

Průvodní listina k SQL projektu

Tento projekt je zaměřen na zpracování dat souvisejících s analýzou vývoje mezd a cen základních potravin v České republice v průběhu několika let, přičemž zahrnuje také porovnání s makroekonomickým ukazatelem, kterým je HDP, pro kontrolu jeho vlivu na vývoj mezd a cen. **Hlavním cílem projektu je připravit robustní datové podklady, které umožní analyzovat trendy ve vývoji mezd, cen potravin a jejich vztah k makroekonomickým ukazatelům, jako je HDP.** To vše je potřeba pro zodpovězení několika klíčových **výzkumných otázek, které se týkají vývoje životní úrovně a dostupnosti potravin.**

Struktura a organizace dat - pro účely projektu byla vytvořena dvě primární datová úložiště, jejichž skripta pro vytvoření lze nalézt v SQL_tab_script:

1. t_michaela_terelya_project_SQL_primary_final

- a. Tabulka obsahuje sjednocená data o mzdách a cenách potravin za Českou republiku v letech 2006–2018. Data byla seskupena podle odvětví a kategorií potravin.

2. t_michaela_terelya_project_SQL_secondary_final

- a. Tabulka obsahuje dodatečná data o makroekonomických ukazatelích, včetně HDP, pro Českou republiku a další evropské země za stejná období.

Pro vytvoření finálních tabulek byly připraveny dvě podpůrné tabulky, které zajišťují efektivní zpracování velkých datových sad a zkrácení doby zpracování dotazů. Tyto tabulky musíme vytvořit jako první. Tímto krokem se zvyšuje přehlednost celého procesu.

1. t_michaela_project_food_prices

- a. Tato tabulka obsahuje informace o cenách potravin, jejich kategoriích a jednotkách. Data jsou získána z tabulek:
 - i. **czechia_price** (obsahující jednotlivé ceny)
 - ii. **czechia_price_category** (obsahující kategorie potravin)
- b. Popis tabulky:
 - i. *price_date*: datum, od kdy je daná cena platná
 - ii. *food_price*: cena potraviny
 - iii. *food_category_name*: název kategorie potravin
 - iv. *price_unit*: jednotka ceny

2. **t_michaela_project_salary**

- a. Tato tabulka obsahuje průměrné mzdy pro jednotlivá odvětví v České republice za roky 2006 až 2018. Tato tabulka pomáhá analyzovat vývoj průměrných mezd v různých odvětvích a porovnat je s vývojem cen potravin. Data jsou získána z tabulek:
- i. **czechia_payroll** (obsahující mzdové údaje)
 - ii. **czechia_payroll_industry_branch** (obsahující informace o odvětvích)
 - iii. **czechia_payroll_unit** (obsahující jednotky mzdy, např. Kč),
 - iv. **czechia_payroll_value_type** (obsahující typy mzdových údajů, např. průměrná mzda)
- b. Popis tabulky:
- i. *payroll_year: rok, ve kterém byla mzda vyplacena*
 - ii. *avg_salary_value: průměrná mzda pro dané odvětví za rok*
 - iii. *salary_unit_name: jednotka mzdy (Kč)*
 - iv. *industry_branch: odvětví, ve kterém byla mzda vyplacena*

Dílič tabulky jsou ve výsledku sloučeny do finálních dvou tabulek:

1. **t_michaela_terelya_project_SQL_primary_final**

- a. Tato tabulka propojuje data o mzdách a cenách potravin na základě společného roku a umožňuje analýzu vztahů mezi nimi. Tabulka kombinuje data o mzdách a cenách potravin na základě společného roku, ke kterému jsou data platná. Zatímco mzdy jsou v poskytnutých datových souborech zaznamenávány ročně, ceny potravin jsou dostupné s přesnějším datem (rok-měsíc-den). Pro účely této analýzy byla kategoriím potravin přiřazena průměrná cena za celý rok. Tabulky obsahuje následující sloupce:
- v. *payroll_year: rok, ve kterém byla vyplacena mzda*
 - vi. *avg_salary_value: průměrná mzda (daný rok a odvětví)*
 - vii. *salary_unit_name: jednotka mzdy*
 - viii. *industry_branch: odvětví práce*
 - ix. *food_category_name: kategorie potravin*
 - x. *avg_food_price: průměrná dané kategorie potravin*
 - xi. *price_unit: jednotka ceny potravin*

2. t_michaela_terelya_project_SQL_secondary_final

- a. Tato tabulka kombinuje údaje o hrubém domácím produktu (HDP), mzdách a cenách potravin v České republice. Cílem je analyzovat ekonomický kontext (HDP) ve vztahu k průměrným mzdám a cenám potravin. Tato tabulka umožňuje provádět hlubší analýzu vztahu mezi ekonomickým růstem (HDP), mzdami a cenami potravin v České republice v letech 2005 až 2018.

