

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

Evidenčné číslo: 103004/I/2021/2041040612

POWER BI A DAX V BUSINESS INTELLIGENCE

Diplomová práca

2021

Bc. Peter Vadovics

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

POWER BI A DAX V BUSINESS INTELLIGENCE

Diplomová práca

Študijný program: Informačný manažment

Študijný odbor: Ekonómia a manažment

Školiace pracovisko: Katedra aplikovanej informatiky

Vedúci záverečnej práce: doc. Dr. Ing. Miroslav Hudec

Bratislava 2021

Bc. Peter Vadovics

Čestné vyhlásenie

Čestne vyhlasujem, že záverečnú prácu som vypracoval samostatne a že som uviedol všetku použitú literatúru

Dátum:

.....

(podpis študenta)

Pod'akovanie

Touto cestou by som sa chcel poďakovať svojmu diplomovému vedúcemu doc. Dr. Ing. Miroslavovi Hudecovi za ochotu, trpezlivosť, cenné rady a pripomienky, ktorými prispel k vypracovaniu tejto diplomovej práce.

ABSTRAKT

VADOVICS, Peter: Power BI a DAX v Business Intelligence – Ekonomická univerzita v Bratislave. Fakulta hospodárskej informatiky; Katedra aplikovanej informatiky. – doc. Dr. Ing. Miroslav Hudec – Bratislava: FHI, 2021, 81s.

Cieľom práce je spracovanie a vizualizácia dát pochádzajúcich z databázy rezervačného systému ubytovania na Slovensku pomocou analytického nástroja Power BI a jazyka DAX.

Práca najprv porovnáva jednotlivé BI nástroje dostupné na trhu. Následne sa pomocou vybraného nástroja v ETL procese dáta očistia, agregujú a vložia do dátového skladu. Presnejšie, dáta rokov 2014 až 2019. Na základe týchto údajov urobia sa reporty a analýzy údajov pre informovanie a podporu rozhodovania.

Práca je rozdelená na teoretickú a praktickú časť. V teoretickej časti sa zaoberáme oblasťou Business Intelligence a analytickými nástrojmi dostupnými na trhu. V praktickej časti sa už venujeme praktickým využitím nástroja Power BI na analýzu dát a vyhodnocujeme naše výsledky.

Kľúčové slová

Business Intelligence, Power Query, OLAP, Dashboard, Report, Power BI, DAX

ABSTRACT

VADOVICS, Peter: Power BI and DAX in Business Intelligence – University of Economics in Bratislava. Faculty of Business Informatics FBI, Department of applied informatics. – The head of the thesis: doc. Dr. Ing. Miroslav Hudec – Bratislava: FHI, 2021, 81p.

The aim of the thesis is the processing and visualization of data from the database of the accommodation reservation system in Slovakia using the analytical tool Power BI and DAX language.

The thesis first compares the individual BI tools available on the market. Subsequently, the data is cleaned, aggregated and inserted into the data warehouse using the selected tool in the ETL process. More precisely, data from 2014 to 2019. Based on these data, reports and data analyzes will be performed for information and decision support.

The thesis is divided into theoretical and practical part. In the theoretical part we deal with the area of Business Intelligence and analytical tools available on the market. In the practical part we deal with the practical use of the Power BI tool for data analysis and evaluate our results.

Keywords

Business Intelligence, Power Query, Dashboard, Report, Power BI, DAX

Obsah

ÚVOD.....	11
1 SÚČASNÝ STAV RIEŠENEJ PROBLEMATIKY DOMA A V ZAHRANIČÍ.....	12
1.1 DEFINÍCIA POJMU BUSINESS INTELLIGENCE	12
1.2 HLAVNÉ KOMPONENTY BUSINESS INTELLIGENCE	14
1.2.1 ETL.....	15
1.2.2 EAI (Enterprise Application Integration).....	15
1.2.3 Dátový sklad (DW - Data Warehouse).....	15
1.2.4 Dočasné úložisko dát (DSA – Data Staging Areas).....	16
1.2.5 Dátová tržnica.....	17
1.2.6 Operatívne úložisko (ODS - Operational Data Store).....	17
1.3 OLAP A MULTIDIMENZIONÁLNE MODELOVANIE.....	17
1.3.1 Fakty a Dimenzie.....	19
1.3.2 Operácie v OLAP kocke.....	21
1.3.3 Schémy tabuliek dimenzií	21
1.4 SAMOOSLUŽNÝ BUSINESS INTELLIGENCE	22
1.5 NÁSTROJE (SOFTVÉRY) SSBI	23
1.6 POROVNANIE BI NÁSTROJOV.....	23
1.6.1 Power BI.....	25
1.6.2 Tableau	26
1.6.3 Qlik.....	27
1.7 POWER BI	28
1.7.1 Komponenty Power BI.....	29
1.7.2 Verzie Power BI	30
1.7.3 Power Query	32
1.7.4 DAX	32
2 CIEĽ PRÁCE, METODIKA PRÁCE A METÓDY SKÚMANIA.....	35
3 VÝSLEDKY PRÁCE A DISKUSIA	36
3.1 DATASET.....	36
3.2 NAČÍTANIE ZDROJOVÝCH SÚBOROV	37
3.3 PRÍPRAVA DÁT V POWER QUERY	38
3.3.1 Tabuľka rezervácií.....	38

3.3.2	Tabuľka ubytovacích zariadení	50
3.4	VYTVORENIE PREHLADOV A VIZUÁLOV.....	58
3.4.1	Vytvorenie nového stĺpca podľa vzorca	59
3.4.2	Vytváranie mierok	60
3.4.3	Destinácie a ubytovacie zariadenia s najvyšším predajom.....	63
3.4.4	Medziročné rozdiely v predaji	64
3.4.5	Predaj v závislosti od dátumu objednávky a dátumu začiatku pobytu	67
3.4.6	Zoradenie krajín klientov podľa výšky tržieb	69
3.4.7	Celkový predaj podľa ceny za osobu	70
3.4.8	Cenové kategórie pobytov	72
3.4.9	Predaj podľa dĺžky pobytov	73
3.4.10	Predaj podľa typu pobytu.....	74
3.5	PUBLIKÁCIA VÝSLEDNÝCH DASHBOARDOV V SLUŽBE POWER BI.....	75
3.6	DISKUSIA	76
	ZÁVER	78
	ZOZNAM POUŽITEJ LITERATÚRY	79
	PRÍLOHY	81

Zoznam obrázkov

Obrázok 1 Hlavné komponenty BI a ich väzby (Novotný, 2005)	14
Obrázok 2 Transformácia OLTP databázy na OLAP	18
Obrázok 3 Multidimenzionálna OLAP kocka s tromi dimenziami	20
Obrázok 4 Schémy tabuliek dimenzií: hviezdnicová schéma a schéma snehovej vločky.....	22
Obrázok 5 Gartner Magic Quadrant pre Business Intelligence a analytické platformy 2021	24
Obrázok 6: Power BI - načítanie údajov (autor, 2021).....	37
Obrázok 7: Náhľad CSV súboru (autor, 2021).....	38
Obrázok 8 Výsledný stĺpec si nazveme TextSplit (autor, 2021)	40
Obrázok 9 Zobrazenie rozšíreného editoru v Power Query (autor, 2021).....	42
Obrázok 10 Transformačná funkcia (autor, 2021).....	44
Obrázok 11 Mazanie použitých krokov v Power Query (autor, 2021).....	44
Obrázok 12 Vyvolanie vlastnej funkcie v Power Query (autor, 2021)	45
Obrázok 13 Transformovaný stĺpec vytvorený pomocou vlastnej transformačnej funkcie (autor, 2021).....	45
Obrázok 14 Extraktovanie textu medzi oddeľovačmi (autor, 2021)	46
Obrázok 15 Pridanie indexovaného stĺpca (autor, 2021).....	47
Obrázok 16 Zlúčenie dotazov (autor, 2021)	48
Obrázok 17 Výsledný stĺpec po zlúčení dotazov (autor, 2021).....	48
Obrázok 18 Rozbalenie stĺpca	49
Obrázok 19 Výsledný stĺpec Country ID po rozbalení (autor, 2021).....	50
Obrázok 20 Stĺpec pre cenu s nesprávnym typom údajov (autor, 2021).....	50
Obrázok 21 Zmena typu údajov stĺpca	51
Obrázok 22 Chybný výsledok zmeny typu stĺpca (autor, 2021).....	51
Obrázok 23 Zobrazenie zmeny typu údajov stĺpca pomocou miestneho nastavenia (autor, 2021)	52
Obrázok 24 Zmena typu údajov stĺpca pomocou miestneho nastavenia (autor, 2021)	52
Obrázok 25 Správny výsledok zmeny typu stĺpca (autor, 2021).....	53
Obrázok 26 Načítanie údajov z Power Query do Power BI	53
Obrázok 27 Zobrazenie dátového modelu v Power BI (autor, 2021).....	54
Obrázok 28 Spravovanie vzťahov v dátovom modeli (autor, 2021)	54

Obrázok 29 Úprava vzťahu v dátovom modeli (autor, 2021).....	55
Obrázok 30 Časová tabuľka Date (autor, 2021)	57
Obrázok 31 Vytvorenie prepojenia medzi časovou tabuľkou a tabuľkou rezervácií (autor, 2021)	57
Obrázok 32 Dátový model rozšírený o časovú dimenziu (autor, 2021)	58
Obrázok 33 Vytvorenie nového stĺpca (autor, 2021).....	60
Obrázok 34 Vytvorený nový stĺpec pre názov ubytovania a lokalitu (autor, 2021).....	60
Obrázok 35 Vytváranie novej mierky (autor, 2021).....	61
Obrázok 37 Stĺpcový graf – porovnanie mierok Celkový predaj v závislosti od časovej dimenzie (autor, 2021)	62
Obrázok 38 Pružový graf - Poradie destinácií a ubytovaní podľa výšky predaja	63
Obrázok 39 Stĺpcový graf – porovnanie mierky Celkový predaj s predchádzajúcim rokom	65
Obrázok 40 Čiarový graf – porovnanie zisku s predchádzajúcim rokom.....	66
Obrázok 41 Nastavenie farby údajov podľa pravidla	67
Obrázok 42 Medziročný rozdiel celkového predaja a zisku podľa mesiacov	67
Obrázok 43 Porovnanie celkového predaja s predchádzajúcim rokom a medziročný rozdiel na mesačnej báze	69
Obrázok 44 Zoradenie krajín klientov podľa výšky tržieb od roku 2014 po 2019.....	69
Obrázok 45 Pomocná tabuľka definovaných intervalov cien za osobu.....	70
Obrázok 46 Celkové tržby podľa cenových kategórií za osobu	71
Obrázok 47 Celkové tržby podľa cenových kategórií za pobyt.....	72
Obrázok 48 Celkové tržby podľa cenových kategórií za pobyt a podľa dĺžky pobytu	72
Obrázok 49 Výška ceny objednávok a pobytov podľa mesiacov v roku	73
Obrázok 50 Celková suma rezervácií v destináciách podľa dĺžky pobytu.....	74
Obrázok 51 Tabuľka: Zoradenie typov pobytov podľa celkovej sumy	75
Obrázok 52 Publikovanie zostavy v službe Power BI.....	75

Úvod

Dostatok kvalitných informácií je dnes pre podnik zásadný predovšetkým z hľadiska kvalifikovaných rozhodnutí. O to dôležitejšie je to hlavne v prostredí stále tvrdšej konkurencie. Efektívne využívanie informácií vo významnej miere prispieva k výkonnosti podnikov. Tieto informácie by mali byť dostupné rýchlo s možnosťou formulovať požiadavky podľa meniacich sa podmienok na trhu.

Vhodné informácie môžu pomôcť nielen pri zlepšovaní vzťahu k zákazníkovi či zvýšení predaja, ale pomáhajú aj pri zlepšovaní spolupráce s vnútorným prostredím podniku a sú základom pre moderné manažérske rozhodovanie.

Firmám sa počas svojej existencie zhromažďujú údaje o predajoch, klientoch, produktoch alebo dodávateľoch. Pri zvyšovaní výkonnosti firmy je potrebné tieto údaje analyzovať a získavať z nich cenné informácie, ktoré poslúžia pri procese plánovania a strategického rozhodovania.

A práve získavaniu cenných informácií pre strategické rozhodovanie sa venuje práve oblasť Business Intelligence, ktorej sa v mojej práci venujem.

Hlavným cieľom diplomovej práce je objasniť problematiku využitia Business Inteligencie reportovacích služieb v prostredí aplikácie Power BI Desktop, pomocou ktorej analyzujem údaje rezervačného systému. Túto tému som si vybral preto, že pracujem v obore cestovného ruchu a veľmi ma fascinuje hĺbková analýza údajov.

1 Súčasný stav riešenej problematiky doma a v zahraničí

1.1 Definícia pojmu Business Intelligence

V dnešnej dobe nájdeme množstvo rôznych definícií pojmu Business Intelligence (BI). Môžeme ho definovať ako využívanie rôznych zdrojov údajov, či už priamo z interných dát podniku alebo tých externých, s cieľom spracovať ich a poskytnúť cenné informácie v takej podobe, aby napomáhali v strategickom rozhodovaní podniku [6]. Môžeme ho chápať ako proces transformácie údajov, ktorého účelom je premeniť veľké množstvo surových dát na poznatky a tieto zas využiť v procese plánovania a rozhodovania [1].

V jednej z množstva definícií sa dozvieme, že Business Intelligence (BI) je proces prijímania „inteligentných“ obchodných rozhodnutí na základe analýzy dostupných údajov [8]. V ďalšej definícii BI sa píše, že systémy podnikového spravodajstva poskytujú informácie použiteľné v správnom čase, na správnom mieste a v správnej forme, aby pomohli osobám s rozhodovacími právomocami. Cieľom je zlepšiť včasnosť a kvalitu vstupov do rozhodovacieho procesu, a tým uľahčiť prácu manažérov [11].

Business Intelligence tvorí veľmi dôležitú súčasť firiem naprieč všetkými odvetvami podnikania. Vďaka získaným informáciám a poznatkom získavajú podnikateľské subjekty významné výhody oproti konkurencii. Doslovne, Business Intelligence znamená byť inteligentný ohľadom vášho biznisu.

Pojem BI zahŕňa sadu metód a rôznych technológií, ktorých cieľom je skvalitniť rozhodovacie procesy vo firme [5]. Patria sem aplikácie a technológie určené pre zber, skladovanie, analyzovanie a prezentovanie dát s cieľom pomôcť riadiacim pracovníkom urobiť lepšie rozhodnutia.

Zo všetkých predchádzajúcich definícií vyplýva, že Business Intelligence využíva informácie na riadenie a rozhodovanie, a nie na základné spracovanie dát a realizáciu bežných transakcií v podniku.

Výkonnosť a konkurencieschopnosť podniku vo veľkej miere závisia aj od toho sa v rámci riadenia využívajú možnosti BI.

Keďže hlavným účelom BI systémov je podporovať procesy rozhodovania manažmentu firiem a vedúcich pracovníkov, nazývame ich aj preto ako systémy na podporu rozhodovania. Získanými informáciami môžeme lepšie porozumieť zmenám predajov, podielov na trhu, správania zákazníkov a podobne. Tieto informácie nám pomáhajú lepšie porozumieť postaveniu podniku na trhu v porovnaní s ostatnými subjektmi a lepšie sa prispôbiť meniacim sa podmienkam. Aby mohli byť lepšie pochopit' získané informácie, sú niektorým z dostupných analytických nástrojov prezentované pomocou dashboardov [8].

V BI systémoch ide o chápanie vzájomných vzťahov skutočností takým spôsobom, aby smerovali k dosiahnutiu požadovaného podnikateľského cieľa [7].

Prvé riešenia, ktoré v procese riadenia podniku smerovali k podpore manažérskych a analytických úloh sa objavili už koncom sedemdesiatych rokov, pojem Business Intelligence definoval až v roku 1989 Howard Dresner, ktorý ho predstavil ako súbor konceptov na zlepšenie rozhodovania sa podnikov. Toto rozhodovanie malo vychádzať z konkrétnych faktov.

Napriek tomu, že táto téma BI nie je príliš stará, využitie nástrojov BI sa v posledných rokoch rýchle rozvíja. Hlavným rozdielom od bežných informačných systémov, je možnosť využitia BI systémov pri podnikateľskom rozhodovaní.

Nástroje BI pristupujú k údajom, analyzujú ich a prezentujú analytické zistenia v reportoch, súhrnoch, dashboardoch, grafoch a mapách, aby riadiacim pracovníkom poskytli podrobné informácie o stave podnikania [13].

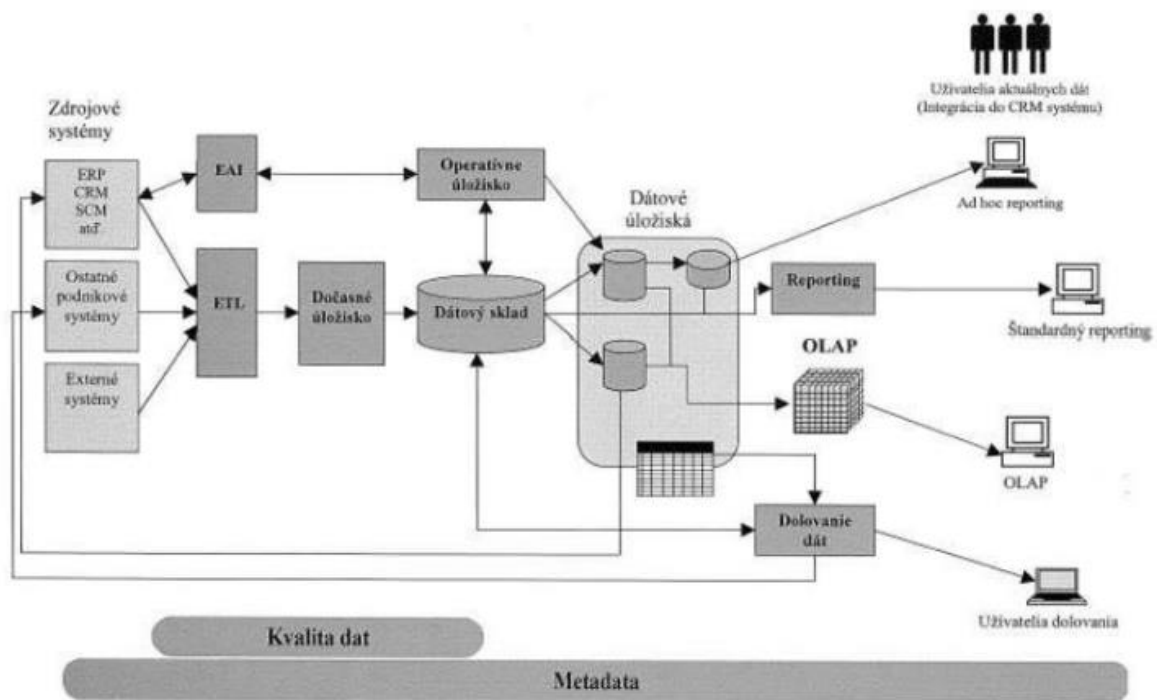
Aplikovanie BI poskytuje analytické a plánovacie funkcie v rôznych segmentoch podnikového riadenia, ako je marketing, controlling, predaj, nákup, ľudské zdroje a podobne. Business Intelligence technológie a procesy sú postavené na princípoch multidimenzionálnych pohľadov na dáta podniku. Princípom multidimenzionality rozumieme možnosť pozerať na realitu z niekoľkých možných uhlov [5]. Z technologického hľadiska BI pokrýva oblasti skladovania dát, reportingu, OLAP, dolovania dát, niektorých vizualizácií dát, a ďalších analytických oblastí [8]. Všetky tieto pojmy si vysvetlíme v ďalšej časti práce.

1.2 Hlavné komponenty Business Intelligence

Počas existencie systémov Business Intelligence sa vyvinula určitá schéma usporiadania jednotlivých komponentov, aj keď konkrétne usporiadanie sa môže aj líšiť podľa potrieb subjektov v závislosti od komplexnosti riešenia .

V rámci funkcionality môžeme komponentov BI rozdeliť do niekoľkých vrstiev [5]:

- **Vrstva pre extrakciu, transformáciu, čistenie a nahrávanie dát** – radíme sem komponenty zamerané na oblasť zberu a prenosu dát zo zdrojových systémov. Do tejto vrstvy patria systémy ETL (systémy na extrakciu, transformáciu a prenos dát) a systémy EAI (systémy na integráciu aplikácií).
- **Vrstva pre ukladanie dát** - databázové komponenty, ktoré sa starajú o procesy ukladania. Vrstva obsahuje: dátové sklady (data warehouses), dátové tržnice (data marts), operatívne dátové úložiská (operational data stores) a dočasné úložiská dát (data staging areas)
- **Vrstva pre analýzu dát** - analytické komponenty, ktoré riešia procesy spojené s analýzou dát. Patrí sem: reporting, systémy OLAP, dolovanie dát (data mining).



Obrázok 1 Hlavné komponenty BI a ich väzby (Novotný, 2005)

1.2.1 ETL

ETL (Extrakcia - Extraction, Transformácia - Transformation, Nahrание - Loading) je jednou z najvýznamnejších častí celého komplexu BI. Úlohou ETL je extrahovať dáta zo zdrojových databáz, následne ich vyčistiť a transformovať do požadovanej podoby a nakoniec ich nahráť do dátového skladu [5]. ETL je jednou z najvýznamnejších častí celého komplexu BI. Prostriedky ETL nazývame aj ako dátová pumpa. Nástroje ETL prenášajú dáta v určitých časových intervaloch, väčšinou na dennej, týždennej alebo mesačnej báze.

1.2.2 EAI (*Enterprise Application Integration*)

V systéme Business Intelligence sa EAI využívajú prevažne v transformačnej vrstve. Ich primárnym účelom je prenos dát hlavne do operatívneho úložiska, kde poskytujú aktuálny zjednotený náhľad na celopodnikové údaje v rámci analytického procesu. Využívajú sa tiež na vytvorenie dátových skladov v reálnom čase. Pracujú v reálnom čase na rozdiel od nástrojov ETL [5].

1.2.3 Dátový sklad (*DW - Data Warehouse*)

Aby mohol podnik analyzovať údaje, je potrebné ich zozbierať rôznych zdrojov a nahráť ich do jedného spoločného úložiska, ktoré nazývame dátový sklad (angl. Data Warehouse).

William Inmon, ktorý je považovaný za otca dátových skladov, definoval dátový sklad, ako „subjektovo orientovaný, integrovaný, stály a časovo rozlíšený súhrn dát určený na podporu procesu manažérskeho rozhodovania“ [14]. Dátový sklad obsahuje podrobné podnikové údaje.

Dátový sklad môžeme označiť ako centrálny podporný systém, obsahuje údaje zhromaždené z veľkého množstva zdrojov v rámci podniku aj mimo neho [9]. Tieto dáta sa v dátovom sklade zhromažďujú, vytvárajú sa medzi nimi vzťahy. Dátový sklad sa používa špeciálne na podporu rozhodovania.

Do dátového skladu sa v pravidelných intervaloch nahrávajú údaje z transakčných databáz, pochádzajúcich z rôznych oddelení podniku alebo popri prípade aj z iných lokalít [1].

Predtým ako sa údaje do dátového skladu nahrajú, sú upravované, pretože dáta sú tu organizované podľa iných pravidiel.

V dátových skladoch sa využíva multidimenzionálny dátový model, ktorý je vysvetlený v ďalšej časti.

Priblížme si teraz, čo znamenajú tie prívlastky v definícii dátového skladu.

- Subjektovo orientovaný - Údaje v dátovom sklade sa ukladajú podľa predmetu záujmu (subjektovo orientovaný) a nie podľa aplikácie, v ktorej boli vytvorené. Predmetom záujmu rozumieme napríklad výrobok, dodávateľ, zákazník, zamestnanec a podobne.
- Integrovaný – údaje sú uložené v rámci celého podniku, nie iba jednotlivých oddelení
- Stály – údaje sa do dátového skladu len zapisujú, je možné ich iba čítať, mazať už nie, dáta sa nemenia
- Časovo rozlíšený – uložené dáta majú sú spojené s časovou dimenziou

1.2.4 Dočasné úložisko dát (DSA – Data Staging Areas)

Dočasné úložisko dát slúži na dočasné uloženie extrahovaných dát z produkčných (transakčných) databáz. Jeho dôvod využitia je prvotné ukladanie ešte netransformovaných dát zo zdrojových systémov. Primárnou úlohou DSA je zefektívniť extrakciu dát z daných zdrojov. Ide o nepovinnú časť, ale v prípade, že je potrebné vyhnúť sa veľkému zaťaženiu produkčných systémov v rámci ich prevádzky a ich potenciálnym výpadkom s čo najmenším dopadom na ich výkonnosť, alebo v prípade potreby konvertovania dát, sa odporúča dočasné úložisko použiť [5]. V dočasnom úložisku sa ukladajú dáta s rovnakou štruktúrou akú mali v produkčných databázach. Dáta neobsahujú históriu, sú detailné (nie sú agregované), a nie sú kontrolované – sú nekonzistentné. Ďalšou ich vlastnosťou je, že sú meniace sa, pretože následne, keď sa upravené uložia do dátového skladu, sa údaje z dočasného úložiska vymažú.

1.2.5 Dátová tržnica

Zatiaľ čo dátový sklad je zdieľaný v celom podniku a obsahuje komplexné informácie, dátová tržnica uchováva údaje o jednom subjekte [9]. Dátová tržnica je v podstate podmnožinou dátového skladu a je určená pre potreby určitého okruhu užívateľov, napríklad oddelenia podniku. Špecializuje sa na jeden predmet, ktorý môže byť napríklad zamestnanci alebo predaj. Neznamená to však, že každé jedno oddelenie má mať vlastné súkromné dátové tržnice. V prípade, že viac oddelení by potrebovali rovnaké údaje, mali byť využívať rovnakú dátovú tržnicu. Údaje v dátovej tržnici sa načítavajú z dátového skladu, čo znamená, že najskôr je vytvorený dátový sklad a až z neho sa vytvárajú dátové trhy. Môžeme povedať, že dátové trhy sú decentralizované dátové sklady [5].

Výhodou vytvorenia dátových tržníc je jednoduchšia a rýchlejšia implementácia, z čoho vyplýva rýchlejší prínos pre koncového užívateľa. Nevýhodou tohto usporiadania je možnosť vytvorenia nekonzistencie medzi jednotlivými tržnicami, ako aj vyššie nároky na údržbu toho systému.

