

Projekt 1.

autor Vladimír Kost'ukovič, email: kostyukovych@gmail.com

databaze: data_academy_content.

Zadání: analytika souvislosti mezd, cen a HDP.

Výzkumná otázka č. 1

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

Postup analýzy

1. Zobrazena tabulka czechia_payroll.
2. Pomocí JOIN byly spojeny tabulky:
 - czechia_payroll_calculation, filtr pouze pro calculation_code = 100 (fyzické přepočtený),
 - czechia_payroll_industry_branch pro získání názvů odvětví.
3. Provedeno čištění dat:
 - odstranění nepotřebných sloupců,
 - odstranění hodnot NULL ve sloupci value,
 - ponechány pouze reálné mzdy větší než 999 Kč.
4. Agregace dat za jednotlivé roky pomocí AVG(value) (údaje jsou čtvrtletní).
5. Výpočet meziročního rozdílu pomocí LAG() – absolutního i procentuálního.
6. Odstraněny první roky v časových řadách (salary_diff IS NOT NULL).
7. Vybrána pouze ta odvětví, kde:
 - MIN(salary_diff) >= 0
 - MIN(percent_diff) >= 0

Výsledek

V těchto odvětvích nedošlo v letech 2001–2021 k meziročnímu poklesu mezd. Mzdy zde rostly každým rokem.

- Administrativní a podpůrné činnosti
- Doprava a skladování
- Ostatní činnosti
- Zdravotní a sociální péče

Ve zbývajících odvětvích byl v některých letech zaznamenán meziroční pokles. Nelze tedy říci, že mzdy rostou v průběhu let ve všech odvětvích.

Použitý SQL dotaz: script otazka 1

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?

Postup analýzy

1. Spojujeme tabulky `czechia_price`, `czechia_price_category` a `czechia_region`.
2. Jelikož **chléb** a **pečivo** lze považovat za jednu skupinu základních potravin, budeme pracovat se třemi kategoriemi:
 - Chléb (Chléb konzumní kmínový),
 - Pečivo (Pečivo pšeničné bílé),
 - Mléko (Mléko polotučné pasterované).
3. Nejstarší dostupný rok v tabulce cen je **2006**, poslední je **2018**.
4. V tabulce mezd (`czechia_payroll`) je dat více, ale pro porovnání použijeme stejné roky — **2006 a 2018**.
5. Vypočítáme:
 - průměrnou mzdu v jednotlivých odvětvích v daných letech,
 - průměrnou cenu daných produktů,
 - kolik jednotek daného produktu si lze za tuto mzdu koupit.
6. Porovnáваме změnu kupní síly mezi lety 2006 a 2018 a vyhodnocujeme, zda mzdy rostly rychleji než ceny.
7. Vytváříme výslednou tabulku (pivot table), která slouží jako analytický přehled rozdílů mezi jednotlivými odvětvími a produkty.

Interpretace výsledků – otázka č. 2

V letech **2006 až 2018** došlo ve většině odvětví k **nárůstu mezd**, nicméně **nárůst kupní síly** (tedy kolik kusů chleba, pečiva či litrů mléka je možné si za průměrnou mzdu koupit) byl různý v závislosti na konkrétním odvětví a typu potravin.

Z výsledné tabulky vyplývá:

- V některých odvětvích (např. **administrativní činnosti, zdravotní péče**) se mzdy zvyšovaly výrazněji než ceny potravin, což vedlo k **zlepšení kupní síly**.
- V jiných sektorech však mzdy rostly pomaleji než ceny nebo stagnovaly, což způsobilo **reálný pokles možnosti nákupu** základních potravin (např. **ve veřejné správě či ubytování a stravování**).
- U produktu **mléko** byl růst ceny mírnější než u pečiva, což se pozitivně projevilo ve větší kupní síle téměř ve všech odvětvích.
- **Pečivo** (zejména bílé pšeničné) zdražovalo rychleji, a v některých oborech si zaměstnanci mohli v roce 2018 za mzdu koupit téměř stejně nebo méně kusů než v roce 2006.

