

# Projekt SQL

Na našem analytickém oddělení jsem se pokusila zodpovědět na pár definovaných výzkumných otázek, které se zaměřují na životní úroveň občanů. Otázky adresují dostupnost základních potravin široké veřejnosti.

Na základě rozsáhlých datových sad jsem vytvořila dvě hlavní tabulky, ze kterých budu v projektu vycházet.

První tabulku jsem nazvala `t_petra_novakova_project_sql_primary_final` a vycházela jsem z těchto datových sad:

- `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.

Potřebovala jsem získat data pouze za roky, proto jsem při tvorbě tabulky použila funkci `YEAR` a vytvořila jsem si dva pohledy, které jsem pak spojila přes `INNER JOIN`, abych získala společná data za mzdy a ceny.

Tato tabulka obsahuje robustní datové podklady, ve kterých je možné vidět porovnání dostupnosti potravin na základě průměrných příjmů za určité časové období. Zodpoví na první 4 otázky, a to:

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 %)?

Jako dodatečný materiál jsem připravila druhou tabulku nazvanou `t_petra_novakova_project_sql_secondary_final`. Tato obsahuje HDP, GINI koeficient a populaci dalších evropských států ve stejném období. Použila jsem tyto datové sady:

- `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.

Abych získala data srovnatelná s primární tabulkou, vytvořila jsem opět dva pohledy, ve kterých jsem omezila období na roky 2006-2018 a zároveň omezila země pouze na evropské. Tyto pohledy jsem pak spojila funkcí LEFT JOIN.

Najdeme zde odpověď na poslední otázku:

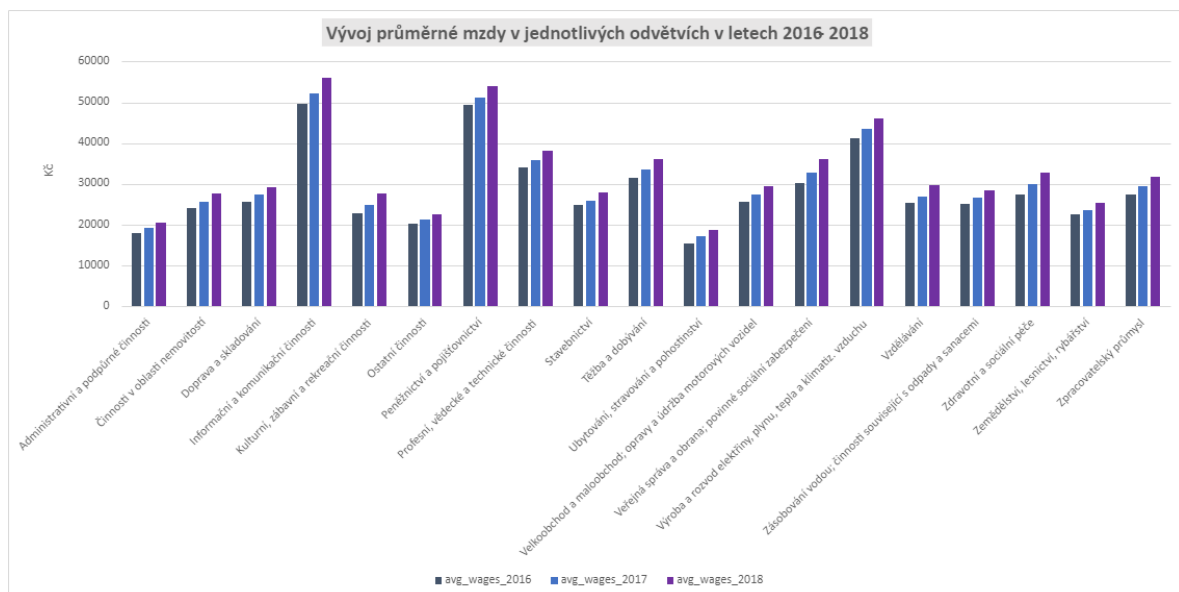
Má výška HDP vliv na změny ve mzdách a cenách potravin? Neboli, pokud HDP vzroste výrazněji v jednom roce, projeví se to na cenách potravin či mzdách ve stejném nebo následujícím roce výraznějším růstem?