1.2.6 Operatívne úložisko (ODS - Operational Data Store)

V operatívnom úložisku sú uložené zjednotené predmetovo orientované celopodnikové dáta, ktoré koncovým užívateľom poskytujú aktuálny a zjednotený pohľad na prevádzku. Údaje v operatívnom úložisku sú konsolidované, konzistentné, predmetovo orientované, aktuálne, nestále a v niektorých prípadoch môžu byť aj agregované. Sú nepretržite aktualizované, takže ich súhrnný pohľad prezentuje aktuálny stav. Operatívne úložisko je tiež nepovinným komponentom BI riešení.

Operatívne úložisko slúži ako databáza podporujúca analytický proces. Je vytvorené s cieľom sprístupniť užívateľom dáta pre analýzy či otázky s minimálnym oneskorením oproti ich obstaraniu.

1.3 OLAP a multidimenzionálne modelovanie

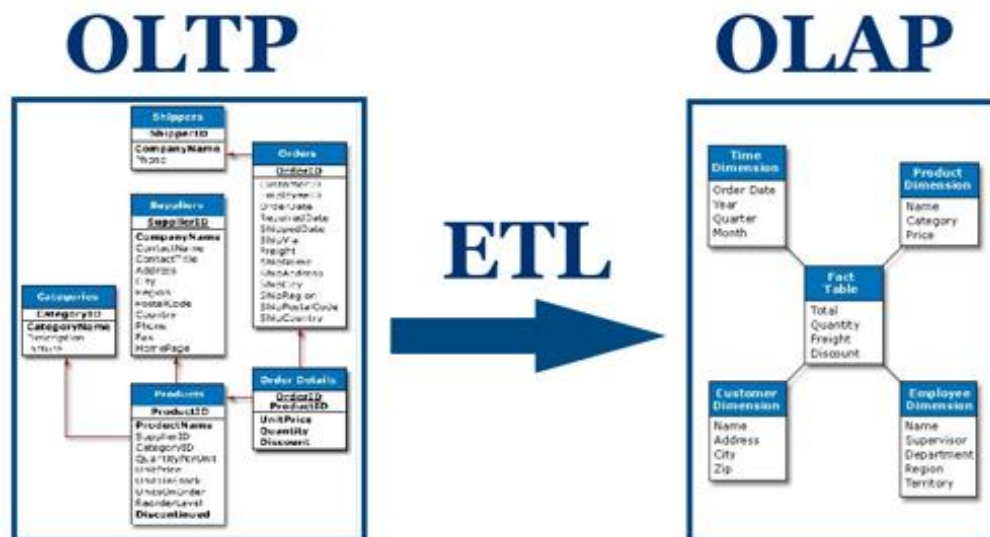
V transakčných databázových systémoch, využívaných na operatívne ukladanie a aktualizovanie dát, sa používajú relačné databázy. Dáta v týchto databázach sa môžu meniť aj niekoľkokrát za deň, pričom realizujú spracovanie v reálnom čase. Tieto systémy

nazývame aj OLTP (On Line Transaction Processing) neboli vytvorené s cieľom analýzy údajov.

Na tento účel s rozvojom Business Intelligence bola nová technológia uloženia dát v databáze - systém OLAP (Online Analytical Processing). Od OLTP sa líšia spôsobom uloženia dát v databáze. Kým v OLTP relačných databázach sa kladie doraz na ukladanie dát v normalizovanej podobe, v OLAP databázach je dôležitá prístupnosť dát pre používateľov zaoberajúcich sa analýzou podnikových dát.

Technológia OLAP využíva koncepciu multidimenzionálnych databáz. OLAP databázy umožňujú realizovať časové porovnávanie vybraných metrik, ako aj analyzovať údaje podľa rôznymi tzv. multidimenzionálnymi pohľadmi. Na tento účel sa používa OLAP kocka, ktorú vytvárame na základe dimenzionálneho modelu. „OLAP databázy predstavujú jednu alebo viacero súvisiacich OLAP kociek“ [5].

Multidimenzionálny dátový model, ktorý sa objavil v priebehu 90. rokov minulého storočia, bol navrhnutý na podporu analýz údajov. Tento dátový model tvorí základ odvetvia Business Intelligence [8].



Obrázok 2 Transformácia OLTP databázy na OLAP

V multidimenzionálnych modeloch sa údaje delia na fakty so súvisiacimi numerickými mierami, a na dimenzie, ktoré charakterizujú fakty. Typickým faktom je nákup alebo rezervácia. Typickými mierami je cena nákupu. Medzi typické dimenzie patrí čas

nákupu miesto nákupu, zakúpený tovar. Multidimenzionálne modely sa využívajú v dátových skladoch.

Multidimenzionálne databázové systémy agregujú veľké množstvo podrobných údajov, aby našli celkové trendy, a prezentujú výsledky multidimenzionálnym spôsobom. Systémy OLAP vždy využívajú multidimenzionálny pohľad na dáta.

Multidimenzionálny dátový model je veľmi výkonný na analýzu údajov. V multidimenzionálnej tabuľke máme možnosť rýchlo a pružne meniť jednotlivé dimenzie a tak poskytovať rôzne pohľady pre používateľa.

V multidimenzionálnom modeli rozdeľujeme dva typy tabuliek - tabuľku faktov a tabuľku dimenzií, ktoré spolu tvoria schému.

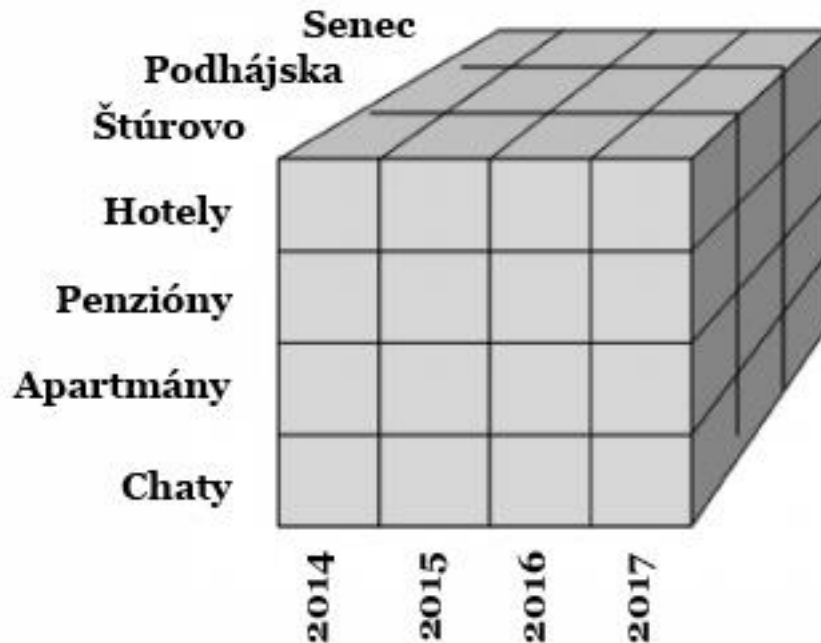
1.3.1 Fakty a Dimenzie

Fakty sú numerické merné jednotky obchodovania, ktoré používatelia analyzujú a sumarizujú s cieľom získať lepšie pochopenie svojho podnikania. Tabuľky faktov obsahujú záznamy obchodných faktov. Tabuľka faktov je hlavná tabuľka, na ktorú sú viazané tabuľky dimenzií. Tabuľky faktov obsahujú kľúče, ktoré zabezpečujú väzbu na tabuľky dimenzií a atribúty, ktoré objasňujú fakty. Každý kľúč v tabuľke faktov predstavuje väzbu. Tabuľka faktov uchováva veľké množstvo dát (má veľa záznamov) a je najväčšia tabuľka v databáze. Tabuľky s faktami sú úplne normalizované.

Dimenzia je konečná diskretná množina s mohutnosťou > 1 . Jednotlivé dimenzie sú funkčne nezávislé, napríklad čas a predajca. Dimenzie sú vo väčšine prípadov konštanty. Tabuľky dimenzií sú prevažne textové informácie o hodnotách, ktoré sú uložené v tabuľke faktov. Tabuľky dimenzií sú naviazané na tabuľku faktov, alebo na inú tabuľku dimenzií. Obsahujú atribúty popisujúce fakty. Tabuľky dimenzií neobsahujú toľko záznamov ako tabuľky faktov a dáta sa v nich nemenia tak často. Priesečník všetkých dimenzií v jednom bode je prvok multidimenzionálnej databázy.

Keďže v dátovom sklade sú uložené historické dáta, každý fakt má svoj časový moment vzniku. Preto je súčasťou schémy dátového skladu aj časová dimenzia. Medzi dimenzie zväčša patria ekonomické ukazovatele, predajca, organizačná jednotka, teritórium, zákazník, komodita a podobne. Tabuľky dimenzií sú nenormalizované.

Na obrázku nižšie môžeme vidieť OLAP kocku s dimenziami destinácia (Senec, Podhájska, Štúrovo), Typ ubytovania (Hotely, Penzióny, Apartmány, Chaty) a časovú dimenziu (2014, 2015, 2016, 2017). Na priesečníkoch môžeme sledovať ukazovatele - fakty, ako napríklad výška predaja, výška provízie, počet rezervácií. To znamená, že riešia napríklad nasledovnú otázku. Aké boli celkové predaje hotelov v Štúrove v roku 2017?



Obrázok 3 Multidimenzionálna OLAP kocka s tromi dimenziami

S rastúcim počtom rozmerov multidimenzionálnej databázy veľmi rýchlo rastú aj požiadavky na úložnú kapacitu. Okrem počtu rozmerov kocky na objem dát vplýva aj granularita dimenzií.

Dimenzie sú usporiadané v stromovej hierarchii, napríklad v časovej dimenzii: deň, mesiac, štvrťrok, rok. Každý člen má jedného rodiča a ani jedného alebo viac podriadených členov. Podradený člen patrí do ďalšej nižšej úrovne hierarchie. Prvý štvrťrok je nadradený člen mesiacu marec, a zároveň je podradeným člen roku 2015.

Granularity definujeme ako úroveň podrobnosti údajov uložených v tabuľke, v našom prípade dimenzie. Podrobnosti údajov závisia od granularity dimenzií. Nízka granularita nám neumožňuje pracovať s detailnými dátami, napríklad by sme v časovej dimenzii nemali dni, ani mesiace ale iba roky. Na druhej strane pri vysokej granularite nám exponenciálne narastá ukladací priestor.

1.3.2 Operácie v OLAP kocke

Základné operácie s dimenziami v OLAP kocke:

- DRILL DOWN - zostupovanie do nižších úrovní konsolidácie; posun v hierarchii na nižšiu úroveň a redukovanie dimenzií (napr. rok -> štvrťrok)
- ROLL UP - zlučovanie do vyšších úrovní konsolidácie; posun v hierarchii o úroveň vyššie alebo pridanie dimenzie - agregácia (napr. štvrťrok -> rok)
- SLICE - rez kockou v príslušnej dimenzii (napríklad jeden rok).
- DICE – výber podmnožiny kocky; zvolenie 2 dvoch a viacerých dimenzií (napríklad zvolíme určité obdobie a nejakú oblasť)

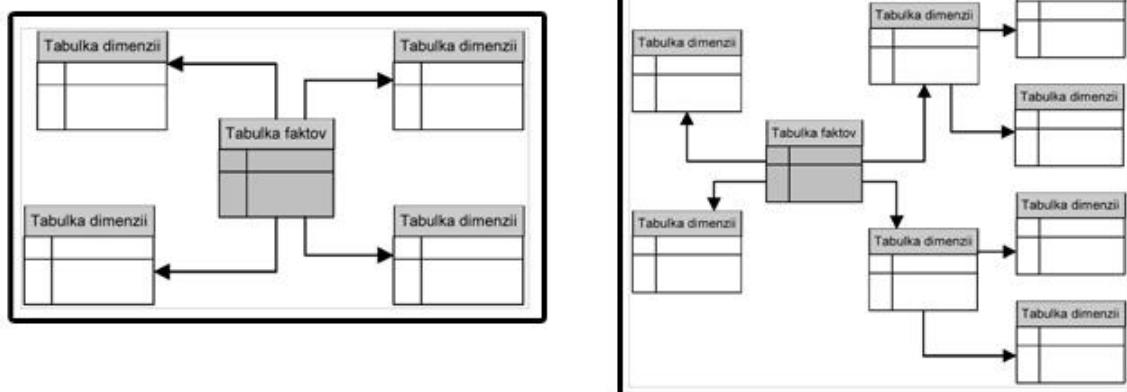
1.3.3 Schémy tabuliek dimenzií

Pri multidimenzionálnom modelovaní existujú tri možné dátové schémy. Ide v podstate o topologické usporiadanie tabuliek vzhľadom na vzťahy medzi nimi.

Základná schéma a zároveň aj najvhodnejšia je hviezdicová schéma (STAR). Pozostáva z jednej tabuľky faktov (môže mať aj viac) a niekoľkých tabuliek dimenzií, ktoré sú priamo napojené na tabuľku faktov. Tabuľky dimenzií nie sú v normalizovanej forme (vyznačujú sa redundanciou dát - duplicita). Je najvhodnejšou schémou spomedzi všetkých, pretože sa vyznačuje rýchlou odozvou na dotazy.

Model snehovej vločky (SNOW FLAKE) sa odlišuje od hviezdicovej tým, že sú niektoré tabuľky dimenzií dekomponované na viac dimenzií. Je výsledkom dekompozície jednej alebo viacerých dimenzií, ktoré majú hierarchie. Uloženie dát tejto schémy v databáze zaberá menej priestoru, nevýhodou je náročnejšia pre administráciu tabuliek a spomalenie rýchlosti odozvy na dotazy.

V schéme Rodina hviezd (FAMILY OF STARS) existuje viac tabuliek faktov, ktoré navzájom zdieľajú viac tabuliek dimenzií.



Obrázok 4 Schémy tabuliek dimenzií: hviezdiová schéma a schéma snehovej vločky

1.4 Samoobslužný Business Intelligence

Zavedenie Business Intelligence systému do podniku stojí nemalé financie a preto malé podniky doposiaľ takéto systémy nemali možnosť využívať. Vznik nových technológií, akým je samoobslužné BI (ang. SSBI - self-service BI), priniesli príležitosti aj menším podnikateľským subjektom.

Samoobslužný Business Intelligence je navrhnutý pre koncových používateľov, ktorí si vytvárajú svoje vlastné reporty a analýzy v rámci podporovanej architektúry a nástrojov [15]. Platformy SSBI sa vyznačujú ľahko použiteľnou funkčnosťou, ktorá podporuje celý analytický pracovný tok - od prípravy údajov až po generovanie prehľadov - s dôrazom na samoobslužné použitie a rozšírenú pomoc používateľa.

SSBI nástroje umožňujú užívateľom byť viac sebestačný a menej závislý na IT oddelení, a ponúkajú jednoduché narábanie s údajmi a relatívne jednoduché možnosti analyzovania a prívetivé užívateľské rozhranie. SSBI predstavuje v súčasnej dobe významný trend v oblasti Business Intelligence.

Samoobslužný BI je relatívne nový prístup k Business Intelligence, ktorý umožňuje menej technicky zdatným koncovým používateľom vykonávať analytiku dát [16]. Takto sa nemusia spoliehať na kvalifikované a skúsené profesionálne tímy.

SSBI umožňujú používateľom vytvárať svoje vlastné systémy dotazy a nastavenia BI analýz. K dispozícii je použitie vizualizácie údajov, ktoré usporiadajú informácie do ľahko použiteľných tabuliek a grafov.

1.5 Nástroje (softvéry) SSBI

Nástroje Business Intelligence uľahčujú úpravu údajov a ich vizualizáciu do podoby, v ktorej ich vieme lepšie pochopiť. Ako jednoduchý je tento proces a aké máme možnosti na vizualizáciu údajov, závisí od nástroja [17]. V súčasnej dobe existuje mnoho BI softvérových nástrojov, dôležitý je výber toho správneho nástroja vzhľadom na naše potreby.

Nástrojmi BI sú typy softvéru používaného na zhromažďovanie, spracovanie, analýzu a vizualizáciu veľkého množstva údajov s cieľom generovať použiteľné obchodné prehľady, vytvárať interaktívne správy a zjednodušiť rozhodovacie procesy [18].

Tieto nástroje nám ponúkajú funkcie, ako je vizualizácia údajov, grafické analýzy, interaktívne dashboards, zobrazovanie KPI ukazovateľov alebo aj prediktívne analytické funkcie. Zobrazené výsledky týchto nástrojov poskytujú užívateľom možnosť rýchlejšie a lepšie sa rozhodnúť a tým zvyšovať predaj, znižovať náklady, sledovať trendy na trhu a podobne, čo vedie k dosahovaniu väčšieho zisku.

Funkcie a výhody SSBI nástrojov:

- Zhromažďujú všetky dôležité údaje z rôznych nesúrodých zdrojov
- Samoobslužný prístup k BI analytike
- Prediktívna analýza
- Automatizácia procesov
- Znižovanie obchodných nákladov
- Sú neustále k dispozícii

1.6 Porovnanie BI nástrojov

Túto časť práce začíname výskumom celosvetovo uznávanej americkej spoločnosti Gartner, ktorá každoročne vykonáva nezávislé analýzy trhu v rôznych segmentoch IT technológií. Výsledky podrobnej analýzy sú zobrazené na sumárnom grafe, nazývanom Magic Quadrant.

Magické kvadranty nám pomáhajú:

- získať rýchly prehľad o konkurenčných poskytovateľoch technológií na trhu a ich schopnostiach plniť to, čo koncoví používatelia dnes a v budúcnosti požadujú
- pochopiť, aké majú poskytovatelia technológií na trhu v konkurenčnom postavení a ich stratégie, ktoré používajú v súťaži o koncových používateľov
- porovnať silné stránky a výzvy poskytovateľov technológie s vašimi konkrétnymi potrebami

Výsledný graf s dvoma dimenziami - kompletnosť vízie a schopnosť realizácie, je rozdelený na 4 časti (vodcovia, vizionári, špecializovaní hráči a vyzývateľia) do ktorých zaraďuje jednotlivé konkurenčné spoločnosti. Najdôležitejšiu úlohu zohrávajú spoločnosti v kvadrante lídri.

V správe Magic Quadrant pre Analytiku a Business Intelligence (Analytics & Business Intelligence), sa v kvadrante Lídri (Leaders) umiestnili 3 spoločnosti: Tableau, Qlik, Microsoft. Najvyššie z nich je umiestnená spoločnosť Microsoft s produktom Power BI.



Obrázok 5 Gartner Magic Quadrant pre Business Intelligence a analytické platformy 2021

Na základe tohto nezávislého porovnania sa v našom porovnaní venovať hlavne BI produktom „vít'azných“ spoločností.

1.6.1 Power BI

Power BI je nástroj Business Intelligence od spoločnosti Microsoft. Poskytuje interaktívne vizualizácie so samoobslužnými funkciami BI. MS Power BI je spoľahlivé a osvedčené riešenie BI, ktoré má veľkú a rastúcu používateľskú základňu. Výkonná cloudová služba Power BI Service umožňuje zdieľať poznatky z údajov.

Power BI ponúka možnosť získať (importovať) dáta z viac ako 40 rôznych zdrojov údajov a databáz. Spoločnosť Microsoft vyvíja Power BI rýchlym tempom, pričom nové verzie a funkcie sa vydávajú väčšinou každý mesiac.

V Power BI je k dispozícii veľké množstvo typov vizuálov, ktoré umožňujú vytvárať pútavé grafy, tabuliek, máp a vizualizácií údajov. Výhodou Power BI je, že sa tieto vizuály používajú veľmi ľahko a pomáhajú netechnickým používateľom rýchlo získať cenné reporty o chode podniku [23].

Microsoft Power BI si získalo pevnú pozíciu v organizáciách, ktoré majú málo ľudí vyškolených v oblasti dát, no napriek tomu majú veľkú potrebu analýzy údajov.

Bezplatná verzia Power BI Desktop vyhovuje samostatným používateľom; verzia Pro uľahčuje spoločnú analýzu za mesačný poplatok a využíva Microsoft Office365, SharePoint a Teams na riadenie prístupu k nespracovaným údajom a zverejneným správam. Pre podniky, ktoré chcú ísť all-in, umožňuje vrstva Premium samoobslužnú prípravu údajov pomocou vopred pripravených konektorov podnikových údajov uchovávaných v Microsoft Dynamics 365, Azure SQL Data Warehouse alebo v zdrojoch tretích strán, ako je napríklad Salesforce [17].

Plusy:

- Extrémne výkonná platforma s množstvom zdrojov údajov.
- Veľmi prívetivé užívateľské prostredie.
- Výnimočné možnosti vizualizácie údajov.
- Možnosť bezplatného používania na neobmedzený čas

Cena: Desktop: zadarmo; Pro: 9,99 dolárov na používateľa, mesačne; Premium: 20 dolárov za užívateľa, 4 995 dolárov za vyhradený cloudový výpočtový a úložný zdroj

1.6.2 Tableau

V súčasnosti jeden najpoužívanejší nástroj, Tableau Desktop bol jedným z prvých hráčov v segmente samoobslužných business intelligence (BI). Ponúka skvelé používateľské rozhranie (UI), ako aj bezplatnú štartovaciu cenu [3].

V súčasnej dobe začína cena od 35 dolárov za používateľa mesačne pre verziu pre stolné počítače a od 42 dolárov pre používateľa mesačne pre verziu online. Aj kvôli konkurencii bol nútený znížiť cenu, ktorá v minulosti (pred rokom 2017) činila 999 dolárov na používateľa na rok. V podstate je to veľmi stabilná aplikácia.

Tableau sa od roku 2018 pýši nástrojom Hyper, ktorý predstavuje nový rýchly dátový stroj. Tableau tvrdí, že poskytne používateľom až 5-krát vyššiu rýchlosť dopytovania v porovnaní s predchádzajúcimi iteráciami a až 3-krát vyššiu rýchlosť vytvárania extrahovania.

Nevýhodou Tableau Desktop je, že si vyžaduje dosť vysokú znalostnú úroveň používateľov. Pre skúseného analytika údajov je použitie aplikácie Tableau Desktop pomerne jednoduché, ale veľa nových používateľov bude vyžadovať dôkladné školenie, kým sa tento nástroj naučí. V konečnom verdikte je Tableau Desktop skvelou aplikáciou s množstvom funkcií.

Plusy:

- Veľký výber dátových zdrojov a vizualizácií.
- Užívateľsky prívetivý dizajn.
- Rýchly dátový stroj (rýchlosť spracovania dotazov).
- Starší produkt so širokou komunitou používateľov

Mínusy:

- Úplné zvládnutie platformy si vyžaduje dlhé školenie.

Cena: Každé nasadenie potrebuje najmenej jedného tvorca (70 dolárov mesačne); pre server Tableau Server (lokálny alebo verejný cloud) môžu byť ďalšími diváci (12 dolárov mesačne, minimálne 100) alebo prieskumníci (35 dolárov mesačne, minimálne 5); alternatívne, s Tableau Online (plne hosťiteľským spoločnosťou Tableau), môžu byť ďalšími diváci (15 dolárov mesačne, minimálne 100) alebo prieskumníci (42 dolárov mesačne, minimálne 5)

1.6.3 Qlik

Qlik Sense, samoobslužný nástroj na prístup k analytickým funkciám, je dostupný v cloudovej a lokálnej verzii.

Spoločnosti tento softvér hodnotia veľmi pozitívne. Najväčšími pozitívami je rýchle nasadenie systému, a vďaka asociatívnemu modulu dokáže Qlik Sense veľmi rýchle zostaviť komplexné dátové modely a vytvoriť dashboardy. Ďalším kladom je, dokáže rýchle načítať údaje zo zdrojov. Na druhej strane cenový a licenčný model nedovoľuje rozšírenia na ďalších používateľov. Najviac funkcií ponúka verzia SaaS. Užívatelia rôzne hodnotia grafické rozhranie, buď tvrdia je intuitívne ale veľmi neprívetivé hlavne pre analytikov a správcov. Softvér dokáže rýchlo vykonať zložité výpočty aj so sadou veľkých údajov.

Qlik dosahuje najlepšie výsledky, keď je potrebné agregovať štruktúrované údaje z mnohých rôznych zdrojov údajov. Celková skúsenosť so samotným Qlikom bola úžasná. Niektorí používatelia ho hodnotia ako dosť náročné na učenie, hlavne pre tých, ktorí predtým pracovali s QlikView. Tento BI nástroj má veľmi strmú krivku učenia.

Plusy:

- Rýchla analýza vďaka asociatívnemu modulu
- Dáta sa čítajú rýchlo a príjemne
- Všetky kanály spoločnosti je možné skontrolovať

Mínusy:

- Pridávanie a odstránenie variantov
- Niektoré správy nie je možné stiahnuť do programu Excel

Cena: Obmedzené verzie: zadarmo; Qlik Sense Business: 30 dolárov na používateľa mesačne; Qlik Sense Enterprise SaaS: Analyzer: 40 dolárov mesačne; Profesional: 70 dolárov mesačne; Balíky kapacity Analyzer, 100 dolárov mesačne

1.7 Power BI

V tejto časti sa ešte bližšie oboznamujeme s aplikáciou MS Power BI, ktorú sme vyhodnotili ako najvhodnejšiu alternatívu pre využitie v praktickej časti tejto práce.

Aplikácia Power BI od spoločnosti Microsoft bola vytvorená ešte v roku 2011 pod iným názvom, v roku 2013 už pod súčasným názvom predstavená ako súčasť balíka Office 365, no pre širokú verejnosť je dostupná až od júla 2015 [2]. Prvé vydanie bolo založené na doplnkoch MS Excelu: Power Query, Power Pivot a Power View. Neskôr boli pridané ďalšie funkcie.

Power BI predstavuje kompletnú platformu samoobslužnej Business Intelligence. V aplikácii je možné dáta importovať zo širokého spektra zdrojov, a následne ich upraviť jednoduchým spôsobom vytvoriť reporty a tieto publikovať. Power BI si nemusíte kupovať, stačí si stiahnuť Aplikáciu Power BI Desktop a nainštalovať v operačnom systéme Windows. Nová verzia aplikácie vychádza každý mesiac, pričom aktualizácie sa prispôbujú potrebám užívateľov. Power BI je stabilná aplikácia, v ktorej môžete pokojne analyzovať veľké množstvo záznamov v tabuľke. Aplikácia je dostupná aj v slovenskej verzii.