Celkově lze říci, že **zvýšení mzdy samo o sobě neznamená automaticky vyšší životní úroveň**. Důležitá je také dynamika cen základních produktů. Analýza ukazuje, že zatímco nominální mzdy rostly, **reálná kupní síla kolísala podle sektoru** a druhu potravin.

Použitý SQL dotaz: script otázka 2

Možné rozšíření

V budoucnu bychom mohli tuto analýzu dále rozvinout například o:

- výpočet **korelace mezi vývojem mezd a cen**,
- identifikaci **nejúspěšnějších odvětví**, kde nikdy nenastala stagnace,
- propojení s daty z kapitálových trhů — jak se firmám z těchto sektorů daří z pohledu účetních výkazů, hospodářských výsledků a dividend,
- zjistit, zda růst mezd **odpovídá růstu příjmů firem** v daném oboru,
- a případně zhodnotit, zda by tyto společnosti mohly být zajímavé pro **investiční rozhodnutí**.
- nebo z kterého do kterého kraje je vyhodně jezdit nakupovat připojením dat podle krajů

Můžeme také doplnit data o **pozici firmy ve vlnové struktuře** Elliota nebo Waikoff, případně rozšířit o zdrojová data z veřejných databází.

Nicméně — to už přesahuje rámec tohoto projektu a nejsme za to placení. Tak to nechme na jindy.

Mezitím vytvoříme kontingenční tabulku (Pivot table) pro Excel a naše virtuální kolegy z noSQL týmu, ve které zobrazíme všechna dostupná data za jednotlivé roky. Na serveru zároveň nainstalujeme Metabase (v Dockeru) a připravíme se na tvorbu GRAFŮ.

Výzkumná otázka č. 3

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

Postup analýzy:

1. Zapojili jsme všechny dostupné tabulky a zobrazili pouze ta data, která jsou pro nás relevantní (např. jsme ignorovali jednotky typu kilogram/litr a regionální kódy, které nejsou důležité pro meziroční porovnání).
 2. Jelikož jsou data k dispozici na měsíční úrovni, bylo potřeba agregovat průměrné ceny za jednotlivé roky v rámci kategorií potravin.
 3. Vytvořili jsme tabulku s průměrnými ročními cenami pro každou kategorii potravin a následně vypočetli meziroční procentuální změnu cen.
-

SQL dotaz: otazka_3_pivot_table

Tento dotaz nám vytvořil **pivotní tabulku**, kterou lze snadno vizualizovat například v Metabase a my to možná udeláme pokud můj soukromý projekt nechá čas I na tohle.

Dalším SQL dotazem (*otazka_cislo_3*) jsme zobrazili pouze **průměrné meziroční nárůsty cen** podle kategorií potravin.

A protože nechceme brát v potaz pouze **jeden produkt s nejnižším růstem** (to by nevypadalo statisticky dobře), vybrali jsme **top 5 nejstabilnějších kategorií** podle průměrného meziročního zdražování.

Výsledek – TOP 5 nejpomaleji zdražujících kategorií (2006–2018):

Kategorie potravin	Průměrný meziroční nárůst (%)
Cukr krystalový	−1,92 %
Rajská jablka červená kulatá	−0,74 %
Banány žluté	0,81 %

Kategorie potravin

Průměrný meziroční nárůst (%)

Vepřová pečeně s kostí

0,99 %

Přírodní minerální voda sycená CO₂

1,03 %

Cukr a jablka to je základ na domácí jablečný džus... nebo rovnou cider.

A pokud jsi to dočetl až sem, tak gratulujeme – máš od nás oficiální analytické povolení otevřít si domácí cider.

Je totiž pátek, 21:00, a tabulky už prostě dneska nemají šanci konkurovat. Na zdraví!

