro vytváření datasetů PBI,
- možnosti nastavení PBI.

9.1 Konektivita

Konektivita k rozdílným datovým zdrojům je jedním z významných efektů PBI. Podporuje **obvyklé standardy**:

- Open Data Protocol (OData),
- Open Database Connectivity (ODBC),
- Object Linking and Embedding Database (OLE DB),
- lze vytvářet i vlastní konektory uživatelů.

Datových zdrojů, z nichž lze čerpat data do datasetu PBI, je značné množství. Lze ale doporučit, pokud to podmínky dovolují, získávat data z datového skladu, protože tato data už prošla kontrolou kvality, respektují datovou integritu a jsou strukturována pro potřeby analytiky.

9.2 Změny v dimenzích – SCD (Slowly Changing Dimensions)

Jedním ze závažných problémů, které se musí v BI i SSBI, a tedy i v projektech Power BI řešit, jsou **změny v dimenzionálních tabulkách**. Tyto změny se musejí **rozlišovat podle významu** zejména z pohledu analýz podle dimenze času. Některé změny v dimenzích mají **pro hodnocení vývoje a srovnatelnost sledovaných ukazatelů zásadní význam**, některé minimální nebo žádný. Principy řešení změn v dimenzích se označují termínem **SCD – Slowly Changing Dimensions**.



SCD – Slowly Changing Dimensions představuje **změny ve struktuře a prvcích dimenzí, tedy číselníků v čase**, přičemž při řešení úloh BI a SSBI existuje reálná potřeba zachovat konzistenci dat z časového hlediska. V tomto kontextu existují následující **možné změny v dimenzích**, tj. jejich prvků, příp. vyšších úrovní v rámci hierarchie dimenze:

- **přidání** nových prvků (např. nová zboží, zákazníci apod.),
- **zrušení** prvků (např. zrušení zboží),
- **změny hodnot atributů** (názvy produktů, jména pracovníků, adresy zákazníků nebo dodavatelů apod.),
- **změny ve struktuře** – zařazení (zboží, pracovníků) do jiných skupin v rámci hierarchické struktury prvků dimenze.

Při řešení problémů, spojených se změnami obsahu dimenzí, je třeba **počítat se složitějším návrhem analytické databáze**, s většími nároky na paměťový prostor dimenzí i s prodloužením běhu datových transformací.

Historii v datech lze uchovávat dvojím způsobem, a to **v rámci tabulky faktů**, nebo tzv. **historickými tabulkami dimenzí**. Řešení historie prostřednictvím tabulky faktů je založeno na principu, kdy se měnící se atribut přesune z tabulky dimenze (číselníku) do tabulky faktů. To je výhodné v těch případech, když je takový atribut součástí dalších kalkulovaných ukazatelů, jako je např. prodejní plocha, cena apod.

V dalším textu je vysvětleno řešení změn přímo v dimenzionálních tabulkách. To je založeno **na různých přístupech**, označovaných v dimenzionálním modelování jako **typ (Type)** a **číslo typu**, např. *Typ 1*. V každém případě **řešení změn** musí být proaktivní, a tedy již **na úrovni hrubého dimenzionálního modelu** musí být pro každý atribut dimenzionální tabulky uveden typ řešení jeho změn.

Typ 0 (SCD 0) určuje, že změny jsou **ignorovány** a do datového skladu se nepromítají. Skutečné promítání změn dimenzí se rozlišuje několika dalšími typy řešení.

9.2.1.1 Typ 1 (SCD 1) – Přepis hodnoty atributu

Řešení *Typu 1* je nejjednodušší, neboť **pouze nahrazuje původní hodnotu atributu hodnotou novou**, změněnou ve stejné řádce dimenzionální tabulky. Předpokládejme, že máme dimenzionální tabulku *Zboží*, kde každý záznam má mimo jiné atributy *Primární klíč* (umělý klíč), *Číslo zboží* (operační klíč z transakčního systému), *Zodpovědný pracovník* (pracovník, který kompletně zodpovídá za prodej daného produktu), viz Tabulka 9-1.

Tabulka 9-1: Původní záznamy dimenzionální tabulky zboží

Primární klíč	Název zboží	Číslo zboží	Číslo skupiny zboží	Název skupiny	Zodpovědný pracovník
111	Acer Travelmate 292	NT-0015	01	Notebook	J. Malá
112	IBM ThinkPad R40	NT-0023	01	Notebook	P. Novák

Pokud dojde např. ke změně zodpovědného pracovníka a místo J. Malé bude produkt NT-0015 zajišťovat F. Klas, pak se tabulka změní pouze přepisem původního jména novým, viz Tabulka 9-2.

Tabulka 9-2: Dimenzionální tabulky zboží s aktualizovaným jménem zodpovědného pracovníka

Primární klíč	Název zboží	Číslo zboží	Číslo skupiny zboží	Název skupiny	Zodpovědný pracovník
111	Acer Travelmate 292	NT-0015	01	Notebook	F. Klas
112	IBM ThinkPad R40	NT-0023	01	Notebook	P. Novák

Pro změny v dimenzi **typu 1** je **charakteristické**:

- **nebyl změněn žádný klíč** (primární, ani operační), a to ani v dimenzionální, ani ve faktové tabulce,
- jde o **nejjednodušší** způsob řešení změn dimenzí,
- **faktová tabulka se nijak nemění**,
- na základě uplatnění typu 1 se však **ztrácí historie sledovaných ukazatelů** podle změn v atributech (nevíme, za jaké prodeje byla ještě zodpovědná J. Malá),
- typ 1 je **vhodný** především tam, kdy **jde o opravu chybných hodnot atributů**, nebo když původní hodnoty nemají podle požadavků uživatelů na výstupy nebo aplikace žádný význam,
- všechny **předchozí výstupy**, přehledy, agregace **se musejí znovu vytvořit** (resp. aktualizovat filtry v reportech, pokud byly pro filtrování využity hodnoty z dimenze, které mají nyní novou hodnotu) s novou hodnotou atributu (tam, kde je daný atribut ve výstupech uplatněn).

9.2.1.2 Typ 2 (SCD 2) – Přidání nové řádky do dimenzionální tabulky

Typ 2 je **nejčastěji používaný způsob** řešení změn dimenzí. Znamená **přidání nového řádku do dimenzionální tabulky**, který obsahuje změněnou hodnotu atributu, případně atributů. Tím se zachovávají historické údaje, ale narůstá objem dimenze. V tomto případě **záznam dimenze obvykle obsahuje**:

- **nový primární klíč** pro nový, resp. doplněný záznam,
- **operační klíč** (původní z transakčního systému) se se změnami nemění, tj. v novém záznamu obsahuje stejnou hodnotu,
- **datum počátku** platnosti záznamu (DATUM_OD),
- **datum konce** platnosti záznamu (DATUM_DO) – u aktuálních záznamů se vyplňuje dostatečně vysokou konstantou (např. 1. 1. 2200).

Způsob řešení předchozí změny u typu 1, dokumentuje pro typ 2 Tabulka 9-3.

Tabulka 9-3: Změna dimenzionální tabulky zboží pro typ 2

Primární klíč	Název zboží	Číslo zboží	Číslo skupiny zboží	Název skupiny	Zodpovědný pracovník	Datum od	Datum do
111	Acer Travelmate 292	NT-0015	01	Notebook	J. Malá	1.1.2011	5.5.2011
321	Acer Travelmate 292	NT-0015	01	Notebook	F. Klas	6.5.2011	1.1.2200
112	IBM ThinkPad R40	NT-0023	01	Notebook	P. Novák	1.1.2011	1.1.2200

Pro změny v dimenzi **typu 2** je, kromě již uvedeného způsobu realizace změny, **charakteristické**:

- **data ve faktové tabulce nebudou** v důsledku změny dimenzionální tabulky **změněny**, tj. záznamy ve faktové tabulce se budou po datu změny (5. 5. 2011) odkazovat na záznam v tabulce dimenze zboží již s novým primárním klíčem, tedy nikoli 111, ale na záznam 321,
- **výběry dat** podle **zodpovědného pracovníka budou probíhat správně** podle různých hodnot daného atributu, tedy dostaneme např. správné objemy prodeje pro J. Malou i pro F. Klase,
- tím, že zůstává stejný operační klíč (NT-0015), dostáváme **správné hodnoty i pro celkové objemy prodeje** tohoto produktu bez ohledu na změnu odpovědného pracovníka,
- **hodnoty atributů Datum_od, Datum_do** slouží **pro pokročilé analýzy** a přesné sledování historie prodeje podle změn atributů, pro základní přehledy (viz předchozí body) se ale nemusejí využívat,
- výhodou typu 2 je zejména **možnost sledování vývoje ukazatelů podle změn atributů** a současně možnost sledovat tolik změn a u tolika atributů v dimenzionální tabulce, kolik je třeba,
- rovněž platí, že **nadřazené (master) tabulky** (ve SNOWFLAKE schématu), **využívající změny typu 2, musí mít i podřazené tabulky (detail) se změnami typu 2**, naopak to ale není nutné,
- vzhledem k tomu, že je přesně dáno, kdy došlo ke změně, pak **není třeba zpětně vytvářet původní výstupy a agregace**.

9.2.1.3 Typ 3 (SCD 3) – Přidání nového sloupce do dimenzionální tabulky

Typ 3 představuje **přidání nového atributu do dimenze** (nového sloupce) a **přesunutí staré hodnoty do tohoto nového atributu**. Do původního (aktuálního) atributu je pak možné vložit novou hodnotu. Příklad řešení změny typu 3 ukazuje Tabulka 9-4.

Tabulka 9-4: Změna dimenzionální tabulky pro typ 3

Primární klíč	Název zboží	Číslo zboží	Číslo skupiny zboží	Název skupiny	Zodpovědný pracovník	Původní zodpovědný pracovník
111	Acer Travelmate 292	NT-0015	01	Notebook	F. Klas	J. Malá
112	IBM ThinkPad R40	NT-0023	01	Notebook	P. Novák	P. Novák

Pro změny v dimenzi **typu 3** je charakteristické:

- lze efektivně **kombinovat výběry** podle původní i nové hodnoty atributu, např. pro srovnání výkonnosti obou pracovníků apod.,
- na rozdíl od typu 2 **může být předchozí i současná hodnota atributu využívána jako platná** současně,
- typ 3 se však **prakticky nepoužívá často**,
- při větším očekávaném počtu změn se může **využít i větší počet nových atributů**, ale to je možné pouze v případech, kdy počet změn lze u dimenze přesněji odhadovat.

9.2.1.4 Typ 4 – Vytvoření nové historické dimenze

Typ 4 je založen na vytvoření nové **historické dimenze, do níž se přesouvají historické údaje**. Tento přístup se používá obvykle tehdy, kdy v dimenzích dochází k rychlým, respektive častým změnám. Jestliže dojde ke změně, je daná **existující dimenze nejdříve zkopírována do historické dimenze** a následně jsou **v původní tabulce dimenze provedeny potřebné změny**.

Při dimenzionálním modelování ve vztahu k SCD je třeba brát v úvahu, že **různé atributy**, které jsou součástí záznamu dimenzionální tabulky, **budou mít různý typ SCD**. Např. datum narození zákazníka se nebude měnit (kromě chyby v datech), bude mít tedy typ 1, zatímco adresa zákazníka se měnit bude a bude účelné ji zpětně sledovat, tedy typ 2, nebo vyšší. V každém případě je nutné typy SCD u jednotlivých atributů definovat a realizovat v transformačních procedurách (ETL), což samozřejmě zvyšuje jejich složitost.

Nejčastěji užívanými typy SCD jsou 1 a 2. Při návrhu analytické databáze se v této souvislosti můžeme rozhodovat **mezi dvěma variantami**:

- **jednu dimenzi lze rozdělit na dvě dimenzionální tabulky**, přičemž v jedné budou atributy typu 1 a v druhé typu 2. Toto řešení je relativně jednoduché a často populární, ale nese s sebou dvě rizika – zvyšuje se počet dimenzí, v dotazech je třeba řešit navíc vazby mezi těmito dvěma dimenzemi, což je tvoří složitějšími a časově náročnějšími,
- **zachovat jednu společnou dimenzi s atributy typu 1 i 2**, což řeší problémy předchozí varianty, ale naopak vede ke složitějším transformacím dat.

Provádění změn v dimenzích **při přípravě dat** je poměrně **časově náročné**. Používají se proto kontrolní součty v každém databázovém záznamu (jako provozní položka). Teprve v případě nesouhlasu kontrolních součtů se kontrolují jednotlivé položky záznamů a zjišťují se změny, které se pak u jednotlivých atributů řeší podle stanoveného typu změny.

9.3 Tabulární model

Tabulární model lze definovat jako sadu metadat zahrnující tabulky, vztahy, metriky, KPI, skupiny kalkulací (*calculation groups*), hierarchie, překlady, bezpečnostní role a mnoho dalších prvků, které tvoří tzv. sémantický model, který zajišťuje systém navigace v klientských nástrojích, jako je realizace reportů v Power BI nebo v Excelu (Pátková 2023).

Model odpovídá sémantickému modelu ve službě Power BI nebo databázi ve službě Analysis Services (*Azure Analysis Services* anebo *SQL Server Analysis Service*). Sémantické modely a databáze jsou

modely, naplněné daty. Definice modelu bez jakýchkoli dat je možné označovat jako „tabulární model“ (Russo 2022).

Sémantické modely nebo databáze, založené na tabulárních modelech, jsou speciální **analytické databáze**, které pro ukládání dat mohou používat dva druhy řešení:

- **In-memory**, což znamená, že vykonávají dotazy na data v paměti aplikace nad celou databází,
- **DirectQuery**, což znamená, že data nejsou kopírována do aplikace, přes kterou k nim přistupujeme (například *Power BI Desktop*, *Microsoft SQL Server Analysis Services*, *Azure Analysis Services*), ale dotazy jsou předávány přímo do zdrojových databází, kde probíhá zpracování a do aplikace je předáván až výsledek dotazu. V tomto případě je tedy tabulární model založen v příslušném nástroji, ale data nejsou jeho součástí.
- **Direct Lake**, obdobně jako Direct Query, ale data leží v datovém jezeře OneLake Microsoft Fabric.

Tabulární modely **kombinují technologii MOLAP a relačních databází** (Rabeler et al. 2018).

Zatímco **in-memory** modely jsou **výchozí**, **DirectQuery režim se používá pro** příliš velké objemy dat na to, aby se vešly do paměti, nebo pokud volatilita dat vylučuje rozumnou strategii zpracování. **DirectQuery** dosahuje shody s in-memory modely prostřednictvím **podpory široké škály datových zdrojů**, schopností zpracovávat kalkulované tabulky a sloupce v DirectQuery modelu, zabezpečení na úrovni řádků prostřednictvím výrazů DAX, které se dostanou do databáze typu back-end, a optimalizace dotazů, jejichž výsledkem je rychlejší průchodnost (propustnost) (Rabeler et al. 2018).

Tabulární modely jsou definovány strukturami a protokoly, dokumentovanými Microsoftem (Russo 2022):

- MS-SSAS-T: SQL Server Analysis Services Tabular Protocol.
- Tabular Object Model (TOM).
- Tabular Model Scripting Language (TMSL).
- XML for Analysis (XMLA) Reference.

Sama definice tabulárního modelu má díky protokolu TOM strukturu JSON souboru, známá jako model.bim a model lze vytvářet prostřednictvím:

- Power BI.
- Microsoft Visual Studio (Analysis Services Project).
- Tabular Editor.
- Případně přímo v SaaS Microsoft Fabric.

Tabulární modely mohou být **nasazeny do služby Azure Analysis Services**, nebo na instance služby **SQL Server Analysis Services (SSAS)**, nakonfigurované na režim Tabular server nebo do služby Microsoft Fabric. Nasazené tabulární modely je možné spravovat i přes SQL Server Management Studio (Rabeler et al. 2018).

Nástroje tabulárního modelu původně podporovaly mezi tabulkami **pouze vazby 1:1 nebo 1:N s jedno-směrnou i obousměrnou filtrací**. Od roku 2020 jsou implementovány i vazby M:N, i když je doporučeno tento typ vazeb používat pouze ve specifických scénářích užití, stejně jako obousměrné vazby 1:N.

9.3.1 Engine

Tato část pojednává o **výpočetních engine**, které jsou **součástí tabulárního modelu**. Dotazy na tabulární databáze jsou v analytických databázích (SSAS, AAS, PowerBI) zpracovávány **ve dvou vrstvách výpočetních engine, formula engine a storage engine**.

Existují **2 druhy storage engine**:

- VertiPaq.
- DirectQuery.

9.3.2 Vertipaq engine

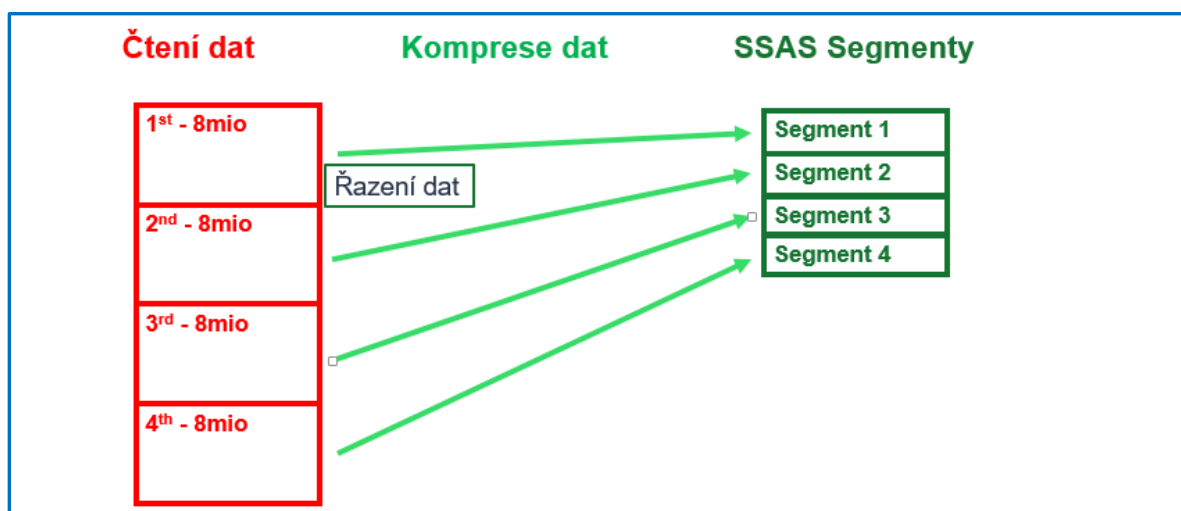
Datový model oficiálně využívá **xVelocity in-memory Analysis Services**, avšak z důvodu, že in-memory engine byl dlouhou dobu pojmenován jako Vertipaq, je známější pod tímto názvem. Navíc interně

je stále zobrazován název Vertipaq a dotazovací engine vykonává Vertipaq dotazy (Haman, 2018). Vertipaq je implementován jak SSAS, AAS (Azure Analysis Services), tak i v Power BI desktop.

Vertipaq engine při načítání dat nejprve **data přečte** a následně je **transformuje do sloupcové struktury** VertiPaq, tedy tabulku si rozdělí na sloupce. **Pro jednotlivé sloupce** tabulky si vytvoří **slovník jedinečných hodnot**, které dále nahrazuje **indexy**. Sloupce tak obsahují pouze číselné datové typy, konkrétně **integer** (Haman, 2018).

V dalším kroku si VertiPaq engine **seřadí duplicitní výskyty indexů** a zpracuje je podle algoritmu RLE (Run Length Encoding). **RLE** je algoritmus **pro komprimaci dat**. Aby bylo dosaženo lepšího kompresního poměru, RLE tedy data nejprve seřadí tak, aby řádky, následující ve sloupci, obsahovaly stejnou hodnotu. Data jsou čtena a **zpracovávána po dávkách (batch)**. Jeden batch může obsahovat až 8 milionů řádků.

Data jsou tedy rozdělena do dávek, zkomprimována a uložena do segmentů, jak je znázorněno na obrázku (Obrázek 9-1). Výsledné řešení je závislé na poměru počtu hodnot a počtu unikátních hodnot, ale především je paměťově úspornější a rychleji prohledatelné, neboť je uložen pouze počet opakování stejných hodnot (Joyful Craftsmen, 2018). Znalost tohoto principu je důležitá při návrhu tabulárních modelů. Stále větší počet sloupců v tabulkách s velmi rozdílnými dlouhými textovými řetězci má vliv na budoucí problémy s velikostí a výkonností modelů, protože nelze data dost dobře komprimovat.



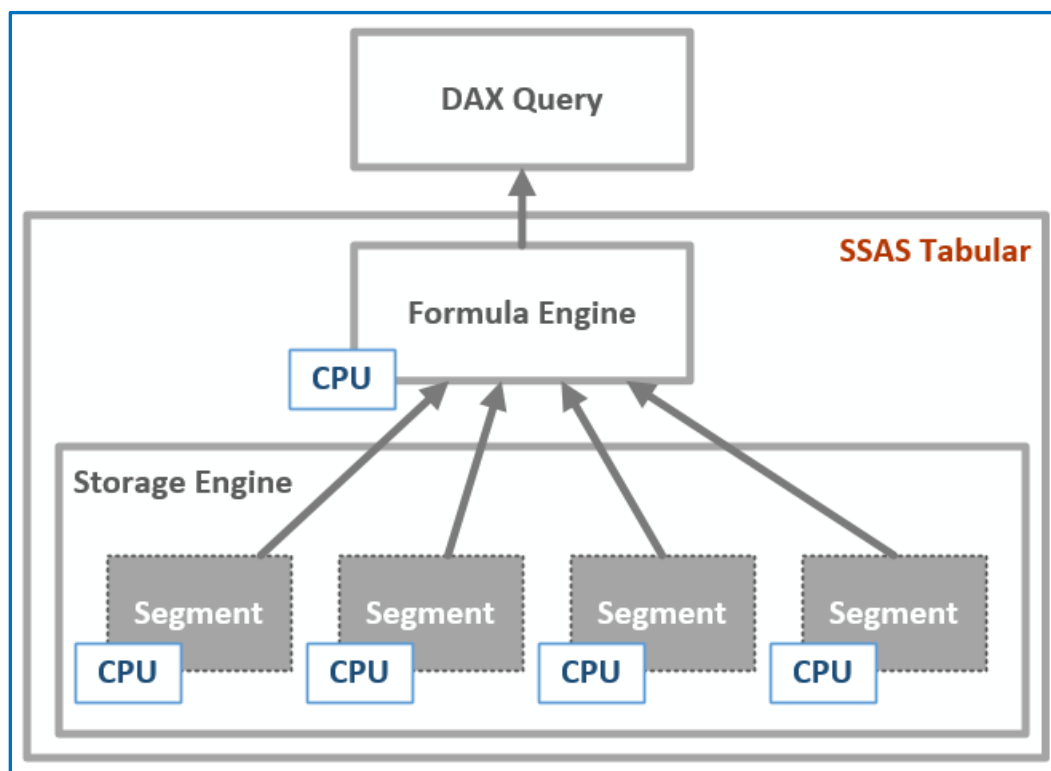
Obrázek 9-1: Komprese dat ve VertiPaq engine (Joyful Craftsmen, 2018)

Komplexnější kalkulace vykonává formula engine po získání výsledků z VertiPaq engine (NisaBI 2017).

Na obrázku (Obrázek 9-2) je nastíněno schéma zpracování dotazů, kde lze vidět, že Formula engine zpracovává dotazy od jednoho uživatele pouze na jednom vlákne, zatímco VertiPaq engine používá více vláken paralelně.

Pro in-memory engine je nutné, aby RAM na serveru, kde je datový model umístěn, byla dvakrát až třikrát větší, než je velikost datové kostky.

Při návrhu struktury tabulárního modelu je doporučeno, aby výsledný model měl podobu „star“. Vazby mezi faktovými a dimenzionálními tabulkami je účelné realizovat pomocí atributů typu integer. Takové modely jsou výkonnější (Russo 2021).