Microsoft Power BI ponúka interaktívne vizualizácie s prepracovaným integrovaným dotazovaním a modelovaním dát. Zostavy a reporty môžeme nielen vytvárať, ale aj publikovať v službe Power BI a zároveň zdieľať s ostatnými užívateľmi.

Aplikácia Power BI Desktop ponúka nasledovné procesy:

- Extrakcia – získanie údajov zo zdrojov
- Transformácia – úprava a čistenie údajov
- Modelovanie – vytváranie relácie medzi tabuľkami
- Kalkulácia – vytváranie kľúčových ukazovateľov výkonnosti

- Vizualizácia – vytváranie zostáv a prehľadov
- Distribúcia – publikovanie zostáv

1.7.1 Komponenty Power BI

Power BI sa skladá z týchto komponentov [2]:

- Power BI Desktop,
- Power BI cloud / služba Power BI,
- Power BI Report Server,
- Power BI Mobile,
- Power BI Data Gateway,
- Power BI Dataflows,
- Power BI for Mixed Reality,
- Power BI Publisher for Excel,
- Power BI Embedded,
- Power BI Premium,
- Power BI API.

Power BI Desktop je centrálna aplikácia určená pre počítače s operačným systémom Windows. V nej sa pripravujú dáta a vytvárajú dátové modely a zostavy a potom následne publikujú výsledky. Ponúka najširšie možnosti práce s údajmi zo všetkých komponentov. Používanie Power BI Desktop je úplne zadarmo. Aplikáciu si môžete stiahnuť webovej stránky Microsoftu.

Aplikácia Power BI Desktop sa skladá z pôvodne excelovských doplnkov Power Query, PowerPivot a Power View [2]. Na používanie Power BI Desktop nie je potrebný MS Office. Aktualizácie sú k dispozícii väčšinou každý mesiac.

Cloudová služba Power BI, v súčasnej dobe na webovej adrese app.powerbi.com, je súčasťou SaaS (softvér vo forme služby) služby Power BI [22]. Pôvodne bola vytvorená ako služba pre publikovanie a zdieľanie reportov Power BI [2]. Vytvorené zostavy v aplikácii Power BI Desktop sú publikované v službe Power BI, kde ich môžete ďalej upravovať. Služba ponúka aj ďalšie funkcionality, ako je vytváranie manažérskych tabúl, automatická aktualizácia dát, streamovanie údajov do reportov a posielanie reportov.

Power BI Report Server pre publikovanie a zdieľanie reportov na lokálne servery s možnosťou prejsť do cloudu. Je súčasťou Power BI Premium. Ponúka možnosť nastaviť prístupy pre užívateľov či ich skupiny.

Mobilná aplikácia Power BI Mobile ponúka možnosť pripájať sa a pracovať na mobilných zariadeniach s údajmi v cloudu. Aplikácia je dostupná zdarma pre Windows 10, Android a iOS.

Power BI Embedded umožňuje vložiť obsah Power BI, ako sú napríklad prehľady, a dashboardy, do webovej aplikácie alebo na web.

Pomocou rozhrania Power BI REST API môžete vytvárať aplikácie, ktoré sa integrujú do zostáv, dashboardov Power BI.

1.7.2 Verzie Power BI

V súčasnosti v Power BI máte 3 možnosti založenia účtov:

- Power BI Free
- Power BI Pro
- Power BI Premium pre užívateľa / pre kapacitu

Power BI Free je účet zadarmo, ktorý si môže založiť každý s firemnou mailovou adresou. Má však obmedzenia - napr. limit dát je 1 GB na užívateľa, možnosť aktualizovať dáta 1- krát denne, reporty nie je možné zdieľať, môžete ale zaslať odkaz na zostavu.

Power BI Pro má vyššie limity a obsahuje viac funkcionality ako Power BI Free. Ponúka možnosť aktualizovať dáta 8 krát denne, 100 GB pre úložný priestor a možnosť publikovať a zdieľať reporty,

Služba Power BI Pro je určená pre používateľov, ktorí zverejňujú zostavy, zdieľajú tabule, spolupracujú s kolegami v pracovných priestoroch a vykonávajú iné súvisiace činnosti, ako napríklad:

- Úprava a uloženie prispôbených zobrazení
- Vytváranie osobných tabulí (pripnutie na novú tabuľu)
- Analýza údajov v Exceli alebo Power BI Desktop

- Zdieľanie s podporou aplikácie Excel Web App
- Zdieľanie tabúl a spolupráca pomocou skupín v Microsoft 365
- Integrácia obsahu so službou Microsoft Teams

Dá sa objednať buď samostatne (cena okolo 9,99 dolárov na užívateľa za mesiac). Nevýhodou je, že svoje údaje máte uložené v cloude aj s ďalšími používateľmi, čo niekedy znamená pomalšiu odozvu.

Power BI Premium ponúka v súčasnosti už dve varianty, a to Premium per Capacity (podľa kapacity) a per User (podľa užívateľa).

Predplatená služba Premium per Capacity je určená pre väčšie podniky a umožňuje zjednodušiť správu údajov a prístup v podnikovej mierke. V službe si prenajímate celé virtuálne servery len pre svoju firmu. Služba ponúka väčší rozsah a výkon na obsah služby Power BI. Každá organizácia má výhradne pre seba pridelenú kapacitu, podporuje vyhradený hardvér spravovaný spoločnosťou Microsoft.

S predplateným Power BI Premium per Capacity získavate [23]:

- väčší rozsah a výkon na zostavy služby Power BI;
- flexibilitu licencovania podľa kapacity;
- špičkové funkcie na vizualizáciu údajov (analýza s podporou umelej inteligencie a pod.),
- zjednotenie samoobslužných a podnikových funkcií BI so širokou škálou možností dostupných len vo verzii Premium;
- vstavanú licenciu na rozšírenie služby o Power BI Report Server;
- podporu umiestnenia údajov podľa oblastí (geografická distribúcia serverov);
- možnosť zdieľať obsah služby Power BI s kýmkoľvek bez zakúpenia licencie na používateľa.

Službu Power BI Premium si môžu objednať aj jednotlivci. Licencia Premium na používateľa (PPU) stojí približne 20 dolárov mesačne a obsahuje všetky funkcie Pro používateľa a ďalšie možnosti, ako je rýchlejší prístup k prehľadom pomocou pokročilých AI, ako aj samoobslužnú prípravu pre veľké údaje.

1.7.3 Power Query

Editor dotazov Power Query je súčasťou aplikácie Power BI Desktop a slúži na transformáciu a prípravu údajov, a to ešte pred ich načítaním do Power BI. To znamená, že Power Query nahrádza fázu ETL, čo znamená že môžete pomocou neho vykonať extrahovanie, transformáciu a načítanie údajov. Umožňuje nám vykonávať úpravy, ako výber riadkov a stĺpcov, zlúčenie alebo rozdelenie riadkov a stĺpcov, mazanie stĺpcov, transponovanie tabuľky a podobne. V Power Query sa môžeme pripojiť k širokej škále zdrojov údajov [21].

Power Query Editor má vlastné grafické rozhranie, ktoré pomocou ponúk a tlačidiel s mnohými mnoho preddefinovanými funkciami transformácie uľahčuje prácu s údajmi. Každý krok transformácie sa automaticky ukladá do zoznamu Použité kroky a s ním aj kód skriptovacieho jazyka M, ktorý sa používa v Power Query na transformáciu údajov. Jednotlivé kroky si môžeme premenovať podľa potreby a upravovať aj spätne. Rovnako môžeme upravovať aj vytvorený kód v jazyku M, čím vlastne upravíme aj proces transformácie. Používame to v prípade, že potrebujeme urobiť pokročilé transformácie, ktoré nám grafické rozhranie neumožňuje. Prístup k skriptu jazyka M sa dostaneme cez funkciu Rozšírený editor.

V súčasnej dobe sú k dispozícii dve prostredia Power Query, a to online verzia prostredníctvom webovej stránky a desktop verzia pre Excel a Power BI Desktop. Obidve verzie ponúkajú rovnakú funkcionality.

1.7.4 DAX

DAX (Data Analysis Expressions) je programovací jazyk, ktorý sa používa v aplikáciách Analysis Services, Power BI Desktop a Power Pivot v programe Excel [20]. Bol vytvorený v roku 2010 s prvým vydaním PowerPivot pre Excel 2010 [10].

Pomocou jazyka DAX je možné vytvoriť tri typy objektov:

- vypočítané tabuľky,
- vypočítané stĺpce,
- mierky.

Ohľadom vypočítaných tabuliek sa tento výpočet väčšinou používa na pridanie tabuľky dátumov. Výrazom jazyka DAX môžeme údaje aj duplikovať alebo transformovať. Vzorce v jazyku DAX obsahujú funkcie, operátory a hodnoty, prevažne údajov súvisiacich tabuliek a stĺpcov z dátového modelu. Takže kedykoľvek píšete výrazy DAX, budú sa odkazovať iba na tabuľky a stĺpce z dátového modelu, a nie na konkrétne bunky ako je to vo vzorcoch Excelu [10].

Jazykom DAX môžete vytvárať jednoduché výpočty, ale aj zložité koncepty, ako sú kontexty hodnotenia, iterácie a prechod kontextu. Jednoduchý vzorec v DAX-e vyzerá nasledovne:

$$\text{Celkový predaj} = \text{SUM}([\text{Suma predaja}])$$

V tomto vzorci sme do mierky Celkový predaj priradili sumu hodnoty stĺpca Suma predaja. Ak túto mierku vložíte do zostavy a vo vizuále ju dáte filtrovať podľa napríklad kategórie produktu, tak sa Celkový predaj vypočíta pre každú kategóriu. Príklad ďalšieho vzorca:

$$\text{Celkový predaj} = \text{Predaj}[\text{Cena Produktu}] * \text{Predaj}[\text{Predané Množstvo}]$$

Týmto vzorcom sme vytvorili mierku Celkový predaj, ktorú sme vypočítali násobením hodnoty v stĺpci Cena Produktu v tabuľke Predaj a hodnoty v stĺpci Predané Množstvo v tabuľke Predaj.

Načo vlastne v Power BI potrebujeme jazyk DAX? V zostavách si vieme pomerne jednoducho vo vizuáloch zobrazit' naše údaje aj bez použitia vzorcov. Problém ale prichádza, keď by sme chceli zobrazit' napríklad medziročný percentuálny nárast celkového predaja, alebo napríklad percentuálny podiel kategórie produktu na celkovom predaji. Tieto hodnoty si vieme pomocou jazyka DAX pomerne jednoducho vypočítat'.

Medzi najbežnejšie funkcie jazyka patria SUM (Sčíta všetky čísla v stĺpci), COUNT (počíta počet buniek v stĺpci), MIN (Vráti najmenšiu číselnú hodnotu), MAX (Vráti najväčšiu číselnú hodnotu). Každá z týchto funkcií má ekvivalentnú iteračnú funkciu: SUMX, COUNTX, MINX, MAXX.

CALCULATE - Vyhodnotí výraz v kontexte upraveného filtra.

DIVIDE - Vykoná delenie a vráti alternatívny výsledok pri delení nulou

ADDCOLUMNS – Pridá vypočítané stĺpce k danej tabuľke

FILTER - Vrátí tabuľku, ktorá predstavuje podmnožinu inej tabuľky alebo výrazu.

CALENDAR, CALENDARAUTO - Vrátí tabuľku s jedným stĺpcom s názvom Date, ktorý obsahuje množinu dátumov nasledujúcich za sebou

USERELATIONSHIP - Určuje vzťah, ktorý sa má použiť v konkrétnom výpočte

RELATED - Vrátí súvisiacu hodnotu z inej tabuľky

DATEADD - Vrátí tabuľku, ktorá obsahuje stĺpec dátumov, posunutých dopredu alebo dozadu v čase o zadaný počet intervalov od dátumov v aktuálnom kontexte.

IF – Podmienený príkaz

DATEDIFF - Vrátí počet ohraničení intervalu s prienikom dvoch dátumov.

2 Cieľ práce, metodika práce a metódy skúmania

Cieľom práce je analyzovať údaje z databázy cestovnej agentúry Hepex-Slovakia, ktorá na webovej stránke ponúka možnosť rezervácie ubytovania a pobytov v ubytovacích zariadeniach na Slovensku. Cestovná agentúra je na trhu už viac ako 20 rokov. V poslednom období je v tomto segmente na trhu stále väčšia konkurencia a v tejto situácii je dôležité poznať priority zákazníkov, aby sa mohla agentúra zamerať na predaj tých najžiadanejších pobytov.

Cestovnú agentúru zaujíma celkový trend predaja pobytov, a to z pohľadu jazykových verzíí, destinácií, krajín klientov a typov pobytov. Pri type pobytov sú to parametre ako napríklad dĺžka pobytu, počet osôb, nástupný deň, alebo aj s akým predstihom si klienti rezervovali pobyt.

Je potrebné vytvoriť reporty na tieto najčastejšie otázky:

- Aká bola výška celkových tržieb a aký bol trend vývoja tržieb podľa krajín klientov, jazykových verzíí ?
- Ktoré destinácie mali najväčšie predaje?
- Ktoré ubytovacie zariadenia sa najviac predávali ?
- Aké boli celkové tržby v rôznych cenových kategóriách za osobu?
- Aké boli tržby v rôznych cenových kategóriách pobytov?
- Aké typy pobytov (počet osôb, dĺžka pobytu, nástupný deň) priniesli najvyššie tržby?

Cestovná agentúra ukladá údaje o rezerváciách do databázy MySQL, z ktorej exportujeme potrebné údaje. Tie následne načítavame a upravujeme do použiteľnej podoby v Power Query. Dashboardy a reporty vytvárame v aplikácii Power BI Desktop aj za pomoci jazyka DAX, ktorým vytvárame pomocné mierky.

Po vytvorení reportov je posledným krokom publikovanie výsledného reportu z Power BI Desktop do Power BI Service, kde je možné report zdieľať s ďalšími osobami.

3 Výsledky práce a diskusia

V tejto kapitole pristúpime k realizácii samotnej implementácie. Postupne popíšeme všetky kroky, ktoré povedli k vytvoreniu dashboardom v aplikácii Power BI Desktop. Začali sme načítaním surových dát, ich úpravou, a pokračovali sme vytvorením dátového modelu, naprogramovaním mierok pomocou DAX-u až nakoniec vytvorením výsledných reportov.

Počas vytvárania reportu sme využívali teoretický základ získaný z predošlých kapitol. Jednotlivé kroky postupu tvorby reportu sú chronologicky usporiadané do podkapitol.

3.1 Dataset

Údaje boli vyexportované z relačnej databázy MySQL a uložené v súboroch do formátu CSV a v tabuľkách do formátu Microsoft Excel. Tieto dáta je potrebné nahrať do aplikácie Power BI Desktop, kde s nimi ďalej pracujeme.

Ešte si povedzme niečo o údajoch, ktoré sme analyzovali. Dataset obsahuje 8 súborov vo formáte CSV a XLSX, z nich každý reprezentuje jednu tabuľku.

- reservations_accommodations.csv (tabuľka rezervácií – tabuľka faktov)
- customers_countries.csv (tabuľka krajín klientov)
- accommodations.csv (tabuľka ubytovacích zariadení)
- accommodations_types.csv (tabuľka typov ubytovacích zariadení)
- cities.csv (tabuľka miest a obcí - destinácií)
- stars.csv (tabuľka kategórií ubytovania - hviezdičky)
- location.xlsx (upravená tabuľka obcí – destinácií, doplnená o okresy a kraje)
- accom_city.csv (pomocný súbor slúžiaci na spárovanie ubytovacích zariadení s obcami)

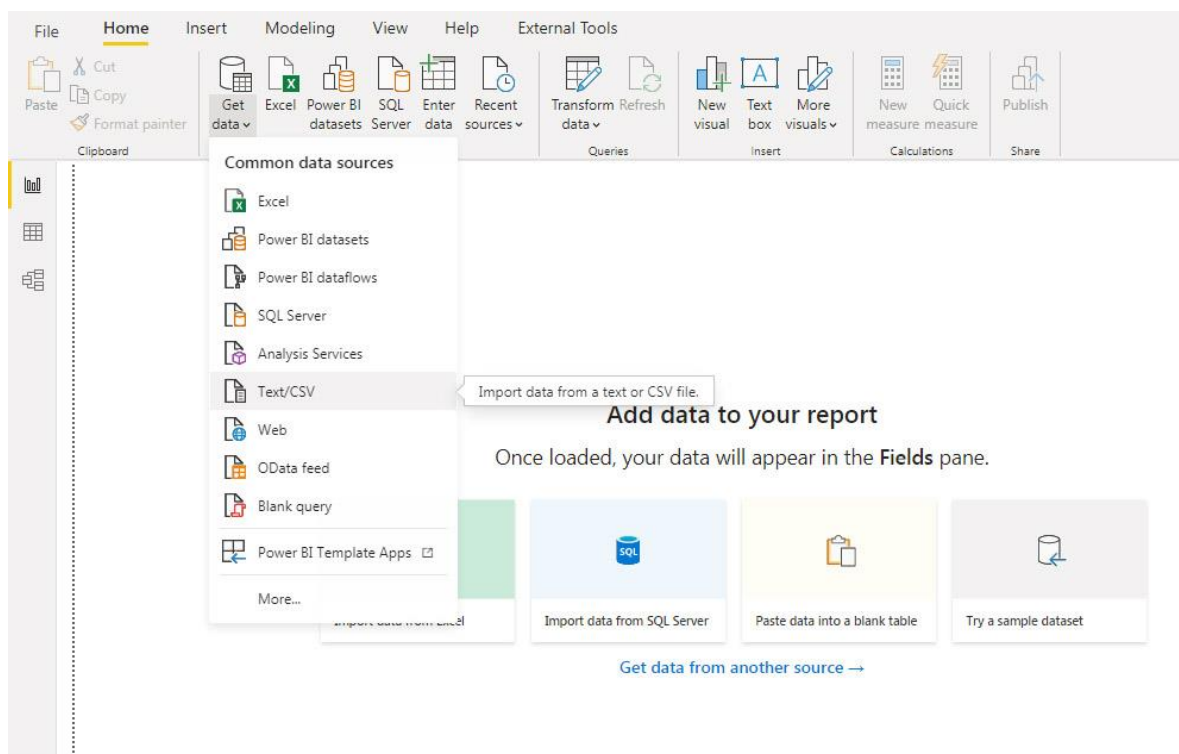
Tabuľka faktov pozostáva z viac ako 165.000 riadkov. V stĺpcoch tabuľky sú údaje ako ID rezervácie, počet dospelých, počet detí, dátum rezervácie, dátum začiatku pobytu,

dátum konca pobytu, typ stravy, ID ubytovacieho zariadenia, krajinu klienta a jazykovú verziu webovej stránky.

Jedna z najhlavnejších tabuliek dimenzií je tabuľka ubytovacích zariadení, ktorá obsahuje nasledovné údaje: ID ubytovania, názov, obec, počet hviezdčiek, typ ubytovania, cena za osobu.

3.2 Načítanie zdrojových súborov

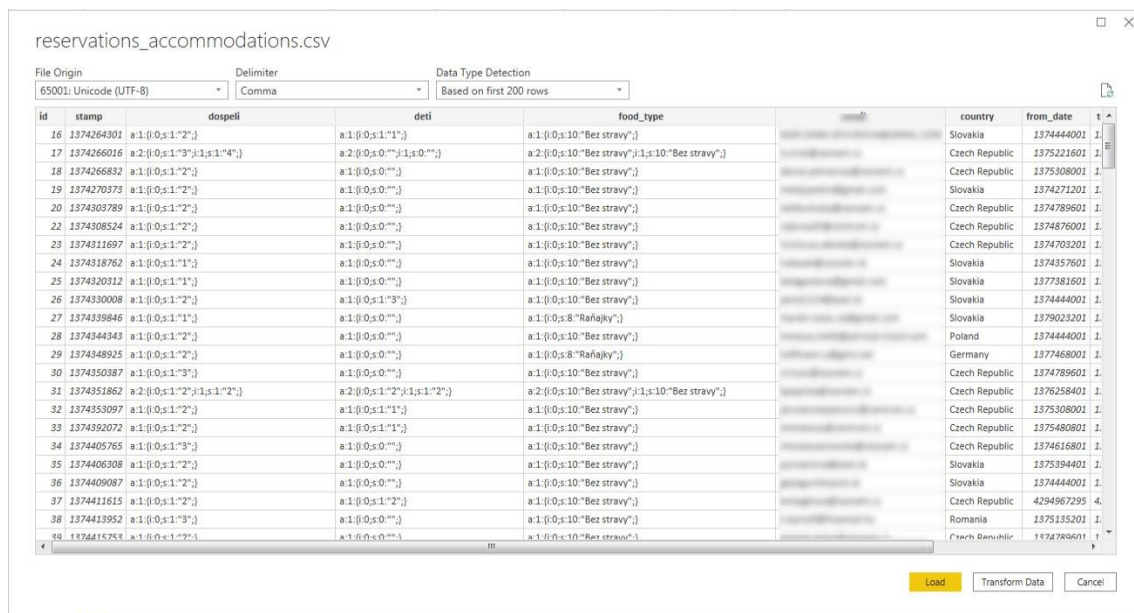
Po spustení aplikácie Power BI Desktop sme vybrali z ponuky možnosť Get Data (Získať údaje). Následne sa nám zobrazil zoznam dostupných zdrojov dát, ktoré Power BI podporuje (vid' Obrázok).



Obrázok 6: Power BI - načítanie údajov (autor, 2021)

Ako prvú si načítame tabuľku rezervácií, ktorá nám bude tvoriť základ pre tabuľku faktov. Po vybraní možnosti Text/CSV a následnom zvolení nášho CSV súboru v sprievodcovi sa zobrazí náhľad údajov vo zvolenom súbore. Podľa potreby si zvolíme parametre načítania súboru (kódovanie súboru, oddeľovač).

Po kliknutí na *Načítať údaje* sa nám načítajú údaje do Dátového modelu v Power BI. No keďže vieme, že tieto údaje je potrebné upraviť pre naše potreby, kliknutím na *Transformovať údaje* sa nám súbor načíta do Power Query, kde ich ďalej upravujeme.



Obrázok 7: Náhľad CSV súboru (autor, 2021)

3.3 Príprava dát v Power Query

V tejto časti si popíšeme transformáciu a prípravu dát exportovaných z databázy. Cieľom tejto prípravy je spracovanie dát do takej podoby, v ktorej s nimi budeme vedieť pracovať v Power Bi a vytvárať potrebné vizuály. Transformovanie dát si ukážeme na tabuľke rezervácií. Hlavná časť transformácie sa bude diať v Power Query.

3.3.1 Tabuľka rezervácií

Naša tabuľka rezervácií pozostáva z nasledovných stĺpcov: ID rezervácie, čas vykonania rezervácie, počet dospelých, počet detí, typ stravy, krajina klienta, dátum nástupu na pobyt, dátum konca pobytu, jazyková verzia webovej stránky, ID ubytovacieho zariadenia.

Hneď po načítaní dát do Power Query si aplikácia automaticky urobí vlastné kroky úprav, a tým je zvolenie si prvého riadku pre názvy stĺpcov a tým druhým je identifikovanie typu údajov pre stĺpce. Je potrebné skontrolovať, či aplikácia zvolila správne typy údajov.

Ako prvý problém budeme riešiť stĺpce počet dospelých a počet detí. Ako môžeme vidieť obidve stĺpce máme uložené ako serializované pole.

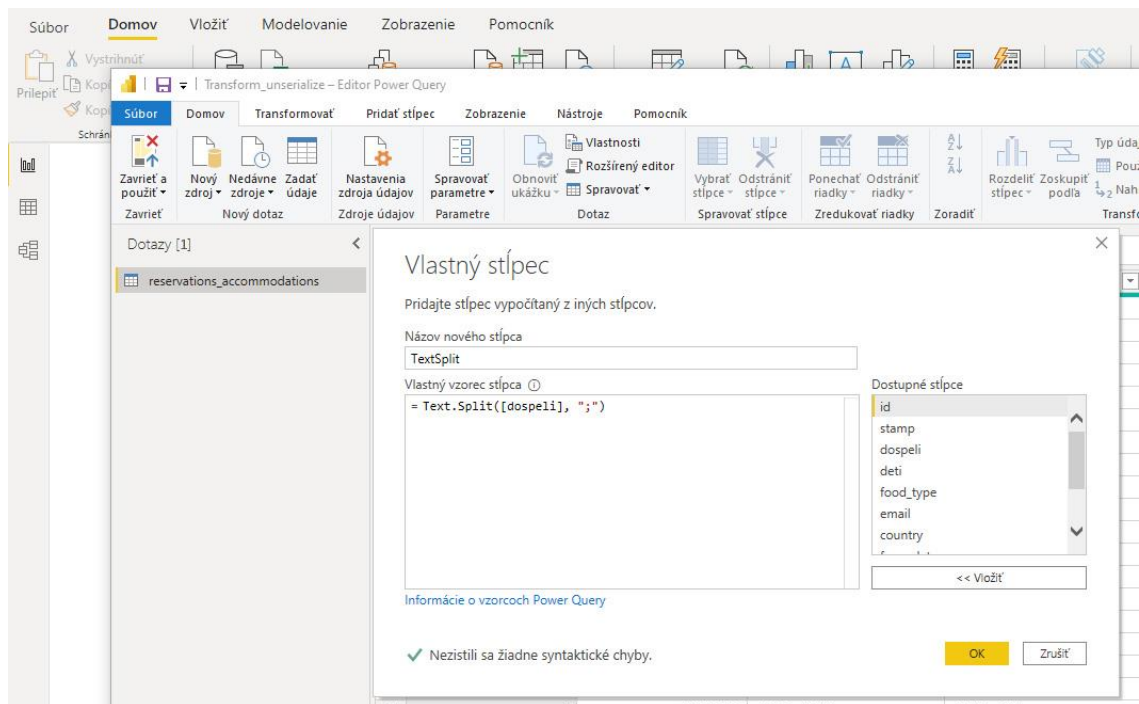
```
a:2:{i:0;s:1:"3";i:1;s:1:"4";}
```

Keďže sme databázu exportovali priamo v PHP adminke bez pomoci skriptovacieho jazyka PHP, nemali sme pri exporte možnosť získať spracované neserializované pole. Toto naše ukážkové serializované pole pozostáva z dvoch položiek 3 a 4, ktoré je potrebné sčítať, čím je náš výsledok 7. Pokúsime sa pomocou jazyka M priamo v Power Query vyčistiť tieto údaje.