Výzkumná otázka č. 4

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

Postup:

1. Vrátili jsme se k datům o mzdách a spočítali jsme meziroční procentní nárůst mezd.
2. Dále jsme doplnili průměrné ceny potravin z otázky č. 3 a vypočítali meziroční procentní nárůst cen.
3. Pomocí SQL dotazu otazka_4.sql jsme získali tabulku s porovnáním těchto růstů a výpočtem rozdílu mezi nimi.

Dostáváme následující výsledky:

2006 — 4.35 — NE

2007 – 6.34 / 7.86 → -1.52 – NE

2008 – 6.41 / 9.33 → -2.92 – NE

2009 – -6.80 / 1.00 → -7.81 – NE

2010 – 1.77 / 2.14 → -0.37 – NE

2011 – 3.50 / 4.42 → -0.92 – NE

2012 – 6.92 / 2.03 → 4.89 – NE

2013 – 5.55 / -1.63 → 7.18 – NE

2014 – 0.89 / 2.56 → -1.66 – NE

2015 – -0.56 / 2.45 → -3.01 – NE

2016 – $-1.12 / 4.94 \rightarrow -6.06$ – NE

2017 – $9.98 / 6.43 \rightarrow 3.55$ – NE

2018 – $1.94 / 7.57 \rightarrow -5.63$ – NE

Závěr:

Na základě dostupných dat nebyl identifikován žádný rok, ve kterém by meziroční růst cen potravin výrazně (tj. o více než 10 %) převýšil růst mezd.

Největší zaznamenaný rozdíl byl v roce 2013, kde rozdíl činil 7,18 %, ale ani tehdy nedošlo k překročení stanovené hranice v 10%.

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?

Postup:

1. Spojili jsme data o meziročním růstu HDP, HDP na osobu, růstu mezd a růstu cen potravin za období 2006–2018.
2. Vytvořili jsme souhrnnou tabulku, ve které sledujeme, jak se mění všechny tyto hodnoty v čase.
3. Porovnáváme jednotlivé roky a hledáme souvislosti mezi růstem HDP a vývojem ostatních proměnných.
4. Vytvoříme Pivot table pro metabase a csv soubor pro noSQL kolegy, pokud neprohrajeme soubor v čase se soukromým projektem.

Zjištění:

1. **Mezi růstem HDP a růstem mezd existuje zjevná souvislost.** Většina let, kdy HDP rostlo výrazněji, vykazuje i růst mezd ve stejném nebo následujícím roce. Například v roce 2007 rostlo HDP o 5,57 % a mzdy o 7,86 %, v roce 2017 HDP o 5,17 % a mzdy o 6,43 %. Lze tedy říct, že silný hospodářský růst se pozitivně promítá do mezd.
2. **Ceny potravin s HDP výrazněji nesouvisí.** Například v roce 2012, kdy HDP kleslo o –0,79 %, vzrostly ceny potravin o 6,92 %. Naopak v roce 2015, kdy HDP vzrostlo o 5,39 %, ceny potravin mírně poklesly (–0,56 %). Z toho vyplývá, že vývoj cen potravin je ovlivněn více jinými faktory (např. globálními cenami, sezonou, vývojem zemědělství), než samotným výkonem domácí ekonomiky.

Graf metabase <http://176.102.66.84:3000/public/dashboard/b31103ad-0a60-434a-9e15-23e0405a2f72>

Závěr:

Ano, růst HDP má vliv na růst mezd. V letech s vyšším HDP obvykle rostou i mzdy obyvatel. U cen potravin však taková souvislost prokázána nebyla — ty reagují na jiné faktory a mohou růst i při stagnaci nebo poklesu HDP.

A mohli bychom pokračovat dál v hledání souvislostí — například zkoumat, jaký vliv má HDP eurozóny nebo celosvětový hospodářský růst na HDP České republiky. Ale jak už bylo řečeno dříve: úkol je splněn a za hlubší makroekonomickou analýzu nejsem placen :)