Obrázek 9-2 Schéma zpracování dotazů v SSAS Tabular (Joyful Craftsmen, 2018)

9.3.3 DirectQuery engine

DirectQuery mód se **dotazuje na data přímo na zdroj**, tedy posílá SQL dotazy do zdrojové relační databáze, které SQL podporují. Je vhodnější použít DirectQuery pro velké datové objemy, jelikož ukládá pouze metadata, a navíc poskytuje nejaktuálnější data (Duncan 2019).

9.3.4 Perspektivy

Perspektivy v tabulárních modelech **umožňují definovat podmnožiny modelu**, které budou zobrazené **specifikovaným uživatelům či skupinám** uživatelů, a zobrazují různé pohledy na konkrétní model tabulární databáze. Primární záměr vytváření perspektiv není zabezpečení tabulárního modelu, nýbrž **usnadněním přístupu** k datům pro uživatele.

Veškeré **zabezpečení pro konkrétní perspektivu je zděděno** ze zabezpečení celého modelu. Není tedy možné poskytnout uživateli pomocí perspektiv přístup k objektům modelu, ke kterým přístup nemá. Perspektivy je možné vytvářet v návrhář modelu pomocí dialogového okna *Perspectives*. Následně, například v Excelu, připojeném na tabulární model v AAS (*Azure Analysis Services*) nebo SSAS, si uživatel může vybrat pouze ty perspektivy, které mu jsou nabízeny.

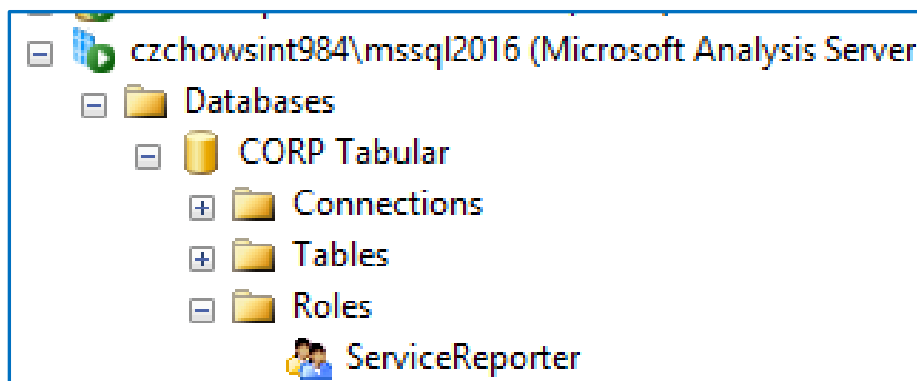
9.3.5 Security handling v tabulárním modelu

Pro zabezpečení tabulárního modelu se používá **definice rolí uživatelů v databázi**. Zabezpečení modelu pomocí rolí lze řešit **až na úroveň řádků (tzv. RLS – Row Level Security)**, které již vyžaduje použití jazyku DAX.

9.3.6 Databázové Role

Jak lze vidět na obrázku (Obrázek 9-3), zabezpečení analytické databáze je **obdobné jako zabezpečení relační databáze**. V databázi nebo přímo v tabulárním modelu jsou vytvářeny **role, pod které jsou přidáváni uživatelé**. Těmto rolím jsou pak **udělována práva**, a to buď plná kontrola nad databází

(Administrator), práva na procesování databáze (Process database), práva pouze na čtení dat z databáze (Read) či práva na čtení a procesování databáze (Read and Process).



Obrázek 9-3 Pohled na role v databázi

Vytvořené role lze zobrazovat a vytvářet filtry jak v aplikaci SSDT, tak v aplikaci SSMS. Filtry lze vytvářet pomocí výrazů v jazyku DAX. K této úrovni zabezpečení lze také využít možnosti proměnných, kdy lze nastavit podmínky za využití těchto proměnných, jako je například *Username()*, která vrací uživatelské jméno přistupujícího uživatele ze služby *Active Directory*.

9.3.7 ACL tabulky

Access Control List (ACL), v překladu **seznam řízení přístupu**, je seznam, obsahující informace o oprávnění a přístupech k objektu, k němuž se váže. **Jednotlivé záznamy** v ACL seznamu lze označit jako **Access Control Entries (ACE)**, tj. položky řízení přístupu, kdy jednotlivé **položky označují jednotlivé uživatele či skupiny uživatelů a jejich oprávnění**, která se mohou týkat nejen přístupu, ale také práce se soubory či daty. Atributy, které se mohou v seznamu ACL objevit, se odvozují od objektu, ke kterému je ACL vázán.

Mohou být dva typy ACL pro zabezpečovaný objekt:

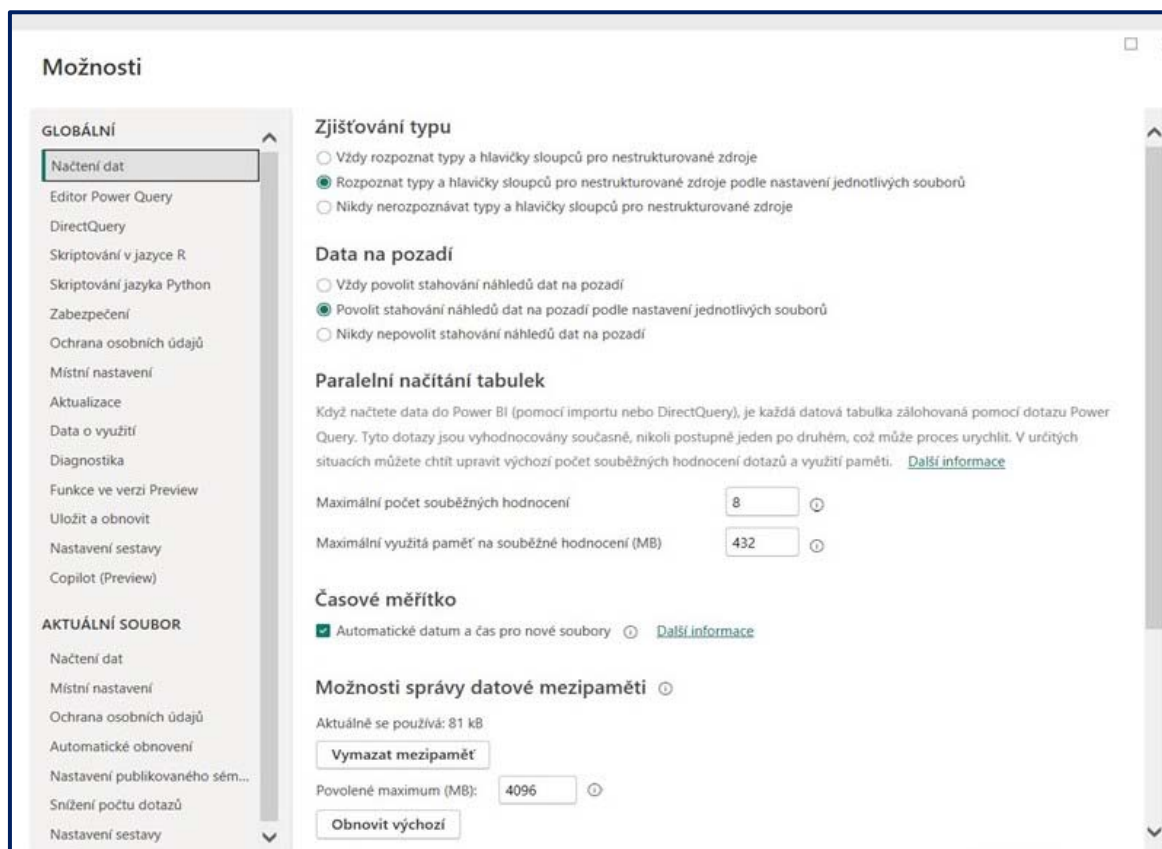
- Discretionary Access Control List (DACL).
- System Access Control List (SACL).

DACL, v překladu volitelný seznam řízení přístupu, pracuje se zabezpečením tak, že jakmile se proces pokusí o přístup k objektu, **systém nejprve zkontroluje, zda tento objekt má DACL**. Pokud objekt DACL nemá, přístup k objektu je tomuto procesu udělen v plné míře. Pokud však objekt má DACL, systém v sekvenci prochází veškeré záznamy, tedy ACE, v tomto seznamu, dokud nenajde alespoň jeden záznam ACE, nebo pokud neprojde celý seznam. Pokud je DACL objektu prázdné, tedy neobsahuje žádné záznamy, systém odepře přístup veškerým procesům. Pokud systém nalezne v seznamu hledané ACE, zachová se následně podle obsahu tohoto záznamu v seznamu.

SACL, v překladu seznam řízení přístupu k systému, slouží k zaznamenávání pokusů o přístup k objektu, ke kterému je SACL vázáno. Pokusy o přístup jsou zaznamenávány do **Security event logu**, tedy do protokolu událostí zabezpečení, kam je možné zaznamenat veškeré přístupy.

9.4 Nastavení možností Power BI desktop

Pro Power BI desktop existuje řada možností a voleb, které lze nastavit. Rozlišují se možnosti globální pro celkové využití PBI a lokální pro aktuální soubor. Příklad okna pro nastavení možností dokumentuje Obrázek 9-4:



Obrázek 9-4: Nastavení možností pro Power BI desktop

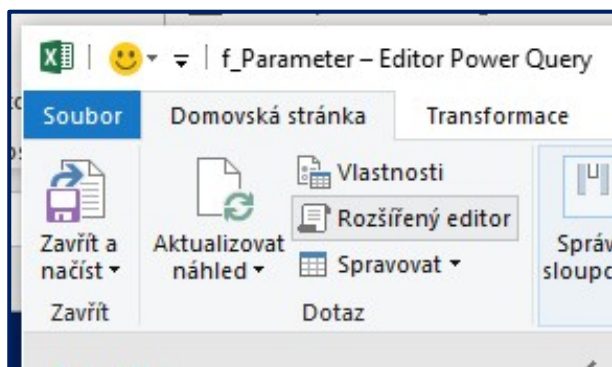
9.5 Pracovní závěry



Z kapitoly vyplývají následující **závěry**:

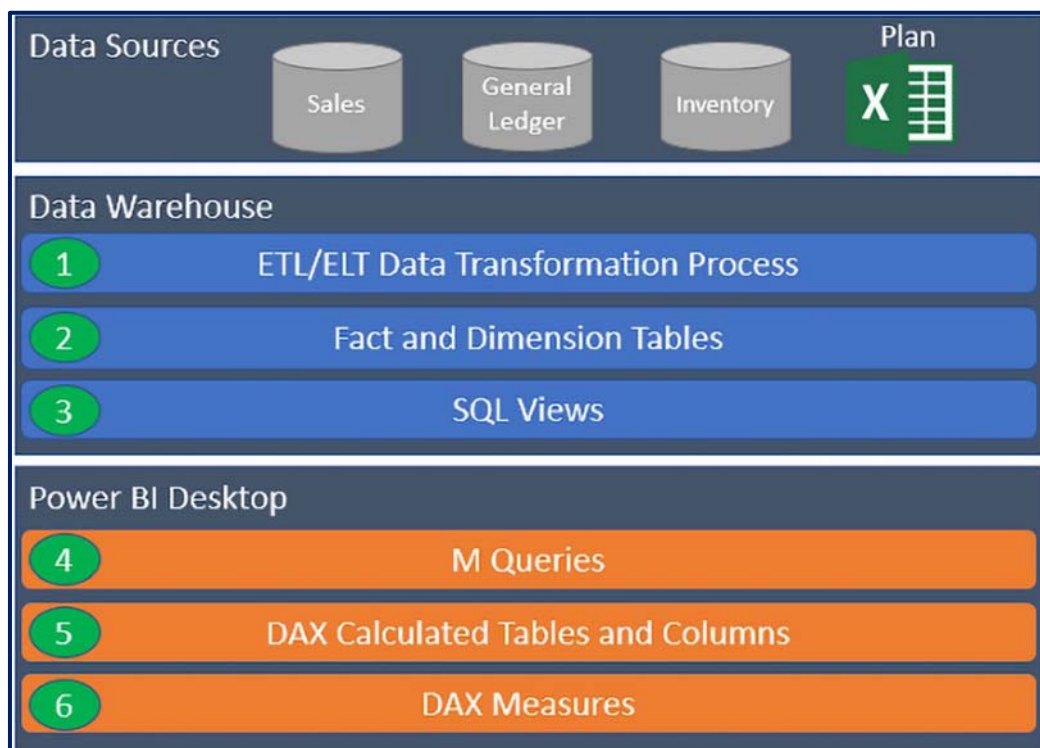
- **Konektivita** k datovým zdrojům, založená na běžných standardech, představuje významný efekt PBI.
- **SCD – Slowly Changing Dimensions** znamenají **změny ve struktuře a prvcích dimenzí v čase**, přičemž při řešení úloh SSBI na bázi PBI existuje reálná potřeba zachovat konzistenci dat z časového hlediska.
- **Různé přístupy k řešení SCD** se identifikují jako typy – (*Type*) a **číslo typu**. Ty zahrnují:
 - **Typ 0** určuje, že změny jsou **ignorovány**.
 - **Typ 1** je nejjednodušší, neboť **pouze nahrazuje původní hodnotu atributu hodnotou novou**, změněnou ve stejné řádce dimenzionální tabulky.
 - **Typ 2** je **nejčastěji používaným způsobem** řešení změn dimenzí, to je **přidání nového řádku do dimenzionální tabulky**, který obsahuje změněnou hodnotu atributu.
 - **Typ 3** představuje **přidání nového atributu do dimenze** (nového sloupce) a **přesunutí staré hodnoty do tohoto nového atributu**.
 - **Typ 4** je založen na vytvoření nové **historické dimenze**, do níž se **přesouvají historické údaje**.
- **Tabulární model** představuje sadu metadat, zahrnující tabulky, vztahy, metriky, KPI, skupiny kalkulací, hierarchie, překlady, bezpečnostní role a mnoho dalších prvků, které tvoří sémantický model.

10. Transformace dat, jazyk M



Účelem kapitoly je alespoň **rámcově charakterizovat jazyk M** pro přístupy k datovým zdrojům, jejich transformace a úpravy dat, vstupujících do sémantických modelů PBI. Jde pouze o stručný přehled, přičemž pro detailní informace k jazyku odkazujeme na portál na adrese <https://bit.ly/3vmFSyr>.

Power Query (M) představuje **dotazy, určené pro načítání a transformace dat** dimenzionálních a faktových tabulek, vstupujících do sémantického modelu. Dotazy mohou být napsány přímo v jazyce M, nebo jsou generovány na bázi funkcí Power Query Editoru. Jazyk M obsahuje cca 700 funkcí, přičemž velká část M dotazů a funkcí je společná pro fyzický transport dat i Direct Query Postavení jazyka M v architektuře, zaměřené na BI projekty, dokumentuje Obrázek 10-1.



Obrázek 10-1: Místo jazyka M v architektuře BI (Zdroj: Deckler, G., Powell, B., 2022)

Často je účelné transformovat logiku M dotazů **do pohledů** – „**view**“ – v rámci zdroje dat, případně datového skladu, z něhož vstupují data do sémantického modelu PBI. Tím se logika dotazů v rámci jazyka M podstatně zjednoduší.

10.1 Typy M dotazů

Existuje několik hlavních typů dotazů pro různé účely v přípravě dat s tímto označením:

- Parametry datového zdroje.
- „Staging“ dotazy.
- Dotazy pro fakta a dimenze.
- Dotazy k parametrům tabulek.
- Dotazy k zajištění bezpečnosti tabulek.
- Dotazy pro customizaci funkcí.

Další podkapitoly představují stručné vymezení jednotlivých typů dotazů.

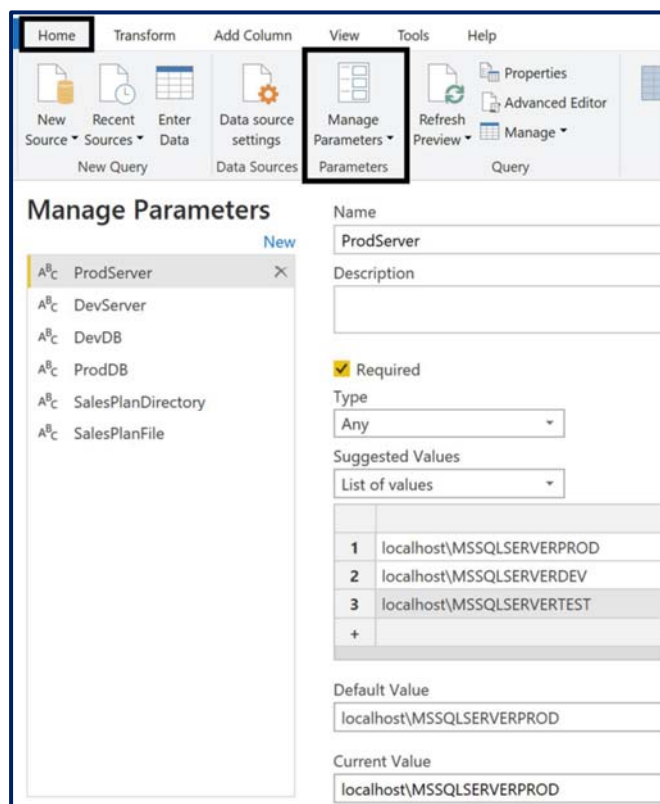
10.2 Parametry datového zdroje

„Parametry datového zdroje“ představují speciální M dotazy, které neznamenaají přístup k externímu datovému zdroji, ale pouze **vracejí vektor podstatných charakteristik datového zdroje**.

Účelem uvedených dotazů je centrálně definovat a spravovat společné a podstatné charakteristiky, jako jsou názvy serverů, názvy databází a další, a promítat je podle potřeby do všech ostatních dotazů. To zvyšuje efektivitu řízení datových setů včetně vazeb na datové zdroje.

Parametry lze vytvářet a modifikovat s pomocí funkcí v okně „Řízení parametrů“, resp. „*Manage Parameters*“, jak dokumentuje Obrázek 10-2.

Parametry jsou často využívány v šablonách pro PBI pro usnadnění uživatelům při customizaci reportů s využitím předdefinovaných dotazů a měr, nebo při definování filtrů v dotazech.

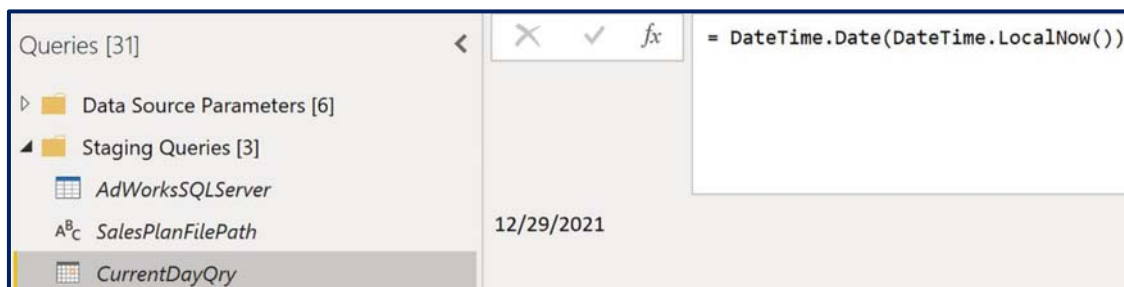


Obrázek 10-2: Správa parametrů v Power Query Editor (Zdroj: Deckler, G., Powell, B., 2022)

10.3 „Staging“ dotazy.

„Staging“ dotazy slouží **pro promítání charakteristik datových zdrojů do dotazů pro faktové a dimenzionální tabulky**. Např. dotaz pro SQL server slouží pro promítnutí parametrů datové báze a serveru do dotazů pro dimenzionální nebo faktové tabulky.

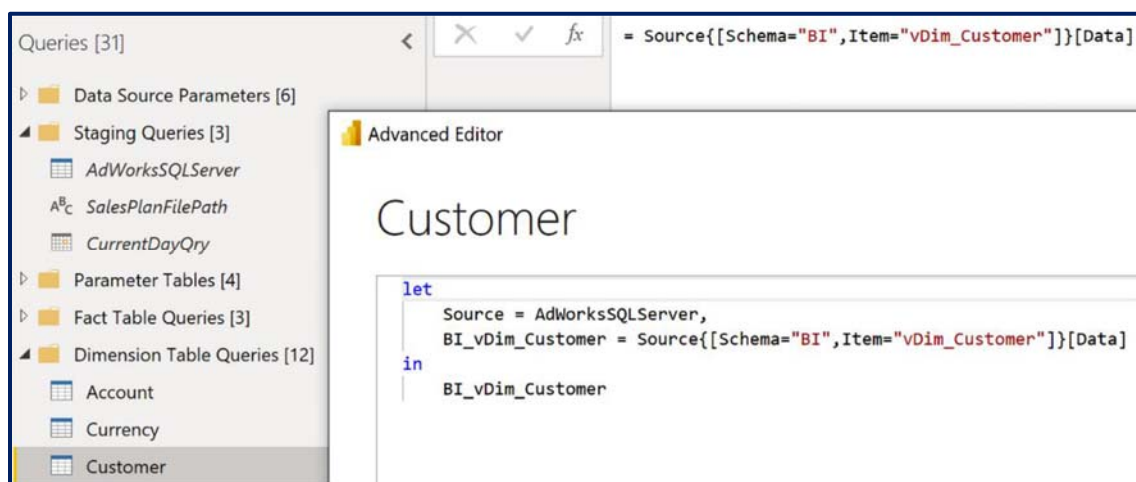
Příklad dotazu pro řešení aktuálního data dokumentuje Obrázek 10-3:



Obrázek 10-3: Dotaz pro běžné datum (Zdroj: Deckler, G., Powell, B., 2022)

10.4 Dotazy pro fakta a dimenze

Dotazy této kategorie slouží **pro připojení a import dat tabulek faktů a dimenzí**. M dotazy jsou vykonávány podle plánovaného obnovení (*refresh*) dat a výsledné dotazy jsou ukládány v komprimovaném sloupcovém formátu. S odpovídající přípravou může většina M dotazů jednoduše přistupovat k jednotlivým SQL pohledům s minimálními nebo žádnými transformacemi. Takové dotazy jsou pokládány za tzv. dotazy „*reference-only*“. Takový dotaz pro dimenzi „Zákazník“ dokumentuje Obrázek 10-4:



Obrázek 10-4: Dotaz pro dimenzi zákazník (Zdroj: Deckler, G., Powell, B., 2022)

Protože se neprovádějí žádné transformace, tak se tento jednoduchý **dotaz provede vždy při dalším obnovení sémantického modelu**. Preferuje se tak přenesení logiky transformace přímo do SQL pohledu na straně datového zdroje.

10.5 Dotazy k parametrům tabulek

Dotazy k parametrům tabulek slouží k efektivnějšímu řízení sémantických modelů. **Konsolidují vzájemně související vytvořené míry v seznamu položek (Field list)**, které se pak současně promítají do reportů. Uživatelé tak usnadňují vytváření reportů.

10.6 Dotazy k zajištění bezpečnosti tabulek.

Dotazy k zajištění bezpečnosti tabulek slouží pro uplatňování **bezpečnosti na úrovni řádků (RLS, row-level security)** v tabulkách sémantického modelu. Tyto dotazy současně umožňují využití principů **UPN (user principal names)** pro řízení přístupů uživatelů k reportům.

10.7 Dotazy pro customizaci funkcí.

Dotazy poskytují možnosti **vytváření znovupoužitelných kódů**, které mohou být využívány pro opakované datové transformace, jako např. *parsing* nebo různé kalkulace.

10.8 Pracovní závěry



Z kapitoly vyplývají následující **závěry**:

- **M dotazy** jsou určené pro načítání a transformace dat dimenzionálních a faktových tabulek, vstupujících do sémantického modelu.
- Je účelné **transformovat logiku M dotazů do pohledů** – „view“ – v rámci zdroje dat, nebo datového skladu
- **V jazyce M** existuje několik **hlavních typů dotazů** pro různé účely v přípravě dat:
 - Parametry datového zdroje.
 - „Staging“ dotazy.
 - Dotazy pro fakta a dimenze.
 - Dotazy k parametrům tabulek.
 - Dotazy k zajištění bezpečnosti tabulek.
 - Dotazy pro customizaci funkcí.