Po úspěšném vytvoření hlavních tabulek, již nejsou dílčí tabulky potřeba, a proto mohou být odstraněny pomocí funkce DROP.

1. DROP TABLE t_michaela_project_salary;

2. DROP TABLE t_michaela_project_food_prices;

VÝZKUMNÁ OTÁZKA Č.1 : Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají?

°Pro získání detailnější odpovědi jsou k dispozici dva skripta (pro přehlednost)

Ze získaných dat **v první tabulce** je patrné, že **u některých odvětví v některých letech došlo k poklesu mzdy**. Nejvíce poklesů průměrných mezd bylo zaznamenáno v letech 2009, 2010, 2011 a **hlavně v roce 2013**, kde mzda klesla u většího množství odvětví naráz. **Nejvýraznější pokles** byl zaznamenán v roce 2013 v odvětví **Peněžnictví a pojišťovnictví**, kdy průměrná mzda klesla až o **8.91 %**.

Druhá tabulka zobrazující celkový trend průměrných mezd ve všech odvětvích ukazuje, že **většina odvětví zaznamenala růst nebo stagnaci**. Klesající mzdy se objevily sporadicky a většinou v letech 2009, 2010, 2011 a 2013. To naznačuje, že **v průběhu let se mzdy většinou zvyšovaly**.

Důležité poznatky k doplnění:

Pokles mezd může být ovlivněn cyklickými změnami v ekonomice, jako jsou hospodářské recese, které ovlivnily celou řadu odvětví. Například celosvětová finanční krize v letech 2007-2008 mohla mít vliv na mzdy v letech 2009 a 2010.

Odvětví jako Těžba a dobývání a Ubytování, stravování a pohostinství se ukázala jako více citlivá na ekonomické poklesy, zatímco některá jiná odvětví jako Informační a komunikační činnosti vykazovala stabilní růst.

VÝZKUMNÁ 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?

°Pro získání detailnější odpovědi jsou k dispozici dva skripta - zvlášť pro chléb a mléko

Podle získaných dat vyplývá, že průměrná mzda vzrostla mezi roky 2006 a 2018, což vedlo k nárůstu kupní síly pro oba potravinové produkty.

- **V roce 2006 si lidé mohli koupit 1,332 kg chleba a 1,475 l mléka, zatímco v roce 2018 si mohli zakoupit 1,355 kg chleba a 1,688 l mléka.**

I přes nárůst cen potravin **zůstala kupní síla relativně stabilní**. V roce 2018 si lidé mohli koupit téměř stejné množství chleba jako v roce 2006, což naznačuje, že **růst mezd pokryl zvýšení cen**. Zajímavé je také to, že průměrná cena chleba vzrostla o přibližně **54 % (z 15.58 Kč/kg na 24.01 Kč/kg)**, zatímco cena mléka o přibližně **37 % (z 14.07 Kč/l na 19.28 Kč/l)**. S velkou pravděpodobností, větší nárůst u ceny chleba je důsledkem **zvýšení cen u surovin na jeho výrobu**, jako jsou máslo, mouka (pšenice) a například také vejce.

Důležité poznatky k doplnění:

Je důležité zdůraznit, že tato analýza se opírá o **průměrné hodnoty za celý rok pro jednotlivé kategorie potravin a průměrné mzdy ve všech odvětvích**, což může skrývat rozdíly mezi jednotlivými sektory. Proto je třeba výsledky interpretovat s jistou opatrností a nadhledem.

VÝZKUMNÁ OTÁZKA Č.3 : Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší procentuální meziroční nárůst)?

°Pro získání detailnější odpovědi jsou k dispozici dva skripta (pro přehlednost)

Na základě získaných dat se zjistilo, že **cukr krystalový** je ta kategorie potravin, která zaznamenala **nejpomalejší zdražování, přičemž průměrný meziroční nárůst cen činil -2,03 %**. Tímto způsobem je možné pozorovat, že v průměru ani moc zdražoval, ale spíše naopak zlevňoval. To lze jednoduše vypočítat také z dat, které získáme doplňujícího skriptu, který ukazuje hodnoty vybrané kategorie (v našem případě cukr) za všechny pozorovatelné roky, kde vidíme ve výsledcích jasnou **převahu záporných hodnot**.

