

Projekt z SQL - Data o mzdách a cenách potravin a jejich zpracování pomocí SQL

1. Zadání projektu: Zjistit dostupnost základních potravin široké veřejnosti a jejich porovnání s příjmy v rámci České Republiky.

Použité datové sady:

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.economies - HDP, GINI, daňová zátěž, atd. pro daný stát a rok

Při vytváření primární tabulky **t_milan_boleslav_project_SQL_primary_final** jsem použil některá omezení, která vycházejí z potřeb zadání:

```
cp.value_type_code = 5958 -> Průměrná hrubá mzda na zaměstnance  
AND cp.calculation_code = 200 -> přepočtené hodnoty mezd  
AND cp.unit_code = 200 -> zajímají nás data o mzdách  
AND cp.industry_branch_code IS NOT NULL -> vyřazení neznámých odvětví  
AND cp.payroll_year BETWEEN 2006 AND 2018 -> sjednocení průniku let v tabulkách  
mezd a cen potravin (czechia_payroll a czechia_price.)
```

Sekundární tabulku **t_milan_boleslav_project_sql_secondary_final** jsem vytvořil pouze pro Českou Republiku, protože ceny potravin a platů známe také jen pro ČR a pro zodpovězení otázek je to dostačující.

2. Výzkumné otázky:

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

Výsledkem je 25 záznamů, ve kterých došlo k poklesu ročního příjmu a to v různých odvětvích. Nejčastěji zasaženým odvětvím byla - Těžba a dobývání. Velkým překvapením pro mě byl pokles příjmu i v odvětví - Informační a komunikační činnosti v roce 2013.

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

Z výsledků je vidět kolik kusů chleba a litrů mléka lze v každém odvětví zakoupit pro první a poslední srovnatelné období tzn. Za rok 2006 a 2018.

Z toho je zřejmé, že i přes zdražení mléka, od roku 2006 do roku 2018 si lze díky růstu platů v každém odvětví zakoupit více litrů mléka. U chleba je situace jiná a to v těchto odvětvích: Ostatní činnosti, Peněžnictví a pojišťovnictví a Těžba a dobývání. Tam růst platů nebyl dostačující na to, aby pokryl zdražení chleba.

3. *Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší procentuální meziroční nárůst)?*

Pokud to budeme brát z hlediska meziročního rozdílu cen tak nejlépe jsou na tom Rajská jablka červená kulatá, která mezi lety 2006 a 2007 zlevnila o více jak 40%. Pokud jde opravdu pouze o zdražení, tak tam došlo k nejmenšímu nárůstu u položky: Rostlinný roztíratelný tuk, který také mezi lety 2006 a 2007 podražil pouze o cca 0,075%. Některé položky se meziročně někdy nezměnily jako např.: Pšeničná mouka hladká mezi léty 2017 a 2018

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

Vzhledem k tomu že máme data o cenách zboží dostupná až od roku 2006, můžeme meziroční porovnání udělat až od roku 2007. Ve sledovaném období takový rok nenastal. Nejvyššího rozdílu mezi růstem cen a mezd bylo dosaženo v roce 2013, kdy ceny potravin vzrostly o zhruba 6% více než ceny mezd. Naopak největší kupní síla obyvatel v ČR byla v roce 2009.

Z detailního výsledku je patrné, že největší nárůst cen potravin byl v roce 2007 a to konkrétně u paprik kde nárůst činil oproti roku 2006 skoro 50%. V témže roce byl nejnižší nárůst platů v odvětví Zásobování vodou, Činnosti související s odpady a sanacemi a to pouze okolo 5%

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?*

V daném vzorku dat nevidím přímou závislost mezi růstem HDP a růstem cen zboží či platů. Co by se dalo říci je, že pokud HDP výrazně vzrostlo tak jak tomu bylo v letech 2006 a 2007, následně vzrostly výrazně i platy a cena zboží. Ke stejné situaci došlo i v roce 2018 kdy po výrazném nárůstu HDP v předchozím roce, došlo opět k výraznému zvýšení platů. Cena zboží v tomto roce vzrostla celkem zanedbatelně, pouze o 2% ale v roce minulém, tedy 2017, vzrostla o skoro 9%, což je nejvíce, v námi porovnatelné období.