11. Závěr

Text tvoří jeden z doplňujících textů k základnímu textu publikace [Podniková analytika] z řady textů „IT a anatomie firmy“, zaměřený na produkt Power BI. V tomto případě **bylo cílem prezentovat celkový pohled na uvedený produkt** a vytvořit předpoklady a podklady pro řešení aplikací na jejich základě.

Účelem tohoto textu bylo vymezení pouze hlavních principů Power BI. **Podrobné tutoriály, videa, příklady a doporučení** k ovládání a pro práci s Power BI Service jsou k dispozici přímo na stránkách <https://powerbi.microsoft.com>.

12. Zdroje

- ASPIN, A., 2016: *Pro Power BI Desktop*. Apress. Staffordshire. 2016. ISBN 978-1-4842-1804-4.
- BI4DYNAMICS, 2017. Sales – Top 30 customer table Report. (online). (29. 08. 2017). Dostupné z: <http://www.bi4dynamics.com/business-intelligence-for-microsoft-dynamics-nav/content/>
- BSC Designer, 2017. Sales Business Unit Scorecard. online. Strategy Maps and KPIs. (online). (29. 08. 2017). Dostupné z: <https://www.webbsc.com/s/sales-kpis>.
- CANVASJS, 2013. JavaScript Range Column & Range Bar Charts. (online). canvasjs.com. Dostupné z: <https://canvasjs.com/javascript-range-column-range-bar-chart/>.
- CIMLER, P., ZADRAŽILOVÁ, D. a kol., 2007: *Retail management*. Praha, Management Press, 2007. ISBN: 978-80-7261-167-6.
- COLLIE, R., SINGH, A., 2016: *Power Pivot and Power BI*, Holy Macro Books, 2016.
- CZSO, 2016. Tab. 02.05 Investice na ochranu životního prostředí (1989-2015). (online). czso.cz. Vydáváme. Česká republika od roku 1989 v číslech – 2016. <https://www.czso.cz/csu/czso/ceska-republika-od-roku-1989-v-cislech-w0i9dxmgghn#03>
- CZSO, 2017a. Česká republika: hlavní makroekonomické ukazatele. Vydáváme. Časové řady. czso.cz. (online). (03. 07. 2017). Dostupné z: https://www.czso.cz/csu/czso/hmu_cr
- CZSO, 2017b. Graf 3 Ceny bytů - ČR (index, 2010 = 100). Ceny bytů. Vydáváme. czso.cz. (online). (07. 07. 2017). Dostupné z: https://www.czso.cz/csu/czso/ceny_bytu
- CZSO, 2017c. Tab. 7 Stravování a pohostinství (CZ-NACE 56). Vydáváme. Obchod, pohostinství, ubytování – časové řady - Základní finanční ukazatele – čtvrtletní – Klasifikace NACE Rev. 2 (CZ-NACE) (online). www.czso.cz (13. 08. 2017). Dostupné z: https://www.czso.cz/csu/czso/1-malzfu_b
- CZSO, 2017d. Peněžní vydání domácností podle počtu vyživovaných dětí. Veřejná databáze. czso.cz. (online). (13. 08. 2017). Dostupné z: https://vdb.czso.cz/vdbvo2/faces/index.jsf?page=vystup-objekt&pvo=ZUR10&z=T&f=TABULKA&katalog=30847&c=v3~8__RP2011&&str=v389
- DECKLER, G., POWELL, B.: *Mastering Microsoft Power BI*. Packt, Birmingham, 2022. ISBN: 978-1-80181-148-4.
- Denver.edu, 2017. Line Graph, Bar Graph, Pie Chart and Scatter Plot. University of Denver. (online). (13. 08. 2017). Dostupné z: <http://www.du.edu/ifs/help/use-online/repeated/general/graph/linebar.html>
- DUNCAN, Owen. 2019. *What is Azure Analysis Services?* [online]. Microsoft Docs. [cit. 2019-04-30]. Dostupné z: <https://docs.microsoft.com/en-us/azure/analysis-services/analysis-services-overview>
- DUNCAN, O., HAMILTON, B., CAI, S., DUNN, J., KUMAR, S., MILENER, G., RABELER, C., 2018. *About SQL Server Analysis Services* [online]. Microsoft Docs. [cit. 2019-04-30]. Dostupné z: <https://docs.microsoft.com/en-us/sql/analysis-services/analysis-services?view=sql-server-2017>
- DUNCAN, O., GUYER, C., RABELER, C., 2018. *Tabular model roles* [online]. Microsoft Docs. [cit. 2019-04-14]. Dostupné z: <https://docs.microsoft.com/en-us/sql/analysis-services/tabular-models/tabular-model-roles-ssas-tabular?view=sql-server-2017>
- ECKERSON, W., 2006. *Deploying Dashboards and Scorecards* [online]. (29. 08. 2017). Dostupné z: http://www.businessobjects.com/pdf/products/performancemanagement/wp_tdw_deploying_dashboards_and_scorecards.pdf
- ECKERSON, W. 2010. *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*, 2. vydání. Wiley. 336 stran. ISBN: 978-0-470-58983-0.
- ECKERSON, W., 2006: *Performance Dashboards*. New Jersey, John Wiley & Sons 2006.
- ENGLISH, L. P., 2003: *Improving Data Warehouse and Business Information Quality: Methods for reducing costs and increasing profits*. New York, John Wiley & Sons 2003. ISBN 0-471-25383-9.
- FEW, S., 2006. *Information Dashboard Design: The Effective Visual Communication of Data*. Sebastopol, O'Reilly Media. ISBN 978-0-596-10016-2.

FEW, S., 2012. Show Me the Numbers: Designing Tables and Graphs to Enlighten. 2nd edition. Burlingame, AnalyticsPress. ISBN 978-0-970-60197-1.

FEW, S., 2013. Information Dashboard Design: Displaying Data for At-a-Glance Monitoring. 2nd edition. Burlingame, AnalyticsPress. ISBN 978-1-938-37700-6.

GALLAGHER, J., 2015. Antibiotic surge revealed by seasonal maps. *bbc.com. News.* (online). (03. 07. 2017). Dostupné z: <http://www.bbc.com/news/health-34790038>

GAPMINDER, 2017. Life expectancy, years. (online). *Gapminder.org.* (13. 08. 2017). Dostupné z: http://www.gapminder.org/tools/#_chart-type=bubbles

HAMAN, Martin. 2018. *Data Day 2018* [online]. Dostupné z: https://www.gopas.sk/getattachment/47219b58-3c18-427e-b304-c940604185a8/Martin_Haman_all.pdf.aspx,.php,.html,

HARINATH, S., PIHLGREN, R., LEE, D., G., SIRMON, J., BRUCKNER, R., M., 2012: *Microsoft SQL Server 2012. Analysis Services with MDX and DAX.* John Wiley and Sons, Inc., Indianapolis, 2012. ISBN 978-1-118-10110-0.

<https://sqldusty.com/2019/02/15/power-bi-architecture-diagram-v4-is-now-available/>

HURSMAN, A., 2010. Effective-dashboard-design-why-your-baby-is-ugly. (online). (13. 08. 2017). Dostupné z: <https://www.slideshare.net/hursman/effective-dashboard-design-why-your-baby-is-ugly>

IMHOFF, C., WHITE, C., 2011: *Self-Service Business Intelligence: Empowering Users to Generate Insights.* Renton, WA : The Data Warehousing InstituteTM, 2011.

INETSOFT.com, 2017. Information about Scorecards and Scorecard Examples. (online). (29. 08. 2017). Dostupné z: https://www.inetsoft.com/info/information_about_scorecards_and_scorecard_examples/

INMON, B., 2002: *Building the Data Warehouse.* Indianapolis, John Wiley and Sons 2002.

JOTHIGANESH, S., 2017. 3Dscatterplot. (online). (13. 08. 2017). Dostupné z: <http://www.jothiganesh.com/category/technical/>

JOYFUL CRAFTSMAN. 2018. *Hitchhiker's Guide to the SSAS Tabular* [online]. [cit. 2019-04-14]. Dostupné z: <https://www.sqlsaturday.com/779/Sessions/Details.aspx?sid=85530>

KIMBALL, R., ROSS, M., 2010: *Relentlessly Practical Tools for Data Warehousing and Business Intelligence.* Wiley Publishing, Inc. 2010. ISBN 978-0-470-56310-6.

KIMBALL, R., CASERTA, J., 2004: *The Data Warehouse ETL Toolkit.* Indianapolis, John Wiley and Sons 2004. ISBN 0-764-56757-8.

KIMBALL, R., ROSS, M., 2002: *The Data Warehouse Toolkit, The Complete Guide to Dimensional Modelling.* Boston, John Wiley 2002. ISBN 0-471-20024-7.

KOŽUŠNÍK, M., KRATOCHVÍL, T., VOZHDAEVA, E.: IT a anatomie firmy (Tvorba dashboardů). Pracovní dokument, dostupný na <https://mbi-af.cz>, 2024

KUTHAN, L. IT a anatomie firmy (Microsoft Fabric). Pracovní dokument, dostupný na <https://mbi-af.cz>, 2024.

ŠEDIVÁ, Z., MARYŠKA, M., POUR, J., STANOVSKÁ, I.: *Self Service Business Intelligence*, Praha. Grada, 2018. ISBN 978-80-271-0616-5.

MICROSOFT, 2013: Power Pivot: Výkonné analýzy a modelování dat v Excelu. MICROSOFT. Office - Office.com [online]. 2013 [cit. 2014-01-02]. Dostupné z: <http://office.microsoft.com/cs-cz/excel-help/power-pivot-vykonne-analyzy-a-modelovani-dat-v-excelu-HA102837110.aspx>

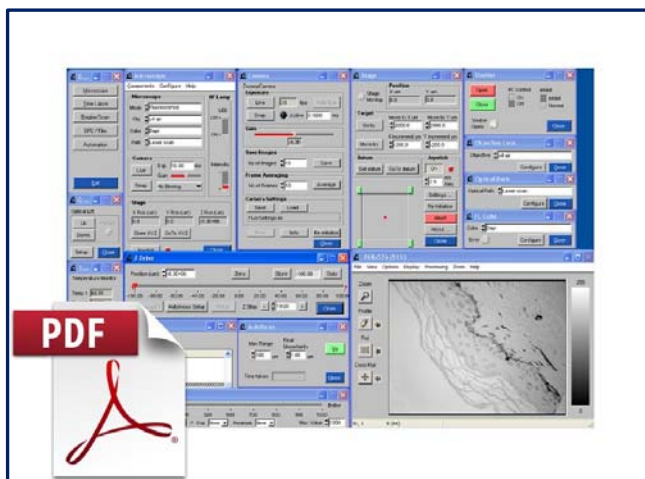
NOVOTNÝ, O., BUCHALCEVOVÁ, A., BRUCKNER, T., STANOVSKÁ, I., ŠEDIVÁ, Z., POUR, J.: IT a anatomie firmy (Podniková analytika). Professional Publishing, 2023. ISBN 978-80-88260-70-7. Dostupné na <https://mbi-af.cz>

OFFICE 1: Typy funkcí jazyka DAX. Office.com [online]. 2013 [cit. 2014-03-14]. Dostupné z: <http://office.microsoft.com/cs-cz/excel-help/typy-funkci-jazyka-dax-HA102836089.aspx>

OFFICE 2: Perspektivy v Power Pivotu. Office.com [online]. 2013 [cit. 2014-03-14]. Dostupné z: <http://office.microsoft.com/cs-cz/excel-help/perspektivy-v-power-pivotu-HA102837427.aspx>

- OFFICE 3: Filtrování a zvýrazňování v Power View. Office.com [online]. 2013 [cit. 2014-03-22]. Dostupné z: <http://office.microsoft.com/cs-cz/excel-help/filtrovani-a-zvyraznovani-v-power-view-HA102834776.aspx>
- KELLE ONEAL, BEYENETWORK, 2012, on-line: http://www.b-eye-network.com/blogs/oneal/archives/2012/02/what_is_the_dif.php
- PÁTKOVÁ, N.: Analýza, návrh a realizace datového katalogu pro velkou nadnárodní logistickou společnost. DP, VŠE, 2023.
- POTANČOK, M., POUR, J., CHRAMOSTOVÁ, V. Podniková analytika pro manažery, Oeconomia, Praha, 2021.
- PROVOST, F., FAWCETT, T., 2013: Data Science for Business. What You Need to Know About Data Mining and Data-Analytic Thinking. O'Reij Media. Sebastopol. 2013. ISBN: 978-1-449-36132-7.
- RABELER, C., GUYER, C., DUNCAN, O. 2018. *Tabular Models*. Microsoft Docs [online]. [cit. 2019-04-14]. Dostupné z: <https://docs.microsoft.com/en-us/sql/analysis-services/tabular-models/tabular-models-ssas?view=sql-server-2017>
- RABELER, C., GUYER, C., DUNCAN, O. 2018. *Perspectives in tabular models* [online]. Microsoft Docs. [cit. 2019-04-29]. Dostupné z: <https://docs.microsoft.com/en-us/sql/analysis-services/tabular-models/perspectives-ssas-tabular?view=sql-server-2017>
- RUSSO, M. 2022 *Development tools for Tabular models in 2021*. [cit. 2022-05-22]. Dostupné z: <https://www.sqlbi.com/articles/development-tools-for-tabular-models-in-2021/>
- RUSSO, M., FERRARI, A. *Header/Detail vs Star Schema models in Tabular and Power BI*. 2021 [cit. 2021-10-11]. Dostupné z: <https://www.sqlbi.com/articles/header-detail-vs-star-schema-models-in-tabular-and-power-bi/>
- RUSSO, M., FERRARI, A.: The Definitive Guide to DAX. Pearsons Education, Inc. 2020. ISBN 13-978-1-5093-0697-8.
- RUSSO, M., FERRARI, A.: DAX Patterns. SQLBI, Corp., 2020. ISBN 978-1-7353652-0-6.
- RUSSO, M., FERRARI, A., 2013: *Microsoft Excel 2013: Building Data Models with PowerPivot*. California, O'Reilly Media, Inc., 2013. ISBN: 978-0-7356-7634-3.
- RUSSO, M., FERRARI, A., 2011: *PowerPivot for Excel 2010. Give Your Data Meaning*. Redmond, Microsoft Press, 2011. ISBN: 978-0-7356-5058-0.
- SIEGEL, E., 2016: Predictive Analytics, The power to predict who will click, buy, lie, or die. John Wiley&Sons. New Jersey. 2016. ISBN: 978-1-119-14567-7.
- SEYFOR: Analytické a prezentační materiály firmy Seyfor, 2024.
- SLÁNSKÝ, D.: Data and Analytics for the 21st Century: Architecture and Governance, Professional Publishing, 2018. ISBN 978-80-88260-16-5.
- SPOFFORD, G., 2001: *MDX Solutions with Microsoft SQL Server Analysis Services*. New York, John Wiley, 2001. ISBN: 0-471-40046-7.
- SPÁL, M.: Dátovo riadená logistika, MBA práce, VŠE, 2021.
- STANOVSKÁ, I., POUR, J., MARYŠKA, M., ŠEDIVÁ, Z.: Self Service Business Intelligence, Praha. Grada, 2018. ISBN 978-80-271-0616-5.
- TANKERSLY, B., 2017. Using Dashboards For a Real-Time View to Productivity. Intuit QuickBooks. (online). (29. 08. 2017). Dostupné z: <https://www.firmofthefuture.com/content/using-dashboards-for-a-real-time-view-to-productivity/>
- TURLEY, P., BRUCKNER, B., SILVA, T., WITHEE, K., PAISLEY, G., 2012: *Microsoft SQL Server 2012. Reporting Services*. John Wiley and Sons, Inc., Indianapolis, 2012. ISBN 978-1-118-10111-7.
- WEXLER, S., SHAFFER, J., COTGREAVE, A., 2017: The Big Book of Dashboards. John Wiley&Sons. New Jersey. 2017. ISBN: 978-1-119-28271-6.

B. DAX, Data Analytics Expressions



[1] Základní principy a užití jazyka DAX

(Základní vymezení pravidel, výpočetní předpis, vytvoření kalkulovaného sloupce a vytvoření kalkulované míry)

[2] Konstrukce jazyka DAX

[3] Kontext vyhodnocování výpočetních předpisů

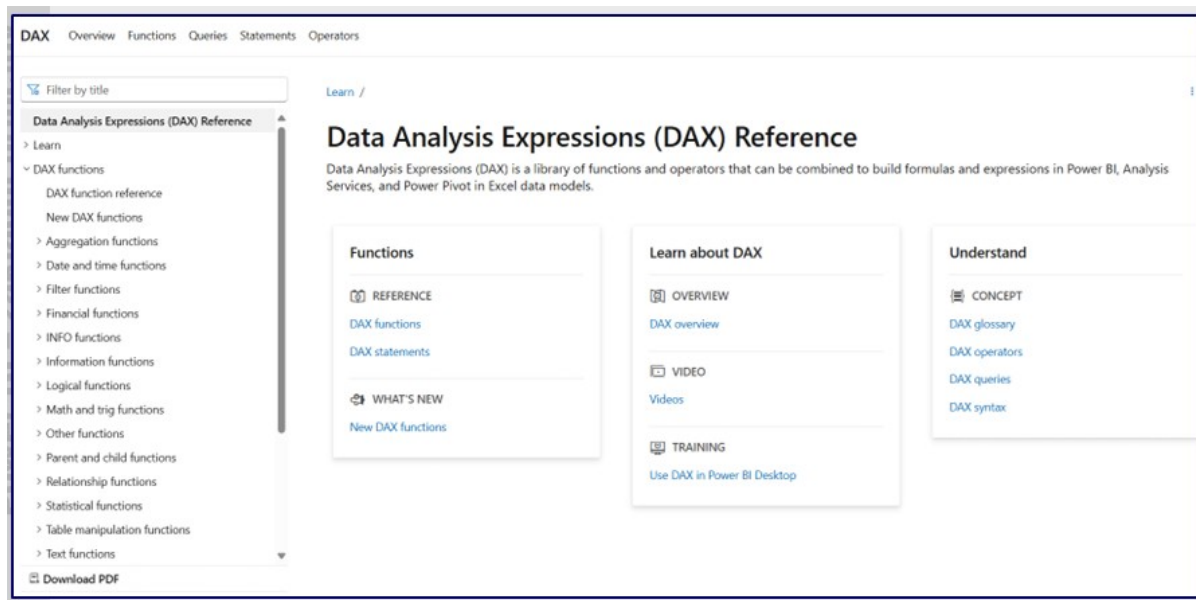
[4] Kalkulované sloupce a kalkulované míry

[5] Řešení vazeb v datech

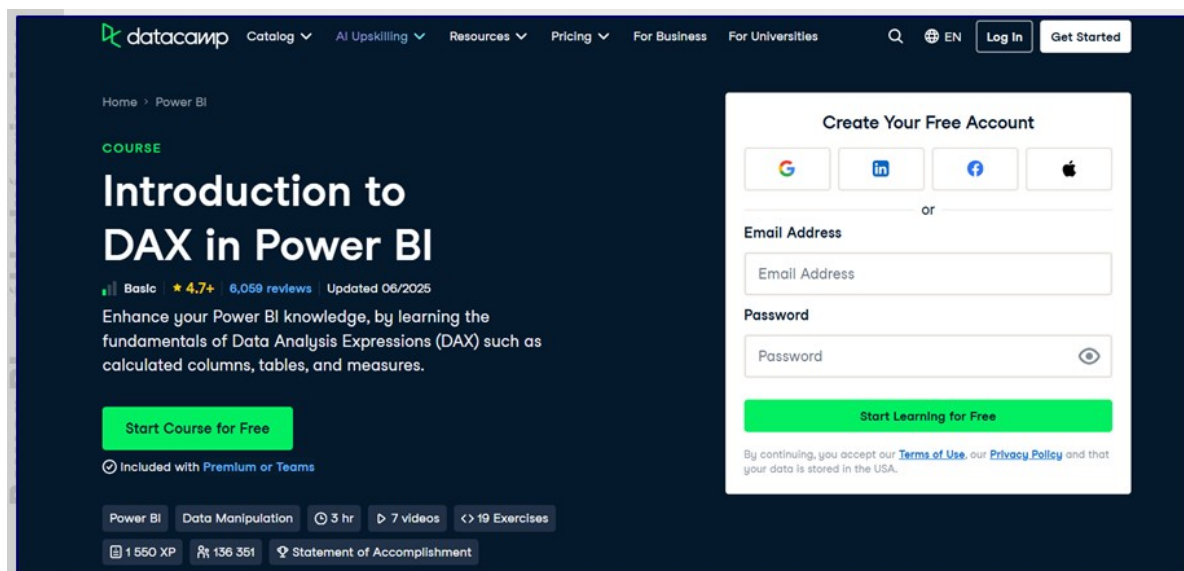
[6] Time Intelligence

Další text oddílu, věnovaný jazyku DAX, je shrnutím a prezentací principů a možností tohoto prostředí. Doplnující informace a detaily nabízejí vybrané aktuálně **dostupné tutoriály**:

Jazyk DAX, přehled: [<https://learn.microsoft.com/en-us/dax/>]



DAX na systému DataCamp: [[Introduction to DAX in Power BI Course | DataCamp](#)]



1. Základní principy a užití DAX



Účelem této vstupní kapitoly je pouze **ve stručné formě poskytnout předpoklad** pro vytváření kalkulač s pomocí jazyka DAX, **bez dalších detailnějších informací**. K těm se vrátíme v dalších kapitolách.

DAX je dotazovací i funkční jazyk. Je to efektivní jazyk **pro práci s multidimenzionálně organizovanými a uloženými daty**. Kromě prostředků pro vytváření analytických aplikací umožňuje uživateli lépe a detailněji pochopit některé důležité principy pro práci v multidimenzionálním prostředí, které mohou být užitečné nejen při vytváření aplikací v Power BI, ale i při návrzích aplikací, realizovaných v jiných produktech.

Na druhou stranu je dobré uvést, že v souvislosti s charakterem Self Service BI aplikací a orientací na samostatnou práci koncových uživatelů, mohou být některé funkce a postupy náročnější. Proto je obsahem tohoto dokumentu určitý **širší základ jazyka DAX** s tím, že je na uživateli, co nakonec pro svoji práci využije, co bude skutečně potřebovat a co nikoli. Navíc je ověřená zkušenost, která platí nejenom pro DAX, že si každý uživatel postupně ověřuje různé možnosti a funkce a tím i rozšiřuje škálu svých možností využití daného prostředku.

DAX je charakterizován jako **funkční jazyk**, což znamená, že výpočty primárně využívají funkce, které generují výsledky. Disponuje aktuálně s více než **200 funkcemi** a dále se rozvíjí. Každá kalkulace využívá jednu nebo více definovaných funkcí, nelze však vytvářet vlastní uživatelské funkce. Funkce **na výstupu poskytují jednotlivé hodnoty nebo tabulky, na vstupu využívají parametry**. Funkce mohou být vnořovány („nested“), tedy výstup jedné funkce je vstupem pro další.

Na rozdíl od SQL jazyka neumožňuje funkce jako INSERT, UPDATE, DELETE. Data v tabulce v Power BI nelze měnit, pouze dotazovat nebo filtrovat.

1.1 Výpočetní předpis v DAX

Formulace kalkulač a dalších operací v DAX je obdobná některým základním pravidlům Excelu, na druhé straně má i výrazné odlišnosti.

Každý **výpočetní předpis v DAX začíná rovnítkem „=“** nebo přiřazením „:=“, jednotlivé prvky předpisu výlučně používají pro identifikaci názvu tabulky a názvu sloupce v tabulce. Například výpočetní předpis pro vynásobení skutečného prodeje v kusech cenou za kus má následující tvar:



= FTQ_Prodej [Prodej_Skut_Ks] * FTQ_Prodej [Cena_Ks]

kde *FTQ_Prodej* je název tabulky a *Prodej_Skut_Ks* a *Cena_Ks* jsou názvy sloupců, které v kalkulačích musí být vždy uzavřeny v hranatých závorkách.

Pokud je název tabulky víceslovný, začíná-li číslicí nebo je shodný s některým rezervovaným slovem DAXu, (např. Date, SUM apod.), je třeba název uzavřít do apostrofů, například *'Data Prodeje'*.

