

# Projekt SQL

## Zadání projektu

Na vašem analytickém oddělení nezávislé společnosti, která se zabývá životní úrovní občanů, jste se dohodli, že se pokusíte odpovědět na pár definovaných výzkumných otázek, které adresují **dostupnost základních potravin široké veřejnosti**. Kolegové již vydefinovali základní otázky, na které se pokusí odpovědět a poskytnout tuto informaci tiskovému oddělení. Toto oddělení bude výsledky prezentovat na následující konferenci zaměřené na tuto oblast.

Potřebuji k tomu **od vás připravit robustní datové podklady**, ve kterých bude možné vidět **porovnání dostupnosti potravin na základě průměrných příjmů za určité časové období**.

Jako dodatečný materiál připravte i tabulku s HDP, GINI koeficientem a populací **dalších evropských států** ve stejném období, jako primární přehled pro ČR.

## Výzkumné otázky

1. Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají?
  2. Kolik je možné si koupit litrů mléka a kilogramů chleba za první a poslední srovnatelné období v dostupných datech cen a mezd?
  3. Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší procentuální meziroční nárůst)?
  4. Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (větší než 10 %)?
  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?
- 

## Příprava datových sad

První fází pro zpracování SQL projektu bylo vytvoření dvou robustních tabulek, z kterých bude následně možné vypracovat odpovědi na výzkumné otázky. První tabulku `t_sabina_sperlova_project_SQL_primary_final` jsem připravila pomocí funkcí JOIN z výchozích

tabulek *czechia\_payroll*, *czechia\_payroll\_industry\_branch*, *czechia\_price* a *czechia\_price\_category*. V nové tabulce jsem selektovala sloupce, které, dle mého uvážení, byly podstatné pro zodpovězení výzkumných otázek. Dále jsem použila funkci průměr, která mi vrátila průměrné ceny dle kategorií a průměrné mzdy dle pracovních odvětví. Druhá tabulka nazvaná *t\_sabina\_sperlova\_project\_SQL\_secondary\_table* byla pro sestavení poměrně jednodušší – vycházela jsem z tabulky *countries*, na kterou jsem navázala tabulku *economies* přes sloupec *country*.

## Odpovědi na výzkumné otázky

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

Meziročně klesly mzdy v několika odvětvích. Jednalo se především o poklesy v roce 2013, 2020 a 2021. První jmenovaný rok byl patrný pokles v 11 odvětvích z 19, především je patrný meziroční pokles v oblasti peněžnictví a pojišťovnictví (necelých 9 %). Tento stav trend zapříčinila pravděpodobně dlouhodobá recese v České republice mezi lety 2012 a 2013. Další poklesy mezd se ve větší míře vyskytovaly v letech 2020 a 2021, které jsou spojeny s pandemií covid-19. Nejvíce ovlivněno bylo odvětví reality (necelých 7 %) a gastronomie a pohostinství (5,87 %). Další poklesy v menší míře během sledovaného období zaznamenala odvětví energetické, zemědělské, těžební, kulturní, vědecké, vzdělávací, stavební, ale i oblast veřejné správy, informatiky či velkoobchodu a maloobchodu.

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

V prvním srovnatelném období, tedy v roce 2006 bylo možné za průměrný plat vypočítaný ze všech odvětví koupit 1.287 kg chleba a 1.437 litrů mléka. V roce 2018 stačil průměrný plat na nákup 1.342 kg chleba a 1.642 litrů mléka. Největší množství mohli nakoupit spotřebitelé pracující v oblasti informačních a komunikačních technologií a peněžnictví a pojišťovnictví, což je shodné pro oba sledované roky.

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

V sledovaném období mezi lety 2006 a 2018 došlo dokonce ke zlevnění některých komodit, a to poměrně znatelně. Nejvíce je to patrné u komodity cukr krystalový, který byl v roce 2018 o 27,5 % levnější, než na začátku sledovaného období v roce 2006, podobný pokles je i rajských jablek (zlevnění o 23 %). U dalších komodit již došlo ke zdražování – nejnižší nárůst můžeme vidět o banánů, které zdražily o 7,4 %.

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

U žádného odvětví nedošlo k výrazně vyššímu nárůstu cen potravin s ohledem na růst mezd. Nejvyšší nárůst ve výši 6,7 % byl zaznamenán v roce 2013 a nárůst o 3,7 % v roce 2012. Obojí opět pravděpodobně v závislosti na silné recesi. Nárůst o 3,35 % evidujeme i v roce 2017.

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?

Dle zobrazených dat nelze jednoznačně říci, zda má změna HDP výrazný vliv na mzdy a ceny potravin. Například v roce 2017 je patrné, že při růstu HDP o 5 % došlo k nárůstu cen (o 10 %) i mezd (o 6,3 %). Naopak v roce 2016, kdy HDP také vzrostlo o 5 %, u cen byl dokonce viditelný pokles o 0,6 % a mzdy vystoupali o 2,5 %. Nelze namítat, že by se nárůst reflektoval až v následujícím roce, protože ceny potravin i nadále klesaly. Přesto, že HDP meziročně kleslo během sledovaného období třikrát (2009, 2012 a 2013, tedy v dobách recesí), na poklesu průměrných mezd se tato skutečnost projevila pouze jednou, a to v roce 2013. Tento trend ovšem není takto patrný u nárůstu a poklesu průměrných cen potravin.