

Efektivní analýza dat a reporting

ictPRO

Efektivní analýza dat a reporting


Efektivní analýza dat a reporting

ictPRO

Co je Power BI

Power BI Desktop

Aplikace určená k vytvoření reportu či šablony reportu, který je následně sdílen jako fyzický soubor, výstup v PDF či nahrán do účtu Power BI



Power Query

Úprava nahrávaných dat před importem do Power Pivot (filtrování, nahrazování a úprava hodnot)

Power Pivot


Uložení nahrávaných dat do operační paměti a následné výpočty

Power View

Vizualizace dat systémovými vizuály či vizuály staženými z Marketplace

Power BI

Online rozhraní, ve které je možné spravovat nahrané reporty, vytvořit si reporty vlastní. Možnost sdílení vlastních reportů dalším uživatelům či čtení, úpravy reportů sdílených jinými uživateli.



Reporty

Vizualizační část nahraného reportu.

Datasetsy

Datová část nahraného reportu, uživatel si nad nimi může vytvořit vlastní report.

Toky dat

Power Query v online podobě určené pro přípravu dat a jejich následné sdílení uživatelům, kteří s toky dat pracují ve svých reportech.

www.ictpro.cz

2

Efektivní analýza dat a reporting

ictPRO

Power Query

www.ictpro.cz 3

3

Efektivní analýza dat a reporting

ictPRO

Power Query

- Slouží pro úpravu dat z datového zdroje před jejich importem do datového modelu
- ETL operace:
 - filtrování dat
 - dopočítávání chybných dat , ošetřování chyb, duplicit apod.
 - změna struktury dat
 - spojování dat z více zdrojů
- Dopočítávání nových sloupců, řádků

www.ictpro.cz 4

4

Základní operace na úrovni řádků a sloupců

- Příkazy v rámci uživatelského rozhraní
- Umělá inteligence
(sloupce z příkladů)
- Jazyk M

Filtrování nadbytečných řádků

[illegible]

Filtrování nadbytečných sloupců

[illegible]

Dopočetní
chybějících
sloupců

[illegible]

Použitý postup

- Lze o něm přemýšlet jako o „záznamu“ našich kroků nad daty
- V rámci použitého postupu lze:
 - měnit pořadí
 - přejmenovávat
 - odstranit jednotlivé kroky
 - vkládat nový krok mezi již vytvořené kroky
- Pokud je k danému kroku přidán komentář na úrovni zdrojového kódu, tento komentář se objeví v popisu kroku

Anatomie použitého postupu

Klíčové slovo uvozující začátek postupu

Jednotlivé kroky zpravidla navazují na sebe, přičemž prvním argumentem metody nového kroku je název předchozího kroku

```
let
    #excel_data = Csv.Document(File.Contents("C:\Users\data1.csv"),[Delimiter=";", Columns=7, Encoding=1250, QuoteStyle=QuoteStyle.None])
    #Headers / zhlavi = Table.PromoteHeaders(#excel_data, [PromoteAllScalars=true])
    // Zdroj obsahuje několik sloupců, které nejsou třeba pro tvorbu reportu, byly tedy odstraněny, zachovány zůstanou pouze níže uvedené sloupce
    // The datasource contains multiple columns that are unnecessary for the report so they are removed, only below mentioned columns are kept
    #vyber sloupce / keep columns = Table.SelectColumns(#Headers / zhlavi,{"TIME", "GEO", "Value"})
    #nahrad hodnoty / replace values = Table.ReplaceValue(#vyber sloupce / keep columns, ":", " ", Replacer.ReplaceValue, {"Value"})
    #zmen typ / change types = Table.TransformColumnTypes(#nahrad hodnoty / replace values,{{"TIME", Int64.Type}, {"GEO", type text}, {"Value", Int64.Type}})
in
    #zmen typ / change types
```

Výpočet
neovlivňující
komentář

Klíčové slovo uvozující konec postupu, za slovem „in“ následuje návratová hodnota

