

PRŮVODNÍ LISTINA K PROJEKTU SQL

Tímto bych Vám ráda představila výsledky analýzy dat, na kterých lze srovnat dostupnost potravin na základě průměrných příjmů v letech 2006–2018. Pracovala jsem s daty za toto období proto, že právě v těchto letech byly dostupné kompletní data potřebná pro zodpovězení pěti předem definovaných otázek. Uvedla jsem nejen odpovědi na tyto otázky, ale také postup, jak jsem se k odpovědi na otázky dopracovala.

Využité tabulky a číselníky sdílených informací o ČR:

- *czechia_payroll* – Informace o mzdách v různých odvětvích za několikaleté období. Datová sada pochází z Portálu otevřených dat ČR.
- *czechia_payroll_calculation* – Číselník kalkulací v tabulce mezd.
- *czechia_payroll_industry_branch* – Číselník odvětví v tabulce mezd.
- *czechia_payroll_unit* – Číselník jednotek hodnot v tabulce mezd.
- *czechia_payroll_value_type* – Číselník typů hodnot v tabulce mezd.
- *czechia_price* – Informace o cenách vybraných potravin za několikaleté období. Datová sada pochází z Portálu otevřených dat ČR.
- *czechia_price_category* – Číselník kategorií potravin, které se vyskytují v našem přehledu.
- *czechia_region* – Číselník krajů České republiky dle normy CZ-NUTS 2.
- *czechia_district* – Číselník okresů České republiky dle normy LAU.
- *countries* - Všechné informace o zemích na světě, například hlavní město, měna, národní jídlo nebo průměrná výška populace.
- *economies* - HDP, GINI, daňová zátěž, atd. pro daný stát a rok.

Tvorba primární tabulky:

V této tabulce jsem si připravila datové podklady, na kterých bude možné vidět srovnání dostupnosti potravin na základě průměrných příjmů. Při tvorbě této tabulky jsem si vytvořila 5 pomocných view:

1. *Czechia_payroll_base*:

Tento view spojuje tabulku *czechia_payroll* s číselníky: *czechia_payroll_industry_branch*, *czechia_payroll_value_type* a *czechia_payroll_calculation*. Z tabulky *czechia_payroll* jsem vyloučila řádky s nulovou hodnotou (*value*, *name*). Z *czechia_payroll_value_type* jsem vybrala pouze kód 5958 – průměrnou hrubou mzdou na zaměstnance.

2. *Czechia_payroll_average_by_quarter*

V tomto view vypočítávám průměr z calculation_type „fyzický a přepočtený“ pro jednotlivé roky – pro jejich jednotlivá čtvrtletí ve všech odvětvích. Používám data z view czechia_payroll_base.

3. *Czechia_payroll_average_by_year*

V tomto view počítám průměr pro celý rok a odvětví. Používám data z view czechia_payroll_base.

4. *Czechia_price_base*

V tomto view spojuji číselník czechia price category s tabulkou czechia price. Vyřazuji řádky s nulovými hodnotami ve sloupci category code, vyřazuji řádky s nulovými hodnotami ve sloupci date_from, vyřazuji řádky s nulovými hodnotami ve sloupci value.

Dále zde převádím datum do formátu year.

5. *Czechia_price_average_by_year*

V tomto view využívám již připravený view czechia_price_base a vypočítám si průměr cen za celý rok pro daný produkt. Poté vytvářím samotnou finální tabulku r_lucie_zackova_project_sql_primary_final, a to spojením view czechia_payroll_average_by_year a czechia_price_average_by_year. Spojením těchto dvou view dostanu finální tabulku.

Pro každý řádek ve finální tabulce přidám ID - AUTO_INCREMENT.

Tvorba sekundární tabulky:

Z tabulky Economies jsem vybrala sloupec country, year a vypočítala jsem HDP z GDP/population. Dále jsem eliminovala řádky s nulovými hodnotami u GDP a u Population. Vybrala jsem společné roky s primární tabulky. Vybrala jsem pouze unikátní roky.

Výzkumné otázky a odpovědi

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

Odpověď: obecně lze říci, že mzdy napříč odvětvími spíše rostou. U některých odvětví lze vidět i pokles, ale to zpravidla jen v jednom roce, v roce následujícím dochází opět k růstu cen.