Nejakým postupným spôsobom si vytvoríme funkciu, ktorá nám poslúži na pretransformovanie údajov. Budeme si pomocou funkcií v jazyku M postupne transformovať stĺpec a z týchto čiastkových funkcií nakoniec vytvoríme funkciu s parametrom, ktorú potom aplikujeme na náš pôvodný stĺpec. Keď sa pozrieme náš serializovaný výraz, vidíme, že čísla, ktoré potrebujeme spočítať, máme v úvodzovkách. Výraz si môžeme najskôr rozdeliť, ako rozdeľovač zvolíme bodkočiarku. Takže si vytvoríme nový stĺpec, ktorý bude obsahovať zoznamy ako polia, ktorého zložky budú vytvorené z nášho serializovaného výrazu rozdeleného bodkočiarkami.

Zvolíme si záložku *Add Column* (Pridať stĺpec) a vyberieme *Custom Column*. Zobrazí sa okno, kde zadáme názov nového stĺpca a funkciu, ktorá nám rozdelí výraz. V tomto prípade použijeme funkciu *Text.Split*. Ako prvý argument bude názov stĺpca, ktorý chceme rozdeliť a druhým argumentom bude oddeľovač, v našom prípade bodkočiarka.

```
= Text.Split([dospeli], ";")
```



Obrázok 8 Výsledný stĺpec si nazveme *TextSplit* (autor, 2021)

Takýto výraz `a:2:{i:0;s:1:"3";i:1;s:1:"4"};` nám pretransformuje na takýto zoznam:

List
<code>a:2:{i:0</code>
<code>s:1:"3"</code>
<code>i:1</code>
<code>s:1:"4"</code>
<code>}</code>

Takto podobne budeme postupovať aj v ďalších krokoch. To znamená, že si zakaždým na záložke *Pridať Stĺpec* zvolíme možnosť *Vlastný stĺpec*.

V zozname si teraz potrebujeme ponechať iba tie položky, ktoré obsahujú úvodzovky, pretože v týchto máme uložené počty, ktoré potrebujeme. Na túto úpravu použijeme funkciu *List.FindText*, ktorá ponechá z pôvodného zoznamu iba tie položky, ktoré obsahujú zadanú hodnotu. Funkcia vyžaduje dva parametre. Prvým parametrom je vstupný zoznam, respektíve názov stĺpca, a druhým parametrom je hľadaná hodnota, v našom prípade úvodzovka. Výsledný stĺpec si nazveme *TextSplitSelected*.

= *List.FindText([TextSplit], """)*

Výsledkom je nasledovný zoznam.

List
s:1: "3"
s:1: "4"

V ďalšom kroku sme si ponechali iba obsah položiek v úvodzovkách. Na túto transformáciu použijeme funkciu *Text.BetweenDelimiters*, ktorá ponechá text medzi dvomi zadanými hodnotami. Tými našimi hodnotami budú úvodzovky. No keďže my máme každý riadok ako zoznam, musíme funkcii povedať aby túto operáciu urobila so všetkými položkami zoznamu. Výsledný stĺpec si nazveme *ListTransform*. Výsledná funkcia bude vyzeráť nasledovne:

= *List.Transform([TextSplitSelected], each Text.BetweenDelimiters(_, """, """))*

Dostaneme zoznam:

List
3
4

V poslednom kroku nám ostáva už len spočítať jednotlivé položky zoznamu. Na to nám posluží funkcia *List.Accumulate*, ktorá nahromadí súhrnnú hodnotu z položiek v zozname. Funkcia vyžaduje tri parametre. Prvým je vstupný zoznam, druhým parametrom je počiatočná hodnota a tretím parametrom je funkcia hromadenia, respektíve transformácie. Aby sme sa predišli chybovým hláseniam z dôvodu chybných vstupných údajov, použijeme funkciu *try* a *otherwise*. V prípade, ak zaznamená funkcia *try* akúkoľvek chybu počas vykonávania funkcie *List.Accumulate*, výsledkom bude číslo 0.

Výsledná funkcia bude vyzeráť nasledovne:

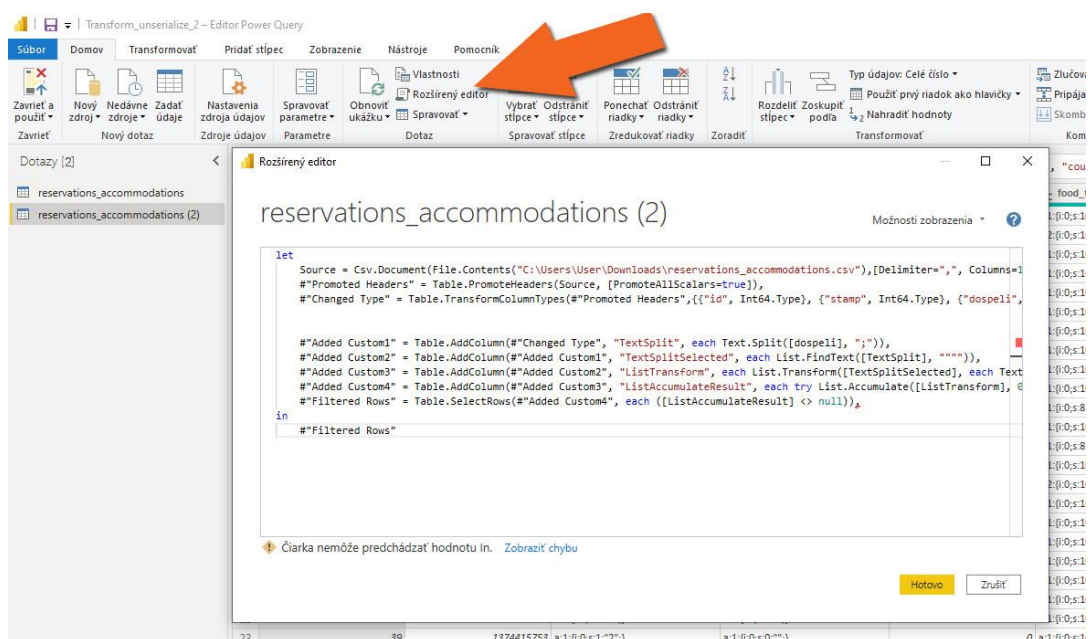
= try List.Accumulate([ListTransform], 0, (state, current) => Number.From(state) + Number.From(current)) otherwise 0

Nakoniec sme si vo výslednom stĺpci ponechali iba hodnoty, ktoré sa nerovnjajú hodnote *null*.

Všetky tieto kroky, ktoré sme vytvorili na transformáciu stĺpca *Počet dospelých*, ktorý bol uložený vo forme serializovaného poľa, môžeme vidieť v ľavej časti, kde sa nám ukladajú *Použité kroky*. V Power Query sa snažíme čím viac zredukovať počet krokov, aby sa skrátil čas transformácie našich údajov. My ale potrebujeme rovnakým spôsobom transformovať ešte jeden stĺpec, a to *Počet detí*.

Aby sa nám nezdvajnosobil počet vykonaných krokov a aby sme sprehľadnili vykonané úkony, vytvoríme si transformačnú funkciu, ktorú budeme aplikovať na tieto dva stĺpce. Tieto stĺpce budú vstupnými parametrami transformačnej funkcie. Na vytvorenie tejto funkcie použijeme kroky, ktorými sme transformovali stĺpec *Počet dospelých*.

Vytvorenie transformačnej funkcie začneme tým, že duplikujeme aktuálny dotaz, ktorý sa nám vytvoril načítaním CSV súboru a následnými úpravami nášho súboru. Nový dotaz ideme pretvoriť na funkciu s parametrom. Na karte *Zobrazenie* klikneme na položku *Rozšírený editor*, ktorý nám zobrazí kód zvoleného dotazu. Rozšírený editor nám umožňuje vytvoriť vlastný kód tvarovania.



Obrázok 9 Zobrazenie rozšíreného editoru v Power Query (autor, 2021)

Celý kód tohto dotazu môžete vidieť v prílohe. Zobrazený kód v jazyku M reprezentuje načítanie a úpravy nášho súboru CSV. Tento kód pozostáva z jednotlivých krokov transformácie, ktoré sú uzatvorené vo výraze *let ... in*. Každý riadok predstavuje jednu úpravu, pričom výsledok každého riadku sa priraďuje do premennej. Tú premennú použijeme v ďalšom kroku ako vstupný parameter. Keďže tieto premenné sa skladajú z viacerých slov, tak začínajú znakom # a sú uzatvorené do úvodzoviek. Konečný výsledok, naša výstupná hodnota, nasleduje po výraze *in*. Je to vlastne názov premennej, do ktorej je priradený výsledok poslednej úpravy.

V našom prípade prvé tri riadky hneď po výraze *let* nebudeme v našom transformačnom dotaze potrebovať. Prvý riadok predstavuje načítanie súboru, kde môžeme vidieť cestu k súboru a nastavenie ďalších parametrov načítania. Druhý riadok určuje prvé riadky ako hlavičky stĺpcov. Tretí riadok vykonáva automatickú zmenu typov údajov. Tieto prvé tri riadky môžeme vymazať.

Ďalšie štyri riadky predstavujú transformáciu nášho stĺpca. Všetky štyri riadky vytvárajú nový stĺpec. My si ich prerobíme tak, aby nový stĺpec nevytvárali, ale aby čiastkový výsledok ukladali do premenných. Funkcii v prvom riadku ale potrebujeme určiť tiež vstupný parameter a tým by mal byť názov stĺpca, ktorý chceme touto funkciou transformovať. Zapišeme to tak, že názov stĺpca bude vstupným parametrom tejto transformačnej funkcie.

Výsledný kód nášho transformačného dotazu bude vyzeráť nasledovne:

```
(InputColumn as text) =>
```

```
let
```

```
TextSplit = Text.Split(InputColumn, ";"),
```

```
TextSplitSelected = List.FindText(TextSplit, """"),
```

```
ListTransform = List.Transform(TextSplitSelected, each Text.BetweenDelimiters( _ ,  
"""" , """")),
```

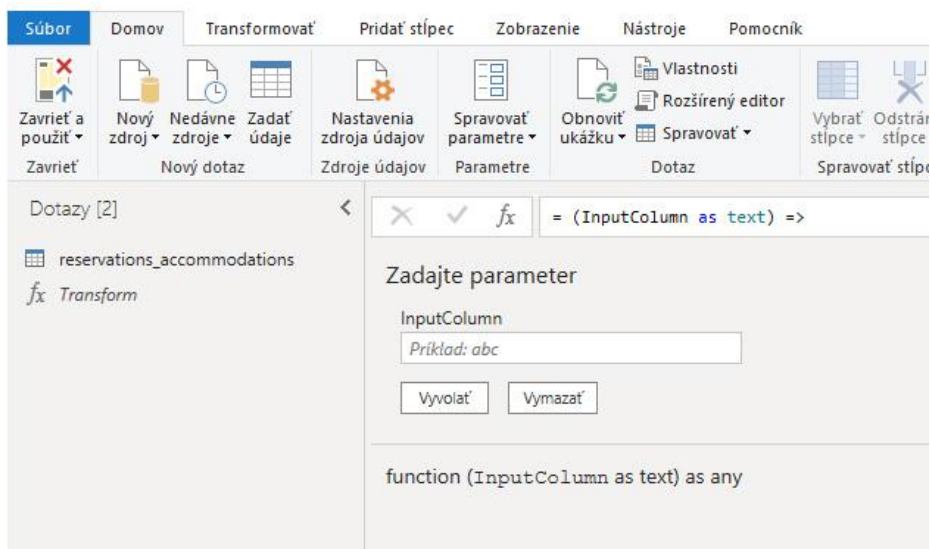
```
ListAccumulatedResult = try List.Accumulate(ListTransform, 0, (state, current) =>  
Number.From(state) + Number.From(current) ) otherwise 0,
```

```
Result = Replacer.ReplaceValue(ListAccumulatedResult, null,0 )
```

```
in
```

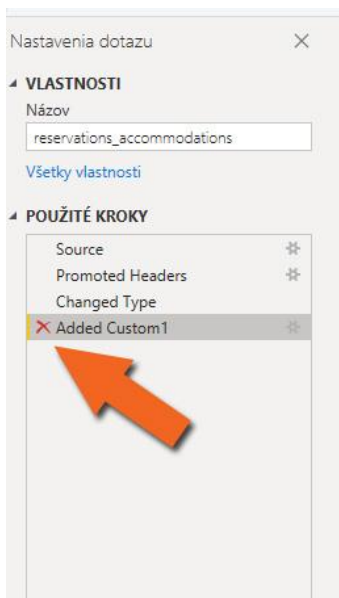
```
Result
```

Nakoniec si ešte premenujeme našu transformačnú funkciu na *Transform*. Hneď ako urobíme tieto, zobrazí sa nám nasledovný výsledok.



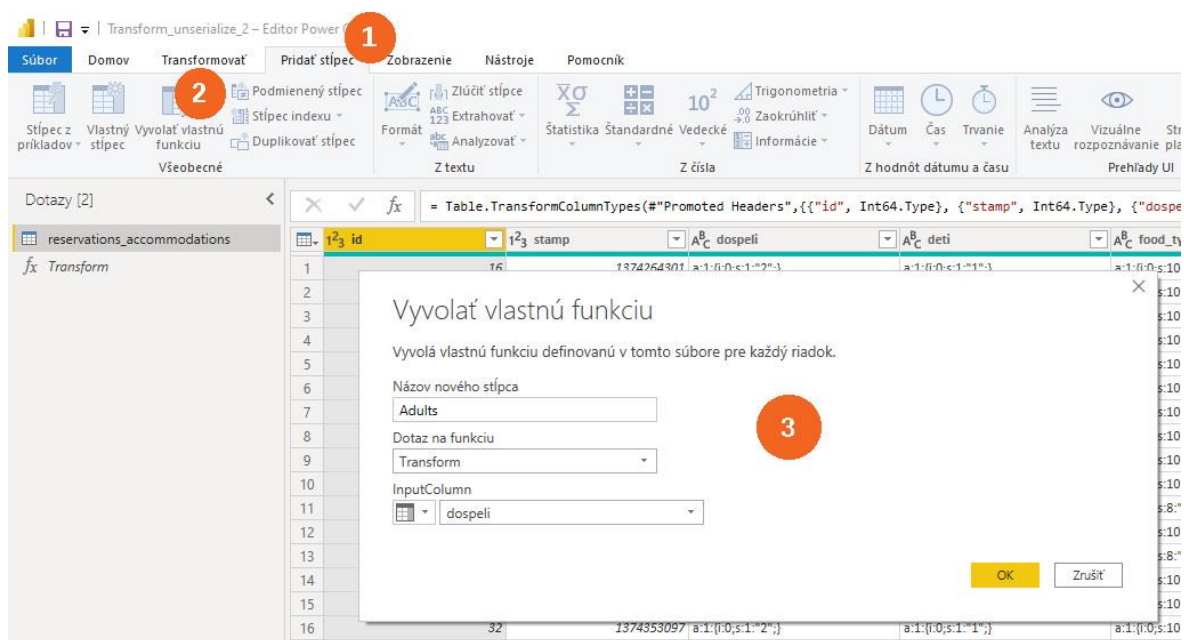
Obrázok 10 Transformačná funkcia (autor, 2021)

Keď už máme hotovú transformačnú funkciu, môžeme si v našom pôvodnom dotaze vymazať kroky, ktoré transformovali stĺpec so serializovaným výrazom. Vymazané kroky budú nahradené transformačnou funkciou, ktorú sme si práve vytvorili.



Obrázok 11 Mazanie použitých krokov v Power Query (autor, 2021)

Ostanú tam iba prvé tri kroky. Následne aplikujeme túto transformačnú funkciu na stĺpec dospelí a stĺpec deti.



Obrázok 12 Vyvolanie vlastnej funkcie v Power Query (autor, 2021)

Zvolíme si záložku Pridať stĺpec a klikneme na Vyvolať vlastnú funkciu. V otvorenom okne zadáme názov nového stĺpca, vyberieme funkciu, ktorú chceme použiť a nakoniec zvolíme stĺpec, ktorý chceme transformovať. Ten bude vstupným parametrom transformačnej funkcie.

Ako vidíme na obrázku nižšie, transformácia prebehla správne.

A ^B C dospelí	A ^B C 123 Adults	
a:1:{i:0;s:1:"2"};		2
a:2:{i:0;s:1:"3";i:1;s:1:"4"};		7
a:1:{i:0;s:1:"2"};		2
a:1:{i:0;s:1:"2"};		2
a:1:{i:0;s:1:"2"};		2
a:1:{i:0;s:1:"2"};		2
a:1:{i:0;s:1:"2"};		2
a:1:{i:0;s:1:"1"};		1
a:1:{i:0;s:1:"1"};		1
a:1:{i:0;s:1:"2"};		2
a:1:{i:0;s:1:"1"};		1
a:1:{i:0;s:1:"2"};		2
a:1:{i:0;s:1:"2"};		2
a:1:{i:0;s:1:"3"};		3
a:2:{i:0;s:1:"2";i:1;s:1:"2"};		4

Obrázok 13 Transformovaný stĺpec vytvorený pomocou vlastnej transformačnej funkcie (autor, 2021)

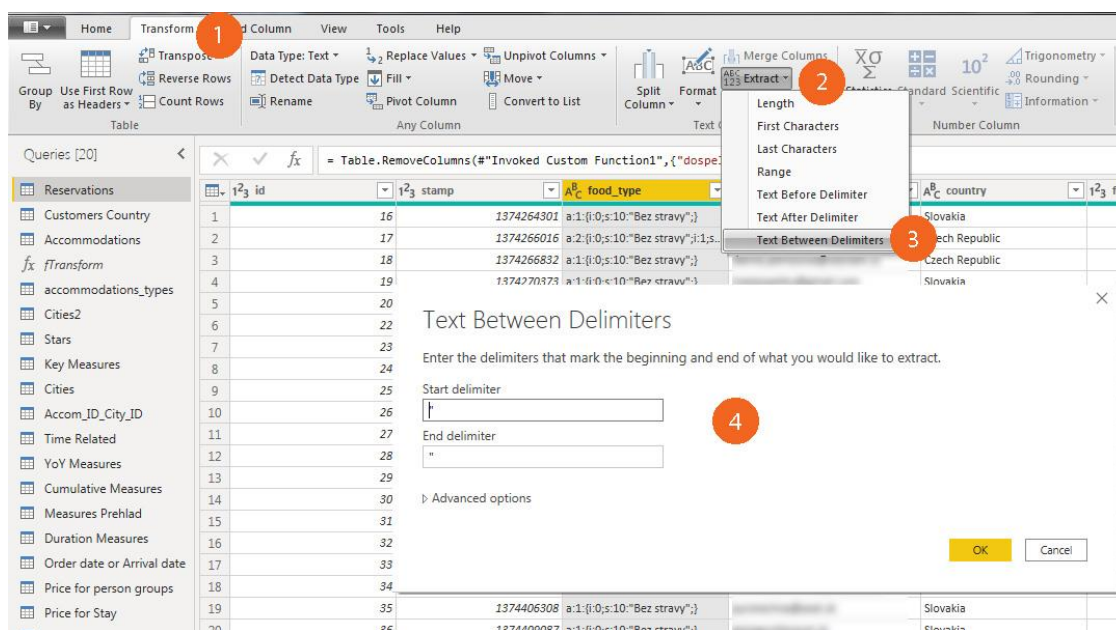
Rovnakým spôsobom si transformujeme aj stĺpec počet detí. Nakoniec pôvodné dva stĺpce môžeme vymazať, keďže tie už nebudeme potrebovať.

Popíšeme si ešte niekoľko úprav tejto tabuľky.

Formu stravy má uloženú tiež vo forme serializovaného poľa, no v tomto prípade nepotrebujeme sčítavať jednotlivé položky. Stačí nám vyexportovať jeden údaj v úvodzovkách.

a:1:{i:0;s:10:"Bez stravy";}

Zvolíme si kartu Transformovať a klikneme na *Extrahovať*. V zobrazenom okne si vyberieme *Text medzi oddeľovačmi*.

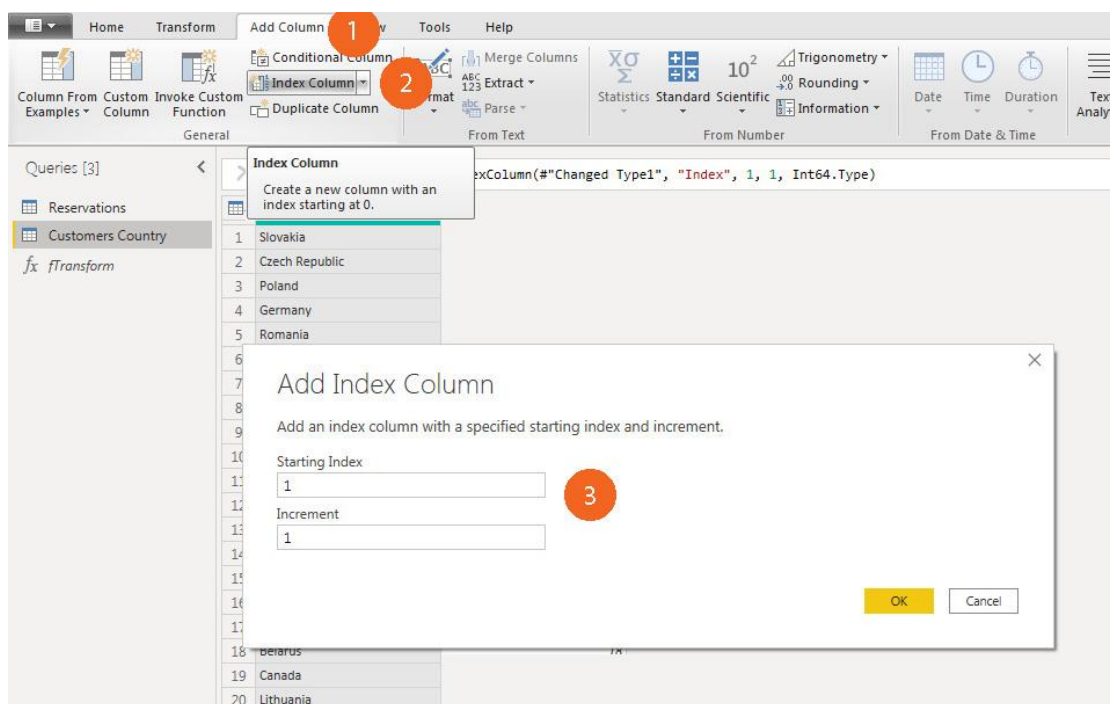


Obrázok 14 Extraktovanie textu medzi oddeľovačmi (autor, 2021)

V tabuľke sú tri stĺpce pre dátumy: Dátum objednávky, Dátum príchodu na pobyt (nástup na pobyt) a Dátum odchodu (koniec pobytu). Všetky tri dátumy máme uložené vo formáte časovej značky (anglicky Timestamp). Táto časová značka predstavuje počet sekúnd, ktoré ubehli od okamžiku 00:00:00 1. januára 1970. Tri stĺpce si pretransformuje na dátum pomocou vzorca:

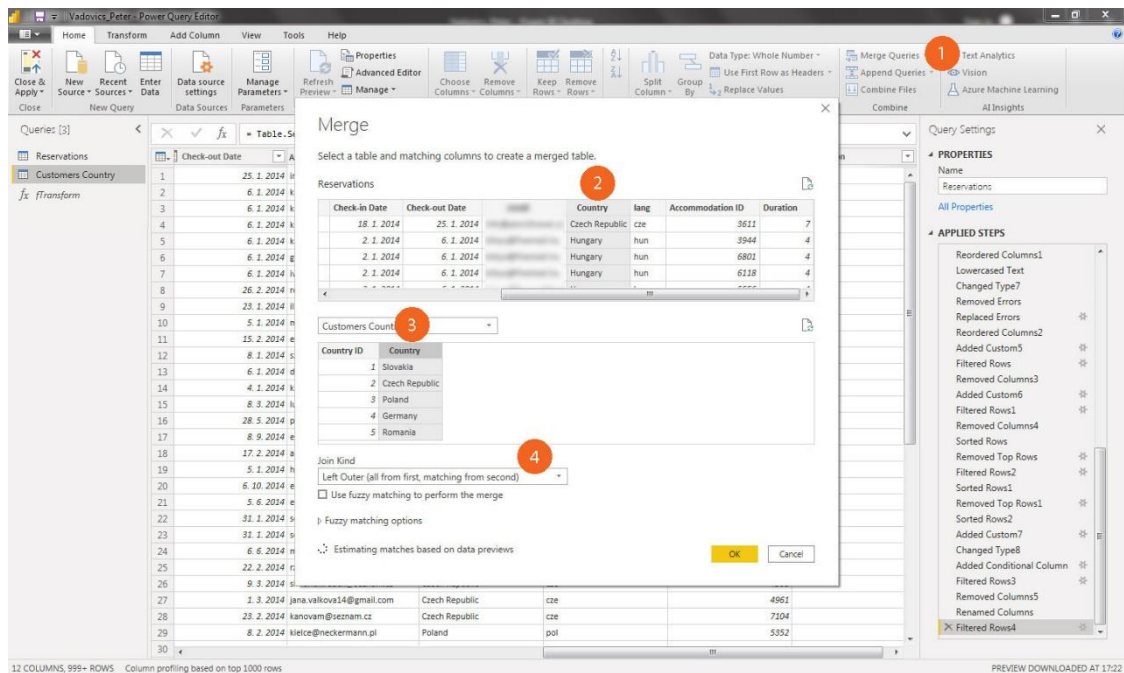
$$= \text{datetime}(1970,1,1,0,0,0) + \text{duration}(0,0,0,[\text{stamp}])$$

Tabuľka rezervácií obsahuje aj stĺpec, ktorý obsahuje názov krajiny klienta, ktorý si vykonal rezerváciu. Tento údaj sa pokúsime oddeliť do tabuľky dimenzií. Súbor CSV, v ktorom máme uložené názvy krajín, si načítame do Power Query rovnakým spôsobom, ako sme načítali prvý súbor. Tabuľka obsahuje iba jeden stĺpec s názvom krajiny. Pridáme druhý stĺpec, ktorým bude obsahovať jedinečné identifikačné číslo krajiny. Na záložke Pridať stĺpec si zvolíme možnosť *Stĺpec Indexu*. V zobrazenom okne zadáme počiatočný index na číslo 1 a prírastok na 1.