Jednotlivé kroky použitého postupu jsou odděleny čárkou s výjimkou kroku předcházejícímu klíčovému slovu in

www.ictpro.cz

7

7

Odebrání řádků

- V záhlaví sloupce jednoho či více sloupců tabulky (podobně jako v MS Excel)
- V hlavní navigaci v kartě Domů / Home části Odebrat řádky / Remove rows
- Možnosti odebrání řádků:
 - na základě toho, jestli hodnota splňuje podmínku
 - odebrání horních x či spodních x řádků
 - odebrání střídavých řádků
 - odebrání duplicit
 - odebrání prázdných řádků
 - odebrání chyb

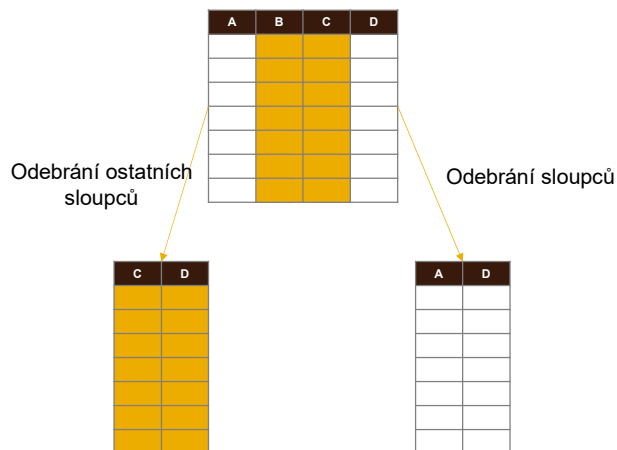
www.ictpro.cz

8

8

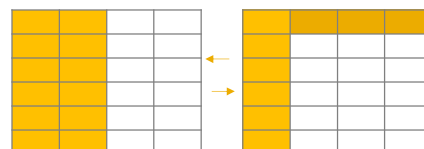
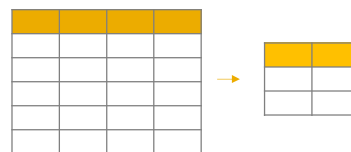
Odebrání sloupců

- Po označení kurzorem a vyvolání kontextové nabídky
- V části hlavní navigace Domů / Home Odebrat sloupce / Remove columns



Agregace a transformace dat

- Agregace data
 - Na základě hodnot jednoho či více sloupců
 - Výsledek obsahuje agregované hodnoty, přičemž lze nad jedním sloupcem použít více agregačních funkcí
- Transformace dat
 - Převedení sloupců do řádků (tzv. unpivotování)
 - Převedení sloupců do řádků (tzv. pivotování), kdy výsledná tabulka obsahuje agregované hodnoty tabulky před pivotováním (podobně jako v kontingenční tabulce v MS Excel)



Spojení dat z více zdrojů – spojení pod sebe

- Dva a více zdrojů jsou spojeny pod sebe
- Nezáleží na pořadí sloupců a jejich typu
- Záleží na názvu sloupců
- Lze připojovat data do již existující tabulky
- Výsledek je rovněž možné nechat vypsat jako novou tabulku

+

Spojení dat z více zdrojů – spojení vedle sebe

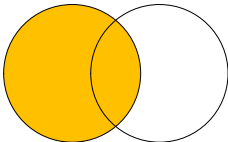
- Spojení tabulek „vedle sebe“ (lze chápat jako dopočítávání sloupců do MS Excel tabulky)
- Výsledek je rovněž možné nechat vypsat jako novou tabulku
- V obou spojovaných tabulkách musí existovat sloupec či kombinace sloupců, přes které lze data dohledat
- Existuje několik možností toho, jaká data jsou na výstupu, viz níže

Efektivní analýza dat a reporting

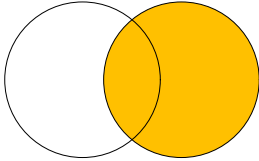
ictPRO

Spojení dat z více zdrojů – spojení vedle sebe

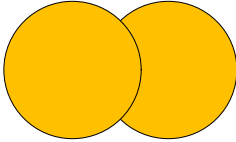
Spojení zleva



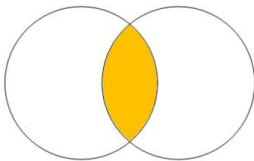
Spojení zprava



Vnější spojení



Vnitřní spojení



www.ictpro.cz

13

13

Efektivní analýza dat a reporting

ictPRO

Základní operace nad datovými typy

- V rámci již existujících sloupců, tj. upravujeme existující hodnoty a počet sloupců zůstává zachován (část hlavní navigace Transformace / Transform)
- V rámci přidání nového sloupce, tj. výsledné hodnoty jsou přidány do nového sloupce (část hlavní navigace Přidání nového sloupce / Add columns)