Postup:

Pro zjištění této odpovědi jsem si připravila dva pomocné sloupce:

- V každém odvětví jsem si zprůměrovala mzdy za jednotlivá čtvrtletí (median_val)
- V dalším sloupečku jsou zobrazeny data s průměrnými mzdami za předchozí rok (previous_year_median)

Poslední sloupeček je nazvaný difference a znázorňuje, zda mzdy v jednotlivých odvětvích v po sobě jdoucích letech rostou nebo klesají.

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

Odpověď: V nejstarším srovnatelné období, v roce 2006, si bylo možné za průměrnou mzdu pořídit 1249 kilogramů chleba a 1424 litrů mléka. V nejmladším srovnatelném období, v roce 2018, si bylo možné za průměrnou mzdu pořídit 1344 kilogramů chleba a 1646 litrů mléka.

Postup: Nejprve jsem si našla společné srovnatelné období pro průměrný plat ve všech odvětvích a pro ceny produktů, konkrétně cenu chleba a cenu mléka. Tímto bylo období 2006-2018. Z nejstaršího srovnatelného tedy z roku 2006 období si vezmu průměrný plat a vydělím ho průměrnou cenou produktu. Tím získám počet kg chleba/litrů mléka, které bylo možné v tomto období koupit. To stejné udělám za rok 2018, což je nejmladší srovnatelné období.

Která kategorie potravin zdražuje nejpomaleji?

Odpověď: Nejpomaleji se zdražovaly banány a v průměru o 0,7 %. Bylo zjištěno, že dva produkty průměrně zlevňovaly a to cukr krystal o 1,99 % a rajská jablka červená kulatá o 0,82 %.

Postup:

Pro každý produkt, jsem si pro každý rok připravila průměrnou cenu produktu a procentuální rozdíl v ceně oproti minulému roku a tyto hodnoty jsem zprůměrovala – vznikla jedna hodnota pro každý produkt – o kolik procent se v průměru zdražil/zlevnil produkt oproti minulým rokům. Tyto hodnoty jsem si seřadila – výsledek začíná produktem který „zdražuje“ nejpomaleji.

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

Odpověď: rozdíl větší než 10%, mezi průměrnou mzdou a průměrnou cenou produktů byl zjištěn v roce 2012.

Postup:

Vytvořila jsem si dva sloupce:

- Průměrné procentuální zvýšení nebo snížení cen ze všech kategorií produktů pro daný rok.
- Průměrné procentuální zvýšení nebo snížení platů ve všech odvětvích pro daný rok.

Dále jsem použila CASE a nastavila jsem, že pokud bude rozdíl průměrné cena produktů a průměrné výše platů větší než 10 – vypiš 1.

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?

Odpověď:

Lze říct, že když se HDP v jednom roce výrazněji zvýší, mzdy v následujících letech stoupnou a ceny se snižují – není to však pravidlem - viz rok 2016, kdy se ceny i přes růst HDP v přechodném roce mírně zvýšily.

Postup:

Připravila jsem si sloupeček s průměrnými mzdami (ze všech odvětví) a sloupeček s procentuálním nárůstem mezd oproti předchozímu roku. Dále jsem připravila sloupeček s průměrnými cenami potravin a sloupeček s procentuálním nárůstem cen oproti přechodnému roku (průměr cen všech produktů) pro jednotlivé roky 2007-2018. Poslední dva sloupečky znázorňují výši HDP v jednotlivých letech a také procentuální nárůst nebo pokles HDP oproti přechodnému roku.

Za roky s výraznějším nárůstem HDP považuji:

- rok 2007, kdy se oproti předchozímu roku HDP zvýšil o 4,96 procent. Průměrný plat v následujícím roce vzrostl o 2,29 % a průměrné ceny klesly o 3,47 %.
- dále rok 2015, kdy se HDP zvýšil o 5,18 procent. Průměrný plat se v následujícím roce zvýšil o 6,44 procent a průměrné ceny se mírně zvýšily také – o 0,19 procent.
- rok 2017 kdy se HDP zvýšil o 4,89 procent. Průměrný plat se v následujícím roce zvýšil o 7,13 procent a průměrné ceny klesly o 9,81 procent.