V rámci DAX se podle daných pravidel uplatňují (viz další kapitoly) tyto typy výpočtů, resp. kalkulač:

- **kalkulovaný sloupec**, realizovaný pouze v rámci jedné (každé) řádky datové tabulky,
- **kalkulovaná míra**, kde se výpočet uskutečňuje na úrovni celé datové tabulky,
- **kalkulovaná tabulka**, mohou být využity pouze v Power BI a SSAS Tabular, v určitých verzích DAX.

Podrobněji se k nim vrátíme v dalších kapitolách, nyní pouze naznačíme jejich využití.

1.2 Vytvoření kalkulovaného sloupce

Kalkulované sloupce jsou v Power BI přidány do tabulek, v nichž jsou definovány. Volbou v menu nebo v seznamu polí tabulky „**Nový sloupec**“ („New Column“) **se zpřístupní příkazová řádka** a uživatel do ní napíše příslušný příkaz DAX.

1.3 Vytvoření kalkulované míry

Míra se počítá na agregační úrovni dat, tj. **za tabulku**, nebo její podmnožinu. Příkladem pro práci s mírou může být i celková marže z prodeje zboží, tj. celkový objem prodeje v Kč – celkové náklady na prodané zboží. K vytvoření míry se využije funkce „Míry“.



```
=SUM (FTQ_Prodej [Prodej_Kc]) - SUM (FTQ_Prodej [Naklady])
```

V každém případě je třeba, aby **byla vybrána dimenzionální nebo faktová tabulka**, do které se příslušný sloupec nebo míra zařadí. **V případě míry** je zřejmé, že **nemusí být v podstatě zařazena do žádné tabulky**, nebo by mohla vzniknout **tabulka obsahující pouze vypočtené míry**, ale pro práci uživatelů je srozumitelnější, aby i míry byly dostupné ze seznamu polí tabulek, k nimž logicky nejvíce patří. Tak například vypočtená míra „**Marže**“ logicky přísluší k prodejem zboží, je tedy zařazena mezi pole faktové tabulky *FTQ_Prodej* (tzv. *Home Table*). Zařazení vypočteného ukazatele (míry) do domovské tabulky je viditelné a zároveň je možné **změnit ho v pohledu Data** v menu *Modeling*.

Jméno vytvořené kalkulované míry musí být v rámci celého sémantického modelu **jedinečné**.

1.4 Vytvoření kalkulované tabulky

Kalkulované tabulky lze vytvořit příslušnou funkcí DAX, která na výstupu vrací tabulku nebo prostým zkopírováním jedné tabulky do nové, jak ukazuje další příklad:



```
Cas_1 = Cas
```

Tabulka *Cas_1* bude prostou a plnou kopií tabulky *Cas* s tím, že **změny tabulky** *Cas* se budou promítat do tabulky *Cas_1*. Na druhé straně změny v tabulce *Cas_1* původní tabulku *Cas* nezmění. To znamená, že do nové tabulky lze přidávat nové sloupce, míry, vazby, aniž by to ovlivnilo původní tabulku *Cas*.

Kalkulované tabulky mohou představovat souhrnné, resp. **agregované hodnoty** z původních tabulek, což pak urychlí a zpřehlední další výpočty (viz další kapitoly). Dalším jednoduchým příkladem vytvoření tabulky je vytvoření modifikované tabulky s využitím funkce *FILTER*, pouze s daty pro rok 2023.



```
Cas_23 = FILTER(  
    Cas,  
    'Cas' [Rok] = 2023  
)
```

2. Konstrukce jazyka DAX



Účelem kapitoly, věnované konstrukci jazyka DAX, je provést **celkovou rekapitulaci** a vymezení definovaných datových typů, operátorů a rovněž proměnných, používaných pro zefektivnění vyjádření kalkulačních předpisů.

2.1 Datové typy

DAX používá **standardní datové typy**, a to:

- *Whole Number, integer.*
- *Decimal Number, Real.*
- *Fixed Decimal / Currency.*
- *Date / Time.*
- *TRUE/FALSE (Boolean).*
- *Text.*

V DAXu platí, že výsledný typ ve výrazu vychází z datových typů jednotlivých částí výrazu, např. pokud je ve výrazu použit datový typ *Date*, pak výsledek bude rovněž mít datový typ *Date*. Pokud je např. třeba zvýšit aktuální datum ve sloupci *Datum* o 3, pak pro výraz

$$= FTQ_Prodej [Datum] + 3$$

budou výsledky vypadat takto: původní: 15/5/2017, nové: 18/5/2017

20/7/2018 23/7/2018 atd.

DAX rovněž automaticky konvertuje řetězce na čísla a čísla na řetězce, ***jak to vyžaduje příslušný operátor*** (tzv. *operator overloading*), např. pro výraz **7 & 2** bude výsledná hodnota 72 (& je operátor konkaténace, spojování řetězců). Na druhé straně pro „6“ + „3“ bude výsledek 9, řetězce jsou převedeny na čísla. Je zřejmé, že zde výsledek závisí na operátorovi, nikoli na typu vstupních dat.

Pokud je třeba vygenerovat v DAX hodnotu datum, využije se k tomu funkce *DATE* s parametry rok, měsíc, den, např. *DATE (2024, 11, 27)*.

2.2 Operátory

Operátory ve výrazech používá DAX standardní, takže dále je pouze jejich stručná rekapitulace:

- aritmetické: +, -, *, /, ^,
- porovnávací: =, <>, >, >=, <, <=,
- konkaténace: &,
- logické: AND nebo &&, OR nebo ||, IN.

Pořadí operátorů podle priorit je následující:

1. ^ (exponent),
2. - (znaménko, kladné nebo záporné),
3. *, /,
4. ! (NOT),
5. + - ,
6. & (konkaténace),
7. = < > <= >= <>.

2.3 Sloupce

Data jsou ukládána ze zdrojů do sloupců tabulek s tím, že jejich **obsah je považován za statický**, tedy nemůže být měněn, pouze jen při dalším vstupu nebo aktualizacích ze zdrojů.

2.4 Vazby

Vazby propojují tabulky s tím, že se připouští **pouze vazby 1:1 nebo 1:M**, ale pouze s využitím pouze jednoho sloupce v každé tabulce. Vazby umožňují, aby výběry řádek na základě funkce *FILTER* na jedné tabulce stejně proběhly v další tabulce s vazbou na původní tabulku. *FILTER*, použitý v tabulce na straně M (many), se neuplatní v tabulce na straně 1. V opačném případě se filtr uplatní, a to i přes více propojených tabulek, ale pouze ve směru 1 : M.

Obdobně se nabízejí možnosti provádět výpočty s daty ve sloupcích různých provázaných tabulek.

2.5 Hierarchie

Hierarchie jsou **seskupení jednoho nebo více sloupců**, které vymezují jednotlivé úrovně hierarchie (pro účely *drill up* a *drill down*).

2.6 Poznámky

Poznámky mohou být zařazovány do výpočtů v DAXu na základě těchto znaků:

- // Text vpravo až do konce řádku je považován za poznámku.
- -- Text vpravo až do konce řádku je považován za poznámku.
- /* */ Text mezi dvojicemi znaků je považován za poznámku i na více řádcích.

2.7 Proměnné (Variables)

Účelem proměnných v DAX je **udělat výpočty jednodušší a některých případech i rychlejší**, využívají se k ukládání výsledků z výrazů DAX. (Seamark, 2018). Syntaxe:



VAR název proměnné = výraz
RETURN výraz

Může být definována jedna nebo více proměnných, ale celý jejich blok musí být uzavřen klíčovým slovem *RETURN*. Název proměnné nemůže obsahovat mezery nebo být uzavřen v apostrofech nebo závorkách. *RETURN* vrací výsledek a může rovněž obsahovat výraz. Následující příklad dokumentuje naplnění proměnné:



```
VAR promenna1 = 7
VAR promenna2 = promenna1 + 3
RETURN promenna2 * 5
```

Proměnné mohou obsahovat nejen numerické hodnoty, ale i texty, např.:



```
VAR promText1 = „Zjistí“
VAR promText2 = „hodnotu prodeje“
RETURN CONCATENATE (promText1, promText2)
```

Proměnné lze vnořovat, ale musí být zajištěno správné nastavení *RETURN*, jak ukazuje další příklad kalkulované míry:



Celkovy_prodej =

VAR Prodej1a = 200

VAR Prodej1b =

VAR Prodej2a = Prodej1a

VAR Prodej2b = Prodej2a + 150

RETURN Prodej2b

RETURN Prodej1b

Proměnné lze využívat v různých situacích včetně kalkulovaných sloupců, kalkulovaných měr i kalkulovaných tabulek. K jejich využití se vrátíme v dalších kapitolách.

3. Kontext vyhodnocování výpočetních předpisů



Kontext vyhodnocování výpočetních předpisů a jeho pochopení představuje **jádro pochopení** konstrukce a využití celého jazyka DAX.

Účelem této kapitoly je objasnit principy a využití dvou tzv. kontextů, tj. **filtr kontextu a řádkového kontextu** a současně i jejich vliv na tvorbu **kalkulovaných sloupců a kalkulovalých měř.**

DAX a na jeho základě definované výpočetní předpisy musí ve svém principu respektovat uspořádání a prostředí databáze Power BI, nebo SSAS Tabular a jejich uložení dat. To znamená, že musí respektovat **kontext**, v němž se definovaný předpis bude vyhodnocovat (*evaluation context*).

Pochopení kontextu vyhodnocování dat **se promítá do dvou základních typů výpočetních předpisů** i následně do využití nejrůznějších funkcí jazyka DAX. Kontext je ve svém základu dán systémem ukazatelů a dimenzí, resp. jejich úrovní a prvků, tedy multidimenzionálním uspořádáním dat v sémantickém modelu Power BI. Z pohledu zobrazení dat v kontingenčních tabulkách je pak kontext dán řádky a sloupci a jejich položkami, filtry a průřezy (*slicers*), na jejichž základě jsou reálná data zobrazena.

DAX rozlišuje **dva kontexty**, a to:

- **filtr kontext (filter context)**, vztahující se k souhrnným hodnotám a aktuálně nastaveným filtrům tabulky,
- **řádkový kontext (row context)**, mající základ ve výpočtech a dalších operacích v rámci jedné řádky.

Většina pravidel kontextu se realizuje **automaticky**, ale v některých případech se nabízejí možnosti, jak pomocí kontextu upravovat požadované výpočty. Kontext představuje filtrovací vrstvu, dynamicky určující způsob výpočtů včetně řádkových i sloupcových součtů.

3.1 Filtr kontext

Filtr kontext je **sada sloupcových filtrů**, využívaných u každých výpočtů, určujících, které řádky tabulky vstoupí do dalších výpočtů. Platí tato pravidla (Seamark, 2018):

- každý filtr může být založen **na jednom nebo i více sloupcích** současně,
- je účelné se dívat **na každou buňku** výstupní tabulky jako na **samostatný výpočet**,
- je rovněž účelné se dívat na kontext jako na **kontejner**, který je buď prázdný, nebo obsahuje jeden nebo více sloupcových filtrů,
- jakmile je příslušný výpočet proveden, kontejner se **vyprázdní**.

3.1.1 Implicitní filtr kontext

Předpokládejme, že máme kalkulované míry tržeb za prodeje skupin zboží v letech (např. v milionech Kč) v následující výstupní tabulce:

	Počítače	Televize	Pračky	Součet
2020	7,3	10,5	8,1	25,9
2021	6,1	8,4	6,7	21,2
2022	8,4	12,6	9,3	30,3
Součet	21,8	31,5	24,1	77,4

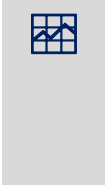
Jednotlivé hodnoty buněk odpovídají výpočtům kalkulované míry, filtrované na základě prvků sloupců a řádek. To představuje efekt **filtr kontextu kalkulované míry**, který **DAX implicitně doplňuje** ke každé z 16 buněk tabulky. To znamená, že každý výpočet kalkulované míry v této tabulce využívá unikátní filtr kontext.

Všechny filtry filtr kontextu jsou založené **na logickém součinu (AND)**. Každá buňka tabulky má tak svůj filtr kontext a každý filtr kontext má svoji sadu sloupcových filtrů.

3.1.2 Explicitní filtr kontext

Explicitní filtr kontext se vztahuje k vyjádření kalkulací, kde specifikuje přidání nebo odebrání sloupcevého filtru k nebo z filtr kontextu. Umožňuje upravovat, resp. měnit předpokládané způsoby výpočtu. Výsledek explicitního filtr kontextu se liší podle toho, zda je uplatněn na kalkulovaný sloupec, nebo kalkulovanou míru. To dokumentuje další příklad:

Kalkulovaný sloupec:



```
Televize_Kalk_Sloupec = CALCULATE (  
    Sum ('Fakt Prodej' [Trzby]),  
    'Dimenze Zboží' [Typ] = „Televize“  
)
```

Kalkulovaná míra:



```
Televize_Kalk_Mira = CALCULATE (  
    Sum ('Fakt Prodej' [Trzby]),  
    'Dimenze Zboží' [Typ] = „Televize“  
)
```

Je evidentní, že kalkulační předpis je zcela identický, ale výstupy se budou vzhledem k odlišnému kontextu lišit. Výpočty probíhají na faktové tabulce *'Fakt Prodej'* a definují filtr na základě sloupce v jiné tabulce *'Dimenze Zboží'* (vazba je definována v sémantickém modelu). Rozdíl ve výsledném zobrazení ukazují následující tabulka.

	Tržby	Televize Kalk Sloupec	Televize Kalk Mira
Počítače	21,8		31,5
Televize	31,5	31,5	31,5
Pračky	24,1		31,5
Součet	77,4	31,5	31,5

Zatímco v případě výpočtu kalkulovaného sloupce je odpovídající hodnota uvedena pouze v případě položky „Televize“, v případě kalkulované míry se opakuje pro všechny položky zboží.

3.2 Řádkový kontext

Řádkový kontext je využíván **při tvorbě kalkulovaných sloupců**, nebo v případě využití tzv. **iterátorů** (viz dále). Jako v případě filtr kontextu se vztahuje k výpočtům pro každou buňku výstupní tabulky. Pokud se např. vytváří kalkulovaný sloupec pro výstupní tabulku o 20 řádcích, pak výpočet se bude opakovat 20x, vždy s mírně odlišným řádkovým kontextem. Platí, že pokud jsou data do sémantického modelu uložena nebo obnovena, už nemohou být dále měněna.

Zásadní rozdíl mezi kalkulovaným sloupcem a mírou je v tom, že **hodnoty buněk kalkulovaného sloupce** se počítají v okamžiku fyzického vstupu nebo obnovy dat do sémantického modelu, zatímco **hodnoty kalkulované míry** se přepočítávají vždy při změně filtru, který ji ovlivňuje.

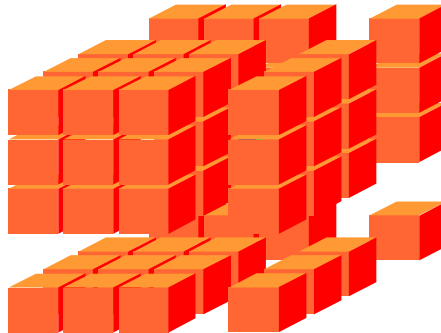
Kalkulované sloupce a míry, oba druhy kontextu a jejich detailní charakteristiky, budou předmětem dalšího textu. Rozdíl mezi nimi dokumentuje Obrázek 3-1.

Kalkulovaný sloupec v kontextu jedné řádky:

A	B	C	D	E	B + D
---	---	---	---	---	-------

Kalkulovaný sloupec –
výpočet položek B +
D v dané řádce

Míra na agregační úrovni v kontextu nastavených filtrů:



Míra – na agregační
úrovni s
respektováním všech
nastavených filtrů

Obrázek 3-1: Rozdíl v kontextu kalkulovaného sloupce a míry

4. Kalkulované sloupce a kalkulované míry



Účelem kapitoly je vrátit se ve větším detailu ke kalkulovaným sloupcům a kalkulovaným mírám a funkcím, které jsou s nimi spojeny, i k příkladům jejich užití.

DAX umožňuje definovat kalkulované sloupce, tedy vlastní výpočty v tabulkách Power BI. **Kalkulovaný sloupec** je vypočítáván **v řádkovém kontextu**, tj. v kontextu jednotlivých řádek tabulky a vytváří v tabulce nový kalkulovaný sloupec z hodnot položek a konstant dané řádky. **Kontext řádky** tak znamená, že definovaný výpočet se vždy provede právě na jedné řádce, a to pouze v jejím rozsahu, tedy v kontextu této řádky. Zpracování postupně přechází z první na další řádky v rámci celé tabulky.

Hodnoty kalkulovaného sloupce jsou vypočítávány v průběhu obnovení, resp. aktualizace tabulky, výpočet **není závislý** na aktivitách uživatele při práci s kontingenční tabulkou, zejména na nastavení filtrů.

4.1 Vytvoření kalkulovaného sloupce a řádkový kontext

K vytváření kalkulovaných sloupců doplníme **několik poznámek**:

- na kalkulované sloupce se lze rovněž **odkazovat jejich názvem**, proto je účelné využívat názvy, které co nejlépe vyjadřují obsah sloupce, resp. kalkulace,
- všechny položky (v řádcích) kalkulovaného sloupce **využívají stejný výpočet** (nikoli jako v listech Excelu) – pokud je nutné zařadit nějaké výjimky, pak pouze s využitím funkce **IF ()**,
- výsledky výpočtů v kalkulovaném sloupci **se ukládají do datasetu Power BI**, tj. do aktualizovaného souboru *nazev_souboru.pbix* – to má pozitivní dopady na výkonnost aplikací Power BI,
- výpočty se provádějí definicí nebo redefinicí** kalkulačního předpisu, případně při spuštění funkce aktualizace, což znamená, že data v kalkulovaném sloupci jsou (oproti *Míře*) statická, nemění se s použitím filtrů),
- v rámci kalkulovaných sloupců lze **využít celou škálu funkcí**, např. **=SUM ([Naklady])** – agregovaná hodnota bude stejná v každém poli, resp. řádku sloupce, nebo **=AVERAGE ([Naklady])**, **=COUNT ([Naklady])**, **=MONTH ([Datum_Objednavky])**, **=YEAR ([Datum_Objednavky])** apod.

4.2 Příklady kalkulovaných sloupců

Příkladem kalkulovaného sloupce je výpočet zisku / ztráty z jednotlivých prodejů zboží:



$$\text{Zisk} = \text{FTQ_Prodej} [\text{Prodej_Kc}] - \text{FTQ_Prodej} [\text{Naklady}]$$

Dalším příkladem je kalkulovaný sloupec pro „Objem_prodeje“, jak ukazuje Obrázek 4-1:

FTQ_Prodej_Id	Cas_Id	Utv_Id	Reg_Id	Zak_Id	Zbo_Id	Cena_Ks	Prodej_Skut_Ks	Naklady	Datum_Objednavky	Rok	Prodej_Kc	Objem_prodeje
553	553	22	1	1	1	1634,178	2	1136,1	pondělí 29. června 2015	2015	3268,356	3268,356
554	554	23	11	93	28	9503,578	2	6607	úterý 30. června 2015	2015	19007,156	19007,156
555	555	24	4	31	55	2247,378	2	1562,4	středa 1. července 2015	2015	4494,756	4494,756
556	556	22	15	120	17	1020,978	2	709,8	čtvrtek 2. července 2015	2015	2041,956	2041,956
557	557	23	8	71	44	877,898	2	610,4	pátek 3. července 2015	2015	1755,796	1755,796
558	558	24	2	11	6	1634,178	2	1136,1	sobota 4. července 2015	2015	3268,356	3268,356
559	559	22	1	1	2	1736,378	3	1207,2	neděle 5. července 2015	2015	5209,134	5209,134
560	560	23	11	94	29	9585,338	3	6663,8	pondělí 6. července 2015	2015	28756,014	28756,014
561	561	24	4	31	56	3780,378	3	2628,2	úterý 7. července 2015	2015	11341,134	11341,134
562	562	22	15	120	18	5313,378	3	3693,9	neděle 5. července 2015	2015	15940,134	15940,134
563	563	23	8	71	45	581,518	9	404,3	pondělí 6. července 2015	2015	5233,662	5233,662
564	564	24	2	11	7	2042,978	9	1420,3	úterý 7. července 2015	2015	18386,802	18386,802
565	565	22	1	1	3	2451,778	5	1704,5	úterý 7. července 2015	2015	12258,89	12258,89

Obrázek 4-1: Příklad kalkulovaného sloupce

Dalším příkladem pro **přidání kalkulovaného sloupce** „Den název“ do tabulky **DATUM** je demonstrace **funkce SWITCH** pro přidání sloupce s českým názvem dne podle hodnoty pořadového čísla dne v týdnu ve sloupci „Den v týdnu“:

	$\text{Den_nazev} = \text{SWITCH} (\text{DATUM} [\text{Den_v_tydnu}],$ $1, \text{"Pondělí"},$ $2, \text{"Úterý"},$ $3, \text{"Středa"},$ $4, \text{"Čtvrtek"},$ $5, \text{"Pátek"},$ $6, \text{"Sobota"},$ $7, \text{"Neděle"})$
---	---

4.3 Vytváření měř a filtr kontext

Míra, resp. *calculated measure*, nebo *počítané pole*, *calculated field*, se počítá na agregační úrovni dat, tj. **za tabulku** nebo její podmnožinu na základě **kontextu**, **daného filtrem** nebo filtry (*filter context*). Data jsou tak vypočítávána na úrovni agregací (nikoli 1 řádky) a s respektováním filtrů, které ovlivňují jednotlivé buňky. *Míry* se tak nemohou vztahovat k jednotlivým řádkům. Míra se může vztahovat i k jednomu nebo více kalkulovaným sloupcům. Je součástí datové sady v Power BI – je k dispozici pro všechny reporty nad datovou sadou. K vytváření a využití *měř* jsou uvedeny **další poznámky**:

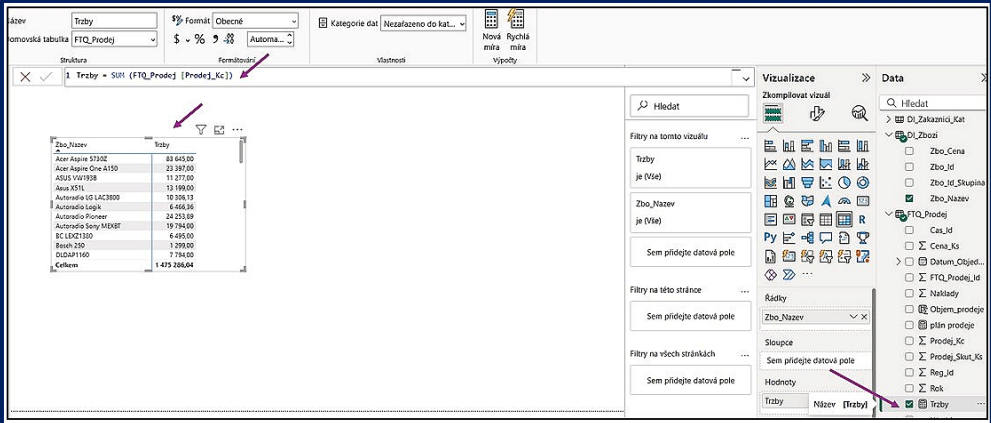
- všechny *míry* **se nabízejí v řídicích oknech**,
- úpravy míry se promítají** do dalších měř, které mají na vstupu danou upravovanou míru (tyto změny mohou probíhat v několika úrovních tak, jak se využívají již vytvořené míry pro další nově vytvářené),
- míry se také označují jako **přenosné výpočty**, resp. *portable formulas*, protože se mohou využívat v mnoha různých situacích,
- každá buňka** pro míru je vypočítávána **nezávisle**.

