

A. Zadání projektu

ZADAVATEL:

Analytické oddělení nezávislé společnosti, která se zabývá životní úrovní občanů.

POŽADAVEK:

- a) **Primární tabulka:** Nutná příprava datových podkladů, ve kterých bude možné vidět porovnání dostupnosti potravin na základě průměrných příjmů za určité časové období.
- b) **Sekundární tabulka:** Jako dodatečný materiál připravit i tabulku s HDP, GINI koeficientem a populací dalších evropských států ve stejném období, jako primární přehled pro ČR.

CÍL:

Získat relevantní odpovědi na 5 definovaných výzkumných otázek, které adresují dostupnost základních potravin. Odpovědi na otázky budou poskytnuty tiskovému oddělení. Které bude následně výsledky prezentovat na konferenci zaměřené na téma dostupnosti základních potravin široké veřejnosti.

PODKLADY:

Primární tabulky:

- 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.

Číselníky sdílených informací o ČR:

- czechia_region – Číselník krajů České republiky dle normy CZ-NUTS 2.
- czechia_district – Číselník okresů České republiky dle normy LAU.

Dodatečné tabulky:

- 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.

B. Výstup projektu - postupy

Primární tabulka:

- a) Pro jaké roky jsem schopna vyhodnotit data? Byl můj první krok pátrání. Jelikož jsem zjistila, že data pro ceny potravin jsou plně dostupné pouze v letech 2006-2018 rozhodla jsem se postupovat tak i u mezd (ikdyž tam byla dostupná data v letech 2001-2021). Tudíž jsem vypracovala dvě primární tabulky v letech 2006-2018, které obsahují trendy, výpočty průměrů cen, rozdílů i poklesů, vyčíslení v procentech jak u mezd tak potravin. Tyto dvě tabulky jsem spojila v jednu primární.
- b) Jaká data budu potřebovat k vyhodnocení 5 výzkumných otázek? Dalším krokem bylo vybrat a spojit pouze sloupčky (hodnoty), které budu potřebovat na zodpovězení otázek (sloupce jsou vidět ve výsledné tabulce).
- c) V tabulce u mezd jsem si všimla, že je nutné vyfiltrovat pouze kód 5958, který představuje přepočtenou hrubou mzdu.
- d) Data jsem transformovala do roků (jak u mezd tak u potravin) a neměla jsem potřebu pracovat se čtvrtletími. Nutnost nevyplývala ze zadaných otázek.

Sekundární tabulka:

- a) Základní tabulka vznikla spojením dat ze 2 tabulek *economies* (country, year, GDP, gini a population) s tabulkou *countries* (continent).
- b) Selektce pouze Evropy.
- c) Opět pracuji s lety 2006-2018 jako u primární tabulky.

Výstupy projektu – script / tabulky uložené na Github

[MichaelaBartu/SQL_project \(github.com\)](https://github.com/MichaelaBartu/SQL_project)

- ✓ **t_michaela_bartu_project_sql_primary_final** (pro data mezd a cen potravin za Českou republiku sjednocených na totožné porovnatelné období – společné roky)
- ✓ **t_michaela_bartu_project_SQL_secondary_final** (pro dodatečná data o dalších evropských státech).
- ✓ **Sada 5 skriptů**, které nacházejí odpovědi na otázky 1-5 (níže uvedených).

C. Odpovědi na výzkumné otázky

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

Odpověď: V letech 2006 – 2018 mzdy (meziročně v řadě za sebou) rostou průměrné mzdy **pouze ve 4 odvětvích**. U ostatních byl zjištěn min 1x meziroční pokles. V odvětvích níže byl zaznamán kontinuální růst bez propadů mezd v letech 2006-2018.

C - Zpracovatelský průmysl

H - Doprava a skladování

Q - Zdravotní a sociální péče

S - Ostatní činnosti

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

Odpověď: Dostupná data u cen potravin jsou v letech 2006 – 2018.

V roce 2006 lze koupit za průměrnou mzdu **1.482 l mléka** nebo **1.297 kg chleba**.*

V roce 2018 lze koupit za průměrnou mzdu **1.626 l mléka** nebo **1.355 kg chleba**.*

**Zaokrouhleno na celá čísla směrem dolů.*

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

Odpověď: Potravina, která má nejnižší meziroční % nárůst je **cukr krystalový** (cena ve sledovaném období min v polovině případů meziročně klesla). Cena 22Kč/kg v roce 2006 klesla na 16Kč/kg v roce 2018.

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

Odpověď: U jednotlivých kategorií potravin eviduji nárůst cen vyšší než 10% v mnohem vyšší míře oproti zvyšování mezd, (jedná se pouze o jednotky nikoli celkové průměry). Nejvyšší nárůst cen potravin byl např. v roce 2007 a to u Paprik (cena vyšší o 49%), v tomto roce byl celkový nárůst mezd pouhých 8%. **POKUD však propojím všechny dostupné výsledky meziročního procentuálního nárůstu/poklesu, trh nevykazuje výrazně vyšší nárůst cen potravin než růst mezd. Naopak v roce 2009 zaznamenávám pokles cen potravin oproti růstu mezd o 11%.**

- 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?**

Odpověď: Ze zjištěných dat růstů a poklesů – cen potravin, průměrných mezd a HDP ... Je pravděpodobné, že má růst HDP přímý vliv na růst mezd (nikoliv však cen potravin). Schodu lze vidět pouze u vývoje průměrných mezd, kdy křivka poklesu i nárůstů se z části kopíruje (když roste HDP, rostou i mzdy). Nikoliv však stejným tempem (měřeno v %) a nikdy ne zcela (vyjímkou je např. rok 2012, kdy HDP je v minusu 0,79% oproti roku 2011 a mzdy rostou z 2% v roce 2011 na 3% v roce 2012).