**Engeto DataAcademy 2025\_01\_23**

**projekt DA SQL**

**Věra Vavrincová**

Odpovědi na výzkumné otázky:

**Primární tabulka** t\_vera\_vavrincova\_project\_sql\_primary\_final

Zdroj: chzechia\_price , czechia\_payroll, czechia\_price\_category,

Postup:pomocí funkce JOIN byly spojeny tabulky chzechia\_price a tabulky czechia\_payroll a jejich vedlejších tabulek. Byly určeny průměrné mzdy za odvětví a rok, a průměrné ceny jednotlivých potravin podle kategorie, období (kvartál a rok)

V jednotlivých úkolech je vždy uvedeno několik skriptů, pro zdroj z tabulek Engeta a poté pro zdroj s primární, resp. Sekundární, tabulky. U skriptů je popis (v poznámce).

**Úkol č.1** – Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají?

Podle analýzy dat je vidět, že s každým dalším rokem je ve většině odvětví vždy vzestup průměrné mzdy. Největší nárůst mezd byl ve Zdravotnictví v r.2021 oproti předchozímu roku.

Pokles mezd je minoritní, je vidět asi u 20-ti analyzovaných položek (rok/odvětví).

Nejvíc u odvětví Těžba a dobývání, kdy pokles mezd byl po 4 roky z analyzovaného období.

Postup:

Funkcí LAG je zjištěna předchozí mzda, tj. ve dvou sloupcích je pak uvedena mzda za rok, a předchozí mzda za minulý rok, a ty jsou pak vzájemně porovnávány. Každé odvětví je porovnáváno zvlášť.

Poté je zjišťován trend růstu nebo poklesu mezd dle odvětví a roků. Provedena kontrola na NULL.

**Úkol č.2** - Kolik je možné si koupit litrů mléka a kilogramů chleba za první a poslední srovnatelné období v dostupných datech cen a mezd?

První srovnatelné období je od 2.1.2006 do 8.1.2006 - tj.7dní -tj.1.Q roku 2006 – v tomto období bylo možné koupit za průměrnou mzdu 1303 kg chleba a 1396 l mléka.

Poslední srovnatelné období je od 10.12.2018 do 16.12.2018 - tj.4.Q r.2018 -

v tomto období bylo možné koupit za průměrnou mzdu 1367 kg chleba a 1727 l mléka.

Postup:

nejprve je určeno první a poslední srovnatelné období. Ceny potravin jsou ve zdrojových datech měněny týdně. Srovnávané období je týden.

Mzdy se ale mění kvartálně, takže je možno uvažovat pro porovnání cen potravin a mezd kvartál v roce. Datum z tabulky czechia\_price je přeformátován na DATE a je extrahován na kvartál a rok.

Poté byla zjištěna průměrná mzda, cena za kvartál pro požadované potraviny mléko a chleba

a poté bylo určena, kolik mléka a chleba bylo možno koupit za mzdu.

**Úkol č.3** - Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší percentuální meziroční nárůst)?

Bylo sledováno 26 potravin (+1 od r.2015) a sledování probíhalo po dobu 12ti let (2006-2018).

Meziroční růst cen potravin je větší než 10% u 68 položek (potravina/rok),

u 95 položek (potravina/rok) došlo ke snížení ceny (1-31% poklesu oproti předchozímu roku)

a u 60 položek (potravina/rok) se cena nezměnila oproti předchozímu roku.

Nejpomalejší zdražení potravin, tedy nejnižší percentuální meziroční nárůst cen potravin, byl zjištěn v letech 2007,2008, 2012, 2015, 2016, 2018 a to u těchto potravin: Rostlinný roztíratelný tuk, Kapr živý, Hovězí maso zadní bez kosti, Šunkový salám, Kuřata kuchaná celá – (šlo o 1% nárůstu ceny).

Ale některé tyto potraviny v jiném roce zdražily mnohem více:

např. Rostlinný roztíratelný tuk roce 2008 až o 20% oproti předchozímu roku

nebo např. Kapr živý v roce 2007 až o 14%, Hovězí maso zadní bez kosti v roce 2012 až o 11%

např. Kuřata kuchaná celá v roce 2007 až o 15%.

Postup:

Byla zjišťována předchozí cena za minulý rok u každé potraviny, a ta byla porovnána s cenou běžného roku. Byl zjištěn průměr růstu ceny, a poté procentuální nárůst.

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

Meziroční nárůst cen potravin byl výrazně vyšší než růst mezd (více než 10%) v roce 2007 u potraviny Papriky až 97% oproti zvýšení mezd o 7% - tj. rozdíl 90%.

Pak v roce 2013 u konzumních brambor byl nárůst ceny oproti mzdám o 61% a

v roce 2012 u vajec byl nárůst ceny oproti mzdám o 53%.

Z analýzy je vidět, že víc než 60 položek (potravina/rok) bylo zvýšeno o víc než 10% oproti mzdám

Postup:

Jako první část pro meziroční nárůst průměrných cen potravin je použit skript z úlohy č.3 a je dopracován. Obdobně je řešena část pro meziroční nárůst průměrné mzdy. Tyto části jsou poté porovnány.

**Úkol č.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?

**Sekundární tabulka**: t\_vera\_vavrincova\_project\_sql\_secondary\_final

– spojení a propracování zdroj. tabulky economies a vytvořené primární tabulky

Výška HDP má jistý vliv na ceny potravin a mzdy.

Při zvyšování HDP se většinou v dalším roce zvýšila průměrná mzda (např.r.2015, 2017).

Ceny potravin se zvedly v závislosti na zvýšení HDP již v tom stejném roce (např.2017).

(Při poklesu HDP se např. v roce 2009 pokles mezd projevil až v dalším roce, ale pokles cen potravin se projevil již v roce s poklesem HDP).

Postup:

Z tabulky economies jsou přebrána data HDP. Jsou použity analýzy předchozích úkolů.

K porovnání jsou použita procentuální data nárůstu mezd, cen a HDP.

Ve druhém skriptu jsou uvažovány všechny zkoumané potraviny.

Ve třetím skriptu je uvažována zprůměrována cena všech zkoumaných potravin za rok, a tato je potom srovnána s HDP a mzdou v daném roce.

**Dodatečná tabulka pro ostatní evropské státy**

Jako dodatečný materiál připravte i tabulku s HDP, GINI koeficientem

a populací dalších evropských států ve stejném období, jako primární přehled pro ČR.

Postup:

Zdrojem je tabulka economies. Jsou vyčíslena data za předchozí rok pro populaci a pro HDP v některých státech EU, a jsou porovnány předchozí roky.

GINI - Giniho koeficient (index) je číslo, které ukazuje nerovnost v příjmech nebo majetku v určité společnosti (pohybuje se mezi 0 a 1 : 0-všichni mají stejně, 1-jeden má všechno, ostatní nic (většinou se udává v %))

Tento projekt mě mnohému naučil. Ještě mám co zlepšovat, ale základy SQL už mám.

**Děkuji firmě Engeto a všem lektorům**.

Věra Vavrincová

tel.737189992