Příklady míry:

- míra pro tržby** (funkce **SUM**), viz Obrázek 4-2:



$$\text{Trzby} = \text{SUM} (\text{FTQ_Prodej} [\text{Prodej_Kc}])$$



Obrázek 4-2: Výpočet míry "Trzby"

- míra pro výpočet **souhrnu nákladů**, viz Obrázek 4-3:

$$\text{Suma_Naklady} = \text{SUM}(\text{FTQ_Prodej}[\text{Naklady}])$$

Zbo_Nazev	Suma_Naklady
Acer Aspire 5730Z	6 691,60
Acer Aspire One A150	6 239,20
ASUS VW1938	3 007,20
Asus X51L	10 559,20
Autradio LG LAC3800	2 396,50
Autradio Logik	2 355,40
Autradio Pioneer	3 383,80
Autradio Sony MEXBT	2 619,20
BC LEI21380	519,60
Bosch 250	1 039,20
DLDAPI160	1 039,20
Celkem	215 405,40

Obrázek 4-3: Míra pro výpočet sumy nákladů

- míry jsou použity **při výpočtu marže**, viz Obrázek 4-4:

$$\text{Marze_2} = \text{CALCULATE}(\text{FTQ_Prodej}[\text{Trzby}] - \text{FTQ_Prodej}[\text{Suma_Naklady}])$$

Zbo_Nazev	Suma_Naklady	Marze_2
Acer Aspire 5730Z	6 691,60	76 933,40
Acer Aspire One A150	6 239,20	17 157,80
ASUS VW1938	3 007,20	6 265,80
Asus X51L	10 559,20	2 639,80
Autradio LG LAC3800	2 396,50	7 909,83
Autradio Logik	2 355,40	4 210,96
Autradio Pioneer	3 383,80	20 870,09
Autradio Sony MEXBT	2 619,20	17 154,80
BC LEI21380	519,60	5 975,40
Bosch 250	1 039,20	259,80
DLDAPI160	1 039,20	6 754,80
Celkem	215 405,40	1 259 880,64

Obrázek 4-4: Výpočet marže

- pro výpočet **počtu prodejních transakcí**, jednotlivých řádků lze využít následující míru, která spočítá počet řádek faktové tabulky, tedy počet prodejních transakcí, viz Obrázek 4-5:



$Pocet_Transakci = COUNTROWS (FTQ_Prodej),$

Zbo_Nazev	Součet hodnot Prodej_Kc	Pocet_Transakci
Acer Aspire 5730Z	83 645,00	1
Acer Aspire One A150	23 397,00	1
ASUS VW1938	11 277,00	1
Asus X51L	13 199,00	1
Autoradio LG LAC3800	10 306,13	2
Autoradio Logik	6 466,36	2
Autoradio Pioneer	24 253,89	2
Autoradio Sony MEXBT	19 794,00	1
BC LEK21380	6 495,00	1
Bosch 250	1 299,00	1
CLDAP1160	7 794,00	1
DWLG510	10 354,66	2
Espresso Delonghi E2100	78 393,00	1
Espresso Delonghi ESAM1300	91 495,00	1
Fujitsu Siemens AMILO Pa 3515	45 597,00	1
HP Compaq 2133	72 793,00	1
HP Pavilion A6611CS	56 893,01	2
HP Photosmart D5460	9 697,00	1
Celkem	1 475 286,04	79

Obrázek 4-5: Výpočet počtu prodejních transakcí

- vytvořené míry lze **využít při definování nových měr a postupně je vkládat do sebe**. To pak umožňuje, že provedení určité změny do jedné míry se automaticky promítne do všech dalších, kde je tato míra využita. Příkladem je **výpočet marže na jednu obchodní transakci**, viz Obrázek 4-6



$Marze_na_transakci = [Marze_2] / [Pocet_Transakci]$

Zbo_Nazev	Tržby	Marze_2	Marze_na_transakci
Acer Aspire 5730Z	83 645,00	76 953,40	76 953,40
Acer Aspire One A150	23 397,00	17 157,80	17 157,80
ASUS VW1938	11 277,00	6 269,80	6 269,80
Asus X51L	13 199,00	2 639,80	2 639,80
Autoradio LG LAC3800	10 306,13	7 909,63	3 954,82
Autoradio Logik	6 466,36	4 210,96	2 705,48
Autoradio Pioneer	24 253,89	20 070,09	10 435,55
Autoradio Sony MEXBT	19 794,00	17 154,80	17 154,80
BC LEK21380	6 495,00	5 975,40	5 975,40
Bosch 250	1 299,00	299,80	299,80
CLDAP1160	7 794,00	6 754,60	6 754,60
Celkem	1 475 286,04	1 259 880,64	15 947,86

Obrázek 4-6: Výpočet marže na transakci

- pro výpočet dnů, kdy byly zpracovány objednávky** na zboží (mohou se v přehledu prodejů opakovat), lze použít následující míru, která vypočítá počet výskytů unikátních, neopakujících se datumů objednávek:



[Pocet_Dnu_Prodeje]:= DISTINCTCOUNT (FTQ_Prodej [Datum_Objednavky])

Zboj_Nazev	Trzby	Marze_2	(Marze na transakci)	Pocet_Dnu_Prodeje
Acer Aspire 5730Z	83 645,00	76 983,40	76 983,40	1
Acer Aspire One A150	23 397,00	17 157,80	17 157,80	1
Astus V10198	11 277,00	8 269,80	8 269,80	1
Astus X31L	13 199,00	2 639,80	2 639,80	1
Autorelio LG LAC1800	10 306,13	7 909,63	3 954,82	2
Autorelio Logik	6 466,96	4 210,96	2 105,48	2
Autorelio Pioneer	24 133,89	20 870,09	10 435,05	2
Autorelio Sony MEXBT	19 784,00	17 154,80	17 154,80	1
8C LEX21303	6 495,00	5 975,40	5 975,40	1
Bosch 250	1 299,00	259,80	259,80	1
OLDAP1160	7 754,00	6 754,00	6 754,00	1
Celkem	1 475 286,04	1 299 880,64	15 947,86	73

4.4 Funkce CALCULATE()

Funkce **CALCULATE()** se považuje **za nejdůležitější, nejužitečnější a nejkomplexnější funkci** jazyka DAX. Podstatnou charakteristikou **CALCULATE()** je to, že umožňuje modifikovat **filtr kontext**, tedy zrušení a nastavení filtrů podle okamžité potřeby. Jinými slovy, **CALCULATE()** nastavuje nový filtr kontext a na jeho základě vyhodnocuje specifikovaný výraz. **Syntaxe** **CALCULATE()** je následující:



CALCULATE (výraz míry, <podmínka 1>, <podmínka 2>, ...)

Pro **CALCULATE()** je **jediný povinný parametr**, a to ten první – **výraz míry**, další – podmínky, resp. filtry jsou nepovinné a jejich počet není nijak omezen. **CALCULATE()** **funguje následujícím způsobem**:

- **převzme aktuální filtr kontext** a vytvoří z něj kopii do nového filtr kontextu,
- **vyhodnocuje každou podmínku** v zadaných parametrech a pro každou z nich platí, že:
 - pokud je použit sloupec, který ještě **nebyl** předmětem filtru v původním kontextu, přidá tuto podmínku do nově vytvářeného filtr kontextu,
 - pokud na druhé straně je použit sloupec, který už **byl** předmětem filtru v původním kontextu, pak nahradí existující filtr novou zadanou podmínkou,
- všechny podmínky jsou v novém filtr kontextu vyhodnocovány **s uplatněním logického součinu**, tedy operátoru **AND**,
- jakmile je nový filtr kontext vytvořen, **výraz míry je na jeho základě vyhodnocen**.

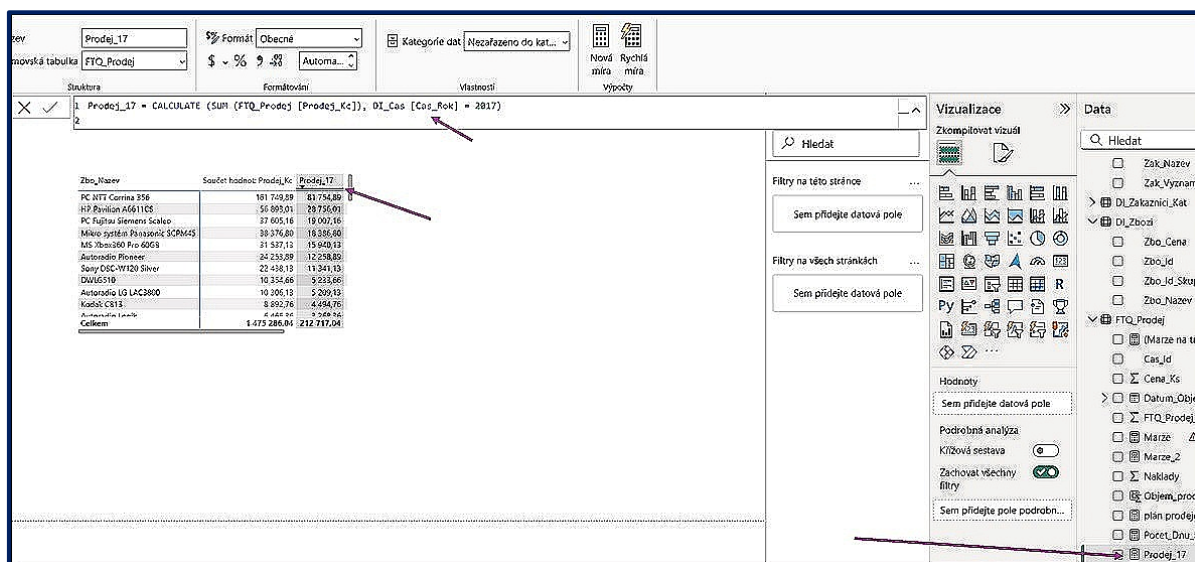
CALCULATE() rozeznává **2 typy filtrů**:

- **booleovské podmínky**, jako např. **DI_Zbozi [Cena] > 2000**, tyto filtry se používají na jednotlivých sloupcích – s výslednou hodnotou **Ano / Ne**, resp. **TRUE / FALSE**,
- **seznamy hodnot, prezentované ve formě tabulky**, které mají být součástí nového filtr kontextu a všechny sloupce této tabulky jsou součástí filtru.

Příklad použití funkce **CALCULATE()** pro **výpočet hodnoty prodeje v roce 2017**, viz Obrázek 4-7:



=**CALCULATE** (SUM (FTQ_Prodej [Prodej_Kc]), DI_Cas [Cas_Rok] = 2017)



Obrázek 4-7: Objem prodeje v roce 2017

Další příklady:

- Prodej zboží pouze v roce 2025 a současně za zboží s identifikátorem 35:



= CALCULATE (SUM (FTQ_Prodej [Prodej_Kc]), DI_Cas[Cas_Rok] = 2025, DI_Zbozi [Zboj_Id] = 35)

- Pokud je potřeba, je ve výrazu použit operátor pro vyjádření logického součtu, tj. ||, nebo OR. Operátor || je možné použít pouze pro porovnání dvou stejných položek (sloupců), např. Zboj_Id.:



=CALCULATE (SUM (FTQ_Prodej [Prodej_Kc]), DI_Zbozi [Zboj_Id] = 60 || DI_Zbozi [Zboj_Id] = 63)

4.5 CALCULATE() v řádkovém kontextu

Další podstatnou úlohou funkce CALCULATE() je **transformace** (v případě potřeby) **řádkového kontextu na filtr kontext**. Předpokládejme, že potřebujeme vytvořit souhrn ceníkových cen a uložit ho do kalkulovaného sloupce (Souhrn_Cen) na základě zápisu:



Souhrn_Cen = SUM (DI_Zbozi [Cena])

V tomto případě se tedy **v řádkovém kontextu vypočítá celkový souhrn ceníkových cen** a ten se uloží do každého řádku kalkulovaného sloupce Souhrn_Cen. Pokud ale v rámci řádkového kontextu požadujeme souhrny ceníkových cen podle jednotlivých druhů zboží, pak použijeme v rámci řádkového kontextu funkci CALCULATE(), takto:



Souhrn_Cen_Calc = CALCULATE (SUM (DI_Zbozi [Cena]))

V tomto případě neobsahuje zápis `CALCULATE()` **žádné filtry** a tedy nemění existující filtr kontext, ale na druhé straně **transformuje existující řádkový kontext na filtr kontext**. Uskutečňuje tak **transformaci kontextu**. Na základě tohoto zápisu výraz `SUM ( DI_Zbozi [Cena] )` je vypočítáván v rámci filtr kontextu vždy pouze pro jednu řádku.

Tento princip se efektivně **uplatňuje ve výpočtech s více provázanými tabulkami**. Jak jsme již zmínili, filtry se v případě řádkového kontextu nepromítají automaticky do dalších provázaných tabulek, zatímco pro filtr kontext platí, že se promítají automaticky ve směru vazby 1 ku M. To znamená, že transformace kontextu takové automatické promítání filtrů do vázaných tabulek zajistí, tedy např.

 `Souhrn_Prodeje_Calc = CALCULATE ( SUM ( FTQ_Prodej [Prodej_Skut_Ks] ) )`

To znamená, že **filtry z tabulky zboží (DI_Zbozi)** se automaticky **promítnou do tabulky faktů (FTQ_Prodej)** – jde o vazbu 1 : M – a dostaneme tak výsledky, odpovídající těmto filtrům, podobně i v obdobných případech.

4.6 Parametr ALLSELECTED

Funkce `CALCULATE()`, jak bylo již uvedeno, umožňuje **měnit, a tedy i zrušit všechny nastavené filtry** v určité tabulce, a to s využitím parametru `ALL`, např. `... ALL (DI_Zbozi [Cena]) ...` – zruší všechny filtry na sloupci ceny zboží. Pokud ale požadujeme tyto filtry **zrušit kromě těch, které jsme aktuálně nastavili** ve výstupní tabulce, pak se pro to využije parametr `ALLSELECTED`, tedy

 `... ALLSELECTED (DI_Zbozi [Cena]) ...`

Funkce `CALCULATE()` má **širokou škálu možností**, zejména může **operativně měnit aktuální filtr kontext a nastavovat nový** a tím měnit prostředí pro realizaci výpočtů a dalších operací. Lze doporučit ověřovat si na dílčích příkladech tyto možnosti a postupně tak rozšiřovat i možnosti využití jazyka DAX pro složitější analytické úlohy a aplikace Self Service Business Intelligence.

4.7 Funkce IF ()

S příkladem, kde počítáme **objem marží na jednu transakci**, tedy podíl objemu marží počtem transakcí, se logicky váže další, často využívaná funkce `IF ()`. Je zřejmé, že **pokud by** v některých případech **byl počet transakcí roven nule**, pak by míra vykazovala chybu. Ošetření tohoto stavu je pak (jako i v jiných produktech) záležitostí funkce `IF ()`. Její **syntaxe** je následující:

 `IF (podmíněný výraz, hodnota pro splnění podmínky, hodnota pro nesplnění podmínky)`

Použití funkce `IF ()` v případě výpočtu marže na jednu transakci dokumentuje další příklad:

 `=IF ([Pocet_Transakci] = 0, BLANK(), [Marze_2] / [Pocet_Transakci])`

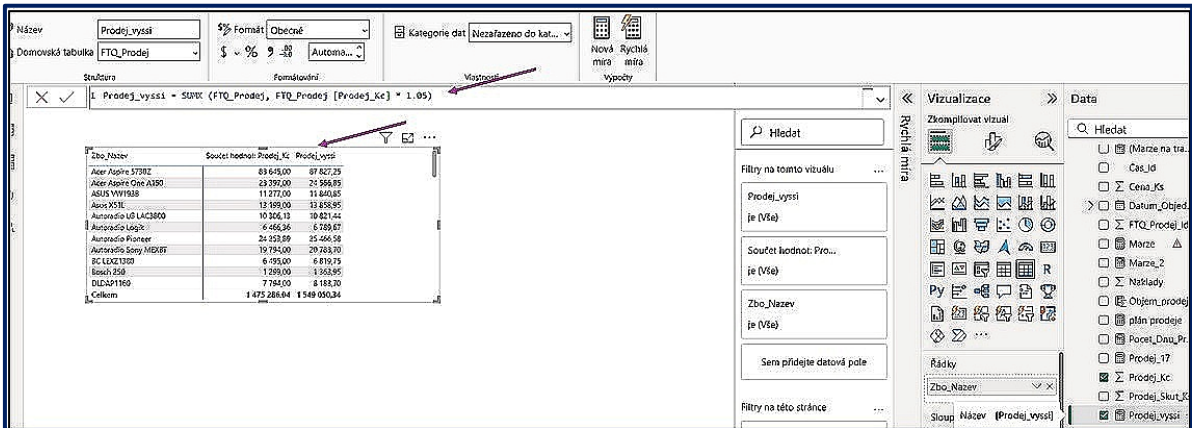
Tedy, pokud je počet transakcí roven nule, vrátí výpočet míry mezeru, resp. funkci `BLANK ()`, pokud ne, vrátí hodnotu příslušného podílu.

4.8 Iterátor

Další možností je využití řádkového kontextu s iteracemi, resp. s použitím *iterátorů*. Všechny funkce DAX, končící na „X“, jsou považovány za iterátory. To znamená, že vyhodnocují výpočty v každé řádce, a nakonec je agregují podle různých algoritmů. Tak např. funkce **SUMX**, viz Obrázek 4-8:



$$=SUMX(FTQ_Prodej, FTQ_Prodej[Prodej_Kc] * 1,05)$$



Obrázek 4-8: Využití iterátoru pro výpočet navýšení objemu prodeje

Zajistí, že v každé řádce tabulky se položka Náklady zvýší o 5 % a nakonec vrátí souhrn těchto přepočítaných hodnot. Funkce **SUMX** tak využije v tabulce **FTQ_Prodej** řádkový kontext v rámci celé iterace, i když zápis odpovídá v tomto případě výpočtu *míry* (obdobně i dále).

Všechny *iterátory* v DAX fungují stejně, a to:

- 1) vytvoří nový řádkový kontext v tabulce, definované prvním parametrem,
- 2) vyhodnotí druhý parametr v řádkovém kontextu, tj. pro každou řádku tabulky,
- 3) vytvoří agregaci z hodnot, vytvořených v průběhu druhého kroku, pokud to specifikuje iterátor. Některé iterátory (**FILTER**, **ADDCOLUMNS**) tento krok neprovádějí.

Další iterátory představuje následující přehled:

- **AVERAGEX**: aritmetický průměr hodnot z vyhodnocených výrazů v řádcích tabulky.
- **COUNTX**: počet hodnot, které jsou výsledkem vyhodnocených výrazů v každé řádce tabulky.
- **GEOMEANX**: geometrický průměr hodnot z vyhodnocených výrazů v řádcích tabulky.
- **MAXX**: maximální numerická hodnota z vyhodnocených výrazů v řádcích tabulky.
- **MEDIANX**: medián z vyhodnocených výrazů v řádcích tabulky.
- **MINX**: minimální numerická hodnota z vyhodnocených výrazů v řádcích tabulky.
- ...

5. Řešení vazeb



Výpočty v DAX se provádějí i **na vzájemně provázaných tabulkách**. Účelem kapitoly je vytvořit představu o tom, jak lze výpočty při různých typech vazeb realizovat.

Úlohy Business Intelligence i Self Service Business Intelligence zahrnují 2 typy tabulek – faktové (*data tables*) a dimenzionální (*lookup tables*). Zatímco **faktové tabulky** představují průběh, obsah a výsledky byznys procesů, **dimenzionální tabulky** vyjadřují podstatné charakteristiky všech faktorů, resp. dimenzí, které tyto procesy ovlivňují a podle kterých je účelné je analyzovat.

Většina úloh se může realizovat automaticky **na základě standardních vazeb** tabulek, definovaných v sémantickém modelu. **Nikoli** ale **ve všech případech**. K tomu se používá **řada funkcí**, zejména:

- GENERATE,
- CROSSJOIN,
- NATURALINNERJOIN,
- NATURALLEFTOUTERJOIN,
- UNION,
- LOOKUPVALUE.

5.1 Řešení ve standardních vazbách

Dimenzionální tabulky se váží na faktové tabulky **ve dvou schématech (STAR a SNOWFLAKE)**. Pro tyto tabulky a jejich vazby platí **následující pravidla**:

- vazby mohou být jen **1 : 1**, nebo **1 : M**,
- u každé tabulky může být pro vazbu využit **pouze 1 sloupec**,
- pro vyjádření vazby může být použit **pouze operátor „=“**,
- **nemůže být** použita vazba na stejnou tabulku („*selfjoins*“),
- **je třeba nedefinovat vazby mezi faktovými tabulkami navzájem**, ale pouze prostřednictvím sdílených dimenzionálních tabulek,
- není dobré vytvářet „**násilně**“ **kombinované faktové tabulky** z několika základních (*flatten tables*), které pak snižují přehlednost řešení a omezují analytické možnosti,
- na druhé straně je efektivním řešením v případě potřeby **kombinovat využití měř, vázaných na různé faktové tabulky**, v jedné výstupní tabulce.

5.1.1 Řádkový kontext s vazbami tabulek

Pokud se **zpracování v řádkovém kontextu** bude vztahovat na dvě nebo více vzájemně provázaných tabulek, pak je třeba využít parametrů **RELATED** nebo **RELATEDTABLE**. Předpokládejme, že chceme vytvořit nový kalkulovaný sloupec s hodnotou prodeje, vypočítané ze skutečného množství prodaného zboží (*Prodej_Skut_Ks*) v tabulce faktů *FTQ_Prodej* a ceníkové ceny (*Cena*) v dimenzionální tabulce *DI_Zbozi*. Zobrazení příkladu vazeb ukazuje Obrázek 5-1. V tomto případě **je podstatná kardinalita a směr vazby**, tedy pro *FTQ_Prodej : DI_Zbozi* je kardinalita vazby **M : 1**.



=COUNTROWS (RELATEDTABLE (FTQ_Prodej))

Zbo_Id	Zbo_Id_Skupina	Zbo_Nazev	Zbo_Cena	Sloupec
1	1	Autoradio Logik	1599	2
2	1	Autoradio LG LAC3800	1699	2
3	1	Autoradio Pioneer	2399	2
4	1	Autoradio Sony MEXBT	3299	1
5	2	Mikro systém Philips MC147	1799	1
6	2	Mikro systém Hitachi AXM717	1599	3
7	2	Mikro systém Panasonic SCPM45	1999	3
8	2	Mikro systém Sony CMTEH25	2199	2

Obrázek 5-3: Řešení vazeb s využitím funkce RELATEDTABLE

Je dobré zdůraznit, že **RELATED** i **RELATEDTABLE** mohou tímto způsobem procházet celý řetěz provázaných tabulek, jejich počet není limitován. Jediné pravidlo je v tom, že **vazby mezi tabulkami musí mít kardinalitu stejného typu**, tedy M : 1 nebo 1 : M a musí být definovány **ve stejném směru**.

Posledním příkladem je ukázka **použití neaktivní vazby mezi tabulkami ve výpočtu míry**, tedy využití funkce **USERELATIONSHIP**. Power BI umožňuje definovat více vztahů mezi dvěma tabulkami. Aktivní však může být pouze jeden. **Vztahy nemají názvy**, pro jejich označení je třeba použít názvy příslušných atributů tabulek na obou stranách vztahu. Aby bylo možné pracovat s počtem dodaných kusů zboží podle data dodání ve vztahu k dimenzi času, je třeba pro výpočet dodaných počtů kusů použít neaktivní vztah. Funkce **USERELATIONSHIP** v DAXu aktivuje vztah mezi tabulkami **DI_Cas** a **FTQ_Prodej** provázáním přes atributy **DI_Cas [Cas_Datum]** a **FTQ_Prodej [Datum_dodani]**, a to právě vždy jen v době použití dané míry:



Počet_dodanych = CALCULATE (FTQ_Prodej [Prodano ks],
USERELATIONSHIP (FTQ_Prodej [Datum_Dodani], DI_Cas [Cas_Datum])
)

5.1.2 Filtrování kontextu s vazbami tabulek

V předchozí části jsme se zabývali využitím řádkového kontextu a vytvářením kalkulovaných sloupců v prostředí více provázaných tabulek. Na tomto místě bude obsahem **řešení měr a filtr kontextu** rovněž v prostředí několika tabulek s definovanými vazbami (viz Obrázek 5-1).

