

PROJEKT Z SQL

ENGETO AKADEMIE

Vzdělávací společnost, která pomáhá zájemcům o IT začít programovat již od roku 2016. Akademie v současné chvíli nabízí kurzy Python Akademie, Datová Akademie, Front-end Developer Akademie, Java Akademie, Testing Akademie a Linux Akademie.

Cílem projektu

Naším úkolem je připravit robustní datové podklady pro analytické oddělení na zodpovězení definovaných výzkumných otázek, které adresují dostupnost základních potravin široké veřejnosti. V podkladech bude možné vidět porovnání dostupnosti potravin na základě průměrných příjmů za určité časové období.

Projekt SQL od akademie ENGETO je zaměřená na prověření znalostí datové analýzy pomocí relační databáze MariaDB. Jako dotazovací jazyk se zde bude používat SQL (Structured Query Language = Standardizovaný strukturovaný dotazovací jazyk).

Výstupy projektu

Jako datový výstup bude tabulka mezd a cen potravin v ČR, a dodatečná tabulka o dalších evropských státech, kde je znázorněno HDP, GINI koeficient a populace států. Z těchto dat je následně možné odpovědět na pět výzkumných otázek.

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?

Vývoj mezd podle odvětví - (otázka č. 1)

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

Z našeho průzkumu lze pozorovat vývoj průměrné mzdy za období od roku 2000 až po rok 2021, a to v 19 pracovních odvětvích. Podle dostupných dat lze vyhodnotit, že s výjimkou některých let byl rostoucí trend napříč všemi odvětvími. Meziroční procentuální růst mezd se pohyboval v intervalu $<4.11 ; 6.77>$.

Množství mléka a chleba – (otázka č. 2)

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

Jako první a poslední srovnatelné období mezi daty a cen potravin je rok 2006 a 2018. V prvním roce byla průměrná cena mléka (polotučné pasterované – kód produktu '114201') 14,44,- Kč a chleba (konzumní kmínový – kód produktu '111301') 16,12,- Kč při průměrné mzdě 20.677,- Kč, která byla měřena ze všech dostupných mezd nezávisle na druhu odvětví. V tomto roce bylo možné nakoupit mléka celkem 893 litrů a chleba 740 kg.

Posledním srovnatelným obdobím je rok 2018, kdy cena mléka stoupla na 19,82,- Kč (dražší o 5,38,- Kč, nárůst o 37,26 %), cena chleba vychází na 24,24,- Kč (dražší o 8,12,- Kč, nárůst o 50,33 %). S novou průměrnou mzdou 32.485,- Kč (vyšší o 11.808,- Kč, nárůst o 57,11 %) lze zakoupit 1.090 litrů mléka a 859 kg chleba.

Nejpomalejší meziroční růst ceny – (otázka č. 3)

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

Zde se z mého pohledu nabízí 2 odpovědi, a to nejnižší procentuální meziroční růst, tedy rok, ve kterém cena potraviny rostla nejméně, a nejnižší průměrný meziroční nárůst, který by mě jako analytik nejspíše zajímal více. Vzhledem k nejistotě v otázce níže popisují oba výsledky.

FINALE 1

V prve řadě bylo zapotřebí zjistit průměrnou cenu jednotlivých produktů za jednotlivé roky, a zároveň zobrazení jejich meziročních procentuálních změn. Některé potraviny však naopak zlevňovaly v některých letech, a proto byly vyfiltrovány jenom pozitivní procentuální změny. Po úpravách se vybraly produkty s nejnižší procentuální hodnotou, která byla 0,02 % v roce 2009. V tomto roce to byly produkty rostlinný roztíratelný tuk (kód produktu - '115201') a vepřová pečeně s kostí (kód produktu - '112201').

FINALE 2

Zde se postupovalo totožným způsobem, jako v první variantě s tím rozdílem, že se vytvářel výpočet průměrného meziročního percentuálního růstu za celé měřené období od roku 2006 až 2018. Z výsledku jsme se dozvěděli, že nejnižší průměrná změna byla u banánů s hodnotou 0,81 % (kód produktu – '116103'). Když bychom se podívali blíže na změny za každý rok u tohoto produktu, docházelo ve většině případů spíše ke snížení ceny než jeho zvýšení.

Porovnání meziročního růstu cen potravin a mzdy (otázka č. 4)

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

Mezi rokem 2007 až 2016 zde nezaznamenáváme rozdíl meziročního nárůstu cen od mezd, který by byl vyšší než 10%. Nejvyšším rozdílem bylo v roce 2013, kde rozdíl činil 7,04 %. Je třeba také přihlížet na to, že v tomto roce průměrná mzda poklesla o -1,49 % oproti cenám potravin, kde hodnota naopak vzrostla o 5,55 %. Ve většině ostatních případech rychleji spíše rostly mzdy. Například v roce 2009 byl zde rozdíl dokonce téměř 10 % (9,97 %).

Vliv změny HDP na mzdy a ceny potravin (otázka 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?

Vzhledem k tomu, že nevíme definici "výraznějšího" růstu, bylo zapotřebí tento údaj zjistit. Rozhodl jsem se vypočítat průměr HDP za celé období od roku 1990 až 2020, které činilo 1,82 %.

Výpočtem meziročních percentuálních změn u HDP, mezd a cen potravin bylo možné začít s porovnáním mezi těmito údaji. Z výzkumné otázky lze vyčíst, že chtějí porovnat růst HDP a jeho vliv ve stejném roce, nebo následující rok. Najdou se období, kdy vliv HDP potvrzuje podmínky, nicméně se nedá říct, že by to tak platilo ve všech případech, protože i když HDP klesne, ceny nebo mzdy vzrostou nebo naopak, při vyšším HDP klesnou ceny, či mzdy. Hrubý domácí produkt tedy nemá výrazný vliv na změnu cen potravin a mezd.