**PROJEKT SQL\_FARADAY MARTIN**

**Úvod**

V rámci zadání projektu mám za úkol zpracovat / odpovědět na pár definovaných výzkumných otázek, které adresují dostupnost základních potravin široké veřejnosti. V zadání mám již definované základní otázky, na které se pokusím odpovědět. Výsledky prezentuji v následujících přílohách.

1. Připravil jsem robustní datové podklady, se kterých je možné vidět porovnání dostupnosti potravin na základě průměrných příjmů za určité časové období.
2. Jako dodatečný materiál jsem připravil i tabulku s HDP, GINI koeficientem a populací dalších evropských států ve stejném období, jako primární přehled pro ČR.

Otázky pro získání informací vyplývající ze zadání projektu:

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 percentuálně 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?

**POSTUP**

1. Vytvoření 1 primární tabulky > t\_Martin\_Faraday\_project\_SQL\_primary\_final2

Nejprve jsem si otevřel tabulky ke mzdám a potravinam. Následně předběžně si vytvořil dva pomocné selecty.

V případě průměrné mzdy, jsem měl problém vůbec se ujistit, zda-li se jedná o průměrnou mzdu. Vkládal jsem AVG a snažil jsem se ke každé pozici zprůměrovat mzdu v daném roce.

Procházel jsem také původní data, v jakém stavu jsou, nemám-li chybu v dotazu.

Následně, vytvořit VIEW: v\_kategorie\_potravin4, v\_pracovni\_pozice4

A v konečné fázi obě VIEW spojil do jedné nové tabulky > t\_Martin\_Faraday\_project\_SQL\_primary\_final2

1. Vytvoření 2 sekundární tabulky > t\_Martin\_Faraday\_ project\_SQL\_secondary\_final

Spojení obou tabulek economies a countries,.

**CREATE** **TABLE** t\_Martin\_Faraday\_project\_SQL\_secondary\_final **AS**

**SELECT**

c.country, c.abbreviation, c.avg\_height, c.calling\_code, c.capital\_city,

c.continent, c.currency\_name, c.religion, c.currency\_code, c.domain\_tld,

c.elevation, c.north, c.south, c.west, c.east, c.government\_type, c.independence\_date,

c.iso\_numeric, c.landlocked, c.life\_expectancy, c.national\_symbol, c.national\_dish,

c.population\_density, c.population, c.region\_in\_world, c.surface\_area,

c.yearly\_average\_temperature, c.median\_age\_2018, c.iso2, c.iso3, e.**year**,

e.GDP, e.population **AS** population\_economist, e.gini, e.taxes, e.fertility, e.mortaliy\_under5

**FROM** countries c

**JOIN** economies e

**ON** c.country = e.country

;

Viz. SQL t\_Martin\_Faraday\_ project\_SQL\_secondary\_final

V jedninou věc, kterou jsem ošetřil byl duplikovaný sloupec populace.

**ZPRACOVÁNÍ ÚKOLŮ**

task\_1 (2)

Otázka *„Kolik je možné si koupit litrů mléka a kilogramů chleba za první a poslední srovnatelné období v dostupných datech cen a mezd?“* Vycházel jsem s dat, že zjistím průměrnou mzdu v roce 2006 a v posledním sledovaném roce, a stejně tak průměrnou cenu Mléka a chleba, která ve sledovaném období byla. Následně jsem upravil dotaz tak, že ve finále zjistím kolik za průměrnou mzdu si člověk mohl koupit mléka a kolik chlebů. Zajímavosti, je že u chlebů, nebyl moc velký nárust. Jelikož mléko v celkovém období skoro vůbec nezdražovalo.

Více v SQL task 2a.

Task\_2 (3)

Otázka „Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší percentuální meziroční nárůst)?“ Nejprve jsem si vyselektoval rozdíly v nárustech procent u jednotlivých potravin v letech.

Následně jsem se zaměřil na to, abych našel potravinu, která za celé sledované období měla nejmenší nárust v procentech. Ve výsledku s toho vyšlo mléko, které mělo cca 17% nárust za sledované období.

Více v SQL task\_3a.

Task\_3 (1)

Otázka „Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají?“ . Po vyselektování průměrných mezd za jednotlivé odvětví v návaznosti na roky jsem následně zkusil doupřesnit dotaz tak abych viděl meziroční nárust, nebo zápor v procentech.

Z tohoto pohledu ne všechna odvětví v průběhu let mají nárust mezd, ba naopak.

Více v SQL task\_1a.

Task\_4

Otázka „Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (větší než 10 %)?“ . Ve výsledku, neexistuje rok, ve ktérém by byl výraznější nárust cen potravin než 10%

Více v SQL task\_4.

Task\_5

Otázka „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 násdujícím roce výraznějším růstem?“ .

Nejprve jsem si připravil dvě VIEW a to získání procentíh rozdílů nárustu mezd, cen potravin a HDP. Následně obě tabulku spojil do jedné abych viděl se zaměřením na roky, výše uvedené tři sloupce pro zhodnocení, jeli v případě zvýšení HDP se zvýší ceny a mzdy.

Kdybych měl zhodnotit tuto otázku, pak ve většině případu má vliv HDP vliv na růst cen a potravin, ALE, ne vždy toto pravidlo platí. Jelikož například v roce 2018, HDP rostlo, ale ceny potravin a mezd šli do záporu. Nejen HDP má vliv zvýšení průměrných mezd a cen, jsou i jiné položky, které toto ovlivnují.

Více v SQL task\_5.