Pro uplatnění filtr kontextu zde platí **následující pravidlo**: Filtrování, aplikované na určitou tabulku, se automaticky **promítá do všech provázaných tabulek**, které jsou **na straně „M“ (many)** ve vazbách **1 : M (one-to-many)**. Na druhé straně, pokud je tabulka na straně „1“ vazby **M : 1**, automatické promítnutí filtru ze strany „M“ do „1“ provázané tabulky neproběhne.

Jako příklady využijeme tyto míry a předpokládáme, že filtr bude nastavený na **DI_Zbozi** (např. Cena = 1000):



$Pocet_Prodeju = COUNTROWS (FTQ_Prodej)$

$Pocet_Zbozi = COUNTROWS (DI_Zbozi)$

$Pocet_Zakazniku = COUNTROWS (DI_Zakaznici)$

Filtr na tabulce DI_Zbozi se promítne takto:

- 1) **Pocet_Prodeju** – filtr je definován na straně „1“ vazby 1 : M, tedy se z tabulky DI_Zbozi **promítne** do tabulky FTQ_Prodej a Pocet_Prodeju tím **bude ovlivněn**.
- 2) **Pocet_Zbozi** – filtr je definován na vlastní tabulce a do výsledku Pocet_Zbozi se **promítne** a výsledek **bude ovlivněn**.
- 3) **Pocet_Zakazniku** – filtr (na DI_Zbozi, tj. Cena = 1000, viz výše) se **nepromítne**, protože tabulka DI_Zakaznici je na straně „1“ vazby M : 1 (FTQ_Prodej : DI_Zakaznici), a to i v případě, že by filtr byl nastaven přímo na FTQ_Prodej, výsledek nebude ovlivněn.

Určitým **řešením posledního příkladu** je **funkce VALUES**. Ta vrací tabulku o pouze jednom sloupci, obsahující všechny hodnoty, odpovídající aktuálnímu filtru. Předchozí příklad by bylo možné takto **modifikovat**:



$Pocet_Zakazniku = COUNTROWS (VALUES (FTQ_Prodej [Zak_Id])$

Jak je patrné, funkce počítá **počty zákazníků z tabulky faktů** (nikoli DI_Zakaznici) s respektováním promítnutého filtru z tabulky DI_Zbozi do tabulky faktů FTQ_Prodej.

Je dobré zdůraznit, že (obdobně jako v řádkovém kontextu) **promítnutí filtrů proběhne i v řetězu tabulek s vazbami stejného typu a ve stejném směru**.

5.2 Parametrické nepropojené tabulky

Jak název napovídá, DAX umožňuje vytvářet **speciální tabulky, které obvykle slouží jako parametry** pro přepočty hodnot ukazatelů podle aktuální situace. Tyto tabulky **nejsou propojené vazbou** k žádné jiné tabulce sémantického modelu, protože neexistuje ani klíčová položka, na jejímž základě by bylo možné vazbu vytvořit.

Příkladem může být **tabulka kursů korun za Euro** pro přepočty cen zboží. Tabulka na obrázku (Obrázek 5-4) představuje možné kursy Kč vzhledem k Euru, která bude dále **sloužit jako parametrická pro přepočty**. Je samozřejmé, že jde pouze o příklad a lze do ní doplnit množství dalších hodnot kursu.

[Kc_Euro]		fx
	Kc_Euro	Přidat sloupec
1	25,51	
2	25,59	
3	25,65	
4	25,72	
5	25,86	
6	25,95	
7	26,09	
8	26,15	
9	26,21	
10	26,25	
11	26,31	

Obrázek 5-4: Parametrická tabulka, kurs Kč za Euro

Tato tabulka bude pak při vytváření výstupních tabulek sloužit jako nabídka kursů v rámci sliceru. V dalším kroku vytvoříme *míru* maximální hodnoty kursu, viz Obrázek 5-5:



$$\text{Korun_za_Euro} = \text{MAX} (\text{CZK_EUR} [\text{Kc_Euro}])$$

Popisky řádků	Průměr Cena_Ks	Korun_za_Euro
Acer Aspire 5730Z	16729	26,31
Acer Aspire One A150	7799	26,31
ASUS VW193B	3759	26,31
Asus X51L	13199	26,31
Autoradio LG LAC3800	1699	26,31
Autoradio Logik	1599	26,31
Autoradio Pioneer	2399	26,31
Autoradio Sony MEXBT	3299	26,31
BC LEXZ1380	1299	26,31
Bosch 250	1299	26,31
Canon Pixma MP190	1799	26,31
DES1005D	389	26,31
DLDAP1160	1299	26,31
DWLG510	569	26,31
Epson Stylus SX100	1199	26,31

Obrázek 5-5: Vytvoření míry – maximální hodnota kursu

Tato *míra* bude poskytovat jednak **maximální hodnotu na kursovním lístku**, nebo bude sloužit jako **hodnota pro přepočítání** ceny zboží.

V dalším kroku vytvoříme míru pro přepočítání průměrné ceny zboží podle zvoleného aktuálního kursu Kč k Euro:



$$\text{Cena_Euro} = [\text{Průměr Cena_Ks}] / [\text{Korun_za_Euro}]$$

Míra Korun_za_Euro funguje v tomto případě tak, že pokud ve sliceru vybereme jednu hodnotu, např. 26,09, pak se do výpočtu doplní tato hodnota. Pokud nevybereme žádnou hodnotu, nahradí se maximální, tedy 26,31, viz

Obrázek 5-6.

Popisky řádků	Průměr Cena_Ks	Cena_Euro	Kc_Euro
Acer Aspire 5730Z	16729	5 770,83	25,51
Acer Aspire One A150	7799	2 690,34	25,59
ASUS VW193B	3759	1 296,70	25,65
Asus X51L	13199	4 553,12	25,72
Autoradio LG LAC3800	1699	586,09	25,86
Autoradio Logik	1599	551,59	25,95
Autoradio Pioneer	2399	827,56	26,09
Autoradio Sony MEXBT	3299	1 138,02	26,15
BC LEXZ1380	1299	398,31	
Bosch 250	1299	398,31	
Canon Pixma MP190	1799	551,63	
DES1005D	389	119,28	
DLDAP1160	1299	398,31	
DWLG510	569	174,47	

Obrázek 5-6: Přepočítání průměrné ceny zboží podle aktuálního kursu Eura

6. Time Intelligence



Účelem této kapitoly je prezentovat úlohy s využitím dimenze času a s pomocí jazyka DAX. Jde zejména o sledování časového vývoje vybraných ukazatelů, hodnocení nárůstu ukazatelů k určitému datu, meziroční porovnání a porovnání hodnot ukazatelů za různá časová období apod.

Analytické úlohy na bázi Business Intelligence a obdobně SSBI pracují v naprosté většině s dimenzí času. To umožňuje důležité analytické funkce jako např. **sledování vývoje firemních ukazatelů, meziroční porovnání jejich hodnot, klouzavé ukazatele** a další velmi široké spektrum nejrozličnějších funkcí, spojených s časem. Souhrnně se celá tato analytická oblast označuje jako **Time Intelligence**. Účelem této části je objasnit základní princip takové funkcionality a ukázat alespoň některé častěji používané případy.

6.1 Základní principy time intelligence

Zcela obecný základ, spojený s uplatněním časové dimenze, byl obsažen již v předchozím textu. Vstupním předpokladem je navržení a **naplnění časové dimenze**, je dobré k ní doplnit několik poznámek:

- **Struktura tabulky časové dimenze** by měla obsahovat kromě údaje o čase i případné další potřebné údaje pro analýzy a podnikový reporting (např. plné názvy měsíců a dnů, určení pracovních dnů atd.).
- Pro **identifikaci jednotlivých období**, většinou dní, je k dispozici primární klíč nebo údaj v definované struktuře, např. 20170101). Tento klíč také slouží k propojení s faktovými, případně některými dalšími tabulkami.
- Vedle primárního klíče se pro většinu funkcí DAX, spojených s časem, používá i **plné datum, označované jako FullDate**, které musí být datového typu. Je možné použít i jiný název, ale s ohledem na četnost použití tohoto označení i v literatuře zůstáváme v tomto případě u anglického názvu, v daném kontextu ji označujeme jako *DI_Cas_DAX*.
- **Zdrojem pro časovou dimenzi** je buď databázová tabulka **podnikového kalendáře**, spravovaná často v centrálních databázích, nebo manuálně vytvořená, resp. vygenerovaná tabulka časové dimenze. Výhodou centrálně spravovaného podnikového kalendáře jsou obvykle jeho systematická aktualizace a použití stejných algoritmů pro některé speciální části dimenze. Otázkou někdy zůstává jeho dostupnost a kvalita.
- Je nutné zajistit, aby **rozsah časové dimenze** pokrýval časový rozsah dat faktových tabulek včetně potřebné rezervy do budoucnosti.

V některých situacích se u analytických úloh **nevyhne vytvoření více tabulek pro časovou dimenzi**. To je v případech, kdy je třeba současně rozlišovat různé časové okamžiky. Například je třeba rozlišit datum objednání zboží, datum požadovaného termínu dodání zboží a datum skutečného dodání zboží. Pro většinu časových funkcí DAX je nezbytné specifikovat, že právě daná tabulka má charakter časové dimenze, a právě daná položka má význam plného data.

6.2 Vygenerování dimenzionální tabulky času

Příklad komplexního výrazu v DAXu patří také do oblasti Time intelligence pro vygenerování nové dimenzionální tabulky času s názvem *DATUM* se všemi potřebnými sloupci, pokrývající dny kalendáře pro všechna data objednávek zboží ve faktové tabulce *FTQ_Prodej*.



DATUM = ADDCOLUMNS

```
( CALENDAR ( DATE ( YEAR ( MIN ( 'FTQ_Prodej' [Datum objednávky] ) ), 1, 1 ),  
    DATE ( YEAR ( MAX ( 'FTQ_Prodej' [Datum objednávky] ) ), 12, 31 ) ),  
    "DateAsInteger", FORMAT ( [Date], "YYYYMMDD" ),  
    "Rok", YEAR ( [Date] ), "Měsíc číslo", FORMAT ( [Date], "MM" ),
```

```

"Rok/měsíc", FORMAT ( [Date], "yyyy/mm" ),
"Měsíc v Roce", FORMAT ( [Date], "yyyy/mmm" ),
"Měsíc Zkratka", FORMAT ( [Date], "mmm" ),
"Měsíc", FORMAT ( [Date], "mmm" ),
"Den v týdnu", WEEKDAY ( [Date],2),
"Den", FORMAT ( [Date], "dddd" ),
"Čtvrtletí", "Q" & FORMAT ( [Date], "Q" ),
"Čtvrtletí v roce", FORMAT ( [Date], "YYYY" ) & "/" & "Q" & FORMAT ( [Date], "Q"
) )

```

Funkce **ADDCOLUMNS** přidá sloupce s uvedenými názvy a vygeneruje tolik řádků, kolik je dnů mezi 1. lednem nejmenšího roku a 31. prosincem nejvyššího roku, určeného datem objednávky v tabulce *FTQ_Prodej*, jak je definováno funkcí *CALENDAR*. **Minimální datum** definuje výraz:



```
DATE ( YEAR ( MIN ( 'FTQ_Prodej'[Datum objednávky] ) ), 1, 1 )
```

maximální datum pak výraz:



```
DATE ( YEAR ( MAX ( 'FTQ_Prodej'[Datum objednávky] ) ), 12, 31 )
```

Den (pořadí) v týdnu definuje výraz (argument 2 říká, že má týden začínat pondělkem, pokud argument chybí nebo je = 1, týden začíná nedělí):



```
"Den_v_tydnu", WEEKDAY ( [Date],2)
```

6.3 Ukazatelé pro pracovní dny

Častým případem funkcí s časovou dimenzí, jsou **výpočty ukazatelů, rozlišujících pracovní dny a svátky** (např. objem prodeje v pracovních dnech oproti svátkům apod.). K tomu je třeba **identifikovat v tabulce** zda jde o svátek či pracovní den, kde sloupec *Cas_Typ_Id* rozlišuje pracovní dny jako „1“ a svátky „0“.

Jako příklad požadujeme **zjistit průměrné denní počty prodaných kusů zboží v pracovních dnech**. Výpočetní předpis pro míru pro tento účel bude mít následující tvar:



```
=SUM (FTQ_Prodej [Prodej_Skut_Ks]) / SUM (DI_Cas_DAX [Cas_Typ_Id]),
```

kde souhrnné hodnoty prodaných kusů (*FTQ_Prodej [Prodej_Skut_Ks]*) se dělí počtem pracovních dnů v roce a v měsíci (*DI_Cas_DAX [Cas_Typ_Id]*). Výsledkem je definování míry „Prodeje – pracovní dny“ a tabulka s průměrnými denními prodejmi, viz Obrázek 6-1:

Popisky řádků	Součet Prodej_Skut_Ks	Prodeje - pracovní dny
2014	5045	20
Leden	482	22
Únor	427	21
Březen	571	27
Duben	486	23
Květen	518	27
Červen	490	23
Červenec	409	18
Srpen	341	16
Září	343	16
Říjen	344	15
Listopad	310	16
Prosinec	324	14
2015	3982	15
Leden	329	15
Únor	266	13
Březen	397	18
Duben	306	14
Květen	331	16

Obrázek 6-1: Průměrné prodeje v pracovních dnech podle měsíců a roků

6.4 Funkce time intelligence – YTD, QTD, MTD

Často je v analýzách různých trendů a sezónních výkyvů nutné **sledovat hodnoty ukazatelů v průběhu času**. Jednou z nejčastějších kalkulací jsou ty, které jsou označeny jako **YTD** (year-to-date) zobrazující **postupný nárůst hodnot od aktuálního data k začátku roku**. Obdobný princip platí pro kalkulace **QTD** (quarter-to-date) pro nárůst hodnot k začátku kvartálu a **MTD** (month-to-date) pro nárůst hodnot k začátku měsíce.

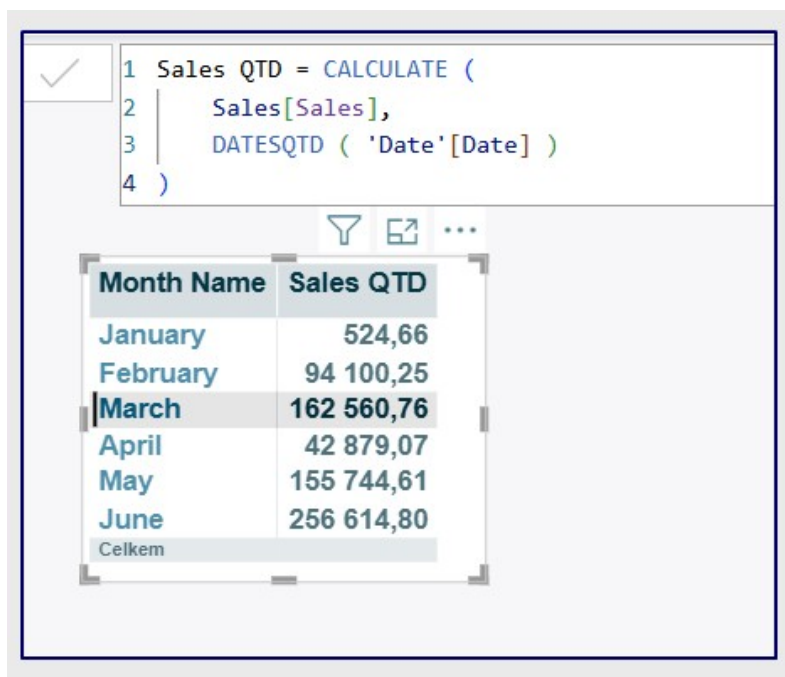
Příklad, dokumentující funkci YTD, představuje Obrázek 6-2

1	Sales YTD = CALCULATE(
2	Sales[Sales],
3	DATESYTD(
4	'Date'[Date]
5	)
6	)

Month Name	Sales YTD
January	524,66
February	94 100,25
March	162 560,76
April	205 439,83
May	318 305,37
June	419 175,56
July	419 175,56
August	419 175,56
September	419 175,56
October	419 175,56
November	419 175,56
December	419 175,56
Celkem	419 175,56

Obrázek 6-2: Uplatnění funkce year-to-date

Další příklad dokumentuje funkci QTD, Obrázek 6-3



Obrázek 6-3: Příklad funkce QTD

Dále předpokládejme požadavek na **sledování nárůstu nákladů pro kvartály a měsíce** vzhledem k začátku roku. K tomu se využívá funkce **TOTALYTD** a výpočetní předpis pro novou míru je následující, viz Obrázek 6-4:

 `=TOTALYTD (SUM (FTQ_Prodej [Naklady] ), DI_Cas_DAX [FullDate] )`

Popisky řádků	Součet Prodej_Skut_Ks	Prodeje - pracovní dny	Naklady YTD
2014	5045	20	1984264,3
Leden	482	22	289420,8
Únor	427	21	449749,5
Březen	571	27	615059,7
Duben	486	23	773850,9
Květen	518	27	986188
Červen	490	23	1218019,7
Červenec	409	18	1421592,3
Srpen	341	16	1540746
Září	343	16	1648124,5
Říjen	344	15	1740276,9
Listopad	310	16	1822128,2
Prosinec	324	14	1984264,3
2015	3982	15	1704670,8
Leden	329	15	161702,1
Únor	266	13	327123,5
Březen	397	18	464421,4
Duben	306	14	575846,1
Květen	331	16	756987,3

Obrázek 6-4: Objem nákladů k počátku roku (YTD)

V některých případech je třeba sledovat **nárůst hodnot nikoli ke standardnímu začátku roku, ale např. k začátku fiskálního roku** a jeho termín tak musíme doplnit do výpočetního předpisu míry, např.:

 `=TOTALYTD (SUM (FTQ_Prodej [Naklady] ), DI_Cas_DAX [FullDate], „30-06“ )`

Kromě funkce *TOTALYTD* nabízí DAX v této souvislosti i **řadu dalších užitečných funkcí**, jako jsou *STARTOFYEAR*, *ENDOFYEAR*, *PREVIOUSYEAR*, *NEXTYEAR*, *DATESYTD*, *OPENINGBALANCEYEAR*, *CLOSINGBALANCEYEAR*.

U většiny těchto funkcí ponecháváme s ohledem na rozsah na čtenáři, aby si je vyzkoušel i s pomocí funkcí nápovědy.

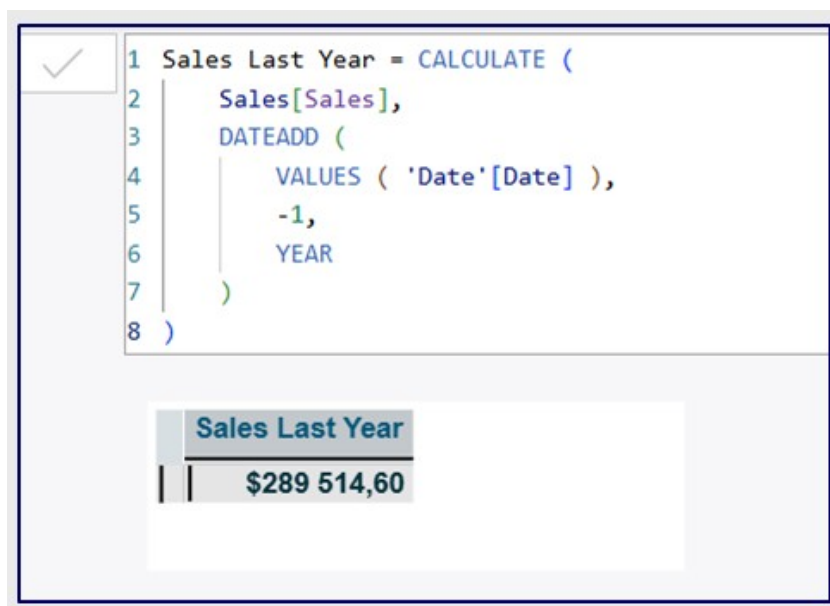
Je účelné ještě doplnit, že realizaci funkce *TOTALYTD* lze řešit **prostřednictvím univerzální funkce CALCULATE()**. V tomto případě by adekvátní zápis pro výše uvedenou míru vypadal následujícím způsobem:

 `= CALCULATE (SUM (FTQ_Prodej [Naklady] ), DATESYTD (DI_Cas_DAX [FullDate] ) )`

Funkce DATESYTD vrací všechna data z tabulky kalendáře, resp. časové dimenze, a to **na základě aktuálně nastaveného filtr kontextu**. To znamená, že funkce *CALCULATE()* zpracovává data pro YTD s respektováním aktuálně nastaveného filtru.

6.5 Sledování hodnot za minulé roky

Podnikové reporty a analytické tabulky často vyžadují hodnoty ukazatelů jak za aktuální období, resp. rok, tak **za odpovídající období v minulých letech** pro účely meziročních porovnání, vývojových trendů apod. Příklad tržeb za minulý rok prezentuje Obrázek 6-5:



Obrázek 6-5: Objem tržeb za minulý rok

V dalším případě požadujeme takové informace za minulá období pro sledování podnikových nákladů. Předpis pro míru bude např. vypadat takto, viz Obrázek 6-6:



=CALCULATE (SUM (FTQ_Prodej [Naklady]), SAMEPERIODLASTYEAR (DI_Cas_DAX [FullDate]))

Popisky řádků	Součet Prodej_Skut_Ks	Naklady YTD	Naklady minulý rok
2014	5045	1984264,3	
Leden	482	289420,8	
Únor	427	449749,5	
Březen	571	615059,7	
Duben	486	773850,9	
Květen	518	986188	
Červen	490	1218019,7	
Červenec	409	1421592,3	
Srpen	341	1540746	
Září	343	1648124,5	
Říjen	344	1740276,9	
Listopad	310	1822128,2	
Prosinec	324	1984264,3	
2015	3982	1704670,8	1984264,3
Leden	329	161702,1	289420,8
Únor	266	327123,5	160328,7
Březen	397	464421,4	165310,2
Duben	306	575846,1	158791,2
Květen	331	756987,3	212227,1

Obrázek 6-6: Náklady v aktuálním a minulém roce

Funkce *CALCULATE()* změní filtr s použitím funkce *SAMEPERIODLASTYEAR()*, která **bude vracet soubor datumů z předchozího roku**. Tato funkce je specifickou verzí obecnější funkce *DATEADD()*, která navíc využívá počtu a typu období, která se mají zpětně sledovat. Ekvivalentní zápis k předchozímu bude vypadat takto:



=CALCULATE (SUM (FTQ_Prodej [Naklady]),
DATEADD (DI_Cas_DAX [FullDate], -1, YEAR))

V některých případech je ale **třeba porovnávat pouze celkovou hodnotu za předchozí rok**, nikoli jednotlivé dílčí, tedy za kvartály nebo měsíce. K tomu slouží funkce *PARALLELPERIOD()*, která je obdobná funkci *DATEADD()*, ale vrací celé období, specifikované třetím parametrem, namísto jednotlivých dílčích období. To znamená, že následující míra bude vždy obsahovat celkový objem nákladů za určený minulý rok:



=CALCULATE (SUM (FTQ_Prodej [Naklady]),
PARALLELPERIOD (DI_Cas_DAX [FullDate], -1, YEAR))

6.6 Klouzavé ukazatele

Častou jsou součástí podnikových reportů a analýz klouzavé ukazatele. Příkladem je **klouzavý roční souhrn (moving annual total, MAT)**, který sleduje souhrnné hodnoty za posledních 12 měsíců (označuje se také *LTM, last twelve month*). Např. MAT pro leden 2017 bude dán souhrnem hodnot od února roku 2016 do ledna roku 2017 apod. Příklad takové funkce dokumentuje následující zápis:



```

Klouzavy_souhrn_hodnot =CALCULATE (
    SUM (FTQ_Prodej [Naklady]),
    DATESBETWEEN
        (DI_Cas_DAX [FullDate],
        NEXTDAY (SAMEPERIODLASTYEAR
            (LASTDATE (DI_Cas_DAX [FullDate] ))),
        LASTDATE(DI_Cas_DAX [FullDate] )
    )
)