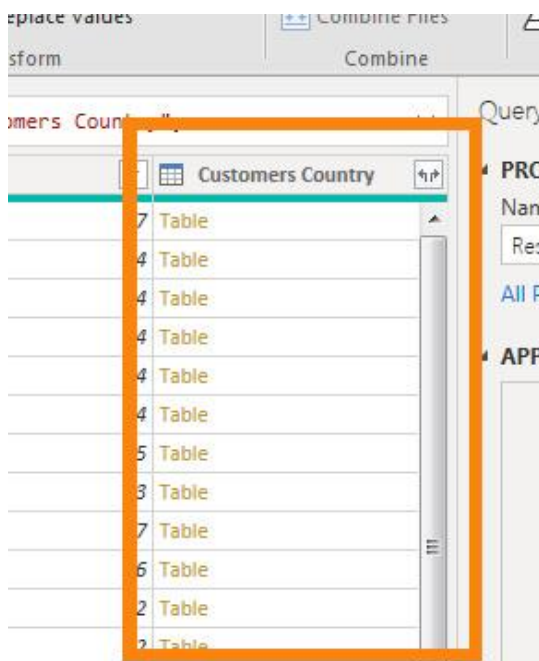
Obrázok 15 Pridanie indexovaného stĺpca (autor, 2021)

Teraz, keď už má každá krajina svoje ID, pokúsime sa názvy krajiny v tabuľke rezervácii nahradiť týmito ID číslami. Zvolíme si dotaz Rezervácie a na karte *Domov* klikneme na možnosť *Zlúčiť dotazy*.



Obrázok 16 Zlúčenie dotazov (autor, 2021)

V zobrazenom okne si vyberieme, ktorý dotaz chceme zlúčiť s našim dotazom Rezervácií. Bude to dotaz s tabuľkou krajín. V každom dotaze zvolíme stĺpec podľa ktorého budeme zlučovať. V našom prípade si v obidvoch zvolíme stĺpec Country, v ktorom sú názvy krajín. Ako posledné vyberieme druh spojenia. V našom prípade to bude vonkajšie spojenie zľava, ktoré ku všetkým riadkom z prvej tabuľky priradí prislúchajúce z druhej. Výsledkom tejto operácie bude nový stĺpec v prvej tabuľke (obrázok).

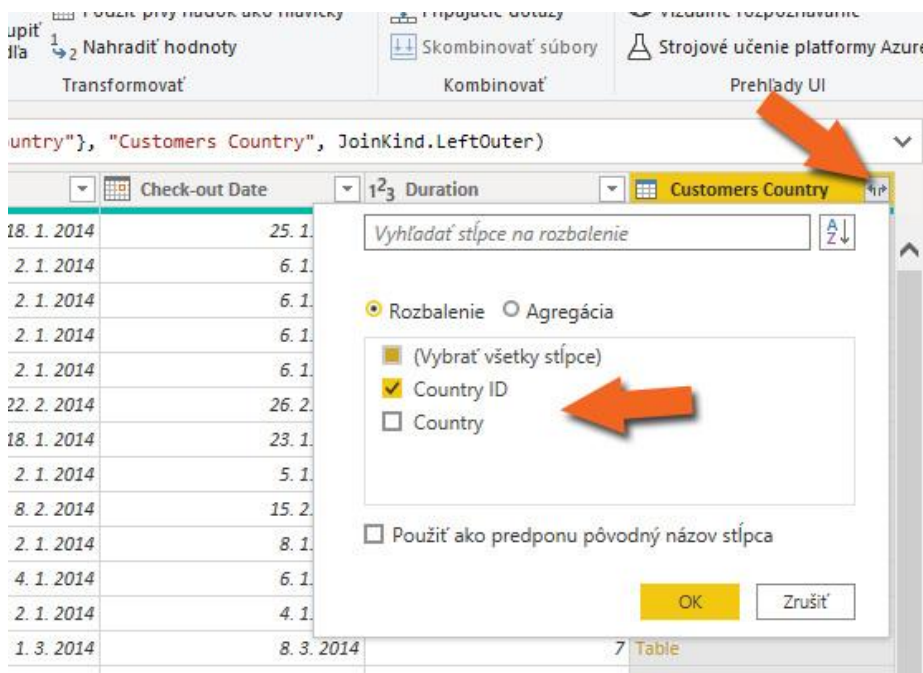


Obrázok 17 Výsledný stĺpec po zlúčení dotazov (autor, 2021)

V každom riadku stĺpca Customers Country je teraz uložená ďalšia malá tabuľka s ID krajiny a názvom krajiny. Jedna bunka v danom stĺpci vyzerá nasledovne:

Country ID	Country
2	Czech Republic

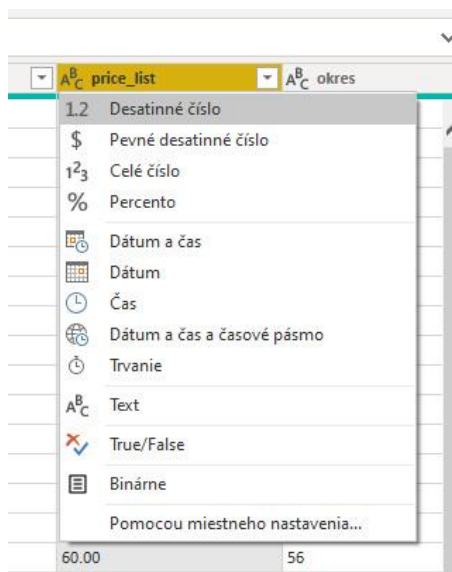
Vidíme, že bunka obsahuje tabuľku s jedným riadkom a dvomi stĺpcami. My si potrebujeme v tabuľke ponechať iba údaj s ID krajiny. Vedľa názvu stĺpca Customers Country klikneme na ikonu, pomocou ktorej rozbalíme tento stĺpec, pričom si vyberiem iba tie údaje, ktoré potrebujeme. Postup môžeme vidieť na obrázku nižšie.



Obrázok 18 Rozbalenie stĺpca

Výsledok tejto úpravy je zobrazený na ďalšom obrázku. Vidíme, že nám ostal iba stĺpec, ktorý sme potrebovali.

V stĺpci price_list (ceny za osobu) sú uložené ceny, no ako vidíme, stĺpec je typu text. Pokúsime sa zmeniť typ údajov na desatinné číslo.



Obrázok 21 Zmena typu údajov stĺpca

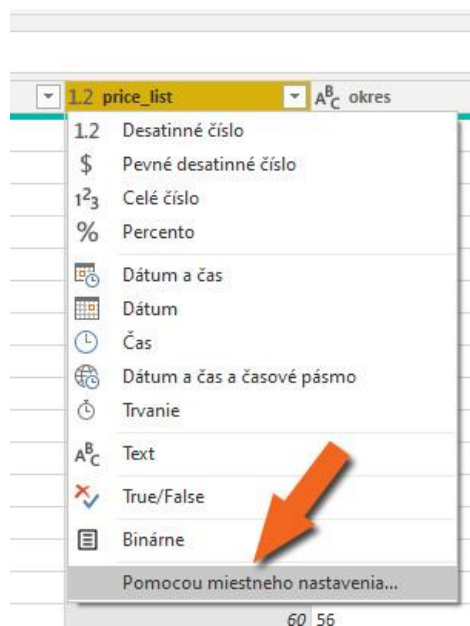
Klikneme na ikonu pri názve stĺpca a z ponuky vyberieme hneď prvú možnosť *Desatinné číslo*. Ako môžeme vidieť na ďalšom obrázku, Power Query hlási chybu v každej bunke tohto stĺpca.



Obrázok 22 Chybný výsledok zmeny typu stĺpca (autor, 2021)

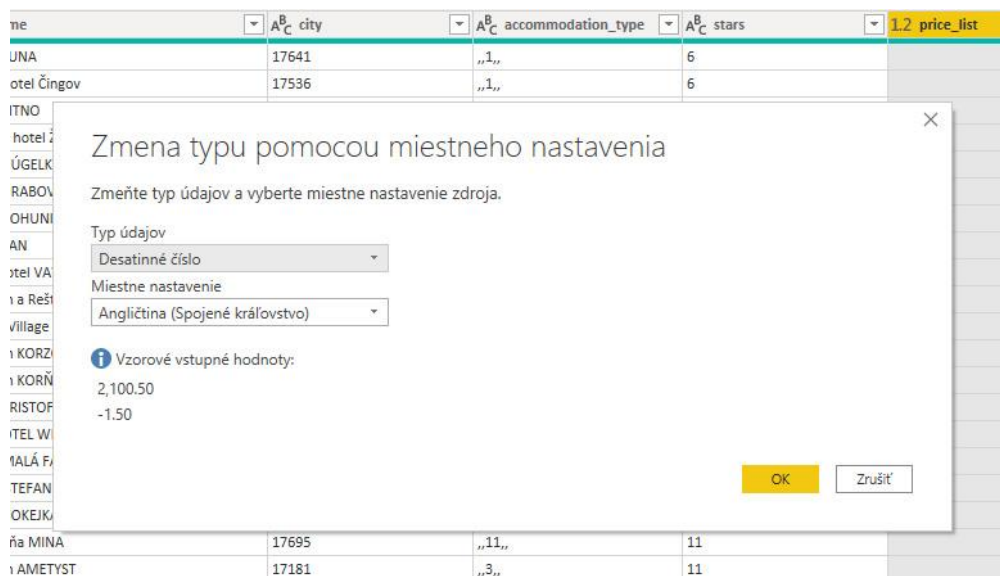
Problémom pri tejto zmene bude v zrejme v miestnom nastavení. Desatinné miestna sú v tabuľke zadané bodkou. V našom miestnom nastavení je ale čiarka. Urobíme zmenu typu pomocou miestneho nastavenia. Klikneme na ikonu vedľa názvu stĺpca a následne

v zobrazenom menu klikneme na poslednú možnosť, na zmenu typu pomocou miestneho nastavenia.



Obrázok 23 Zobrazenie zmeny typu údajov stĺpca pomocou miestneho nastavenia (autor, 2021)

Po kliknutí na možnosť „Pomocou miestneho nastavenia“ sa zobrazí dialógové okno.



Obrázok 24 Zmena typu údajov stĺpca pomocou miestneho nastavenia (autor, 2021)

Výsledok našej zmeny môžeme vidieť na ďalšom obrázku.

1.2 price_list
26,2
22
60,5
10
15
29,5
0
13,5

Obrázok 25 Správny výsledok zmeny typu stĺpca (autor, 2021)

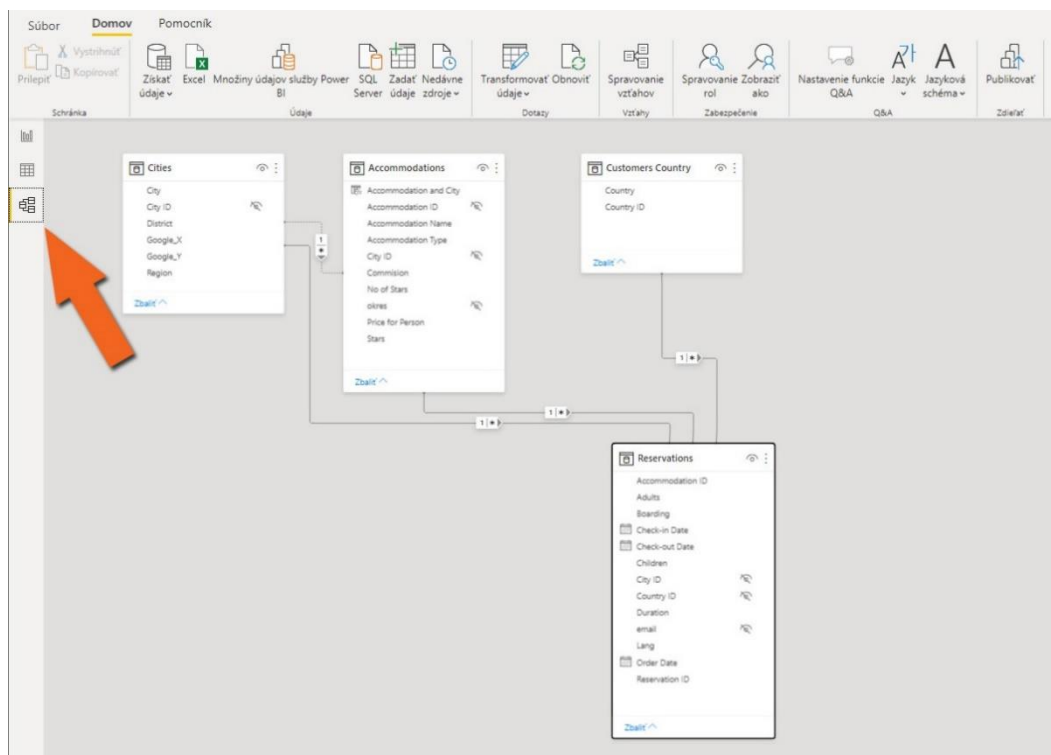
Načítame si aj ďalšie potrebné súbory. Úpravám týchto súborov sa nebudeme detailnejšie venovať, keďže sú to veľmi podobné úpravy. Keď máme načítané všetky tabuľky, môžeme sa vrátiť do Power BI pomocou tlačidla *Zavrieť a použiť*.

The screenshot shows the Power BI Desktop interface. The ribbon is set to the 'Transform' tab. An orange arrow points to the 'Zavrieť a použiť' button. Below the ribbon, the 'Dotazy [5]' pane shows a list of queries: Reservations, Customers Country, Accommodations, fTransform, and Cities. The main view displays a data table with the following data:

1-2-3	Reservation ID	Order Date	ABC 123	Adults
1	93	23. 7. 2013		
2	4000	27. 1. 2014		
3	4061	29. 1. 2014		
4	4230	5. 2. 2014		
5	1202	26. 8. 2013		
6	2107	21. 10. 2013		
7	2133	21. 10. 2013		
8	2825	26. 11. 2013		
9	2876	29. 11. 2013		

Obrázok 26 Načítanie údajov z Power Query do Power BI

Na ľavej strane aplikácie Power BI Desktop sa nachádzajú ikony troch zobrazení aplikácie Power BI Desktop: **Zostava**, **Údaje** a **Model** zhora nadol. Kliknutím na spodnú ikonu sa nám zobrazí náš model. V tomto zobrazení každý blok predstavuje tabuľku a čiary medzi nimi predstavujú vzťahy.



Obrázok 27 Zobrazenie dátového modelu v Power BI (autor, 2021)

Ak chceme vzťah odstrániť, klikneme pravým tlačidlom myši na vzťah a vyberieme položku Odstrániť. Ak chceme vzťah vytvoriť, presunieme polia, ktoré chcete prepojiť, medzi tabuľkami. Podrobnejší pohľad na vzťahy medzi údajmi zobrazíme tak, že na karte *Domov* vyberieme položku *Spravovať vzťahy*.

Spravovanie vzťahov

Aktívne	Z: Tabuľka (Stĺpec)	Do: Tabuľka (Stĺpec)
<input type="checkbox"/>	Accommodations (City ID)	Cities (City ID)
<input checked="" type="checkbox"/>	Price for person groups (Min Price)	Price for stay groups 3 (Min Price)
<input checked="" type="checkbox"/>	Reservations (Accommodation ID)	Accommodations (Accommodation ID)
<input checked="" type="checkbox"/>	Reservations (City ID)	Cities (City ID)
<input checked="" type="checkbox"/>	Reservations (Country ID)	Customers Country (Country ID)
<input type="checkbox"/>	Reservations (Check-in Date)	Date (Date)
<input checked="" type="checkbox"/>	Reservations (Check-in Date)	Date Check-in (Date)
<input checked="" type="checkbox"/>	Reservations (Check-out Date)	Date Check-out (Date)
<input type="checkbox"/>	Reservations (Order Date)	Date (Date)

Nové... Automaticky zistiť... Upraviť... Odstrániť

Zavrieť

Obrázok 28 Spravovanie vzťahov v dátovom modeli (autor, 2021)

Vzťahy sa v tomto dialógovom okne zobrazia v podobe zoznamu. Môžete vybrať položku *Automaticky zistiť*, vďaka čomu sa vyhľadajú vzťahy v nových alebo aktualizovaných údajoch. Vzťahy môžeme upraviť aj manuálne kliknutím na položku *Upraviť*.

Úprava vzťahu

Vyberte navzájom súvisiace tabuľky a stĺpce.

Reservations

Reservation ID	Order Date	Boarding	Lang	Accommodation ID	Check-in Date
4100	piatok, 31. januára 2014	Bez stravy	svk	7004	sobota, 1. februára 2014
7352	streda, 4. júna 2014	Bez stravy	svk	4382	štvrtok, 12. júna 2014
4394	utorok, 11. februára 2014	Bez stravy	svk	4381	pondelok, 17. februára 2014

Accommodations

Accommodation ID	Accommodation Name	City ID	Price for Person	okres	Commision	Stars	No of
3625	Penzión AMETYST	17181	0	18	0,09	0	0
3653	Penzión PRACHÁREŇ	17716	0	18	0,12	0	0
3698	Penzión ŠPORT RUDINA	17286	0	18	0,15	0	0

Kardinalita: N:1 (*:1)

Smer krížového filtrovania: Jeden

Aktivovať tento vzťah

Používanie filtrov zabezpečenia v oboch smeroch

Predpokladať použitie referenčnej integrity

OK Zrušiť

Obrázok 29 Úprava vzťahu v dátovom modeli (autor, 2021)

V zobrazenom dialógovom okne môžeme nastaviť rozšírené možnosti, ako sú kardinalita a smer krížového filtrovania vzťahu.

Ako vidíme na našom modeli, máme jednu tabuľku faktov a tri tabuľky dimenzií. Tabuľku faktov predstavujú jednotlivé rezervácie. Medzi tabuľky dimenzií zatiaľ patria tabuľka ubytovacích zariadení, tabuľka miest a obcí, ako aj tabuľka krajín, odkiaľ pochádzajú klienti. Chýba nám ale jedna dôležitá dimenzia, a tou je časová dimenzia. Preto si vytvoríme novú tabuľku pomocou jazyka DAX.

Na ľavom paneli sme klikli na záložku Údaje, v ktorej môžeme vidieť všetky naše načítané tabuľky. Tu si môžeme napríklad skontrolovať, či sa načítali všetky tabuľky a či sú v nich údaje správne naformátované. Na záložke *Nástroje tabuľky* klikneme na *Nová tabuľka*. Pomocou nasledovného kódu si vytvoríme tabuľku Date.

```

Date =
VAR MinDate = Year ( MIN ( Reservations[Dátum príchodu] ) )
VAR MaxDate = Year ( MAX ( Reservations[Dátum objednávky] ) )
RETURN
ADDCOLUMNS (
    FILTER (
        CALENDARAUTO(),
        YEAR ( [Date] ) >= MinDate &&
        YEAR ( [Date] ) <= MaxDate
    ),
    "Year", YEAR ( [Date] ),
    "Month", FORMAT ([Date], "MMM"),
    "MonthLong", FORMAT ( [Date], "mmmm"),
    "Month Number", MONTH ( [Date] ),
    "Quarter", FORMAT ( [Date], "\QQ" ),
    "Day of Month", DAY([Date]),
    "Year-Month Sort", (YEAR([Date]) * 100) + MONTH([Date]),
    "Month & Year", FORMAT ([Date], "MMM") & " " & YEAR ( [Date] ),
    "Quarter & Year", FORMAT ( [Date], "\QQ" ) & " " & YEAR ( [Date] ),
    "Year-Quarter Sort", (YEAR([Date]) * 10) + QUARTER([Date]),
    "Weekday", FORMAT([Date], "dddd"),
    "Weekday No", WEEKDAY([Date], 2)
)

```

Náš novú tabuľku dátumov môžeme vidieť na ďalšom obrázku.

Date	Year	Month	MonthLong	Month Number	Quarter	Month & Year	Year-Month Sort	Quarter & Year	Year-Quarter Sort	Weekday
1. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Tuesday
2. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Wednesday
3. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Thursday
4. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Friday
5. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Saturday
6. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Sunday
7. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Monday
8. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Tuesday
9. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Wednesday
10. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Thursday
11. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Friday
12. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Saturday
13. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Sunday
14. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3 2014	20143	Monday

Obrázok 30 Časová tabuľka Date (autor, 2021)

Následne sme prešli do záložky Model a prepojili našu novú časovú tabuľku s tabuľkou rezervácií. Medzi týmito dvomi tabuľkami urobíme dve prepojenia. Jedno prepojenie bude s dátumom objednávky a druhé prepojenie s dátumom nástupu na pobyt, pričom obidve nastavíme ako neaktívne. V ďalších častiach si podľa potreby pomocou DAX funkcie určujeme, ktoré prepojenie sa má vo vzorcoch využívať.

Úprava vzťahu ✕

Vyberte navzájom súvisiace tabuľky a stĺpce.

Reservations

Reservation ID	Order Date	Boarding	Lang	Accommodation ID	Check-in Date
4100	piatok, 31. januára 2014	Bez stravy	svk	7004	sobota, 1. februára 2014
7352	streda, 4. júna 2014	Bez stravy	svk	4382	štvrtok, 12. júna 2014
4394	utorok, 11. februára 2014	Bez stravy	svk	4381	pondelok, 17. februára 2014

Date

Date	Year	Month	MonthLong	Month Number	Quarter	Month & Year	Year-Month Sort	Qua
1. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3
2. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3
3. 7. 2014	2014	Jul	July		7 Q3	Jul 2014	201407	Q3

Kardinalita

N:1 (*:1)

Aktivovať tento vzťah

Predpokladať použitie referenčnej integrity

Smer krížového filtrovania

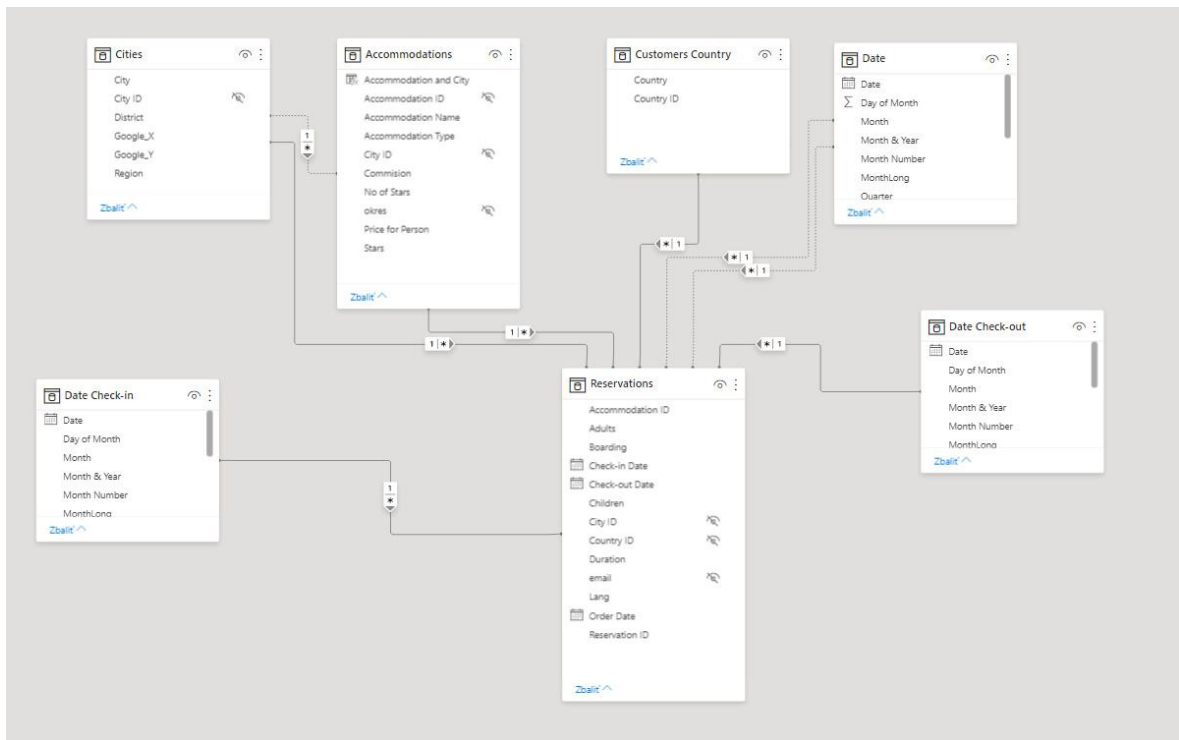
Jeden

Používanie filtrov zabezpečenia v oboch smeroch

OK Zrušiť

Obrázok 31 Vytvorenie prepojenia medzi časovou tabuľkou a tabuľkou rezervácií (autor, 2021)

Ďalej sme si vytvorili ešte dve kópie našej tabuľky dátumov s názvami Date Check-in (Dátum príchodu na pobyt) a Date Check-out (Dátum odchodu z pobytu). Obidve tabuľky prepojíme s tabuľkou Rezervácií. Jednu prepojíme so stĺpcom dátumu príchodu a druhú so stĺpcom dátumu odchodu z pobytu. Tieto tabuľky dátumov aj spolu s prepojeniami sme si vytvorili z toho dôvodu, že v jednom vizuály v nasledovnej časti práce používame dva rôzne dátumy. Celý model je zobrazený na ďalšom obrázku.



Obrázok 32 Dátový model rozšírený o časovú dimenziu (autor, 2021)

3.4 Vytvorenie prehľadov a vizuálov

V tejto časti už vytvárame dashboards a reporty pomocou vizuálov – rôznych typov grafov a tabuliek, s pridaním rýchlych filtrov, pomocou ktorých filtrujeme naše vizuály podľa potreby.