A
1.1.2023
2.1.2023
3.1.2023
4.1.2023
5.1.2023

→

A
2023
2023
2023
2023
2023

A	B
1.1.2023	2023
2.1.2023	2023
3.1.2023	2023
4.1.2023	2023
5.1.2023	2023

www.ictpro.cz

14

14

7

Efektivní analýza dat a reporting

ictPRO

Základní operace nad datovými typy

- Text
 - Odstranění nadbytečných znaků (mezery, netisknutelné znaky)
 - Úprava velikosti písmen (vše malé, vše velké, velká počáteční písmena slov)
 - Extrakce znaků dle pozice či oddělovače
- Čísla
 - Základní matematické operace (sčítání, odečítání, násobení, dělení apod.)
 - Zaokrouhlování
 - Datum
 - Dopočtení hodnoty z datumu (poslední datum či první datum v měsíci, čtvrtletí, roce apod.)
 - Extrakce hodnoty z datumu (rok, čtvrtletí, měsíc, den)

www.ictpro.cz15

15

Efektivní analýza dat a reporting

ictPRO

Interaktivita v Power BI

www.ictpro.cz16

16

Vizuály = filtry

- Každý 1 bod vizuálu je filtrem tabulky, jejíž data jsou použita
- Každý 1 vybraný bod jiného vizuálu vycházející ze stejné tabulky je filtr
- Každý 1 vybraný bod jiného vizuálu vycházejícího z tabulky propojené aktivní relací je filtr
- Všechny sloupce a míry dosazené do karty Filtry působí jako filtry

Vizuály = filtry

- Filtry se mohou přenášet napříč stránkami prostřednictvím:
 - synchronizace průřezů
 - nastavením filtru na úrovni „Filtry na všech stránkách“
 - v rámci podrobné analýzy či popisu

Efektivní analýza dat a reporting

ictPRO

Ovlivňování interakcí

- Po vybrání vizuálu v části „Formát – Upravit interakce“
- Změnou relací v datovém modelu
- Vlastním výpočtem: kalkulovaná míra, sloupec, tabulka

www.ictpro.cz19

19

Efektivní analýza dat a reporting

ictPRO

Formátování

www.ictpro.cz20

20

Efektivní analýza dat a reporting

ictPRO

Formátování stránky

- Na úrovni celé stránky
- Skládající se z více objektů (tvary, obrázky)
- Za pomoci obrázku (např. šablona Power Point exportovaná jako obrázek)

www.ictpro.cz21

21

Efektivní analýza dat a reporting

ictPRO

Formátování vizuálu

- Ručně v části „Naformátujte vizuál“
- Ze šablony v části Zobrazení
- Kopírováním formátu jiných vizuálů

www.ictpro.cz22

22

Efektivní analýza dat a reporting

ictPRO

Datový model a relace

www.ictpro.cz 23

23

Efektivní analýza dat a reporting

ictPRO

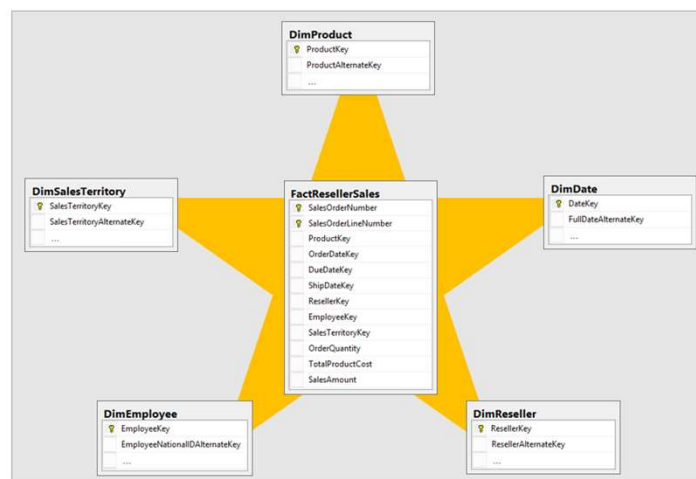
Datové modelování

- Kritická část tvorby reportu, která rozhodne o tom:
 - jestli budeme vůbec schopni vytvořit dotaz vracející správné hodnoty
 - jestli budeme schopni výpočet dokončit s ohledem na dostupnost operační paměti