Důležité poznatky k doplnění:

Je zajímavé **porovnat cukr s jinými “běžnými” potravinami**, které narozdíl od něj vykazovaly stabilní nebo rostoucí ceny. Tak například dříve zkoumané mléko nebo chléb prokázaly **výrazné zvýšení cen**. Tady je důležité mít na paměti, že **dostupnost dat není ve všech letech rovnoměrná**. Například **u vína** jsou ceny zaznamenány **pouze od roku 2015**, což může ovlivnit srovnání a interpretaci trendů v cenách.

VÝZKUMNÁ OTÁZKA Č.4 : Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (vyšší než 10 %)?

°Pro získání detailnější odpovědi jsou k dispozici dva skripta (pro přehlednost)

Ne, v žádném roce tak velký rozdíl v meziročním nárůstu zaznamenán nebyl. O tom nám svědčí výstup z prvního skripta, který ve výsledku po spuštění vydá **prázdnou tabulku.**

Z druhého doplňujícího skripta lze vidět, **že nejvyšší rozdíl mezi růstem cen potravin a růstem mezd byl zaznamenán v roce 2013**, kdy ceny potravin rostly o 5,74 % a mzdy naopak poklesly o 1,56 %, **což vytvořilo rozdíl 7,3 %**. V žádném jiném roce **nebyl** rozdíl mezi růstem cen potravin a mezd **větší než 5 %**. To znamená, že i v letech s vyšším růstem cen potravin tento nárůst nikdy dramaticky nepřekonal růst mezd.

Důležité poznatky k doplnění:

Z dat je patrné, že **ceny potravin jsou mnohem citlivější na ekonomické podmínky než mzdy**. Například v roce **2009**, během ekonomické krize, ceny potravin poklesly o **-5,8 %**, zatímco mzdy stále rostly o **3,16 %**. Naopak v roce **2017**, kdy ekonomika rostla, ceny potravin prudce vzrostly o **10,45 %**. **Mzdy mají tendenci být ale méně pružné** a reagují na tyto ekonomické změny s menším zpožděním oproti potravinám. To znamená, že mzdy často rostou i během krize, zatímco ceny potravin mohou kolísat výrazněji v závislosti na ekonomickém cyklu.

VÝZKUMNÁ 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?

Z dat lze vyvodit, že **výraznější růst HDP často vede k růstu mezd i cen potravin, ale tento vztah není vždy přímočarý ani pravidelný.** Na růst mezd a cen potravin mohou mít značný vliv i jiné faktory, jako jsou **náklady na suroviny, situace na trhu práce a globální ekonomické podmínky.** V některých případech se také zdá, že růst HDP má vliv až s určitým zpožděním, kdy se projeví až v následujících letech.

Ze získaných dat lze vypožorovat několik skutečností:

- Obecně platí, že **roky s výraznějším růstem HDP jsou spojeny i s růstem mezd**, například v letech 2007 a 2017, kdy HDP vzrostlo o 5,57 % a 5,17 %, mzdy se zvýšily o 6,84 % a 6,28 %. **Na druhou stranu, ceny potravin nejsou s růstem HDP tak úzce spjaty.** V roce 2017, kdy HDP rostlo jen o necelá 3 %, se ceny potravin výrazně zvýšily o 10,45 %, což může být způsobeno **dalšími faktory, jako je růst cen surovin.** Naopak v roce 2008, kdy HDP rostlo pomaleji (o 2,69 %), ceny potravin vzrostly také o pouhé 5,3 %, což ukazuje, že vztah mezi HDP a cenami potravin je složitější a méně předvídatelný než u mezd.
- V některých letech lze také pozorovat **zpožděný vliv HDP** na růst mezd a cen potravin. Například v roce 2016, kdy HDP rostlo jen mírně o 2,54 %, byl růst mezd i cen potravin pomalejší, zatímco **v roce 2017 došlo k jejich výraznému nárůstu**, což naznačuje, že dopady růstu HDP se mohou projevit až s časovým odstupem.
- Zatímco **ceny potravin vykazují větší volatilitu** (rychlé změny směrem nahoru i dolů), mzdy jsou obecně **méně pružné.** Například i během ekonomické krize v roce **2009, kdy HDP pokleslo o -4,66 %**, mzdy stále rostly, i když mírně. Ceny potravin naopak rychleji reagují na ekonomické změny, jak ukazuje pokles v letech s negativním HDP.