3.4.1 Vytvorenie nového stĺpca podľa vzorca

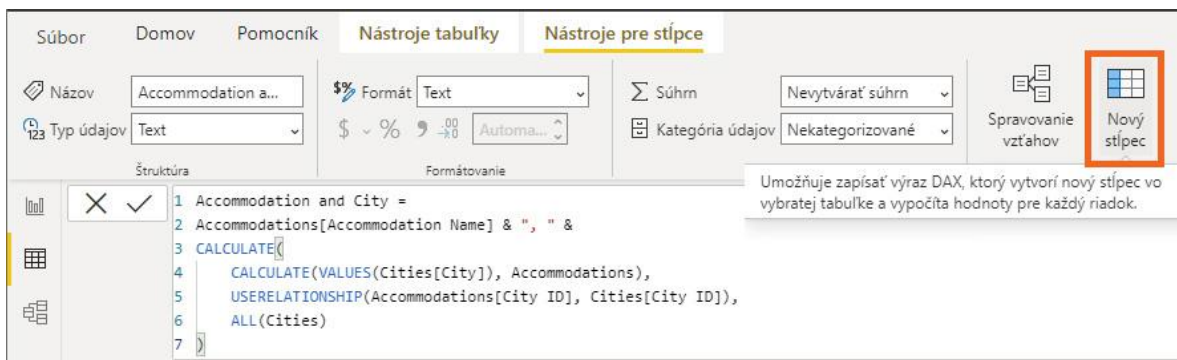
Aby sme mohli unikátne identifikovať v reportoch ubytovacie zariadenia, vytvorili sme si pre každé z nich ďalší údaj, ktorý bude vytvorený spojením názvu zariadenia a názvu obce, v ktorej sa nachádza. Je to preto, lebo určite existujú aj viaceré zariadenia s rovnakým názvom a v niektorých reportoch by sme nevedeli, o ktoré presne zariadenie sa jedná. Na predchádzajúcom obrázku vidieť, že, v našom modeli sme si vytvorili neaktívne prepojenie medzi tabuľkou Ubytovacích zariadení a tabuľkou obcí. Toto prepojenie teraz využijeme v našom vzorci, ktorý priradí hodnotu každému riadku v novom stĺpci. Prešli sme do záložky Údaje, vybrali sme tabuľku ubytovacích zariadení a na záložke *Nástroje pre stĺpce* sme zvolili možnosť *Nový stĺpec*.

Do riadku pre vzorec sme napísali nasledovný vzorec na vytvorenie stĺpca pre „Ubytovanie a mesto“:

Accommodation and City =

```
Accommodations[Accommodation Name] & ", " &  
CALCULATE(  
    CALCULATE(VALUES(Cities[City]), Accommodations),  
    USERELATIONSHIP(Accommodations[City ID], Cities[City ID]),  
    ALL(Cities)  
)
```

Výraz na ľavej strane je názov nového stĺpca. Vzorec na pravej strane zlúči názov zariadenia s názvom obce. Naše neaktívne prepojenie tabuliek použijeme funkciou *USERELATIONSHIP*. Vo vzorci použijeme aj najviac používanú funkciu jazyka DAX, funkciu *CALCULATE*, ktorej prvým parametrom je vyhodnocovaný výraz a ďalšími parametrami sú filtre.



Obrázok 33 Vytvorenie nového stĺpca (autor, 2021)

Výsledkom tohto vzorca je nový stĺpec „Accommodation and City“ (Ubytovanie a mesto) v tabuľke.

Comission	Stars	No of Stars	Accommodation and City	Accommodati
0,09		0	Penzión AMETYST, Bojnice	Penzióny
0,12		0	Penzión PRACHÁREŇ, Trnava	Penzióny
0,15		0	Penzión ŠPORT RUDINA, Rudina	Penzióny
0,13		0	Penzión U ODEHNALU, Bodice	Penzióny
0,19		0	Penzión POD ČERTOVICOU, Malužiná	Penzióny
0,07		0	Penzión PRI KOŠICKEJ BRÁNE, Levoča	Penzióny
0,06		0	Penzión U MEDVEĎA, Párnica	Penzióny
0,2		0	Penzión TATRIA, Podbiel	Penzióny
0,1		0	Penzión STEINGRUBE, Banská Štiavnica	Penzióny
0,19		0	Penzión SKLENÁR, Močenok	Penzióny
0,13		0	Penzión ŠTART, Jezersko	Penzióny
0,16		0	Chata STRÁŇAVY, Stráňavy	Penzióny
0,1		0	Penzión VILA TEREZA, Trenčianske Teplice	Penzióny
0,2		0	Penzión POD STRÁŇOU, Hričovský štátok	Penzióny

Obrázok 34 Vytvorený nový stĺpec pre názov ubytovania a lokalitu (autor, 2021)

3.4.2 Vytváranie mierok

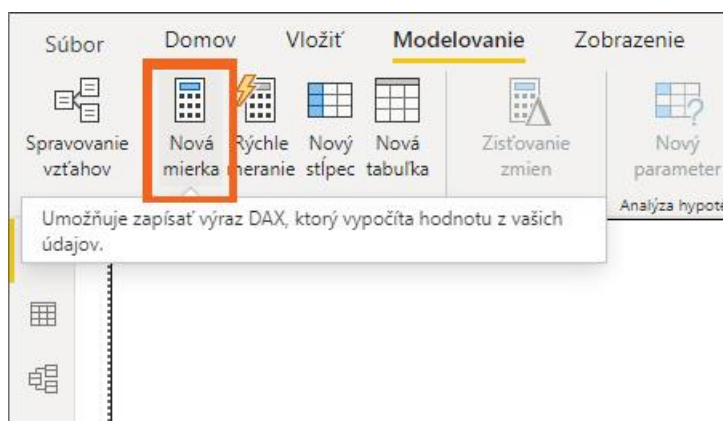
Aby sme mohli vytvárať vizuály, je potrebné vypočítať mierky. Mierky v aplikácii Power BI predstavujú definované výpočty založené na našich údajoch, ktoré sa vykonávajú zároveň s dotazom. Mierky sa vypočítavajú pri interakcii so zostavami a neukladajú sa do databázy. Pomocou mierok vytvárame účinné riešenia analýzy údajov. Mierky nám pomáhajú tak, že vykonávajú výpočty s údajmi počas práce so zostavami.

Vlastné mierky nám umožňujú vykonať zložitejšie a jedinečné výpočty. V Power BI Desktop si vlastné mierky vytvárame pomocou jazyka DAX (Data Analysis Expressions).

Po vytvorení vlastnej mierky, sa táto mierka pridá sa do zoznamu Polia a zaradí sa do vybranej tabuľky. Mierkam môžete priradiť ľubovoľný názov, čo nám pomôže ich vedeli ľahšie identifikovať, a môžeme ich ďalej použiť ako argumenty v ďalších výrazoch jazyka DAX.

Ešte predtým, ako sme začali vytvárať mierky, vytvorili sme si novú tabuľku. Táto tabuľka nám bude slúžiť ako priečinok, do ktorého si budeme ukladať mierky. Budeme mať takto prehľadnejšie zoradené údaje.

Keď máme vytvorenú pomocnú tabuľku, už nám nič nebráni v tom, aby sme sa pustili do vytvárania mierok. V zobrazení zostavy sme vybrali na karte Modelovanie položku *Nová mierka*.



Obrázok 35 Vytváranie novej mierky (autor, 2021)

Prvá mierka, ktorú sme si vytvorili, predstavuje sumu celkového predaja (Celkový predaj).

Celkový predaj =

SUMX(Reservations,

*(Reservations[Adults] + Reservations[Children] / 2) * Reservations[Duration] **

RELATED(Accommodations[Cena za osobu])

)

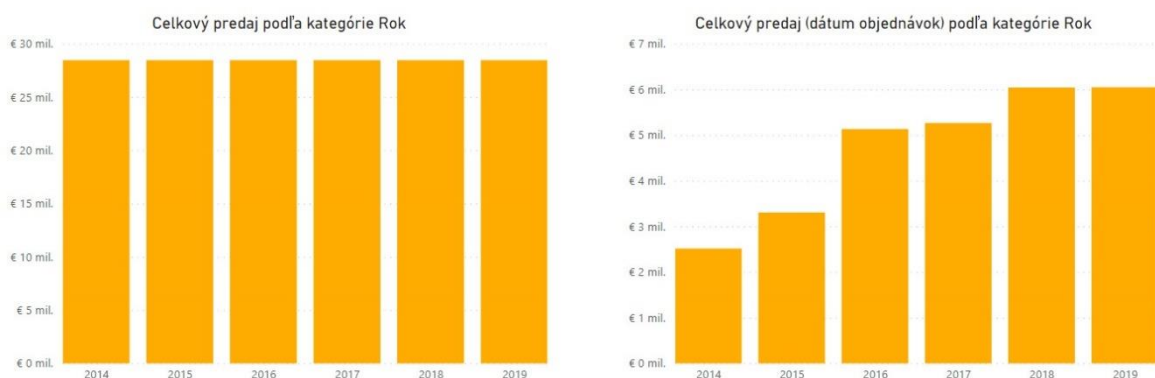
Celkový predaj dostaneme tak, že vynásobíme počet osôb s cenou za osobou a počtom dní. Pre deti orientačne počítame polovičnú cenu. Súčet vypočítame pomocou funkcie SUMX, ktorá nám vráti súčet výrazu pre každý riadok tabuľky. Prvým vstupným parametrom funkcie je názov tabuľky a druhým parametrom je výraz, ktorý sa má vyhodnotiť. Problémom je ale tá skutočnosť, že údaj o cene za osobu sa nachádza v inej tabuľke. Pomôžeme si funkciou RELATED, ktorá nám vráti súvisiacu hodnotu z inej tabuľky.

Aby sme mohli použiť túto mierku v grafe alebo v akomkoľvek vizuály a vidieť ju ako sa vyvíjala v čase, potrebujeme použiť jedno z prepojení, ktorými sme prepojili tabuľku rezervácii s časovou tabuľkou. Vytvoríme si preto ďalšiu mierku, v ktorej si pomocou funkcie USERELATIONSHIP určíme ktoré z našich neaktívnych prepojení použijeme na výpočet. V tomto prípade využijeme prepojenie na dátum objednávky.

Celkový predaj by Dátum objednávky =

CALCULATE([Celkový predaj], USERELATIONSHIP(Reservations[Dátum objednávky], 'Date'[Date]))

Aby sme si vedeli predstaviť rozdiel medzi týmito dvomi mierkami, vytvoríme si dva grafy, kde použijeme naše dve mierky. Pokúsime sa vytvoriť graf, kde budeme vidieť celkový predaj podľa jednotlivých rokov.



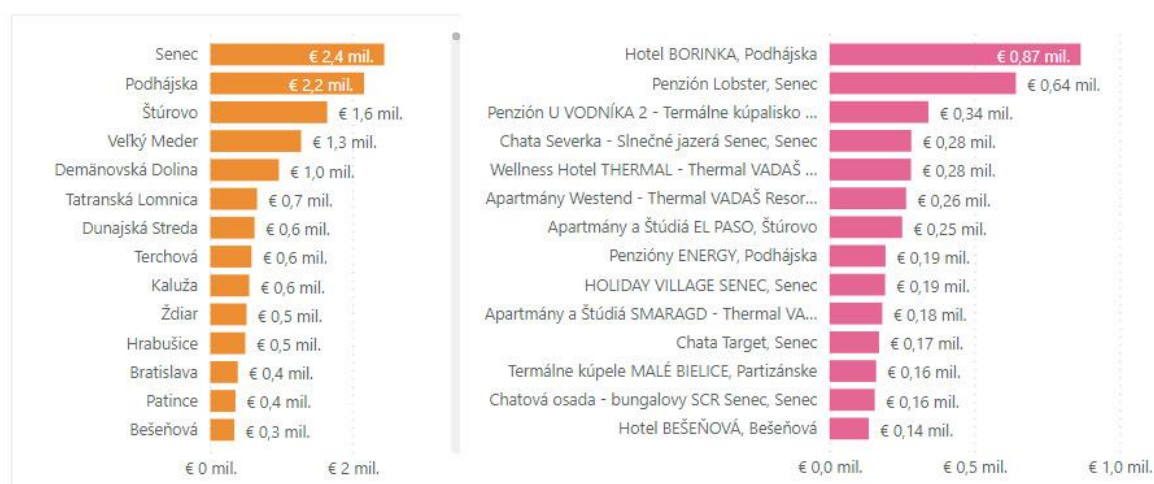
Obrázok 36 Stĺpcový graf – porovnanie mierok Celkový predaj v závislosti od časovej dimenzie (autor, 2021)

Naľavo je graf s mierkou *Celkový predaj* a vpravo je graf s mierkou *Celkový predaj s dátumom objednávky*. Ako môžeme vidieť na ľavom grafe, v každom roku nám ukazujú rovnakú hodnotu predaja, keďže neexistuje prepojenie medzi danou mierkou a časovou tabuľkou.

V období od roku 2014 do roku 2019 sa celkový predaj neustále zvyšoval. Najväčší nárast bol medzi rokmi 2015 a 2016, na druhej strane najmenší nárast nastal medzi rokmi 2016 a 2017 a takisto v roku 2019 bol relatívne malý nárast v predaji oproti roku 2018. V slovenskej verzii vidieť najvyšší nárast v posledných dvoch rokoch, ktorý činil medziročne okolo 50%. Naopak v českej verzii badáme nárast predaja do roku 2016 a v nasledovných rokoch už iba pokles. Všetky ostatné jazykové verzie stránky evidovali nárast predaja až do roku 2018 a v roku 2019 zaznamenali pokles.

3.4.3 Destinácie a ubytovacie zariadenia s najvyšším predajom

V ďalšom kroku zistíme, po ktorých destináciách a konkrétnych ubytovacích zariadeniach bol najväčší dopyt. Tentoraz pre obidva vizuály sme použili pruhový graf. Do hodnoty pre graf sme vložili Celkový predaj a pre os sme použili názov destinácie. Výsledný graf je na ďalšom obrázku.



Obrázok 37 Pruhový graf - Poradie destinácií a ubytovaní podľa výšky predaja

Medzi najobľúbenejšie miesta patria letné destinácie s termálnym kúpaliskom (Podhájska, Veľký Meder, Štúrovo) a jazerom (Senec). Najžiadanejšie ubytovacie zariadenia ležia práve v týchto destináciách.

3.4.4 Medziročné rozdiely v predaji

Aby sme vedeli nárast respektíve pokles predaja vyčíslit' percentuálne ako aj v absolútnych číslach, musíme si vypočítať ďalšie mierky. Ďalšou mierkou, ktorú si vytvoríme bude celkový predaj za predošlý rok. Postup pri vytváraní bude rovnaký ako pri predošlých mierkach. Vo vzorci použijeme mierku *Celkový predaj by Dátum objednávky*, ale tentoraz pomocou funkcie DATEADD odkážeme obdobie posunuté jeden rok dozadu. Výsledný vzorec bude vyzerat' takto:

Celkový predaj LY by Dátum objednávky =

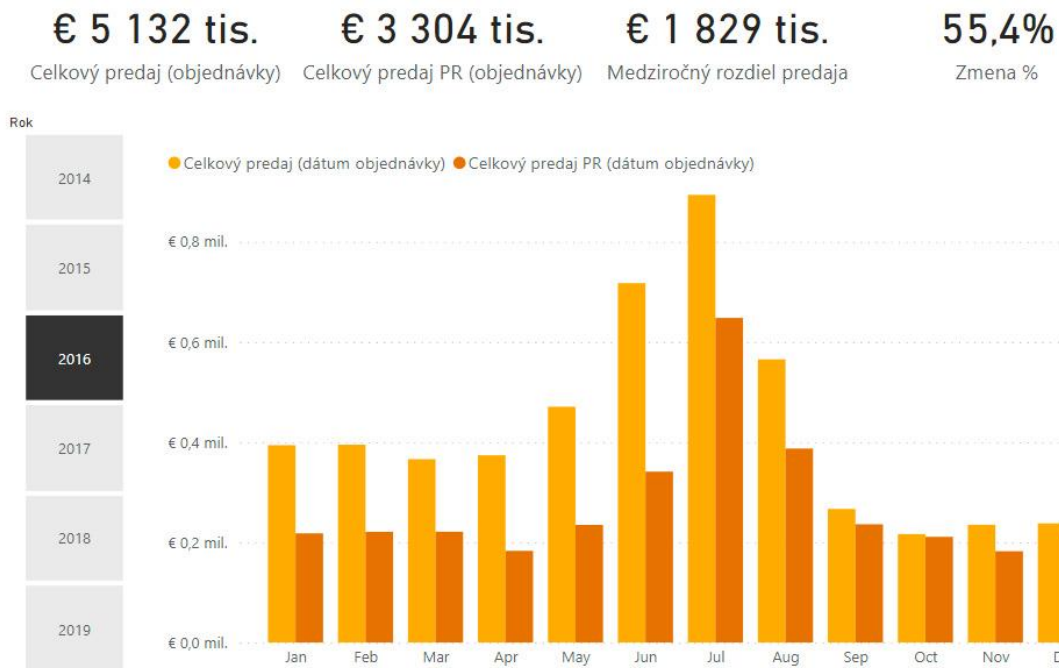
$CALCULATE([Celkový predaj by Dátum objednávky], DATEADD('Date'[Date], -1, YEAR))$

Vypočítame si rozdiel medzi dvomi rokmi.

Celkový predaj YoY (Dátum objednávky) =

$[Celkový predaj by Dátum objednávky] - [Celkový predaj LY by Dátum objednávky]$

Na zobrazenie predajov počas vybraného a predchádzajúceho roka použijeme čiarový graf. Na osi X budeme mať tentoraz mesiace a k našej zostave pridáme aj rýchly filter, v ktorom si môžeme zvoliť rok. Okrem toho pridáme do zostavy aj tri karty, na ktorých si zobrazíme hodnoty našich mierok. Na výber roku, ktorého výsledky chceme vidieť, použijeme rýchly filter.



Obrázok 38 Stĺpcový graf – porovnanie mierky Celkový predaj s predchádzajúcim rokom

Dôležitým ukazovateľom predaja je aj zisk, ktorý je vlastne províziou ubytovacieho zariadenia cestovnej agentúre. Do tabuľky Key Measures si uložíme ďalšiu mierku Zisk.

$$Zisk = SUMX(Accommodations, [Celkový predaj] * Accommodations[Commision])$$

Vytvoríme si mierku pre Zisk (Zisk), ktorá bude prepojená na časovú tabuľku cez Dátum objednávky (dátum objednávky).

Zisk by Dátum objednávky =

$$CALCULATE([Zisk], USERELATIONSHIP(Reservations[Dátum objednávky], 'Date'[Date]))$$

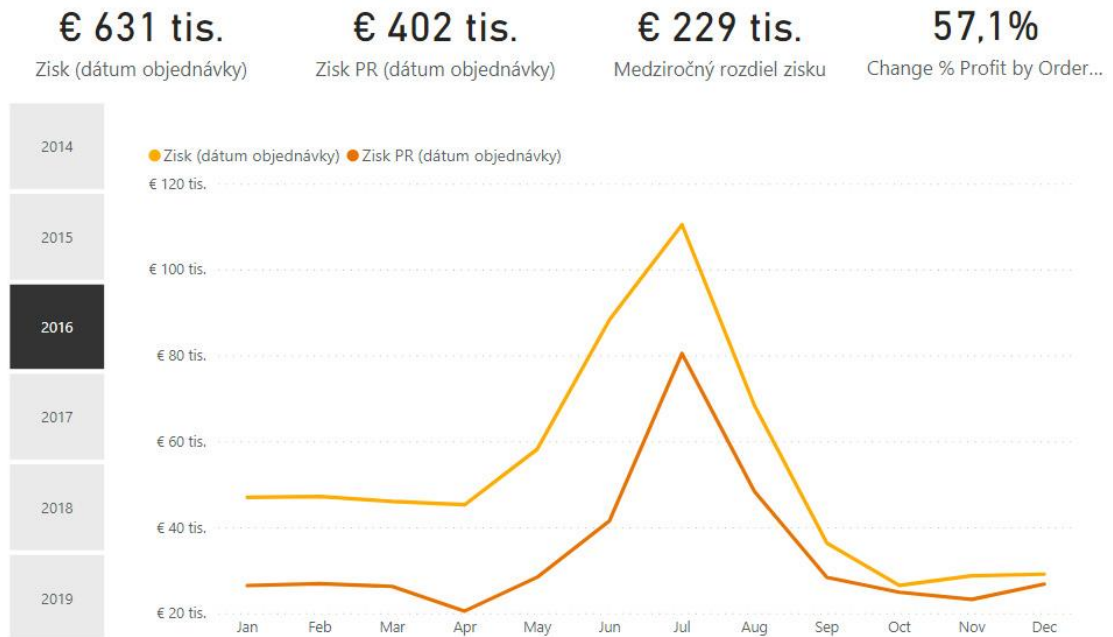
Pre porovnanie s predchádzajúcim rokom vytvoríme mierku Zisk LY by Dátum objednávky (Zisk podľa dátumu objednávky pre predchádzajúci rok).

Zisk LY by Dátum objednávky = CALCULATE([Zisk by Dátum objednávky],

$$DATEADD('Date'[Date], -1, YEAR))$$

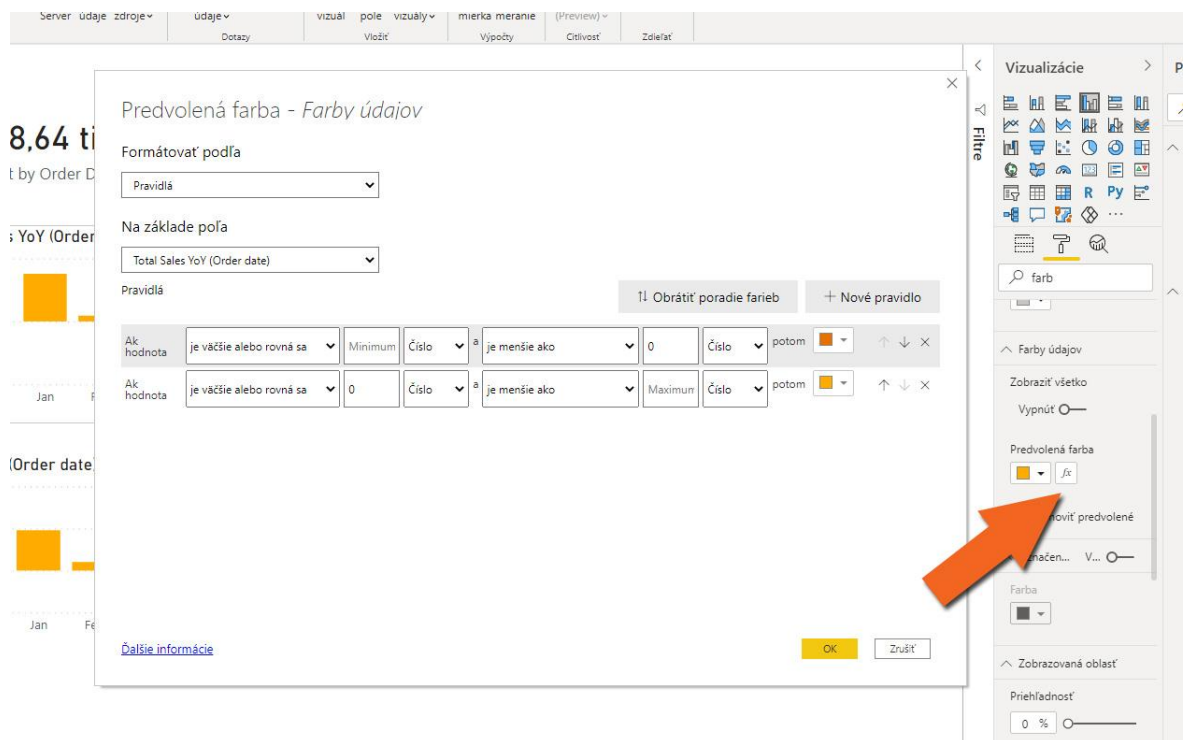
Rozdiel medzi dvomi rokmi nám bude ukazovať ďalšia mierka Zisk YoY (Dátum objednávky).

$Zisk\ YoY\ (Dátum\ objednávky) = [Zisk\ by\ Dátum\ objednávky] - [Zisk\ LY\ by\ Dátum\ objednávky]$



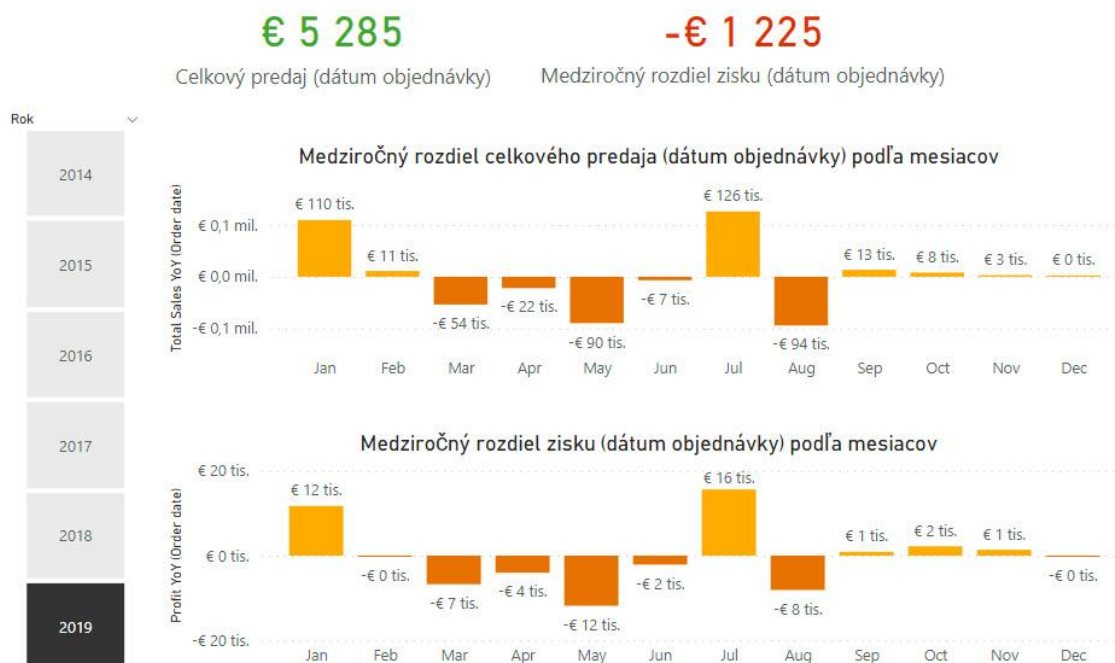
Obrázok 39 Čiarový graf – porovnanie zisku s predchádzajúcim rokom

Na ďalšom vizuály si zobrazíme iba rozdiel celkových tržieb a zisku vo vybranom roku oproti vlaňajšiemu roku. Tentoraz použijeme stĺpcový graf. Aby bol vizuál prehľadnejší, nastavíme si farbu stĺpcov podľa toho či je kladná alebo záporná hodnota.



Obrázok 40 Nastavenie farby údajov podľa pravidla

Výsledný graf je na ďalšom obrázku.



Obrázok 41 Medziročný rozdiel celkového predaja a zisku podľa mesiacov

3.4.5 Predaj v závislosti od dátumu objednávky a dátumu začiatku pobytu

Doteraz sme v našich vizuáloch analyzovali výsledok predajov v rámci času iba z pohľadu dátumu zaslania objednávky. Znamená to, že hodnota pobytu bola priradená k dátumu objednávky. Chceli by sme vidieť vo vizuáloch hodnotu tržieb aj v čase nástupu na pobyt. Takto by sme mohli vidieť, na ktorý termín si v akom množstve klienti rezervujú pobyty.