www.ictpro.cz 24

24

Datové modelování – schéma hvězdy



www.ictpro.cz

25

25

Relace

- Základní možnosti propojení tabulek relacemi:
 - tabulky nejsou propojeny
 - tabulky jsou propojeny neaktivní relací, která je zapínána v rámci výpočtu
 - tabulky jsou propojeny aktivní relací
- Rozdělení v závislosti na směru:
 - jednosměrné, kdy jedna tabulka filtruje druhou a ne naopak
 - obousměrné, kdy se tabulky ovlivňují navzájem

www.ictpro.cz

26

26

Relace

- Rozdělení relací v závislosti na kardinalitě:
 - **Relace M:1 (*:1)**: Sloupec v dané tabulce může obsahovat více než jeden výskyt nějaké hodnoty a druhá související obsahuje jen jeden výskyt nějaké hodnoty.
 - **Jedna k jedné (1:1)** : V relaci 1:1 obsahuje sloupec v jedné tabulce jen jeden výskyt konkrétní hodnoty a jiná související tabulka obsahuje jen jeden výskyt konkrétní hodnoty.
 - **Jedna k mnoha (1:*)** : V relaci 1:N obsahuje sloupec v jedné tabulce jen jeden výskyt konkrétní hodnoty a jiná související tabulka může obsahovat více než jeden výskyt nějaké hodnoty.
 - **Mnoho až mnoho (*:*)**: U složených modelů můžete mezi tabulkami vytvořit relaci M:N, která odstraňuje požadavky na jedinečné hodnoty v tabulkách.

Power Pivot

Power Pivot

- Nástroj pro práci s velkým množstvím dat
- Data jsou uložena v operační paměti
- Po importu z Power Query, tabulky jsou rozděleny na jednotlivé sloupce, komprimovány a následně uloženy
- K datové analýze slouží jazyk DAX
- Při analýze pracujeme se sloupci tabulek uložených v operační paměti, tyto objekty lze chápat jako databázové tabulky, případně jako Excel tabulky

DAX

Efektivní analýza dat a reporting

ictPRO

Syntaxe jazyka DAX

- Nástroj pro práci s velkým množstvím dat
- Data jsou uložena v operační paměti
- Po importu z Power Query, tabulky jsou rozděleny na jednotlivé sloupce, komprimovány a následně uloženy
- K datové analýze slouží jazyk DAX
- Při analýze pracujeme se sloupci tabulek uložených v operační paměti, tyto objekty lze chápat jako databázové tabulky, případně jako Excel tabulky

www.ictpro.cz

31

Efektivní analýza dat a reporting

ictPRO

Operátory

Typ operátoru	Symbol	Použití
Aritmetické operátory	+	Sčítání
	-	Odčítání
	*	Násobení
	/	Dělení
	^	Umocnění
Operátory porovnávání	=	je rovno
	>	Větší než
	<	Menší než
	>=	Větší či rovno než
	<=	Menší či rovno než
	<>	Nerovná se
Logické operátory	&&	A zároveň
		Nebo

www.ictpro.cz

32

Efektivní analýza dat a reporting

ictPRO

Názvy parametrů funkcí

Pojem	Definice
expression	Libovolný výraz DAX, který vrací jednu skalární hodnotu a který se má vyhodnotit několikrát (pro každý řádek/kontext)
hodnota	Libovolný výraz DAX, který vrací jednu skalární hodnotu a který se má vyhodnotit právě jednou před všemi ostatními operacemi
tabulka	Libovolný výraz DAX, který vrací tabulku dat.
tableName	Název existující tabulky pomocí standardní syntaxe jazyka DAX. Nemůže se jednat o výraz.
columnName	Název existujícího sloupce pomocí standardní syntaxe jazyka DAX, obvykle plně kvalifikovaný. Nemůže se jednat o výraz.
name	Řetězcová konstanta, která se použije k zadání názvu nového objektu.
pořadí	Výčet sloužící k určení pořadí řazení.
vazby	Výčet sloužící k určení zpracování hodnot vazeb.
typ	Výčet sloužící k určení datového typu pro PathItem a PathItemReverse.