```

V rámci tohoto předpisu je použita **funkce DATESBETWEEN**, která **vrací data ze sloupce, uvedeného mezi dvěma uvedenými daty**, a to mezi prvním a posledním dnem požadovaného intervalu. Poslední den se získá na základě funkce **LASTDATE**, která vrácí poslední datum ze zadaného sloupce (*FullDate*) a vždy přitom respektuje aktuální filtr kontext. Od tohoto data se získá první den intervalu požadavkem na následující den funkcí **NEXTDAY** z odpovídajícího posledního data, o rok předtím získaného funkcí **SAMEPERIODLASTYEAR**, viz Obrázek 6-7:

Popisky řádků	Naklady minuly rok	Klouzavy souhrn hodnot
2014		1984264,3
Leden		289420,8
Únor		449749,5
Březen		615059,7
Duben		773850,9
Květen		986188
Červen		1218019,7
Červenec		1421592,3
Srpen		1540746
Září		1648124,5
Říjen		1740276,9
Listopad		1822128,2
Prosinec		1984264,3
2015	1984264,3	1704670,8
Leden	289420,8	1856545,6
Únor	160328,7	1861638,3
Březen	165310,2	1833626
Duben	158791,2	1786259,5
Květen	212337,1	1755063,6

Obrázek 6-7: Klouzavý souhrn hodnot

7. Závěr

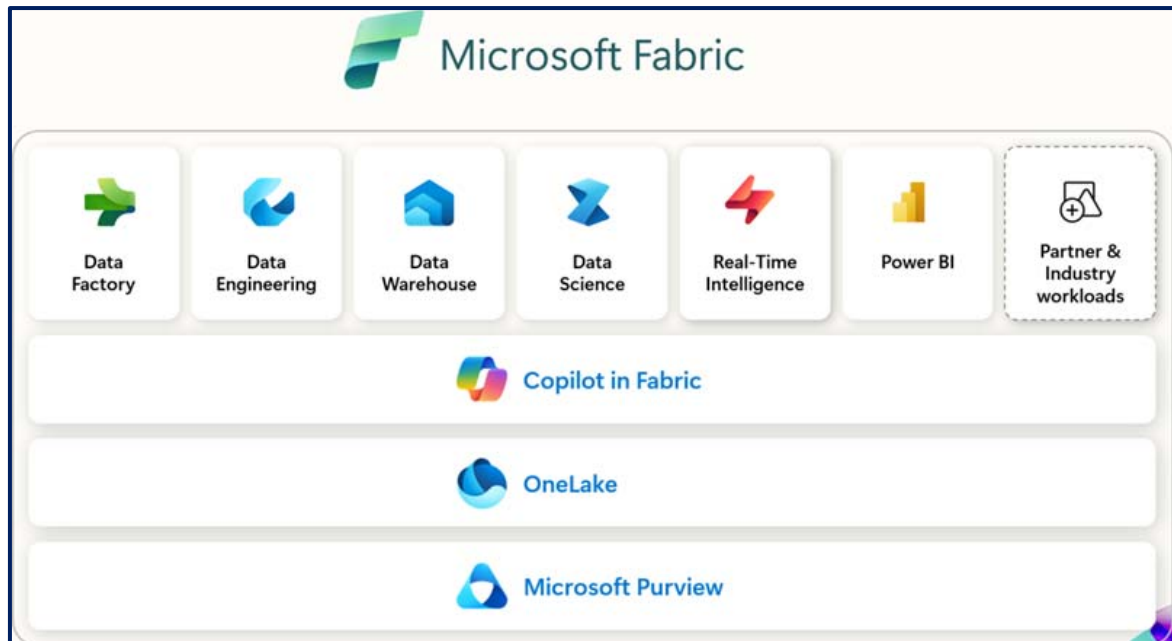
Jazyk DAX je pro řešení a využití analytických úloh velmi důležitý prostředek. Vztahuje se jak na úlohy SSBI, zejména ve vztahu na Power BI, tak i na řešení komplexních BI úloh, např. v prostředí MS SQL Serveru. Je velmi účelné, pokud analytik principy a jednotlivé funkce a možnosti ovládá, neboť tak ve větším rozsahu může připravovat zadání analytických úloh, které budou přesněji odpovídat potřebám jednotlivých sfér řízení firmy. Např. s využitím funkcionality Time Intelligence lze sledovat vývojové trendy v obchodu, marketingu, finančních zdrojích atd.

8. Zdroje

- ASPIN, A., 2016: *Pro Power BI Desktop*. Apress. Staffordshire. 2016. ISBN 978-1-4842-1804-4.
- BI4DYNAMICS, 2017. Sales – Top 30 customer table Report. (online). (29. 08. 2017). Dostupné z: <http://www.bi4dynamics.com/business-intelligence-for-microsoft-dynamics-nav/content/>
- BSC Designer, 2017. Sales Business Unit Scorecard. online. Strategy Maps and KPIs. (online). (29. 08. 2017). Dostupné z: <https://www.webbsc.com/s/sales-kpis>.
- CANVASJS, 2013. JavaScript Range Column & Range Bar Charts. (online). canvasjs.com. Dostupné z: <https://canvasjs.com/javascript-range-column-range-bar-chart/>.
- CIMLER, P., ZADRAŽILOVÁ, D. a kol., 2007: *Retail management*. Praha, Management Press, 2007. ISBN: 978-80-7261-167-6.
- COLLIE, R., SINGH, A., 2016: *Power Pivot and Power BI*, Holy Macro Books, 2016.
- CZSO, 2016. Tab. 02.05 Investice na ochranu životního prostředí (1989-2015). (online). czso.cz. Vydáváme. Česká republika od roku 1989 v číslech – 2016. <https://www.czso.cz/csu/czso/ceska-republika-od-roku-1989-v-cislech-w0i9dxmghn#03>
- CZSO, 2017a. Česká republika: hlavní makroekonomické ukazatele. Vydáváme. Časové řady. czso.cz. (online). (03. 07. 2017). Dostupné z: https://www.czso.cz/csu/czso/hmu_cr
- CZSO, 2017b. Graf 3 Ceny bytů - ČR (index, 2010 = 100). Ceny bytů. Vydáváme. czso.cz. (online). (07. 07. 2017). Dostupné z: https://www.czso.cz/csu/czso/ceny_bytu
- CZSO, 2017c. Tab. 7 Stravování a pohostinství (CZ-NACE 56). Vydáváme. Obchod, pohostinství, ubytování - časové řady - Základní finanční ukazatele - čtvrtletní - Klasifikace NACE Rev. 2 (CZ-NACE) (online). www.czso.cz. (13. 08. 2017). Dostupné z: https://www.czso.cz/csu/czso/1-malzfu_b
- CZSO, 2017d. Peněžní vydání domácností podle počtu vyživovaných dětí. Veřejná databáze. czso.cz. (online). (13. 08. 2017). Dostupné z: https://vdb.czso.cz/vdbvo2/faces/index.jsf?page=vystup-objekt&pvo=ZUR10&z=T&f=TABULKA&katalog=30847&c=v3-8__RP2011&&str=v389
- Denver.edu, 2017. Line Graph, Bar Graph, Pie Chart and Scatter Plot. University of Denver. (online). (13. 08. 2017). Dostupné z: <http://www.du.edu/ifs/help/use-online/repeated/general/graph/linebar.html>
- ECKERSON, W., 2006. Deploying Dashboards and Scorecards [online]. (29. 08. 2017). Dostupné z: http://www.businessobjects.com/pdf/products/performancemanagement/wp_tdw_deploying_dashboards_and_scorecards.pdf
- ECKERSON, W. 2010. *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*, 2. vydání. Wiley. 336 stran. ISBN: 978-0-470-58983-0.
- ECKERSON, W., W., 2006: *Performance Dashboards*. New Jersey, John Wiley & Sons 2006.
- ENGLISH, L., P., 2003: *Improving Data Warehouse and Business Information Quality: Methods for reducing costs and increasing profits*. New York, John Wiley & Sons 2003. ISBN 0-471-25383-9.
- FEW, S., 2006. *Information Dashboard Design: The Effective Visual Communication of Data*. Sebastopol, O'Reilly Media. ISBN 978-0-596-10016-2.
- FEW, S., 2012. *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. 2nd edition. Burlingame, AnalyticsPress. ISBN 978-0-970-60197-1.
- FEW, S., 2013. *Information Dashboard Design: Displaying Data for At-a-Glance Monitoring*. 2nd edition. Burlingame, AnalyticsPress. ISBN 978-1-938-37700-6.
- GALLAGHER, J., 2015. Antibiotic surge revealed by seasonal maps. bbc.com. News. (online). (03. 07. 2017). Dostupné z: <http://www.bbc.com/news/health-34790038>
- GAPMINDER, 2017. Life expectancy, years. (online). Gapminder.org. (13. 08. 2017). Dostupné z: http://www.gapminder.org/tools/#_chart-type=bubbles.
- HARINATH, S., PIHLGREN, R., LEE, D., G., SIRMON, J., BRUCKNER, R., M., 2012: *Microsoft SQL Server 2012. Analysis Services with MDX and DAX*. John Wiley and Sons, Inc., Indianapolis, 2012. ISBN 978-1-118-10110-0.

- HURSMAN, A., 2010. Effective-dashboard-design-why-your-baby-is-ugly. (online). (13. 08. 2017). Dostupné z: <https://www.slideshare.net/hurzman/effective-dashboard-design-why-your-baby-is-ugly>
- IMHOFF, C., WHITE, C., 2011: Self-Service Business Intelligence: Empowering Users to Generate Insights. Renton, WA : The Data Warehousing InstituteTM, 2011.
- INETSOFT.com, 2017. Information about Scorecards and Scorecard Examples. (online). (29. 08. 2017). Dostupné z: https://www.inetsoft.com/info/information_about_scorecards_and_scorecard_examples/
- INMON, B., 2002: *Building the Data Warehouse*. Indianapolis, John Wiley and Sons, 2002.
- JOTHIGANESH, S., 2017. 3Dscatterplot. (online). (13. 08. 2017). Dostupné z: <http://www.jothiganesh.com/category/technical/>
- KIMBALL, R., ROSS, M., 2010: Relentlessly Practical Tools for Data Warehousing and Business Intelligence. Wiley Publishing, Inc. 2010. ISBN 978-0-470-56310-6.
- KIMBALL, R., CASERTA, J., 2004: *The Data Warehouse ETL Toolkit*. Indianapolis, John Wiley and Sons 2004. ISBN 0-764-56757-8.
- KIMBALL, R., ROSS, M., 2002: *The Data Warehouse Toolkit, The Complete Guide to Dimensional Modelling*. Boston, John Wiley 2002. ISBN 0-471-20024-7.
- MICROSOFT, 2013: Power Pivot: Výkonné analýzy a modelování dat v Excelu. MICROSOFT. Office - Office.com [online]. 2013 [cit. 2014-01-02]. Dostupné z: <http://office.microsoft.com/cs-cz/excel-help/power-pivot-vykonne-analyzy-a-modelovani-dat-v-excelu-HA102837110.aspx>
- OFFICE 1: Typy funkcí jazyka DAX. Office.com [online]. 2013 [cit. 2014-03-14]. Dostupné z: <http://office.microsoft.com/cs-cz/excel-help/typy-funkci-jazyka-dax-HA102836089.aspx>
- OFFICE 2: Perspektivy v Power Pivotu. Office.com [online]. 2013 [cit. 2014-03-14]. Dostupné z: <http://office.microsoft.com/cs-cz/excel-help/perspektivy-v-power-pivotu-HA102837427.aspx>
- OFFICE 3: Filtrování a zvýrazňování v Power View. Office.com [online]. 2013 [cit. 2014-03-22]. Dostupné z: <http://office.microsoft.com/cs-cz/excel-help/filtrovani-a-zvyraznovani-v-power-view-HA102834776.aspx>
- KELLE ONEAL, BEYENETWORK, 2012, on-line: http://www.b-eye-network.com/blogs/oneal/archives/2012/02/what_is_the_dif.php
- PROVOST, F., FAWCETT, T., 2013: *Data Science for Business. What You Need to Know About Data Mining and Data-Analytic Thinking*. O'Reilly Media. Sebastopol. 2013. ISBN: 978-1-449-36132-7.
- RUSSO, M., FERRARI, A., 2013: *Microsoft Excel 2013: Building Data Models with PowerPivot*. California, O'Reilly Media, Inc., 2013. ISBN: 978-0-7356-7634-3.
- RUSSO, M., FERRARI, A., 2011: *PowerPivot for Excel 2010. Give Your Data Meaning*. Redmond, Microsoft Press, 2011. ISBN: 978-0-7356-5058-0.
- SEAMARK, P.: *Beginning DAX with Power BI*. Apress, 2018. ISBN: 978-1-4842-3477-8.
- SIEGEL, E., 2016: *Predictive Analytics, The power to predict who will click, buy, lie, or die*. John Wiley&Sons. New Jersey. 2016. ISBN: 978-1-119-14567-7.
- SPOFFORD, G., 2001: *MDX Solutions with Microsoft SQL Server Analysis Services*. New York, John Wiley, 2001. ISBN: 0-471-40046-7.
- TANKERSLY, B., 2017. Using Dashboards For a Real-Time View to Productivity. Intuit QuickBooks. (online). (29. 08. 2017). Dostupné z: <https://www.firmofthefuture.com/content/using-dashboards-for-a-real-time-view-to-productivity/>
- TURLEY, P., BRUCKNER, B., SILVA, T., WITHEE, K., PAISLEY, G., 2012: *Microsoft SQL Server 2012. Reporting Services*. John Wiley and Sons, Inc., Indianapolis, 2012. ISBN 978-1-118-10111-7.
- WEXLER, S., SHAFFER, J., COTGREAVE, A., 2017: *The Big Book of Dashboards*. John Wiley&Sons. New Jersey. 2017. ISBN: 978-1-119-28271-6.

C. Microsoft Fabric



Tento oddíl je věnovaný přehledu platformy Microsoft Fabric, další informace jsou **dostupné např. na této stránce:**

Microsoft Fabric: [Co je Microsoft Fabric – Microsoft Fabric | Microsoft Learn]



1. Úvod

Společnost Microsoft v roce 2023 představila novou platformu **Microsoft Fabric, sloužící pro end-to-end datově analytická řešení** jako reakci na stále zvyšující se nároky na komplexitu a různorodost v datové sféře (Zdroj: <https://learn.microsoft.com/en-us/fabric/fundamentals/microsoft-fabric-overview>).

U většiny firem v kontextu Business Intelligence zůstává základním kamenem klasický přístup **budování datových skladů spolu s reportingovým řešením**, který je vhodný pro práci **se strukturovanými daty** a pomáhá tradičně businessu při plánovacích, rozhodovacích a optimalizačních činnostech. Avšak jedním z hlavních trendů poslední doby je například **analýza nestrukturovaných dat**, která vyžaduje jiné přístupy už k samotnému ukládání a celkovému zpracování pro smysluplné výstupy (Antunes et al., 2022).

Business analytika může být buď deskriptivní, diagnostická, prediktivní nebo preskriptivní. Každá úroveň, stejně jako forma dat, vyžaduje jiné přístupy, znalosti a technologie pro její ovládnutí. Například **pokročilé úrovně analytiky** typicky vyžadují datovou architekturu a technologie, které zvládnou ukládat a zpracovávat velké množství dat různé formy a umožní aplikovat **algoritmy strojového učení a umělé inteligence**. Všechny úrovně analytiky a s nimi spojené přístupy mají však z pohledu firem jeden společný cíl – ovlivňovat pozitivním způsobem kvalitu a výkonnost firem (Mallam et al. 2021).

Výše zmíněné situace často vyžadují **více nástrojů a technologií** od různých dodavatelů, což má za důsledek například duplicitní uložení týchž dat, rozdílné licenční podmínky, odlišné ekosystémy a především **vysoké nároky** (znalostní, časové, finanční) **na samotnou integraci** všech komponent. Tuto situaci se snaží řešit právě **Microsoft Fabric** prostřednictvím unifikované analytické platformy, pokrývající **veškeré datové potřeby firem od úvodní integrace a uložení dat až po samotnou prezentační vrstvu** spolu s celkovou správou (*data governance*). Platforma nebyla vytvořena na zelené louce, ale tvoří ji například známé produkty Microsoftu pro práci s daty jako **Azure Data Factory, Azure Synapse Analytics anebo Power BI**. Hlavní výhody této platformy představují široké low-code i no-code možnosti integrace rozmanitých datových zdrojů, podpora pokročilé analytiky, škálovatelnost a flexibilita pro aktuální potřebu, podpora umělé inteligence a celková unifikace všech komponent do jednoho prostředí (Anon, 2023).

Fabric umožňuje firmám i jednotlivcům efektivně pracovat s daty napříč celým životním cyklem – od sběru a transformace, přes ukládání a správu, až po pokročilou analytiku, strojové učení a reporting. Primárním cílem je zjednodušit datovou architekturu, odstranit datová sila a umožnit rychlou adopci nových technologií včetně AI.

Hlavní zdroj: KUTHAN, L. IT a anatomie firmy (Microsoft Fabric). Pracovní dokument, dostupný na <https://mbi-af.cz>, 2024.

Poznámka: S ohledem na přesnost výrazů, spojených s jednotlivými komponentami Microsoft Fabric, je zde ponecháváme v anglické verzi.

2. Microsoft Fabric

Microsoft Fabric dokáže pokrýt potřeby všech datových rolí od datového inženýra a vědce až po datového / business analytika spolu se samotnými business uživateli a konzumenty datové analytických výstupů (Ghoshal, 2023). Microsoft Fabric je tedy možné definovat, jako moderní cloudovou platformu pro komplexní datovou analytiku, která sjednocuje nástroje pro integraci, správu, zpracování, analýzu a vizualizaci dat v rámci jednoho prostředí. Od svého uvedení v roce 2023 prošla platforma rychlým vývojem a v průběhu let **2024–2025** se stala **standardem pro datové projekty** v prostředí Microsoftu.

Microsoft Fabric je **tvořen jednotlivými komponentami**, které mají vždy specifické zaměření a účel. Jinými slovy **každá komponenta** nebo sada komponent je **určena pro jinou úlohu** při práci s daty, respektive pro jinou roli datového charakteru. V základním pojetí je Microsoft Fabric **cloudová sada SaaS** nástrojů, která umožňuje zpracovávat velké množství různorodých dat s vysokou mírou škálovatelnosti. Celá platforma má za cíl pokrýt **požadavky jak vývojářských rolí, tak i ty ze strany business uživatelů**. Microsoft Fabric integruje vybrané nástroje a produkty z rodiny Azure a Power Platform spolu s jejich optimalizací, aby tvořily unifikovanou platformu a jedno místo pravdy pro zpracování, ukládání a prezentování podnikových dat (Bisson, 2023).

Unifikace a jednotnost nově vzniklé platformy by se dala popsat z více úhlů pohledu. Nalezneme jak technologické aspekty, tak i ty, které souvisejí již s výše zmiňovanou komplexitou pro různé role a scénáře. K těmto aspektům se řadí i **oblast Data Governance**, která se podílí na logice vybudování a užívání platformy Microsoft Fabric (Bisson, 2023).

Fabric je od poloviny roku 2024 dostupný v režimu GA (*General Availability*), což znamená plnou podporu, SLA a integraci do Microsoft 365 ekosystému.

2.1 Scénáře využití Microsoft Fabric

Využití SaaS Microsoft Fabric je typické pro následující scénáře datové analytiky:

- **End-to-end analytika:** Fabric umožňuje pokrýt celý datový životní cyklus – od sběru dat, přes jejich transformaci, ukládání, správu, až po reporting a pokročilou analytiku v jednom prostředí.
- **Self-service BI:** uživatelé z byznysu (např. controlling, marketing, obchod) mohou sami vytvářet reporty, analýzy a vizualizace bez závislosti na IT oddělení.
- **Pokročilá analytika a strojové učení (AI a ML):** Fabric umožňuje datovým vědcům a analytikům vyvíjet, trénovat a nasazovat modely strojového učení přímo nad daty v Lakehouse nebo Warehouse.
- **Data mesh, Data Governance a doménová architektura:** Fabric umožňuje zavést moderní správu dat (Data Governance) a podporuje doménovou architekturu (Data Mesh), kde jednotlivé týmy spravují svá data autonomně, ale v rámci centrální platformy.
- **Real-time analytics a monitoring:** Fabric podporuje zpracování a analýzu dat v reálném čase (*streaming*), což je klíčové pro IoT, výrobu, logistiku, bezpečnost nebo finanční monitoring.
- **Hybridní a multicloudová analytika:** Fabric umožňuje napojení na externí zdroje (Azure, AWS S3, *on-premises databáze*) a práci s daty napříč cloudy bez nutnosti kopírování (*Shortcuts*).

2.2 Architektura a komponenty Microsoft Fabric

Microsoft Fabric integruje několik komponent / oblastí, které **slouží specifickým úlohám či datovým rolím**. Každá komponenta pak nabízí možnost vytvářet různé artefakty a používat jiné technologie pro co možná nejlepší pokrytí daných potřeb. Hlavní komponenty Fabric jsou stručně popsány následujících subkapitolách.

2.2.1 OneLake

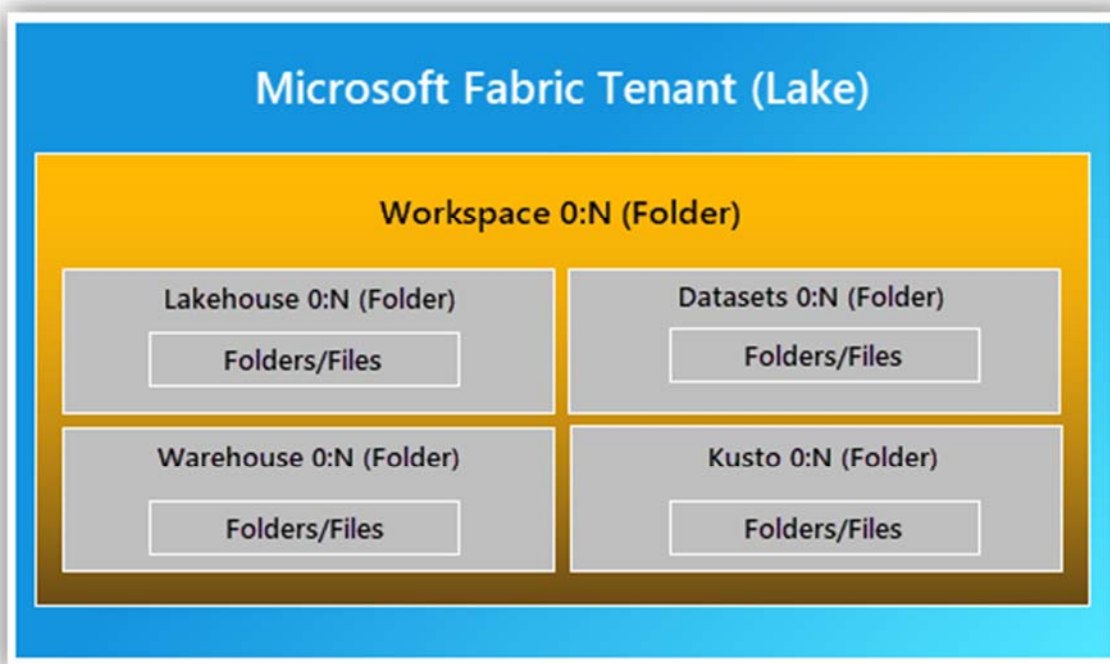
Microsoft Fabric by měl být standardně implementován **na úrovni jednoho Microsoft tenantu**, který typicky představuje jednu firmu či společnost. **Jednomu tenantu** pak přísluší vždy právě **jedno centrální úložiště všech dat – OneLake**. Jedná se o úložiště **typu data lake**, které ukládá různé formáty dat, podporuje Delta Lake formát a umožňuje přímý přístup z různých analytických nástrojů. Lze si ho představit jako „OneDrive pro data“.

OneLake je elegantním řešením v situacích, kdy firmy a společnosti typicky disponují několika datovými sily. Ta si lze představit jako data lake úložiště či relační databáze, které vznikají například na úrovni jednotlivých oddělení. Ve výsledku vzniká situace, která může být poměrně náročná na správu a řízení. Tato náročnost se stává realitou v situacích, kdy se používají různé nástroje, technologie a přístupy, které jednotlivá úložiště aplikují.