Aby sme mali vizuály prehľadnejšie a nemuseli vytvárať nové mierky, ktoré by boli závislé iba od termínu začiatku pobytu, pridáme do vizuálov možnosť voľby, v ktorom časovom momente nás zaujíma hodnota tržby - buď termín objednávky alebo termín začiatku pobytu. Následne nám vizuál ukáže hodnotu tržby podľa vybraného termínu.

Na tento účel si vytvoríme pomocnú tabuľku *Objednávky or Pobytu*. Do vizuálov vložíme rýchly filter s hodnotami stĺpca našej pomocnej tabuľky. Pomocou vizuálu budeme vyberať dátum.

Ďalej sme si vytvorili novú mierku Zvolený predaj (zvolený typ predaja), ktorá má hodnotu buď Celkového predaja v čase objednávky alebo Celkového predaja v čase nástupu na pobyt, v závislosti od toho, ktorá z dvoch možností je zvolená.

Zvolený predaj =

IF(HASONEVALUE('Dátum objednávky or Dátum príchodu'[Name]),

IF(SELECTEDVALUE('Dátum objednávky or Dátum príchodu'[Name]) = "Dátum objednávky",

[Celkový predaj by Dátum objednávky],

[Celkový predaj by Dátum príchodu]),

[Celkový predaj by Dátum objednávky]

)

Rovnakým spôsobom sme si i vytvorili mierku pre predchádzajúci rok Celkový predaj LY, ako aj pre zisk.

Diff Celkový predaj YoY = [Zvolený predaj] - [Zvolený predaj LY]

Zmena % Celkový predaj = DIVIDE([Diff Celkový predaj YoY], [Zvolený predaj LY], 0)

Týmito výpočtami sme umožnili, že v danom časovom intervale už môžeme tržby sledovať z pohľadu objednávok alebo realizovaných pobytov. Aj v ďalšom vizuály môžeme vidieť, že máme možnosť prepínať medzi objednávkami a pobytmi.

Vo vizuály sme použili stĺpcový graf, v ktorom sme si dali zobrazit' celkový predaj v danom roku a celkový predaj v predchádzajúcom roku. Ako tretiu hodnotu sme si zvolili medziročný rozdiel v predaji (červená čiara).

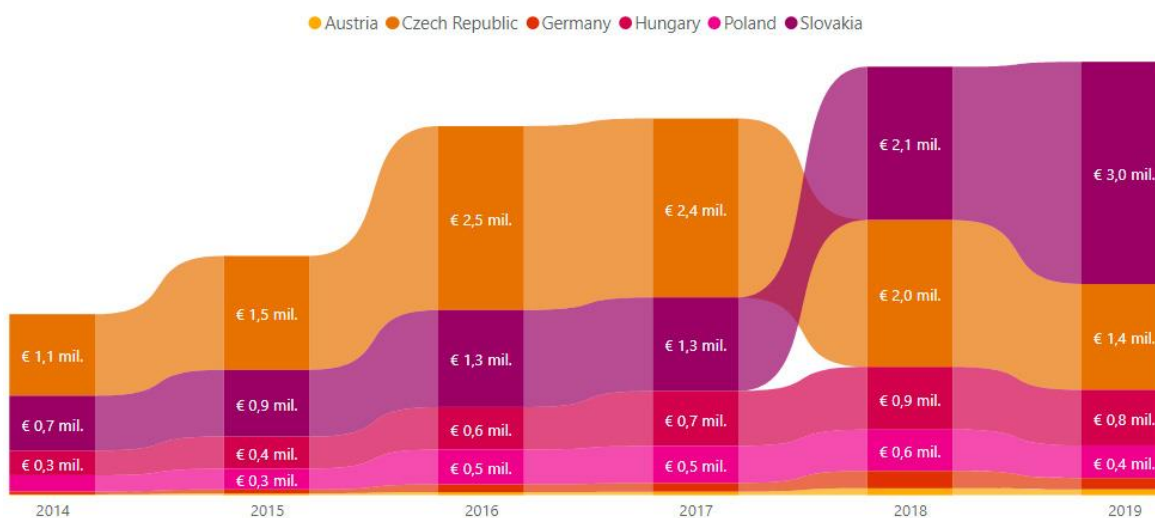


Obrázok 42 Porovnanie celkového predaja s predchádzajúcim rokom a medziročný rozdiel na mesačnej báze

V tomto reporte si vieme nastaviť sledovaný rok a jazykovú verziu. Zároveň si vieme prepínať, či chceme v danom čase sledovať počet objednávok alebo počet realizovaných pobytov.

3.4.6 Zoradenie krajín klientov podľa výšky tržieb

Našou ďalšou úlohou je preskúmať, z ktorých krajín si klienti najviac rezervovali pobyt. Zaujímá nás celková výška sumy pobytov a zároveň ako sa predaj vyvíjal v čase od rokov 2014 až po rok 2019.



Obrázok 43 Zoradenie krajín klientov podľa výšky tržieb od roku 2014 po 2019

V tomto vizuály sme použili pásový graf, aby sme videli zmeny v poradí krajín v čase. Okrem stúpajúceho trendu výšky tržieb vidieť aj to, že v rokoch 2014 až 2017 bolo na prvých šiestich miestach nasledovné poradie: Česká republika, Slovensko, Maďarsko, Poľsko, Nemecko a Rakúsko. Zásadná zmena nastala v roku 2018, keď Česká republika začala klesať a Slovensko naopak začalo rapídne stúpať. V rokoch 2018 a 2019 je Slovenská republika na prvej pozícií a Česká republika na druhej pozícií.

3.4.7 Celkový predaj podľa ceny za osobu

Zaujímalo by nás aké tržby sme mali v jednotlivých cenových kategóriách. To znamená, že chceme zistiť celkové tržby za pobyty rozdelených podľa ceny za osobu.

Začneme tým, že si v Power BI vytvoríme pomocnú tabuľku *Skupiny pre cenu za osobu*, v ktorej si definujeme cenové intervaly cien za osobu. Tabuľku môžete vidieť na obrázku.

Cena za osobu	Cena od	Cena do
0 - 10 €	0	10
11 - 15 €	10	15
16 - 25 €	15	25
26 - 50 €	25	50
nad 50 €	50	1000

Obrázok 44 Pomocná tabuľka definovaných intervalov cien za osobu

Hodnoty v stĺpci Cena za osobu použijeme v rýchlom filtri a ďalšie dva stĺpce nám pomôžu vo vzorcoch na výpočet tržieb. V ďalšom kroku si vytvoríme mierku pre celkový predaj podľa ceny za osobu, ktorá nám vypočíta hodnotu tržieb v jednotlivých cenových kategóriách.

Predaj podľa ceny za osobu =

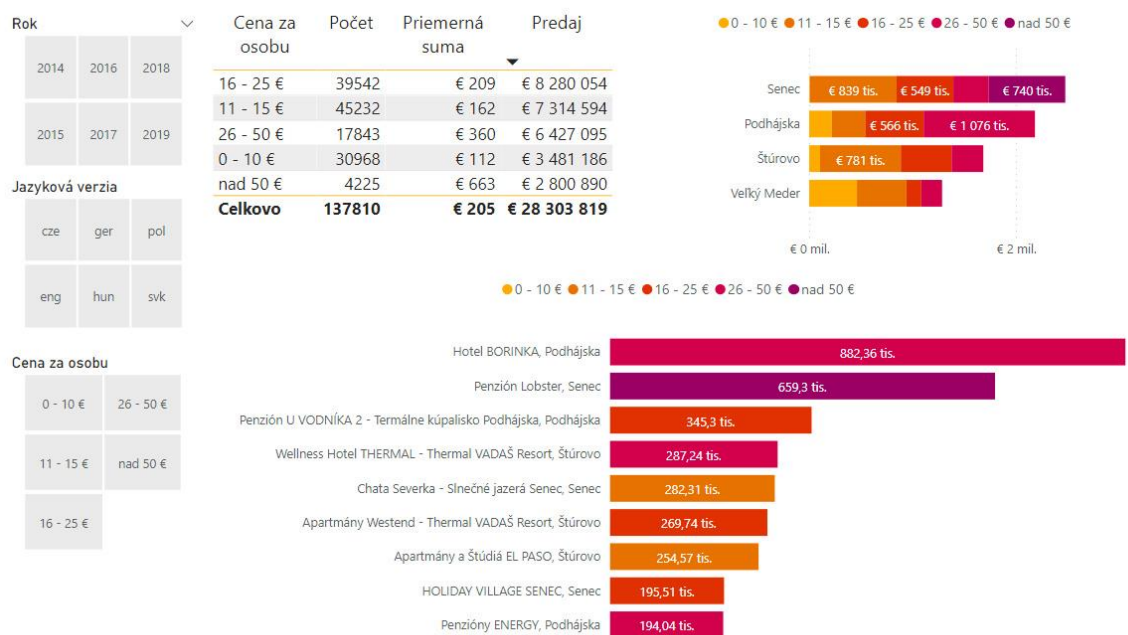
`VAR MinCena = IF(HASONEVALUE('Skupiny pre cenu za osobu'[Cena od]),
SELECTEDVALUE('Skupiny pre cenu za osobu'[Cena od]), 0)`

`VAR MaxCena = IF(HASONEVALUE('Skupiny pre cenu za osobu'[Cena do]),
SELECTEDVALUE('Skupiny pre cenu za osobu'[Cena do]), 10000)`

`RETURN`

CALCULATE([Predaj], FILTER(Accommodations, AND(Accommodations[Cena za osobu] > MinCena, Accommodations[Cena za osobu] <= MaxCena)))

Vytvorenú mierku potom môžeme využívať v jednotlivých vizuáloch. Ako môžeme vidieť na ďalšom obrázku, mierku sme použili v rýchlom filtri, stĺpci tabuľky a v stĺpcovom grafe ako legendu v pruhových grafoch. Takto sme farebne rozdelili jednotlivé ubytovania a hneď na prvý pohľad z grafu vidieť do ktorej cenovej kategórie patria.



Obrázok 45 Celkové tržby podľa cenových kategórií za osobu

Z výsledného reportu vidieť, že najvyššie tržby boli z cenovej kategórie 16 až 25 € za osobu a najnižšie boli v kategórii nad 50 €. Do reportu sme pridali aj ďalšie vypočítané mierky pre počet rezervácií a priemernú sumu za pobyt v každej kategórii. Z tabuľky je vidieť aj podľa očakávania, že najvyššia priemerná cena je v cenovej kategórii nad 50 € za osobu a najnižšia je zase v kategórii do 10 €.

Report je plne interaktívny, takže keď si v rýchlom filtri zvolíme napríklad cenovú kategóriu 11-15 €, tak sa nám v jednotlivých grafoch zobrazia iba ubytovacie zariadenia v danej kategórii a rovnako sa vypočíta aj celková suma predaja v tejto cenovej kategórii pre destinácie.

3.4.8 Cenové kategórie pobytov

Radi by sme vedeli, aké typy pobytov si klienti najviac rezervujú. Najskôr chceme zistiť v akej cenovej kategórii sú najobľúbenejšie pobyty. Znova je potrebné vytvoriť skupiny pre cenové kategórie pobytov, podobne, ako sme to robili v predchádzajúcej časti pri cenách za osobu. Výslednú tabuľku môžeme vidieť na ďalšom obrázku.

Cena za pobyt	Celkový predaj	Počet rezervácií	Počet nocí (priemer)	Cena pobytu (priemer)
0 - 100 €	€ 3 209 403	60042	1,95	€ 53
100 - 250 €	€ 7 164 845	43658	3,71	€ 164
250 - 500 €	€ 7 819 858	22682	4,99	€ 345
500 - 750 €	€ 3 945 231	6533	5,55	€ 604
750 - 1000 €	€ 2 111 370	2443	5,84	€ 864
nad 1000 €	€ 4 053 112	2452	6,22	€ 1 653
Celkovo	€ 28 303 819	137810	3,32	€ 205

Obrázok 46 Celkové tržby podľa cenových kategórií za pobyt

Najviac tržieb bolo z pobytov cenovej kategórie od 250 do 500 € a hneď za ňou je cenová kategória od 100 do 250 €. Do tabuľky sme pridali aj stĺpec pre priemerný počet nocí, ale z tohto údaju nevieme presne ktorý počet nocí pobytu bol najviac zastúpený v danej kategórii. Pomôžeme si preto s ďalšou tabuľkou, kde sú v riadkoch cenové kategórie za pobyt a v stĺpcoch počet nocí.

Cena za pobyt	1	2	3	4	5	6	7	Celkovo
250 - 500 €	€ 128 729	€ 922 273	€ 1 515 477	€ 1 190 272	€ 770 900	€ 737 256	€ 1 990 548	€ 7 255 454
100 - 250 €	€ 341 009	€ 1 690 607	€ 1 835 400	€ 1 223 324	€ 657 661	€ 452 224	€ 778 997	€ 6 979 223
500 - 750 €	€ 50 750	€ 276 843	€ 530 470	€ 654 451	€ 420 046	€ 465 847	€ 1 127 838	€ 3 526 243
nad 1000 €	€ 35 217	€ 282 106	€ 422 149	€ 517 287	€ 340 868	€ 375 934	€ 1 343 397	€ 3 316 958
0 - 100 €	€ 952 943	€ 1 270 525	€ 593 266	€ 240 192	€ 82 069	€ 33 449	€ 22 106	€ 3 194 550
750 - 1000 €	€ 30 583	€ 126 419	€ 234 216	€ 327 343	€ 223 095	€ 270 626	€ 617 114	€ 1 829 396
Celkovo	€ 1 539 231	€ 4 568 774	€ 5 130 977	€ 4 152 868	€ 2 494 639	€ 2 335 336	€ 5 880 000	€ 26 101 825

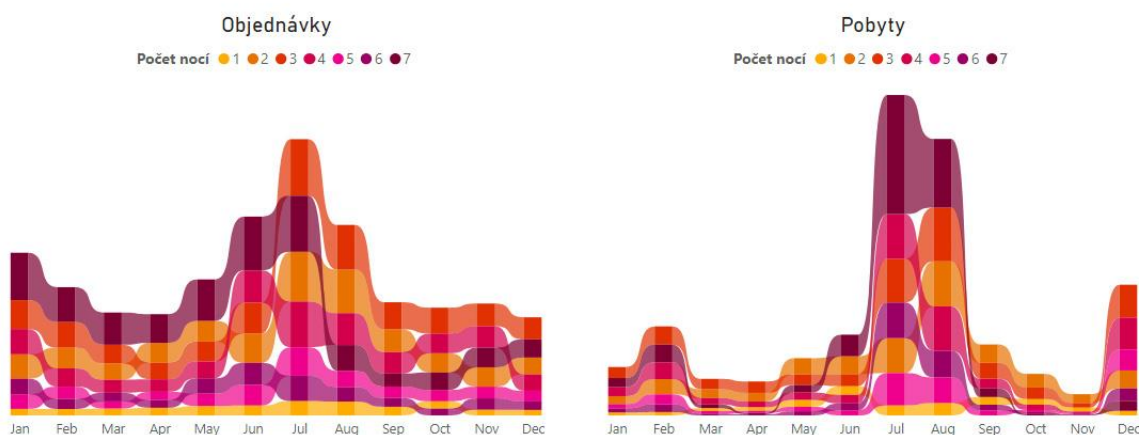
Obrázok 47 Celkové tržby podľa cenových kategórií za pobyt a podľa dĺžky pobytu

V tabuľke sme si dali zobrazit' iba prvých 7 počet nocí, ktoré majú najväčší predaj, preto nám aj presne nesedia celkové súčty. Z uvedenej tabuľky vidíme, že v kategórii 250 až 500 € majú najväčší predaj pobyty na 7 nocí a hneď za nimi sú na druhom mieste pobyty

na 3 noci. V ďalšej cenovej kategórii 100 až 250 € je najvyššia suma pri pobytoch na 3 noci, na 2 noci a potom na 4 noci.

3.4.9 Predaj podľa dĺžky pobytov

Zaujímavá je aj podrobná analýza vývoja dĺžky pobytu z časové pohľadu mesiacov v roku. Vytvorili sme si dva pásové grafy. Na X-ovej osi oboch grafov máme mesiace v roku a v riadkoch je dĺžka pobytu. V jednom grafe sledujeme výšku ceny objednávok v daných mesiacoch podľa dĺžky pobytov a v druhom grafe sledujeme výšku ceny pobytov.



Obrázok 48 Výška ceny objednávok a pobytov podľa mesiacov v roku

Na výsledných grafoch vidíme, že od januára až do júna sa najvyššie sume objednávajú pobyty na 7 nocí. Od júla až do decembra zase pobyty na 3 noci. Počas letných prázdnin v mesiacoch júl a august jednoznačne dominujú pobyty na 7 nocí. Sú to tie pobyty, ktoré boli objednávané počas predošlých mesiacov. Z toho usudzujeme, že letnú dovolenku si klienti začínajú objednávať až od januára. Pokiaľ v júli bolo poradie dĺžky realizovaných pobytov 7, 4 a 3 noci, v mesiaci august bolo poradie 7, 3 a 2 noci.

Detailnejšie sa vieme na túto problematiku pozrieť v tabuľke, v ktorej máme v riadkoch mesiace, v ktorých bola vykonaná objednávka a v stĺpcoch máme mesiace, v ktorých sa začína pobyt. Túto tabuľku môžeme vidieť v prílohe C. Z tabuľky vyčítame, že najvyššiu celkovú sumu majú rezervácie vykonané v mesiaci júl na mesiac júl a august. Druhú najvyššiu celkovú sumu predstavujú rezervácie v júni na mesiac júl a rezervácie v mesiaci august na mesiac august. V reporte je možné filtrovať rezervácie podľa dĺžky pobytu, ako aj podľa jednotlivých rokov. Vieme si tak pozrieť kedy boli v akej celkovej výške rezervované napríklad pobyty na 7 nocí alebo na 3 či 4 noci.

V rámci analýzy celkovej sumy rezervácii podľa dĺžky pobytov sme vytvorili ešte posledný vizuál, kde môžeme vidieť aká je suma pobytov v najžiadanejších destináciách podľa dĺžky pobytov.

Mesto, obec	1	2	3	4	5	6	7	Celkovo
Senec	€ 42 409	€ 214 596	€ 283 985	€ 351 866	€ 224 678	€ 242 755	€ 910 352	€ 2 270 640
Podhájska	€ 30 796	€ 250 015	€ 357 457	€ 300 347	€ 233 057	€ 330 599	€ 491 881	€ 1 994 154
Štúrovo	€ 20 361	€ 101 745	€ 146 266	€ 139 816	€ 91 831	€ 133 701	€ 882 826	€ 1 516 545
Veľký Meder	€ 22 491	€ 139 451	€ 273 807	€ 230 458	€ 110 441	€ 92 547	€ 351 818	€ 1 221 014
Demänovská Dolina	€ 7 946	€ 86 027	€ 188 502	€ 220 302	€ 97 861	€ 84 740	€ 232 285	€ 917 662
Tatranská Lomnica	€ 13 404	€ 76 967	€ 125 354	€ 107 356	€ 75 750	€ 76 364	€ 133 061	€ 608 256
Dunajská Streda	€ 29 289	€ 96 046	€ 143 780	€ 81 131	€ 45 668	€ 51 271	€ 146 820	€ 594 006
Terchová	€ 17 929	€ 108 791	€ 128 579	€ 91 320	€ 55 927	€ 53 403	€ 103 905	€ 559 854
Kaluža	€ 23 733	€ 100 580	€ 87 093	€ 82 353	€ 39 596	€ 50 036	€ 100 331	€ 483 721
Ždiar	€ 9 946	€ 47 365	€ 86 127	€ 65 426	€ 99 229	€ 62 961	€ 93 366	€ 464 419
Hrabušice	€ 25 036	€ 99 796	€ 103 387	€ 88 073	€ 37 120	€ 34 601	€ 54 958	€ 442 970
Bratislava	€ 68 628	€ 114 911	€ 53 583	€ 41 228	€ 15 598	€ 21 805	€ 28 326	€ 344 079
Patince	€ 7 964	€ 29 100	€ 48 792	€ 36 679	€ 27 589	€ 27 918	€ 159 053	€ 337 094
Bešeňová	€ 11 808	€ 68 799	€ 73 577	€ 48 303	€ 35 038	€ 31 947	€ 55 792	€ 325 264
Donovaly	€ 15 535	€ 65 504	€ 78 074	€ 81 968	€ 21 877	€ 19 602	€ 39 708	€ 322 267
Celkovo	€ 435 735	€ 1 816 341	€ 2 463 767	€ 2 206 882	€ 1 371 511	€ 1 436 447	€ 4 053 978	€ 13 784 660

Obrázok 49 Celková suma rezervácií v destináciách podľa dĺžky pobytu

Bezkonkurenčne sú na prvých miestach v predaji mestá Senec a Štúrovo s pobytmi na 7 nocí. Na druhej strane v Podhájskej sú relatívne vyrovnaneršie sumy pobytov od 2 do 7 nocí.

3.4.10 Predaj podľa typu pobytu

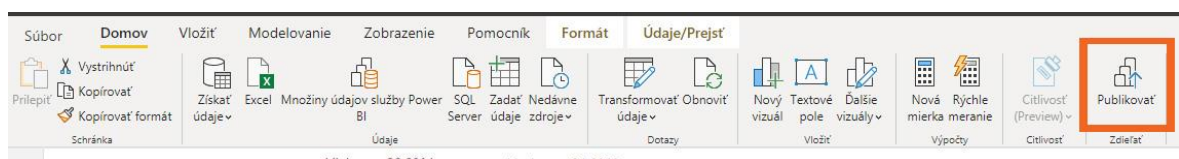
Pobyty si ešte môžeme rozdeliť aj podľa počtu osôb, nástupného dňa na pobyt (deň v týždni) a dĺžky pobytu. Podľa tohto rozdelenia je vidieť v tabuľke, že najvyššiu sumu majú pobyty na 7 nocí pre dvoch dospelých a dve deti s nástupným dňom sobota. Na druhom mieste sú rovnaký typ pobytov, akurát s tým rozdielom, že bez detí. Na treťom mieste sú víkendové pobyty na 2 noci pre dve dospelé osoby.

Dospelí	Deti	Deň v týždni	Počet nocí	Celkový predaj
2	2	Saturday	7	€ 837 642
2	0	Saturday	7	€ 619 032
2	0	Friday	2	€ 615 166
2	1	Saturday	7	€ 396 072
2	0	Thursday	3	€ 367 054
4	0	Saturday	7	€ 281 029
4	0	Friday	2	€ 268 029
4	0	Thursday	3	€ 207 914
2	0	Monday	6	€ 201 716
2	3	Saturday	7	€ 191 579
2	0	Monday	4	€ 190 371
2	2	Friday	2	€ 183 143
2	0	Saturday	1	€ 182 601
2	0	Friday	3	€ 177 511
Celkovo				€ 27 679 735

Obrázok 50 Tabuľka: Zoradenie typov pobytov podľa celkovej sumy

3.5 Publikácia výsledných dashboardov v službe Power BI

Aby sme mohli sprístupniť naše výsledné dashboardy a reporty, je potrebné ich publikovať v službe Power BI. Aby ste svoje výsledky vedeli publikovať, je potrebné mať v službe Power BI registrovaný účet s firemným emailom. Súkromné emaily služba neakceptuje. Svoju zostavu v Power BI Desktop publikujeme tak, že na karte Domov vyberieme na páse s nástrojmi možnosť Publikovať.



Obrázok 51 Publikovanie zostavy v službe Power BI

V čase publikovania je potrebné byť prihlásený v službe Power BI. Po dokončení publikovania sa zobrazí prepojenie na zostavu. Kliknutím na prepojenie otvoríte zostavu na lokalite Power BI. Aby ste mohli zdieľať Vaše zostavy, musíte mať konto Power BI Pro. V opačnom prípade môžete zaslať prepojenie na zostavy ostatným užívateľom s existujúcim kontom.

3.6 Diskusia

Vytvorenými dashboardmi a reportmi sme dospeli k nasledovným záverom. Celkové tržby za pobyty od roku 2014 do roku 2019 sa neustále zvyšovali, napriek tomu ten trend sa v poslednom roku veľmi spomalil. Najväčší podiel na predaji tvoria slovenská, česká a maďarská verzia.

Za posledné dva roky bol najväčší nárast predaji zaznamenaný od klientov zo Slovenska a to hlavne vďaka letným destináciám Štúrovo, Podhájska a Senec. V roku 2018 to 57% nárast predajov a v roku 2019 bol zaznamenaný 38% nárast. Predaj českej klientele klesá už od roku 2017, čo je veľká škoda, pretože v roku pokiaľ v roku 2016 tvorila česká klientela takmer 50% na tržbách, v roku 2019 to bolo necelých 25%.

Ponuky na letnú sezónu je potrebné pripraviť už začiatkom roka, keďže si klienti rezervujú letné dovolenky už od januára. Cestovná agentúra by sa mala zamerať hlavne na letné pobyty na 7 nocí, prevažne od soboty do soboty.

Čo sa týka počtu osôb, tak najžiadanejšie sú pobyty na 7 nocí pre 2 dospelé osoby a dve deti ako aj pobyty iba pre 2 dospelé osoby.

U maďarskej klientele tvoria významný podiel na predaji popri letných pobytoch aj zimné pobyty. Medzi najobľúbenejšie destinácie patria Ždiar, Demänovská dolina a Hrabušice, ktoré patria medzi horské lokality. Najviac rezervované sú pobyty na 3 noci s nástupom vo štvrtok a piatok, pobyty na 2 noci s nástupom v piatok a pobyty na 4 noci. V cenových kategóriách boli najviac rezervované pobyty s cenou za osobu od 16 do 25 €.

Podrobnejšie skúmanie by mohlo byť predmetom ďalšej analýzy. Analyzovať by sme mohli rezervácie nielen z pohľadu destinácie, ale aj z pohľadu služieb, ktoré tieto ubytovacie zariadenia poskytujú. Pozornosť by sme mohli upriamiť na obľúbené služby, ako je napríklad wellness, sauny, bazén, služby pre deti, poprípade liečebné či relaxačné procedúry.