www.ictpro.cz33

Efektivní analýza dat a reporting

ictPRO

Kalkulované tabulky

- Tabulku lze do datového modelu přidat:
 - ručně (část zadat data)
 - za pomoci funkcí s návratovou hodnotou tabulku, např. ALL(), FILTER()
 - agregační funkcí, např. SUMMARIZE()
 - funkcí CALCULATETABLE()
 - funkcí ROW() vytvářející 1 řádek tabulky

www.ictpro.cz34

Kalkulované sloupce

- Využívají tzv. kontextu řádku, kdy výpočet je proveden pro každý jeden řádek tabulky
- Data jsou uložena v operační paměti a aktualizována při aktualizaci datového zdroje
- Je vhodné je použít pouze v těchto případech:
 - výsledek chceme použít do průřezu či osu vizuálu
 - vytváříme segmenty dat
 - výsledek skutečně potřebujeme pro každý jeden řádek

Kalkulované míry

- Využívají tzv. kontextu filtru, kdy reagují na filtry aplikované vizuály
- K aktualizaci dochází při změně kontextu filtru, například interakci uživatele s průřezem

Závorky v DAX

- Hranaté závorky pro identifikaci sloupce, např. zápis `data[Value]` odkazuje na tabulku `data` a její sloupec `[Value]`
- Kulaté závorky pro funkce, např. `SUM()` volá funkci pro součet hodnot ve sloupci
- Složené závorky pro vytvoření množiny, např. `{2022,2023}` vytváří množinu obsahující hodnoty 2022 a 2023

Závorky v DAX - příklad

```

1 value 2019 cesko =
2
3 // krok4: Funkce SUMX() prochází řádky tabulky vracené funkcí FILTER() a následně
4 // agreguje hodnoty ve sloupci [Value]
5 SUMX(
6     //krok3: Funkce filtr vrací novou tabulku
7     FILTER(
8         // krok1: Funkce filtr přebírá tabulku data se všemi filtry, které jsou aplikovány
9         data,
10        // krok2: Funkce filtr aplikuje další filtry SUMX(FILTER(table,condition),expression)
11        data[TIME] = 2019
12        && data[Country] = "Česko"
13    ),
14    data[Value]
15 )

```

Čárky oddělují jednotlivé parametry funkcí

Každá funkce je uvozena levou kulatou závorkou a uzavřena pravou kulatou závorkou

Kalkulované míry

- Využívají tzv. kontextu filtru, kdy reagují na filtry aplikované vizuály
- K aktualizaci dochází při změně kontextu filtru, například interakci uživatele s průřezem

Příklad výpočtu v jazyku DAX

```

1 01 Oblast rok 2020 =
2
3
4 SUMX(
5     --krok3: funkce FILTER() vraci tabulku
6     FILTER(
7         --krok1: Funkce FILTER() prevezme tabulku filtrovanou vizuály
8         cesko,
9         --krok2: Funkce FILTER() dale aplikuje filtr nad sloupcem
10        cesko[Rok] = 2020,
11        --krok4: Pro kazdy jeden radek tabulky vracene
12        -- funkci FILTER() je pouzita funkce agregacni
13        -- funkce suma
14    cesko[Hodnota])

```

Jednotlivé argumenty jsou odděleny čárkou (v USA lokalizaci středníkem)

Funkce začíná a končí kulatou závorkou

Příklad výpočtu v jazyku DAX

```

4 SUMX(
5
6   --krok4: Funkce FILTER() filtruje tabulku vrácenou
7   -- funkci ALL() a vrací tabulku
8   FILTER(
9     --krok3: Funkce ALL() vrátí tabulku
10    -- se všemi jejími hodnotami
11    ALL(cesko),
12    cesko[Rok]=
13      -- krok2: Funkce MINX() najde minimální
14      -- hodnotu v tabulce vrácenou funkcí ALL()
15      MINX(
16        --krok1: Funkce ALL() vrátí tabulku
17        -- se všemi jejími hodnotami
18        ALL(cesko),
19        cesko[Rok])),
20 -- funkce SUMX() prochází tabulku vrácenou
21 -- funkcí FILTER() a počítá hodnoty ve sloupci
22 cesko[Hodnota])