S tím často souvisí i specifické **požadavky na sdílení těchto dat** mezi jednotlivými odděleními (úložišti) nebo požadavky na celkové zabezpečení a správu dat. Při samotném sdílení dat dochází ke kopírování dat a tatáž data se ukládají na více místech najednou a zvyšují nároky na datovou infrastrukturu a potřeby datových kontrol (Bisson, 2023).

OneLake se snaží reagovat na výše zmíněné situace tím, že v **kontextu jednoho tenantu (firmy) existuje striktně právě jednou a vzniká tak centrální úložiště veškerých dat pro účely datové analýzy**. Systém OneLake má hierarchickou strukturu, což usnadňuje správu dat napříč celou organizací. Platforma Microsoft Fabric obsahuje OneLake automaticky, takže není potřeba žádná instalace. Každý tenant obsahuje jeden sjednocený OneLake s jednotným namespace souborového systému, který se rozprostírá napříč uživateli, geografickými oblastmi i cloudovými prostředími.

OneLake organizuje data do tzv. kontejnerů (Obrázek 2-1), což umožňuje jejich snadnou manipulaci a správu. Tenant představuje kořenovou úroveň OneLake a nachází se na vrcholu hierarchie. V rámci jednoho tenantu lze vytvořit více pracovních prostorů (*workspaces*), které fungují podobně jako složky. Následující schéma znázorňuje, jak Fabric ukládá data v rámci OneLake. Každý tenant může obsahovat několik pracovních prostorů, přičemž každý pracovní prostor může zahrnovat více lakehouse instancí. **Lakehouse představuje kolekci souborů, složek a tabulek**, které společně fungují jako databáze nad datovým jezerem (data lake) (Microsoft, 2025).



Obrázek 2-1: Ukládání dat ve Fabric OneLake. Zdroj: (Microsoft 2025)

Microsoft Fabric také přímo **nabízí doménové uspořádání architektury**. Pracovní prostory shlukují datové artefakty například podle jejich nadřazené domény či obecné funkcionality. Pracovní prostory poté poskytují možnost vlastní správy veškerého obsahu a řízení přístupu k němu. Tento koncept umožňuje následnou práci s daty, která už může být distribuována **do jednotlivých domén** a firmy tak mohou jít naproti **konceptu doménového vlastnictví dat**.

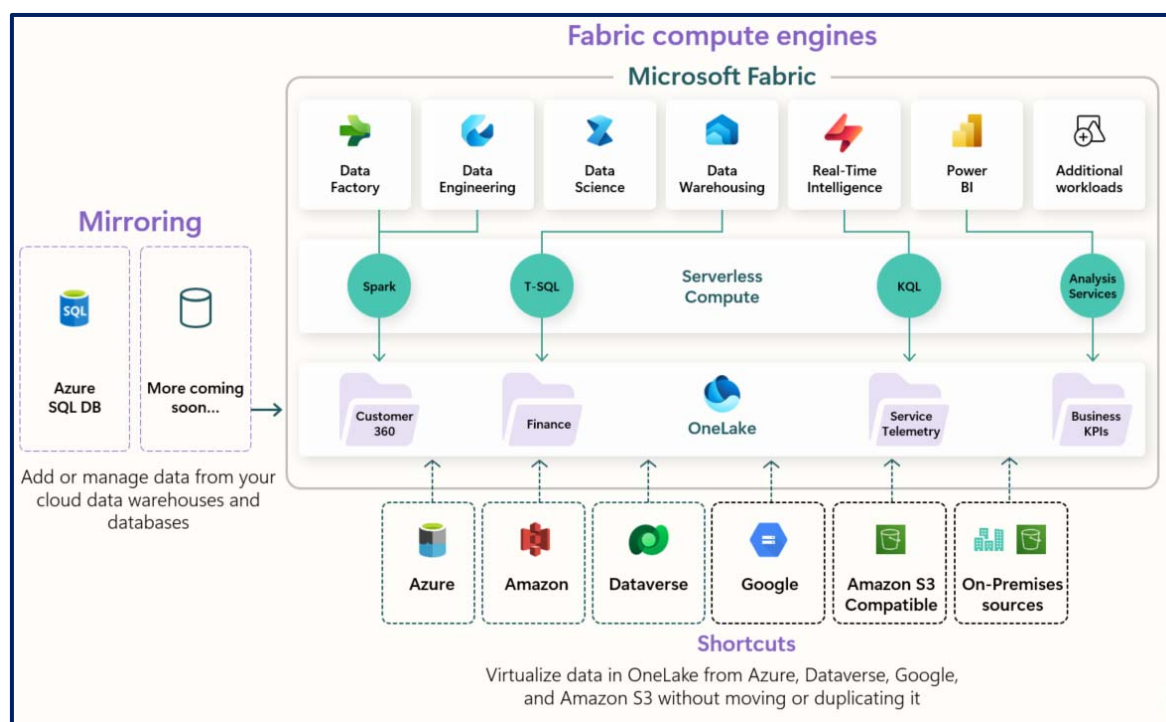
2.2.2 Data Factory

Komponenta Data Factory, známá již z prostředí Azure, nabízí více než **150 konektorů na datové zdroje pro extrakci dat** z různorodých datových zdrojů. Primárním účelem je **orchestrace ETL/ELT** procesů, tvorba datových pipelines spolu s mechanismy pro error handling. Data Factory umožňuje aplikovat **Spark notebooks**, využívající Apache Spark engine nebo jednodušší **low-code transformační kroky** a integrovat DataFlow gen2 pro Power Query transformace (Berge, 2023).

2.2.3 Synapse Data Engineering

V této komponentě je umožněno provádět složitější a více rozmanité transformace za použití **Spark Jobs a Notebooků, které využívají pro práci s daty například rozhraní Pythonu pro Apache Spark – PySpark**. Klíčovým artefaktem pro uložení strukturovaných a nestrukturovaných dat je **Lakehouse**, který kombinuje principy data lake a datového skladu. Zde je možné ukládat data jako jednotlivé soubory a převádět je poté do tabulkové struktury a formátu **Delta Lake**. **Lakehouse** také umožňuje využití SQL endpointu pro read-only operace spolu s možností vytváření view objektů (Berge, 2023).

Aby se však data od svého úvodního načtení dostala do požadované lokace a podoby, Microsoft Fabric podporuje **čtyři typy enginů**, které umožňují s daty pracovat. Podle Obrázek 2-2 se jedná o **Spark, T-SQL, KQL a Analysis Services**. Jejich použití závisí na požadovaném výstupu či roli, která s daty pracuje. Těmito enginy jsou pokryty nejenom potřeby pro klasickou analýzu dat pomocí jazyka SQL s výstupem v Power BI, ale i scénáře v oblastech strojového učení nebo real-time analytiky, které generují specifické požadavky na výkon a možnost práce s velkými objemy dat (Microsoft, 2025).



Obrázek 2-2: Operace s daty v Microsoft Fabric pomocí různých enginů a shorecuts.
Zdroj: (Microsoft, 2025)

Zmiňované enginy podporují práci s **Delta Parquet formátem**, který je zde nativním standardem pro uložení a práci s daty. Jinými slovy, pokud jsou data uvnitř platformy transformována, dřív nebo později se do tohoto formátu data dostanou. Typicky se tomu tak děje například při transformaci dat z originálních formátů, jako jsou například csv nebo json soubory, do optimální tabulkové struktury právě typu Delta **v kontextu Lakehouse a Warehouse** úložišť. Delta Parquet nebo také někdy Delta Lake představuje formát, který je tvořen parquet soubory s tím, že jsou k nim **přidány i transakční historie, důležitá metadata** (například o schématu dat) a **optimalizační mechanismus**, díky kterému dochází k významné komprimaci dat (Armbrust et al., 2020).

2.2.4 Synapse Data Science

Tato komponenta utváří prostředí pro vývoj, testování a **nasazování modelů strojového učení**. Umožňuje bezproblémovou integraci populárních knihoven pro strojové učení (ML) a s využitím enginu **Apache Spark** je optimalizován pro práci s velkými objemy dat. Nejčastěji se pro vývoj modelů využívá již zmíněný PySpark anebo SparkR (Berge, 2023).

2.2.5 Synapse Data Warehousing

V této komponentě Fabric implementuje plnohodnotné řešení cloudového datového skladu. Jedná se zde především o možnost podporující **DDL a DML dotazy jazyka T-SQL** nad hlavním artefaktem této komponenty – Warehouse, což není možné realizovat v Lakehouse artefaktu z komponenty **Synapse Data Engineering**. Jazykem pro zpracování a manipulaci s daty je zde **T-SQL**. Warehouse typicky umožňuje ukládat pouze plně strukturovaná data (Berge, 2023).

2.2.6 Synapse Real-Time Analytics

Synapse Real-Time Analytics řeší v Microsoft Fabricu **analýzu a zpracování streamovaných dat** s poskytnutím vhodných mechanismů pro optimalizaci a škálování těchto úloh. Klasickým scénářem zde může být zpracování dat, která jsou například vysokofrekvenčně **generována z IoT** čidel. Zásadním nástrojem je zde **jazyk KQL**, uzpůsobený pro **dotazování se na streamovaná data**.

2.2.7 Power BI a Direct lake

Power BI v kontextu Microsoft Fabric řešení figuruje primárně jako **vizualizační prvek pro reproting a analytiku**. S příchodem Microsoft Fabric se rozšířila nabídka možností pro připojení na data o **Direct Lake mód**, který je ve své podstatě přímo napojen na datové struktury platformy Microsoft Fabric a dokáže data přenášet do vizualizační části v téměř reálném čase. Na pozadí funguje optimalizační mechanismus, který loguje často dotazovaná data a načítá je stále rychleji. (Berge, 2023). Výhodu pro uživatele také představuje Copilot, který umožňuje dotazovat sémantické modely pomocí přirozeného jazyka včetně generování vizuálních výstupů.

2.2.8 Data Activator

Poslední komponentou je Data Activator, který slouží pro **automatizované úkony**, které se mají provést při splnění definovaných podmínek. Jedná se o **low-code nástroj**, který například umožňuje kontrolovat uživatelem určená data a při splnění jistých kritérií posílat **oznámení do Microsoft Teams** nebo Outlooku. Komponenta také nabízí **integraci Power Automate** pro vytváření vlastních mechanismů podmínek a následných úkonů (Berge, 2023).

2.3 Další koncepty Microsoft Fabric

2.3.1 Shortcuts

Velmi užitečným konceptem platformy Microsoft Fabric jsou takzvané „Shortcuts“. **Shortcuts** by se daly doslovně přeložit jako **zkratky**, ale výstižnější by bylo použít spíše slovo **odkaz nebo pointer**. Shortcuts řeší problém několikanásobného kopírování a ukládání těchto dat. Podle obrázku (Obrázek 2-2) je vidět jejich princip fungování. V zásadě umožňují získat přístup k datům bez nutnosti je přesouvat nebo kopírovat, což **šetří úložiště a redukuje počet operací** při manipulaci s daty. Shortcuts umožňují odkazovat se na data uvnitř prostředí Microsoft Fabric i mimo něj. První varianta tedy například umožňuje sdílení dat mezi doménami a jejich datovými sklady, neboli Warehouses (podle originálního názvosloví Microsoft Fabric). Druhá varianta umožňuje získávat data z externího prostředí, a to konkrétně z úložišť Azure Data Lake Storage Gen2 a Amazon S3 (Bisson, 2023).

2.3.2 Integrace umělé inteligence

Microsoft Fabric aspiruje na globální integraci umělé inteligence do valné většiny svých komponent. Skrze zabudování služeb **Azure OpenAI a Copilot** cílí na širší a rychlejší adopci celé platformy, kde je na základě přirozeného jazyka možné **vygenerovat potřebný kód pro transformaci dat** či celý ETL mechanismus, nebo vytvořit **kompletní Power BI report** (Ghoshal, 2023).

Od roku 2025 je Copilot v Microsoft Fabric již dostupný napříč všemi komponentami (Data Factory, Synapse, Power BI). Umožňuje generovat kód, transformace, DAX výrazy, SQL dotazy i celé reporty na základě přirozeného jazyka. AI je využívána i pro automatizaci datových pipeline, detekci anomálií a prediktivní analytiku.

2.3.3 Data Governance

Microsoft Fabric **pro oblast Data Governance vytváří velmi solidní podmínky**. Základem je **centrální úložiště dat**, které umožňuje jeho globální správu, jednotné technologické přístupy a snazší kontrolu oproti decentralizovaným celkům s rozdílnými specifiky. Centrální úložiště však reálně figuruje spíše na pozadí, protože vzniklé datové artefakty spadají do pracovních prostorů, se kterými přicházejí do styku jak vývojáři, tak konzumenti obsahu.

Pracovní prostory disponují vlastní správou a nastavením. Hierarchicky je důležité zmínit, že Microsoft Fabric zavádí i **domény neboli Domains**, které mohou shlukovat více pracovních prostorů a taktéž nabízí možnost správy a vlastního řízení dat. Z pohledu Data Governance tak vzniká **víceúrovňová správa**, která na té nejvyšší úrovni definuje naprostý základ, ale například vlastnictví dat je delegováno jednotlivým doménám.

Oblast kvality dat podporují výše definované Shortcuts, které redukuje počet potřebných kopií a extrakčních operací nad daty. Tím se **snižuje prostor pro vznik chyb** vedoucích k nekvalitě a zároveň se eliminuje situace, kdy více domén / oddělení potřebuje část stejných dat, která se tak musí několikrát namnožit a na výstupu už nemusí být vůbec stejná a validní podle původního zdroje poskytnutí, a to vzhledem k jiným transformacím a přístupům. Pro plnohodnotné zavedení Data Governance Microsoft doporučuje využít i **dedikovaný produkt Microsoft Purview** pro tuto oblast (Microsoft, 2024).

2.4 Opodstatnění použití Microsoft Fabric

Microsoft Fabric se využívá jako platforma pro vytvoření datové architektury a reportingových řešení z několika **hlavních důvodů**:

- Dosavadní architektura a reportingová řešení využívají **produkty z rodiny Microsoft Azure a Power Platform**, což znamená jednodušší integraci a implementaci do prostředí Microsoft Fabric.
- Jednoduchá **škálovatelnost** je nespornou výhodou této platformy. Bezproblémová škálovatelnost platformy eliminuje případné starosti s konfigurací potřebné infrastruktury.
- Jednotná **Data Governance** nad celým datovým spektrem je dalším validním důvodem pro volbu této platformy. Redukuje se tak počet platformů či systémů, které je třeba spravovat a efektivně řídit.
- Budování **best-practise přístupů** v ekosystému Microsoft Fabric představuje další důvod, který má za cíl prakticky zdokonalit firemní znalosti a zkušenosti s touto platformou. Tento fakt povede k možnosti nabízet kvalitnější, osvědčená a stabilní řešení. S tímto bodem souvisí i udržení kroku s nekonečným rozvojem informačních technologií na poli datových nástrojů a produktů.
- Posledním bodem je **soulad s doménovým pohledem na vlastnictví a správu** dat, kdy firma bude těžit z definované struktury domén uvnitř platformy Microsoft Fabric a bude tak možné reflektovat v této platformě odraz reality organizační struktury firmy.

3. Zdroje

KUTHAN, L. IT a anatomie firmy (Microsoft Fabric). Pracovní dokument dostupný na <https://mbi-af.cz>, 2024.

Anon., 2023. Build 2023: Microsoft launches Fabric, a new data and analytics platform [GADGETS NEWS]. *The Times of India* [online]. [vid. 2024-02-10]. Dostupné z: <https://www.proquest.com/docview/2817793506/citation/B9343615FA2D4168PQ/1>

ANTUNES, António Lorvão, Elsa CARDOSO a José BARATEIRO, 2022. Incorporation of Ontologies in Data Warehouse/Business Intelligence Systems – A Systematic Literature Review. *International Journal of Information Management Data Insights* [online]. 2(2), 100131. ISSN 2667-0968. Dostupné z: doi:10.1016/j.jjime.2022.100131

ARMBRUST, Michael, Tathagata DAS, Liwen SUN, Burak YAVUZ, Shixiong ZHU, Mukul MURTHY, Joseph TORRES, Herman VAN HOVELL, Adrian IONESCU, Alicja ŁUSZCZAK, Michał ŚWITAKOWSKI, Michał SZAFRAŃSKI, Xiao LI, Takuya UESHIN, Mostafa MOKHTAR, Peter BONCZ, Ali GHODSI, Sameer PARANJPYE, Pieter SENSTER, Reynold XIN a Matei ZAHARIA, 2020. Delta lake: high-performance ACID table storage over cloud object stores. *Proceedings of the VLDB Endowment* [online]. 13(12), 3411–3424. ISSN 2150-8097. Dostupné z: doi:10.14778/3415478.3415560

BERGE, Eirik, 2023. An Overview of Microsoft Fabric Going Into 2024. *Medium* [online] [vid. 2024-03-24]. Dostupné z: <https://towardsdatascience.com/an-overview-of-microsoft-fabric-going-into-2024-cd51a18c07b5>

BISSON, Simon, 2023. Understanding OneLake and lakehouses in Microsoft Fabric. *InfoWorld.com* [online]. [vid. 2024-03-24]. Dostupné z: <https://www.proquest.com/docview/2848449805/abstract/6DCAD662D62A4DA1PQ/1?sourcetype=Trade%20Journals>

BUTTE, Vijay Kumar a Sujata BUTTE, 2022. Enterprise Data Strategy: A Decentralized Data Mesh Approach. In: *2022 International Conference on Data Analytics for Business and Industry (ICDABI): 2022 International Conference on Data Analytics for Business and Industry (ICDABI)* [online]. s. 62–66 [vid. 2024-03-26]. Dostupné z: doi:10.1109/ICDABI56818.2022.10041672

CIRKOVSKÝ, 2018. Co je to strategické řízení nákupu? – I. část. *Benefico* [online] [vid. 2024-02-24]. Dostupné z: <https://benefico.cz/co-je-to-strategicke-rizeni-nakupu-i/>

GARTNER a GUPTA, 2021. Using Data Fabric Architecture to Modernize Data Integration. *Gartner* [online] [vid. 2024-03-26]. Dostupné z: <https://www.gartner.com/smarterwithgartner/data-fabric-architecture-is-key-to-modernizing-data-management-and-integration>

GHOSHAL, Anirban, 2023. Why Microsoft is combining all its data analytics products into Fabric. *InfoWorld.com* [online]. Dostupné z: <http://ezproxy.techlib.cz/login?url=https://www.proquest.com/trade-journals/why-microsoft-is-combining-all-data-analytics/docview/2817696905/se-2?accountid=119841>

HAELLEN, Bennie a Dan DAVIS, 2023. *Delta Lake: Up and Running*. B.m.: O'Reilly Media, Inc. ISBN 978-1-09-813968-1.

HALKJÆR, Mathias, 2023. *Microsoft Fabric deep dive – Data Factory* [online]. 2023. Dostupné z: <https://www.fellowmind.com/globalassets/denmark/downloadable-content/microsoft-fabric-deep-dive---data-factory-v.2.pdf>

HARIHARAN, 2023. *Adopt Data Mesh in Microsoft Fabric – Hari BI Academy* [online] [vid. 2024-03-29]. Dostupné z: <https://haribiacademy.com/2023/07/adopt-data-mesh-in-microsoft-fabric/>

HROUDA, Adam, 2023. Import SharePoint data to Fabric. *Adam's Bubble* [online] [vid. 2024-04-08]. Dostupné z: <https://adamsbubble.com//posts/import-sharepoint-data-to-fabric>

HRŮZA, Tomáš, 2007. *Jak poznat vyspělé procesy?* [online] [vid. 2024-02-25]. Dostupné z: <https://www.systemonline.cz/sprava-it/cmmi-model-hodnoceni-vyspelosti-procesu-1.htm>

<https://sqldusty.com/2019/02/15/power-bi-architecture-diagram-v4-is-now-available/>

ILIJASON, Robert, 2020. *Beginning Apache Spark Using Azure Databricks: Unleashing Large Cluster Analytics in the Cloud*. B.m.: Apress. ISBN 978-1-4842-5781-4.

- JOHNSON, Aimee, 2023. Data Lakehouse vs Data Warehouse in Microsoft Fabric. *Advancing Analytics* [online] [vid. 2024-04-06]. Dostupné z: <https://www.advancinganalytics.co.uk/blog/2023/6/13/data-lakehouse-vs-data-warehouse-in-microsoft-fabric>
- KLOPFENSTEIN, 2024. *Data Factory in Microsoft Fabric documentation – Microsoft Fabric* [online] [vid. 2024-04-06]. Dostupné z: <https://learn.microsoft.com/en-us/fabric/data-factory/>
- LAPLANTE, A. a an O'Reilly Media Company SAFARI, 2020. *The Modern Cloud Data Platform* [online]. B.m.: O'Reilly Media, Incorporated. Dostupné z: <https://books.google.cz/books?id=0D1GzgEACAAJ>
- L'ESTEVE, Ron, 2022. *The Azure Data Lakehouse Toolkit: Building and Scaling Data Lakehouses on Azure with Delta Lake, Apache Spark, Databricks, Synapse Analytics, and Snowflake* [online]. Berkeley, CA: Apress [vid. 2024-04-02]. ISBN 978-1-4842-8232-8. Dostupné z: doi:10.1007/978-1-4842-8233-5
- MALLAM, Pooja, Ashu ASHU a Baljeet SINGH, 2021. Business Intelligence Techniques Using Data Analytics: An Overview. In: *2021 International Conference on Computing Sciences (ICCS): 2021 International Conference on Computing Sciences (ICCS)* [online]. Phagwara, India: IEEE, s. 265–267 [vid. 2024-02-10]. ISBN 978-1-66549-445-8. Dostupné z: doi:10.1109/ICCS54944.2021.00059
- MAZUMDAR, Dipankar, Jason HUGHES a J. B. ONOFRE, 2023. *The Data Lakehouse: Data Warehousing and More* [online]. 12. říjen 2023. B.m.: arXiv. [vid. 2023-11-25]. Dostupné z: <http://arxiv.org/abs/2310.08697>
- MICROSOFT, 2023. *What is OneLake? - Microsoft Fabric* [online] [vid. 2024-03-24]. Dostupné z: <https://learn.microsoft.com/en-us/fabric/onelake/onelake-overview>
- MICROSOFT, 2024. *Governance and compliance in Microsoft Fabric – Microsoft Fabric* [online] [vid. 2024-03-24]. Dostupné z: <https://learn.microsoft.com/en-us/fabric/governance/governance-compliance-overview>
- MICROSOFT, 2055. *What is Microsoft Fabric?* [online] [vid. 2025-09-06]. Dostupné z: <https://learn.microsoft.com/en-us/fabric/fundamentals/microsoft-fabric-overview>
- REŠL, Štěpán, 2023. *Copy Activity, Dataflows Gen2, and Notebooks vs. SharePoint Lists – Data Meerkař* [online] [vid. 2024-04-04]. Dostupné z: <http://localhost:4000/copy-activity-dataflows-gen2-and-notebooks-vs-sharepoint-lists>
- RUSSO, Marco, 2024. *Direct Lake vs. Import mode in Power BI – SQLBI* [online] [vid. 2024-04-20]. Dostupné z: <https://www.sqlbi.com/blog/marco/2024/04/06/direct-lake-vs-import-mode-in-power-bi/>
- SPARK DOCUMENTATION, 2024. *Cluster Mode Overview – Spark 3.5.1 Documentation* [online] [vid. 2024-04-02]. Dostupné z: <https://spark.apache.org/docs/latest/cluster-overview.html>
- STRAKOVÁ, Jarmila a Jan VÁCHAL, 2018. *Situační Analýza, Strategické Řízení a Rozhodování Podniků. Mladá Věda*. 6(2), 205–215.
- ŠVIRK, Martin, 2017. *Strategie podniku* [online]. Jindřichův Hradec. Diplomová práce. Vysoká škola ekonomická v Praze. Dostupné z: https://vskp.vse.cz/69512_strategie-podniku?title=Strategie+podniku&page=1
- WITCOMBE, Jordan, 2023. Microsoft Fabric: An introduction to OneLake. *Advancing Analytics* [online] [vid. 2024-03-24]. Dostupné z: <https://www.advancinganalytics.co.uk/blog/2023/5/24/microsoft-fabric-an-introduction-to-onelake>