### Vizualizace a výsledná analýza:

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

Ano, mzdy rostou v průběhu let ve všech odvětvích. Pro analýzu jsem si vybrala poslední 3 roky z dostupných dat, tj. roky 2016, 2017 a 2018.

Graf č. 1: Růst mezd v období 2016 – 2018



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

Tabulka č. 1: Dostupnost vybraných komodit v letech 2006 a 2018

category_name	avg_price	avg_wages	price_vs_wage	payroll_year
Chléb konzumní kmínový	16,12	20 753,78	1 287,16	2006
Chléb konzumní kmínový	24,24	32 535,86	1 342,32	2018
Mléko polotučné pasterované	14,47	20 753,78	1 437,46	2006
Mléko polotučné pasterované	19,82	32 535,86	1 641,78	2018

V roce 2006 bylo možné si koupit 1287 kg chleba a 1437 l mléka.

V roce 2018 bylo možné si koupit 1342 kg chleba a 1641 l mléka.

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

Percentuálně nejnižší meziroční nárůst evidujeme u Rajských jablek červených a kulatých mezi lety 2006 a 2007.

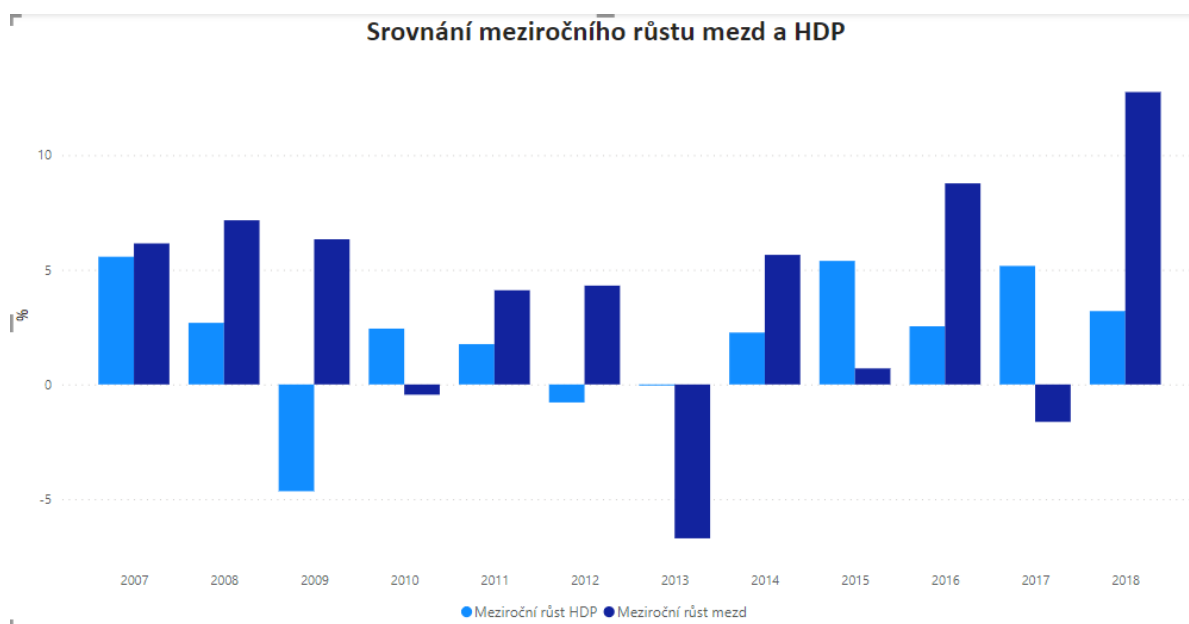
**Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (větší než 10 %)?**

Ano, v letech 2013 a 2017 byl meziroční nárůst cen potravin vyšší než růst mezd o více než 10 %.

**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?**

Výška HDP nemá zásadní vliv na změny ve mzdách a cenách, což dokazují následující grafy.

Graf č. 2: Porovnání meziročních změn HDP a mezd



Graf č. 3: Porovnání meziročních změn HDP a cen potravin