```

Funkce ALL() vytváří kopii tabulky data se všemi jejími hodnotami. Původní tabulka data zůstává zachována

Funkce ALL() vytváří kopii tabulky data se všemi jejími hodnotami. Původní tabulka data zůstává zachována

	data	

	ALL(cesko)	

	ALL(cesko)	

Proměnné

- Zlepšují výkon
- Zpřehledňují kód
- Zjednodušují ladění chyb
- Jsou definovány klíčovým slovem VAR
- Požadavky na název jsou:
 - podporovaná sada znaků: a–z, A–Z, 0–9
 - znaky 0–9 nejsou platné jako první znak
 - jako předpona názvu identifikátoru je povoleno _ (dvojitě podtržítka)
 - nejsou podporované žádné jiné speciální znaky
 - nejsou povolená rezervovaná klíčová slova
 - nejsou povolené názvy existujících tabulek
 - nejsou povolené prázdné mezery

Efektivní analýza dat a reporting

ictPRO

Proměnné

- V rámci skriptu nelze měnit jejich hodnotu
- V rámci jednoho skriptu můžeme mít více proměnných
- Nemůže na ně být odkazováno z jiných kalkulovaných měř či sloupců

Pozn. Možné řešení je použití celé kalkulované míry jako proměnné

- Pokud pracujeme s proměnnými, musíme použít klíčové slovo RETURN pro určení návratové hodnoty

www.ictpro.cz

43

43

Efektivní analýza dat a reporting

ictPRO

Shrnutí

www.ictpro.cz

44

44

Efektivní analýza dat a reporting

ictPRO

Sumarizace

- **V rámci Power Query jsme se naučili:**
 - vyčistit si data z datových zdrojů
 - agregovat a transformovat data
 - spojit data z více datových zdrojů
 - určit, která data budou načtena do datového modelu
 - práci s použitým postupem

www.ictpro.cz

45

45

Efektivní analýza dat a reporting

ictPRO

Sumarizace

- **V rámci Power BI části „plátno“ jsme se naučili:**
 - vytvářet základní dashboards a řídit interakci mezi jednotlivými vizuály
 - formátovat vizuály
 - tvořit report z více typů stránek (podrobná analýza, popis)
 - práci se záložkami
 - základní práci s tlačítky

www.ictpro.cz

46

46

Efektivní analýza dat a reporting

ictPRO

Sumarizace

- **V rámci Power BI části „DAX“ jsme se naučili:**
 - přidávat do datového modelu objekty typu tabulka, sloupec, míra
 - chápat kontext filtru, kontext řádku a kontext výpočtu
 - ovlivňování kontextu výpočtu za pomoci základních agregačních funkcí, tj. SUMX(), MAXX(), MINX(), AVERAGEX() při práci s jednou tabulkou
- **V rámci Power BI části „model“ jsme se naučili:**
 - propojovat tabulky relacemi
 - rozlišovat mezi funkčností jednotlivých typů relací

www.ictpro.cz 47

47

Efektivní analýza dat a reporting

ictPRO

Na co se můžete těšit na pokročilém kurzu?

- **Power Query**
 - ošetřování chyb ve výpočtech
 - úprava dat pro rychlejší a snadnější výpočty
 - hromadné načítání dat z více souborů v jedné složce
 - optimalizaci použitého postupu
- **Plátno**
 - pokročilá práce s tlačítky
 - tvorba vlastních vizuálů

www.ictpro.cz 48

48

Efektivní analýza dat a reporting

ictPRO

Na co se můžete těšit na pokročilém kurzu?

- **DAX:**
 - funkce CALCULATE() – nejdůležitější funkce v DAX
 - ovlivňování kontextu výpočtu napříč více tabulkami datového modelu
 - optimalizace zápisu jazyka DAX pro snadnější tvorbu a rychlejší výpočet
 - ovlivňování nastavení relací v rámci samotného výpočtu
 - ovlivňování výpočtu v závislosti na interakci uživatele
 - virtuální tabulky pro výpočty nad daty s různou granularitou
 - segmentace dat

www.ictpro.cz 49

49

Efektivní analýza dat a reporting

ictPRO

Děkuji za pozornost

50