Ďalšia analýza by sa mohla týkať nielen celkového predaja, ale aj výšky provízie, ktorú cestovnej agentúre poskytujú ubytovania. Takto by sme mohli zistiť, ktoré pobyty alebo ktoré konkrétne ubytovania prinášajú najvyššie zisky. Takto by sa cestovná agentúra

mohla zamerať na tie ubytovania, ktoré sú žiadané klientmi a zároveň poskytujú vyššiu províziu.

Podrobnejšia analýza sa môže týkať aj klientov a to z pohľadu či ide o nového klienta, alebo vracajúceho sa klienta.

Záver

Hlavným cieľom práce bolo analyzovať údaje z databázy rezervačného systému ubytovania na Slovensku pomocou aplikácie Power BI.

V prvej kapitole sme si priblížili jednotlivé pojmy z oblasti Business Intelligence, samoobslužného BI a porovnali najpoužívanejšie BI nástroje. Pri hodnotení sme vychádzali aj z nezávislého porovnávania spoločnosťou Gartner. Najlepšie sa umiestnil softvérový nástroj Power BI od spoločnosti Microsoft. Sme si ho bližšie popísali, ako aj jeho komponent Power Query a programovací jazyk DAX, pomocou ktorého vieme v Power BI lepšie analyzovať údaje.

V druhej kapitole sme si stanovili ciele a metodiku práce ako tieto ciele dosiahnuť. Určili sme si hlavné ukazovatele v predaji, ktoré nás v analýzach zaujímajú. Predovšetkým to bola výška predaja podľa rôznych parametrov rezervovaných pobytov, či už sa jednalo o destináciu, obdobie rezervácie, dĺžku pobytu či pôvod krajiny klienta. Zaujímal nás aj trend predaja jednotlivých krajín.

V tretej kapitole sme na základe týchto stanovených cieľoch začali spracovávať údaje. Najskôr sme vyexportované dáta z relačnej databázy pomocou nástroja Power Query očistili, upravili do potrebného formátu a načítali do Power BI. Ďalej sme vytvorili ďalšie pomocné tabuľky a pomocou vzťahov medzi tabuľkami sme vytvorili dátový model. V modeli sme jednu tabuľku faktov a niekoľko tabuliek dimenzií. Do modelu sme postupne pomocou jazyka DAX pridávali vypočítané mierky pre naše kľúčové ukazovatele. Tieto mierky nám pomohli zostaviť dashboards a reporty, ktoré nám dali odpovede na naše otázky.

Naše výsledné analýzy môžu slúžiť cestovnej agentúre na podporu rozhodovania ohľadom vytvorenia ponúk pobytových balíkov s cieľom zamerať sa na určitý typ pobytov pre vybranú klientelu vo vybranom období.

Zoznam použitej literatúry

- [1] LACKO, Ľuboslav. Business Intelligence v SQL Serveru 2008, Reportovací, analytické a ďalšie dátové služby, Praha: Computer Press, 2009. 456 s. ISBN 978-80-251-2887-9.
- [2] CHMELÁR, Michal. Reporting v Power BI, Power Pivot a jazyk DAX. 1. vyd., Bratislava: Smart People, spol. s r.o., 2018. 566 s. ISBN 978-80-9773078-0-6.
- [3] NĚMEC R., Principy projektování a implementace systému Business Intelligence. Ostrava: VŠB-TU Ostrava, 2014. 113 s. ISBN 978-80-248-3452-8.
- [4] LACKO, Ľuboslav. Databáze: datové sklady, OLAP a dolování dat s příklady v Microsoft SQL serveru a Oracle. 1. vyd., Brno: Computer Press, 2003. 488 s. ISBN 80-2-26-969-0.
- [5] NOVOTNÝ, Ota – POUR, Jan – SLÁNSKÝ, David. Business Intelligence: Jak Využít bohatství ve vašich datech. 1. vyd., Praha: Grada Publishing, 2005. 256 s. ISBN 80-247-1094-3.
- [6] SABHERWAL R., BECERRA-FERNANDEZ I., Business Intelligence: Practices, Technologies, and Management. Hoboken: John Wiley & Sons Inc., 2011. 456 s. ISBN 978-0-470-46170-9.
- [7] GROSSMANN W., RINDERLE-MA S, Fundamentals of Business Intelligence, Berlin Heidelberg: Springer-Verlag. 2015. 348 s. ISBN 978-3-662-46531-8.
- [8] VAISMAN A., ZIMÁNYI E., Data Warehouse Systems, Berlin Heidelberg: Springer-Verlag, 2014. 379 s. ISBN 978-3-642-54655-6.
- [9] JENSEN C., PEDERSEN B., THOMSEN C., Multidimensional Databases and Data Warehousing, San Rafael: Morgan & Claypool, 2010. 112 s. ISBN 9781608455386.
- [10] FERRARI, Alberto – RUSSO Marco. The Definitive Guide to DAX: Business intelligence with Microsoft Excel, SQL Server Analysis Services, and Power BI. 1. vyd., Redmond, WA: Microsoft Press, 2015. 1182 s. ISBN 978-0-7356-9835-2.
- [10] NEGASH Solomon. Business intelligence. Communications of the Association for Information System 13(1):177–195. 2004. 20 s. Dostupné na: https://www.academia.edu/400082/Business_Intelligence?auto=download
- [11] BAKER, Pam. Tableau Desktop Review [online]. 2018. [cit. 2021-07-25]. Dostupné na: <https://www.pcmag.com/reviews/tableau-desktop>

- [12] VOJTYLA, Michal. Čo je business intelligence a ako môže pomôcť podnikateľom?. 2014, [online], Dostupné na: <https://www.podnikajte.sk/informacne-technologie/business-intelligence>
- [13] PRATT, Mary. FRUHLINGER, Josh. What is business intelligence? Transforming data into business insights. [online], 2019. [cit. 2021-07-18], Dostupné na: <https://www.cio.com/article/2439504/business-intelligence-definition-and-solutions.html>
- [14] IMMON, William. Building the Data Warehouse. 4. vyd. Indianapolis: Wiley Publishing, Inc. 2005. 576 s. ISBN-10: 0-7645-9944-5.
- [15] GARTNER. Glossary self-service business intelligence [online]. 2013. [cit. 2021-07-21]. Dostupné na: <http://www.gartner.com/it-glossary/self-service-business-intelligence>
- [16] TECHNOPEdia. Self-Service Business Intelligence (SSBI) [online]. 2017 [cit. 2021-07-21]. Dostupné na: <https://www.techopedia.com/definition/30486/self-service-business-intelligence-ssbi>
- [17] SAVER, Peter. OLAVSRUD, Thor. Top 12 BI tools of 2021 [online]. 2021. [cit. 2021-07-22]. Dostupné na: <https://www.cio.com/article/3322749/top-business-intelligence-bi-tools.html>
- [18] DATAPINE. Business Intelligence Tools: Discover a List of The 12 Best BI Tools On The Market [online]. 2021. [cit. 2021-07-22]. Dostupné na: <https://www.datapine.com/articles/best-bi-tools-software-review-list>
- [19] RICHARDSON, James et al. Magic Quadrant for Analytics and Business Intelligence Platforms [online]. 2021. [cit. 2021-07-23]. Dostupné na: <https://www.gartner.com/doc/reprints?id=1-26QIXY7E&ct=210707&st=sb>
- [20] DUNCAN, Owen. . Prehľad jazyka DAX [online]. 2021. [cit. 2021-07-25]. Dostupné na: <https://docs.microsoft.com/sk-sk/dax/dax-overview>
- [21] LLOPIS Miguel. Čo je Power Query? [online]. 2020. [cit. 2021-07-25]. Dostupné na: <https://docs.microsoft.com/sk-sk/power-query/power-query-what-is-power-query>
- [22] HART, Michele. Základné pojmy pre návrhárov v službe Power BI [online]. 2021. [cit. 2021-07-25]. Dostupné na: <https://docs.microsoft.com/sk-sk/power-bi/fundamentals/service-basic-concepts>
- [23] THOMAS, Steve. Power BI Review Disadvantages Advantages Pros Cons [online]. 2021. [cit. 2021-07-25]. Dostupné na: <https://www.businessintelligencesoftware.co/power-bi-review.html>

Prílohy

Príloha A: Kód skriptovacieho jazyka M v Power Query

let

Source =

Csv.Document(File.Contents("C:\Users\User\Downloads\reservations_accommodations.csv"),[Delimiter=";", Columns=11, Encoding=65001, QuoteStyle=QuoteStyle.None]),

#"Promoted Headers" = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),

#"Zmenad Type" = Table.TransformColumnTypes(#"Promoted Headers",{{"id", Int64.Type}, {"stamp", Int64.Type}, {"dospeli", type text}, {"deti", type text}, {"food_type", type text}, {"email", type text}, {"country", type text}, {"from_date", Int64.Type}, {"to_date", Int64.Type}, {"lang", type text}, {"accommodation", Int64.Type}}),

#"Added Custom1" = Table.AddColumn(#"Zmenad Type", "TextSplit", each Text.Split([dospeli], ";")),

#"Added Custom2" = Table.AddColumn(#"Added Custom1", "TextSplitSelected", each List.FindText([TextSplit], "")),

#"Added Custom3" = Table.AddColumn(#"Added Custom2", "ListTransform", each List.Transform([TextSplitSelected], each Text.BetweenDelimiters(_, "" , ""))),

#"Added Custom4" = Table.AddColumn(#"Added Custom3", "ListAccumulateResult", each try List.Accumulate([ListTransform], 0, (state, current) => Number.From(state) + Number.From(current)) otherwise 0),

#"Filtered Rows" = Table.SelectRows(#"Added Custom4", each ([ListAccumulateResult] <> null))

in

#"Filtered Rows"

Príloha B – Celkový prehľad predajov (hlavný dashboard)

Celkový prehľad predajov

Minimum: Q1 2014

€ 419 394

Celkový predaj

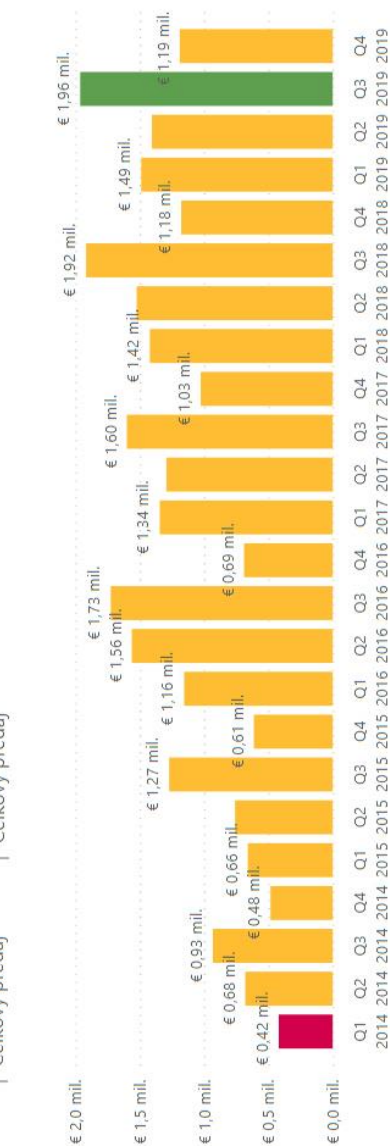
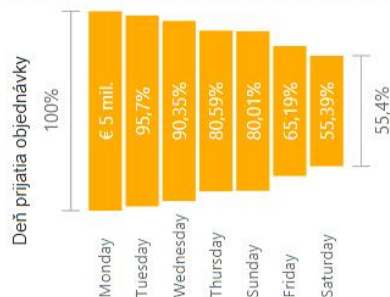
Maximum: Q3 2019

€ 1 963 111

Celkový predaj

cze eng ger hun pol svk

Objednávky Pobyty



- Kraj
- Banskobystrický kraj
 - Bratislavský kraj
 - Košický kraj
 - Nitriansky kraj
 - Prešovský kraj
 - Trenčiansky kraj
 - Trnavský kraj
 - Žilinský kraj

- Mesto, obec
- Bešeňová
 - Bratislava
 - Demänovská Dolina
 - Donovaly
 - Dunajská Streda
 - Hrabušice
 - Kaluža
 - Patince
 - Podhájska
 - Senec
 - Štúrovo
 - Tatranská Lomnica

Typ ubytovania	Celkový predaj
Hotely	€ 8 548 333
Penzióny	€ 8 396 884
Chaty a chalupy	€ 3 629 820
Apartmenty	€ 3 541 157
Priváty	€ 1 531 155
Bungalovy	€ 805 359
Ubytovne	€ 727 474
Rekreačné zariadenia	€ 518 705
Kúpeľné hotely a liečebné domy	€ 225 307
Hostely	€ 194 207
Rekreačné domy	€ 91 650
Motely	€ 65 038
Farmy a ranče	€ 28 732
Celkovo	€ 28 303 819



Príloha C: Analýza predaja podľa dĺžky pobytu a dňa nástupu (dashboard)

Predaje podľa dĺžky pobytu a dňa nástupu

Deň nástupu na pobyt

Friday	Monday	Saturday	Sunday	Thursday	Tuesday	Wednesday
--------	--------	----------	--------	----------	---------	-----------

Počet nocí

1	2	3	4	5	6	7	8	9	10	11	12	13	14
---	---	---	---	---	---	---	---	---	----	----	----	----	----

Objednávky / Pobyt (Celkový predaj)

Celkové predaje podľa dĺžky pobytu a nástupného dňa na pobyt

Poč. nocí	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Celkovo
7	€ 310 633	€ 144 895	€ 113 456	€ 572 354	€ 106 839	€ 113 456	€ 5 698 052	€ 5 698 052
3	€ 556 725	€ 318 076	€ 500 130	€ 318 076	€ 358 881	€ 500 130	€ 4 989 432	€ 4 989 432
2	€ 312 781	€ 406 141	€ 230 154	€ 188 948	€ 230 154	€ 266 716	€ 4 518 841	€ 4 518 841
4	€ 1 014 040	€ 620 914	€ 413 520	€ 398 556	€ 413 520	€ 908 861	€ 4 047 737	€ 4 047 737
5	€ 751 377	€ 334 357	€ 173 761	€ 463 504	€ 328 838	€ 141 015	€ 2 448 211	€ 2 448 211
6	€ 773 431	€ 479 167	€ 104 716	€ 599 343	€ 78 661	€ 137 327	€ 2 299 087	€ 2 299 087
1	€ 92 678	€ 681 509	€ 157 877	€ 71 986	€ 89 574	€ 115 374	€ 1 531 070	€ 1 531 070
8	€ 68 088	€ 348 184	€ 34 464	€ 41 826	€ 26 928	€ 40 830	€ 637 314	€ 637 314
9	€ 93 922	€ 69 714	€ 32 039	€ 29 831	€ 35 019	€ 27 643	€ 334 398	€ 334 398
14	€ 16 879	€ 32 278	€ 201 704	€ 41 674	€ 10 969	€ 11 494	€ 325 993	€ 325 993
10	€ 52 344	€ 64 809	€ 43 365	€ 43 202	€ 49 392	€ 24 306	€ 320 776	€ 320 776
11	€ 13 843	€ 72 250	€ 36 793	€ 22 802	€ 8 034	€ 14 174	€ 190 603	€ 190 603
12	€ 15 681	€ 65 444	€ 24 815	€ 39 152	€ 7 560	€ 14 636	€ 186 365	€ 186 365
13	€ 6 299	€ 49 191	€ 27 408	€ 45 074	€ 7 557	€ 7 984	€ 151 856	€ 151 856
Celkovo	€ 5 293 552	€ 4 233 439	€ 7 896 568	€ 2 789 693	€ 3 370 456	€ 1 772 081	€ 2 323 945	€ 27 679 735

Destinácie zoraďené podľa predaja

Jazyková verzia

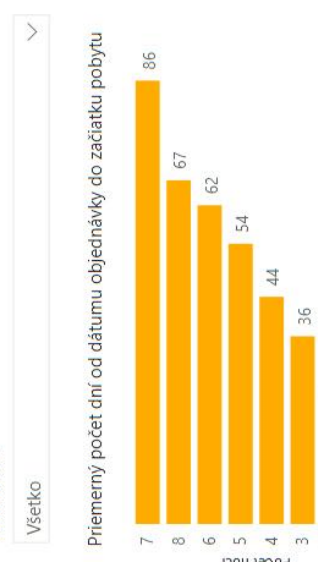
cze	hun
eng	pol
ger	svk

City

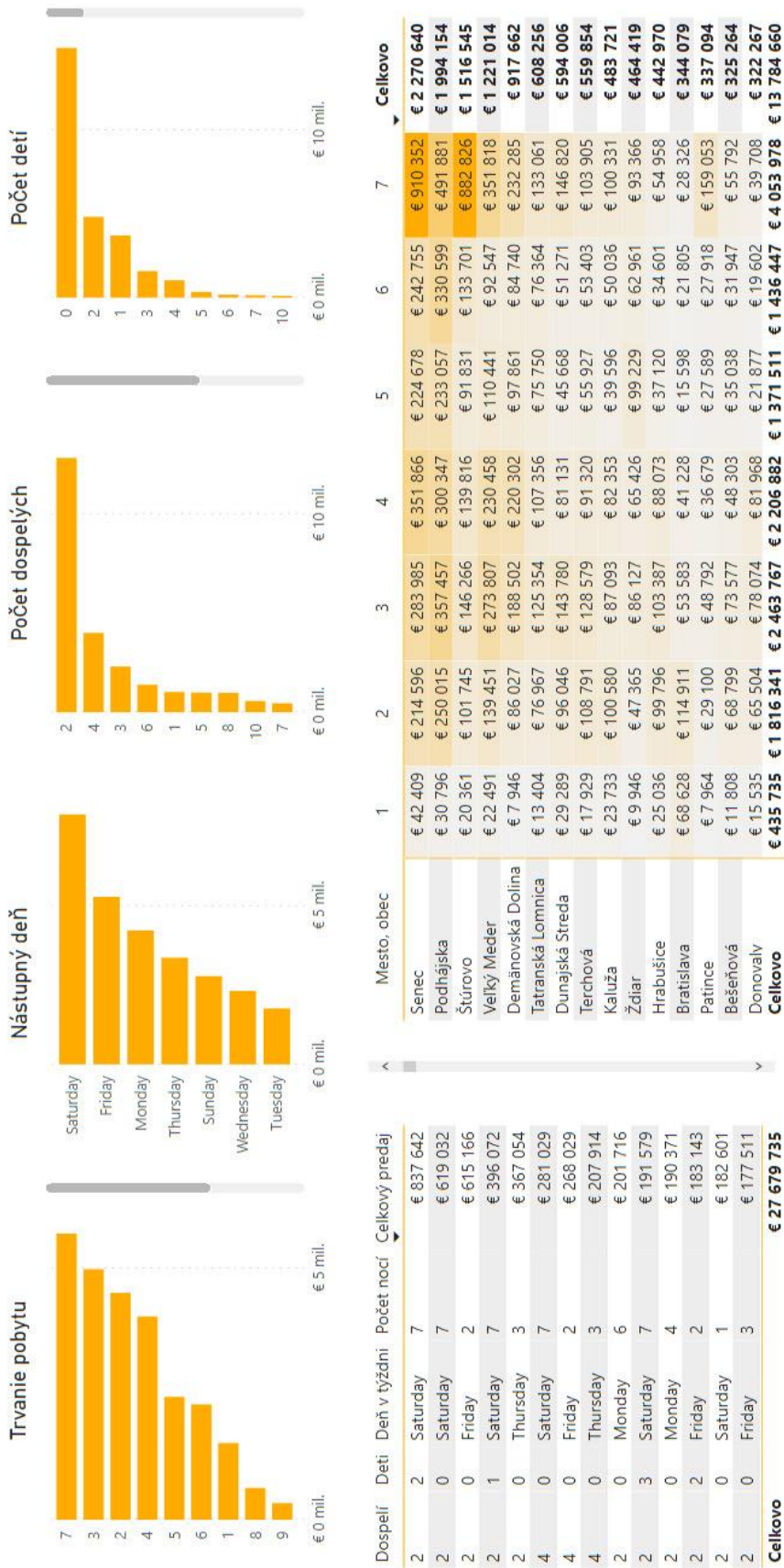
Jan	May	Sep
Feb	Jun	Oct
Mar	Jul	Nov
Apr	Aug	Dec

Príloha C: Mesiac objednávky vs mesiac pobytu (dashboard)

		2014	2015	2016	2017	2018	2019	cze	eng	ger	hun	pol	svk	
Mesiac objednávky vs mesiac pobytu (počet rezervácií a suma pobytov)		Počet rezervácií v mesiaci (riadok) na mesiac (stĺpec)												
Počet nocí	Destinácia	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Celkovo
1	Všetko	2630	3350	633	307	434	532	1387	579	71	12	80	10015	
2		7	6	2152	1436	901	620	1457	646	90	12	2	37	7366
3		1	4	6	2402	2027	977	1857	846	153	23	2	13	8311
4		3	6	1	2663	2677	3380	1427	225	33	7	23	10445	
5		10	4	6	12	4383	8997	2949	374	49	9	32	16825	
6		16	19	2	1	12	26	14427	10979	892	153	25	91	26643
7		60	38	9	7	64	22	72	15169	3193	452	59	463	19608
8		125	151	18	10	64	37	30	30	4870	2080	270	994	8686
9		240	259	44	25	113	82	109	63	8	3476	1197	1687	7303
10		458	422	59	41	90	97	171	69	22	2097	3044	6570	136613
11		1205	498	110	47	124	121	293	114	14	1	3771	6298	136613
12	Celkovo	4745	8084	4499	4768	6982	10087	33598	33565	10043	6300	3668	10274	136613
Celková suma rezervácií vykonaných v mesiaci (riadok) na mesiac (stĺpec)		Celková suma rezervácií vykonaných v mesiaci (riadok) na mesiac (stĺpec)												
Mesiac	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Celkovo	
Jan	€ 393 736	€ 807 322	€ 176 375	€ 85 131	€ 133 203	€ 208 097	€ 566 594	€ 222 832	€ 32 800	€ 4 548	€ 46 443	€ 2 677 081		
Feb	€ 482 219	€ 482 219	€ 282 283	€ 122 282	€ 176 938	€ 168 508	€ 499 556	€ 254 852	€ 52 825	€ 3 949	€ 46 222	€ 2 089 634		
Mar	€ 3 715	€ 2 251	€ 232 209	€ 254 028	€ 167 776	€ 181 418	€ 553 703	€ 251 070	€ 38 426	€ 4 495	€ 1 350	€ 1 714 963		
Apr	€ 295	€ 1 792	€ 2 422	€ 249 137	€ 297 192	€ 217 827	€ 559 758	€ 270 149	€ 53 864	€ 6 485	€ 280	€ 1 666 335		
May	€ 2 361	€ 2 423	€ 450	€ 278 883	€ 425 408	€ 973 335	€ 481 241	€ 78 044	€ 14 692	€ 2 275	€ 11 554	€ 2 270 665		
Jun	€ 5 248	€ 5 092	€ 2 480	€ 1 262	€ 5 267	€ 477 349	€ 1 916 125	€ 730 543	€ 97 248	€ 9 988	€ 4 107	€ 3 268 681		
Jul	€ 52 295	€ 7 223	€ 14 247	€ 825	€ 3 384	€ 13 403	€ 2 132 478	€ 2 067 397	€ 211 301	€ 38 126	€ 7 551	€ 4 531 200		
Aug	€ 61 685	€ 28 027	€ 6 355	€ 1 807	€ 42 157	€ 14 046	€ 42 034	€ 1 949 432	€ 515 184	€ 109 343	€ 13 938	€ 3 008 305		
Sep	€ 107 325	€ 94 806	€ 7 101	€ 8 238	€ 23 006	€ 20 430	€ 17 886	€ 21 821	€ 527 119	€ 364 045	€ 61 904	€ 507 052	€ 1 715 093	
Oct	€ 158 626	€ 122 304	€ 18 090	€ 11 554	€ 64 406	€ 34 229	€ 57 811	€ 35 651	€ 8 354	€ 358 576	€ 199 604	€ 585 296	€ 1 603 199	
Nov	€ 237 987	€ 173 900	€ 23 988	€ 18 117	€ 45 754	€ 44 786	€ 80 418	€ 39 592	€ 13 747	€ 195 235	€ 800 079	€ 1 594 240		
Dec	€ 1 023 273	€ 1 891 119	€ 36 194	€ 17 631	€ 46 585	€ 52 245	€ 133 889	€ 49 875	€ 4 843	€ 280	€ 617 489	€ 1 386 136		
Celkovo	€ 1 023 273	€ 1 916 475	€ 801 744	€ 770 461	€ 1 284 550	€ 1 857 744	€ 7 533 590	€ 6 374 454	€ 1 633 754	€ 914 527	€ 486 243	€ 2 928 719	€ 27 525 533	



Príloha D: Prehľad predajov podľa typu pobytu

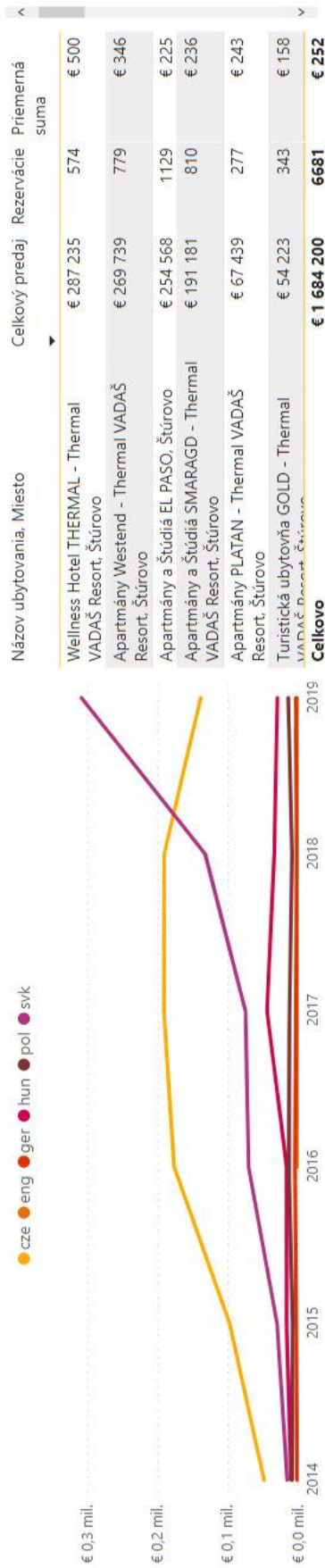


Príloha E: Podrobná analýza destinácie (mesto/obec)



Štúrovo

Dátum objednávky



Príloha F: Podrobná analýza ubytovacieho zariadenia

