

Vypracovanie projektu SQL

Cieľom tohto projektu bolo pripraviť robustné dátové podklady a odpovedať na vopred stanovené výskumné otázky. K vypracovaniu projektu som mala dostupné dátové sady z „Portálu otvorených dát ČR“ a tabuľky s dátami.

Postup pri tvorbe primárnej tabuľky

Najnáročnejšia bola tvorba primárnej tabuľky, kde bolo potrebné si vopred definovať, čo konkrétne v tejto tabuľke budem potrebovať, aby som vedela vytvoriť selecty a odpovedať na výskumné otázky. Vytvorila som si pomocné tabuľky, ktoré som potom spojila do jednej veľkej primárnej tabuľky.

Zo základnej tabuľky „*Czechia_payroll*“ som si vybrala potrebné stĺpce a pripojila číselník pre lepšiu orientáciu v tabuľke pri hodnotení výsledkov. Prvý plán bol „*value*“ spočítať pomocou SUM agregátnej funkcie ale pri kontrole dát som prišla na to, že v poslednom roku chýbajú údaje k 2 kvartálom, tak som sa vybrala cestou priemeru (AVG). Ešte som si overila, že v žiadnom roku nie je v kvartály NULL.

Ako ďalší krok som si upravila pomocou agregátnych funkcií aj tabuľku „*Czechia_price*“. Urobila som priemer cien a upravila roky, aby som pomocou rokov mohla tabuľku prepojiť.

Vrátila som sa k úprave mnou vytvorenej tabuľky

„*t_zuzana_salamon_SQL_primary_payroll*“ a pridala som stĺpec s hodnotou predošlého roku, stĺpec s rozdielom a stĺpec s rozdielom vyjadrený v percentách. Pre kontrolu správnosti som si overila počet riadkov s predošlou tabuľkou.

Podobne som si upravila aj tabuľku „*Czechia_price*“ a tabuľku s HDP („*economies*“), kde som si podmienkou vyselektovala len Českú republiku.

V poslednom kroku som spojila upravené pomocou agregátnych funkcií tabuľky *price*, *payroll* a *GDP* a zaokrúhlila stĺpec, kde som vypočítala množstvo potravín.

Postup pri tvorbe sekundárnej tabuľky

Pri tvorbe sekundárnej tabuľky som si overila s akými rokmi budeme pracovať pomocou funkcií SUM a MAX z primárnej tabuľky. Spojila som dve tabuľky s potrebnými stĺpcami.

Vypracovanie otázok

Otázka č. 1

„Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají?“

V nami sledovaných rokoch nerastú všetky mzdy. V priebehu rokov v niektorých odvetviach mzdy klesli. Najvyšší pokles mzdy bol v odvetví „*Peněžnictví a pojišťovnictví*“, roku 2013 a jednalo sa o pokles 9,68 %.

Otázka č.2

„Kolik je možné si koupit litrů mléka a kilogramů chleba za první a poslední srovnatelné období v dostupných datech cen a mezd?“

Prvom kroku pri tejto otázke som si vyseletovala prvé a posledné porovnateľné obdobie. Boli to roky 2006 a 2018. Ak by sme brali v úvahu, že hodnoty, kde sú uvedené v Industry hodnoty NULL, sú priemerné mzdy za všetky odvetvia, tak môžeme povedať nasledovné:

V roku 2006 sme mohli za ročnú mzdu kúpiť 1 211,64 kilogramov „chleba konzumného kmínového“ a 1 616,9 litrov „mlieka“.

V roku 2018 sme mohli za ročnú mzdu kúpiť 1 321,99 kilogramov „chleba konzumného kmínového“ a 1 353,11 litrov „mlieka“.

Vypracovanie otázky č.3

„Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší percentuální meziroční nárůst)?“

Najpomalšie zdražil “Rostlinný roztíratelný tuk” a to konkrétne o 0,016%.

Vypracovanie otázky č.4

„Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (větší než 10 %)?“

Áno, v roku 2008, 2009 a 2018 bol medziročný nárast cien výrazne vyšší ako nárast miezd.

Vypracovanie otázky č.5

„Má výška HDP vliv na změny ve mzdách a cenách potravin? Neboli, pokud HDP vzroste výrazněji v jednom roce, projeví se to na cenách potravin či mzdách ve stejném nebo následujícím roce výraznějším růstem?“

Pri poslednej otázke som hľadala súvis ako vplýva výška HDP na ceny potravín a výšku miezd. Porovnávala som samostatne v selectoch najvyššiu zmenu HDP za jeden rok a najvyššiu zmenu HDP za uplynulé roky samostatne so mzdami a cenami potravín.

Významnú koreláciu som nenašla. Mohli by konštatovať, že najvyšší nárast HDP v nami dostupných dát, bol v roku 2006 (6,338%) a na hrubý odhad má výška vplyvom klesajúceho HDP, taktiež klesajúcu tendenciu.

Koreláciu medzi výškou HDP a cien produktov som nenašla.