

Průvodní zpráva k projektu v 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.

Potřebují 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.

Datové sady, které je možné použít pro získání vhodného datového podkladu

Primární tabulky:

1. [czechia_payroll](#) – Informace o mzdách v různých odvětvích za několikaleté období. Datová sada pochází z Portálu otevřených dat ČR.
2. [czechia_payroll_calculation](#) – Číselník kalkulací v tabulce mezd.
3. [czechia_payroll_industry_branch](#) – Číselník odvětví v tabulce mezd.
4. [czechia_payroll_unit](#) – Číselník jednotek hodnot v tabulce mezd.
5. [czechia_payroll_value_type](#) – Číselník typů hodnot v tabulce mezd.
6. [czechia_price](#) – Informace o cenách vybraných potravin za několikaleté období. Datová sada pochází z Portálu otevřených dat ČR.
7. [czechia_price_category](#) – Číselník kategorií potravin, které se vyskytují v našem přehledu.

Číselníky sdílených informací o ČR:

1. [czechia_region](#) – Číselník krajů České republiky dle normy CZ-NUTS 2.
2. [czechia_district](#) – Číselník okresů České republiky dle normy LAU.

Dodatečné tabulky:

1. **countries** - Všechné informace o zemích na světě, například hlavní město, měna, národní jídlo nebo průměrná výška populace.
2. **economies** - HDP, GINI, daňová zátěž, atd. pro daný stát a rok.

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?

Výstup projektu

Úvod:

Tato průvodní zpráva slouží k prezentaci výsledků a závěrů z analýzy dat týkajících se mezd, cen potravin a jejich vztahu k výšce HDP. Projekt se zaměřil na důkladné zpracování dostupných dat o mzdách a cenách potravin z různých let s cílem odhalit klíčové trendy a vzájemné souvislosti mezi těmito faktory. V následujících sekcích představíme odpovědi na pět výzkumných otázek, které nás vedly k analýze dat.

Pro každý úkol je vytvořený jednotlivý soubor (Task1 - Task5). Zároveň je samostatný soubor pro vytváření tabulek (tables.sql)

1. *Růst mezd v odvětvích*

- Prozkoumali jsme data a zjistili jsme, že mzdy ve většině odvětvích průběžně rostou v průběhu let. Nicméně ve vybraných odvětvích jsme také zaznamenali pokles mezd.
- Ve většině odvětví bylo nalezeno období ve které mzda meziročně klesala. To si můžeme nádherně zobrazit ve sloupci growing - pokud je tento sloupec vyplněn textem znamená to že mzda meziročně klesala. V ostatních případech mzda meziročně rostla.

2. *Kolik litrů mléka a kilogramů chleba lze koupit za první a poslední srovnatelné období?*

- Analyzovali jsme dostupná data o cenách mléka a chleba za první a poslední srovnatelné období.
- Na základě těchto dat jsme vypočítali, že v prvním období (rok 2006) jsem si mohli koupit 1192 kg chleba a 1331 litrů mléka.
- V posledním období (rok 2018) jsem si mohli koupit 1300 kg chleba a 1590 litrů mléka, tedy více než v orvním období

3. *Která kategorie potravin zdražuje nejpomaleji*

- Identifikovali jsme různé kategorie potravin a analyzovali jsme jejich ceny v průběhu let. Na základě meziročních změn jsme

zjistili že kategorii, která zdražuje nejpomaleji jsou **banány žluté**. Meziroční procentuální nárůst této kategorie je průměrně 0,81%

- Zároveň jsme vypořizovali, že mezi sledovanými daty jsou dokonce dvě kategorie, u kterých cena meziročně klesala. A to jsou: **Rajská jablka červená kulatá** a **Cukr krystalový**

4. *Porovnání růstu cen potravin a mezd*

- Prozkoumali jsme data o meziročních změnách cen potravin a mezd a hledali rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (větší než 10 %). Výsledky ukazují, že takový rok neexistuje

5. *Má výška HDP vliv na změny ve mzdách a cenách potravin?*

- Prostudovali jsme data o výšce HDP v jednotlivých letech a analyzovali jsme její vztah k mzdám a cenám potravin. Zajímalo nás, zda výrazný nárůst HDP v jednom roce ovlivní mzdy a ceny potravin ve stejném nebo následujícím roce výraznějším růstem. Jelikož nebylo přesně zadáno co znamená výrazný růst, rozhodli jsme se jako tuto hranici použít 5%
- Výrazný růst HDP byl zaznamenán mezi lety 2006 - 2007, 2014 - 2015, 2016 - 2017.
- Mezi lety 2006 - 2007 rostly mzdy o 7,23% a cena potravin rostla o 8,26%
- Mezi lety 2007 - 2008 rostly mzdy o 7,91% a cena potravin rostla o 8,91%
- Mezi lety 2014 - 2015 rostly mzdy o 3,18% a cena potravin klesala o 0,7%
- Mezi lety 2015 - 2016 rostly mzdy o 4,43% a cena potravin klesala o 1,4%
- Mezi lety 2016 - 2017 rostly mzdy o 6,84% a cena potravin rostla o 7,06%
- Mezi lety 2017 - 2018 rostly mzdy o 8,1% a cena potravin rostla o 2,41%
- Na základě poskytnutých informací můžeme vidět, že existuje určitá korelace mezi vývojem HDP, mzdy a cenami potravin, ale

není to přímočarý vztah. Změny ve výšce HDP mohou ovlivnit změny ve mzdách a cenách potravin, ale nemusí nutně vést k okamžité reakci ve stejném nebo následujícím roce.

Závěr:

Na základě analýzy dat jsme získali odpovědi na výzkumné otázky týkající se mezd, cen potravin a jejich vztahu k výšce HDP.

Výstupy z úkolů lze získat z tabulek:

t_jan_kaminsky_project_SQL_primary_final

t_jan_kaminsky_project_SQL_secondary_final

Vypracoval: Jan Kaminský

Dne 6.7.2023