

## O čem je projekt?

Cílem projektu je zodpovědět na pět otázek týkajících se vývoje cen, mezd a HDP v České republice. Jsou použita následující data:

- czechia\_payroll – Informace o mzdách v různých odvětvích za několikaleté období.  
Datová sada pochází z Portálu otevřených dat ČR.
- czechia\_payroll\_industry\_branch – Číselník odvětví 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.

## Tvorba tabulek

Nejdříve bylo zapotřebí vytvořit dvě hlavní tabulky, a to primární a sekundární. První tabulka s názvem `t_katerina_pekarova_project_SQL_primary_final` je spojením dvou dílčích pomocných tabulek:

Pomocná tabulka payroll je složena z:

- payroll\_year – rok
- industry\_branch\_code – kód odvětví
- industry\_branch\_name - připojené názvy odvětví z tabulky `czechia_payroll_industry_branch`
- payroll\_value – mzdy

Pomocná tabulka price je složena z:

- year – rok
- category\_code – kód kategorie potravin
- category\_name – připojené názvy potravin z tabulky `czechia_price_category`
- price\_value - ceny

Tyto dvě dílčí tabulky byly vytvořeny z důvodu optimalizace a zrychlení při zpracovávání dat.

Sekundární tabulka má název `t_katerina_pekarova_project_SQL_secondary_final` a obsahuje údaje o vývoji HDP v České republice v průběhu let 1990 až do současnosti.

## Výzkumné otázky a závěr

Otázka 1 - Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají? Na základě vstupních dat byl vytvořen přehled s názvy jednotlivých odvětví a k nim byla přiřazena průměrná hodnota mezd v prvním sledovaném roce (rok 2000) a dále průměrná hodnota mezd v posledním sledovaném roce (rok 2021). Z výsledků je patrné, že se mzdy zvýšily v každém odvětví.

Otázka 2 - Kolik je možné si koupit litrů mléka a kilogramů chleba za první a poslední srovnatelné období v dostupných datech cen a mezd? Pro zjištění odpovědi na tuto otázku byl vytvořen přehled obsahující rok 2006 a 2018 (jedná se o první a poslední srovnatelné období) a následně byly dopočteny litry mléka a kilogramy chleba. Výsledkem je množství potravin, které se dají za průměrnou cenu daného roku koupit za sumu všech mezd.

Otázka 3 - Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší procentuální meziroční nárůst)? Pro získání odpovědi byl opět porovnán první a poslední sledovaný rok (roky 2006 a 2018). Byly vytvořeny průměry cen za danou kategorii potravin a porovnána procentuální změna. Z výsledků je patrné, že cukr krystalový a rajska jablka mezi lety zlevnily, naopak nejvíce zdražilo maslo.

Otázka 4 - Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (větší než 10 %)? Pro získání odpovědi byl vytvořen kód, který analyzuje změny průměrných cen potravin a mezd za jednotlivé roky. Výsledkem je seznam, který zobrazuje roční hodnoty změn, zaokrouhlené na 4 desetinná místa, a indikátor, zda je rozdíl mezi změnou cen potravin a mezd větší než 10 %. Vstupní data týkající se cen obsahují informace od roku 2006 do roku 2018 a data týkající se mezd informace od roku 2000 až 2021. Signifikantnějšího rozdílu můžeme zaznamenat v roce 2013.

Otázka 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? Byl vytvořen kód, který umožňuje analyzovat vztah mezi HDP, cenami potravin a mzdami v průběhu let a zkoumá meziroční změny v těchto proměnných. Jak již bylo řečeno, můžeme porovnávat pouze data od roku 2006 do 2018, neb údaje o cenách máme pouze z tohoto rozmezí. Z námi vytvořeného přehledu vyplývá, že mezi změnami HDP, cen a mezd neexistuje prokazatelná závislost.