

Zadanie projektu

Cieľom projektu je vypracovať odpovede na 5 výzkumných otázok, týkajúcich sa dostupnosti základných potravín a to hlavne na základe vývoja cien a miezd.

Odpovedi na tieto otázky musia byť vypracované z dvoch tabuliek vytvorených za týmto účelom z dostupných dátových zdrojov, pričom z prvej tabulky vzniknú odpovede na otázky 1- 4 a druhá tabuľka slúži na odpoveď č. 5.

Tabuľka č.1

Aby bolo pomocou tabulky č.1 možné odpovedať na prvé 4 otázky, musí obsahovať výšku mzdy pre každý hospodársky sektor a cenu každej sledovanej potraviny. To všetko pre všetky roky, kedy sú dostupné dáta pre mzdy aj ceny potravín, v tomto prípade obdobie 2006-2018.

Ukážka časti tabulky:

	123 year_sum	abc industry_name	123 average_salary	abc food_name	123 price
3	2,006	Administrativní a podpůrné činnosti	14,585	Eidamská cihla	110.95
4	2,006	Administrativní a podpůrné činnosti	14,585	Hovězí maso zadní bez kosti	166.34
5	2,006	Administrativní a podpůrné činnosti	14,585	Chléb konzumní kmínový	16.12
6	2,006	Administrativní a podpůrné činnosti	14,585	Jablka konzumní	30.71
7	2,006	Administrativní a podpůrné činnosti	14,585	Jogurt bílý netučný	5.83
8	2,006	Administrativní a podpůrné činnosti	14,585	Kapr živý	69.35
9	2,006	Administrativní a podpůrné činnosti	14,585	Konzumní brambory	12.07
10	2,006	Administrativní a podpůrné činnosti	14,585	Kuřata kuchaň celá	47.47
11	2,006	Administrativní a podpůrné činnosti	14,585	Máslo	104.39
12	2,006	Administrativní a podpůrné činnosti	14,585	Mléko polotučné pasterované	14.44

Pomocou Common Table Expresion som vytvoril dve medzitable czechia_payroll_by_year a czechia_price_by_year pre informácie o mzdách a cenách potravín pre daný rok. V oboch prípadoch bolo potrebné urobiť join s pomocnými tabuľkami, pretože ani jedna neobsahovala názvy hospodárskych odvetví, resp. potravín, len kódy. Nakoniec som obe CTE spojil pomocou left join na základe stĺpcov s údajmi o rokoch.

Otázka č. 1: Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají?

Odpoveď na túto otázku sa môže líšiť v závislosti na vyložení si slovného spojenia „v průběhu let“. Svoj skript som sa snažil postaviť tak, aby na túto neurčitost čo najlepšie reagoval.

Najprv som urobil self join tak, aby bolo možné porovnať dáta v rôznych rokoch. Tento self join som potom vyseletoval aby ukazoval sledované hospodárske odvetvie, určité obdobie a výsledok, či v tomto období a odvetví mzda rastie alebo nie. Toto som uložil do medzitable ako CTE. Z CTE som potom urobil jednoduchý selekt, ktorý počíta obdobia, kedy nerastú mzdy pre každé hospodárske odvetvie.

Pokiaľ je obdobie nastavené na jeden rok, vyzerá výsledná tabuľka takto:

	ABC name	123 not_growing_years
1	Administratívni a podpúrné činnosti	2
2	Činnosti v oblasti nemovitostí	3
3	Doprava a skladování	1
4	Informační a komunikační činnosti	2
5	Kulturní, zábavní a rekreační činnosti	3
6	Ostatní činnosti	2
7	Peněžnictví a pojišťovnictví	3
8	Profesní, vědecké a technické činnosti	2
9	Stavebnictví	2
10	Těžba a dobývání	4
11	Ubytování, stravování a pohostinství	3
12	Velkoobchod a maloobchod; opravy a údržba motorových vozidel	2
13	Veřejná správa a obrana; povinné sociální zabezpečení	3
14	Výroba a rozvod elektřiny, plynu, tepla a klimatiz. vzduchu	3
15	Vzdělávání	1
16	Zásobování vodou; činnosti související s odpady a sanacemi	2
17	Zdravotní a sociální péče	2
18	Zemědělství, lesnictví, rybářství	2
19	Zpracovatelský průmysl	0

To znamená, že medziročný pokles mzdy zaznamenali v sledovanom období všetky odvetvia okrem zamestnávateľského priemyslu. Mzdy v odvetví ťžba a dobývání, poklesli 4 krát.

Pokiaľ ale obdobie zmeníme na 5 rokov (stačí u self joinu upraviť číslo podmienky), zmení sa aj výsledná tabuľka:

	ABC name	123 not_growing_years
3	Doprava a skladování	0
4	Informační a komunikační činnosti	0
5	Kulturní, zábavní a rekreační činnosti	0
6	Ostatní činnosti	0
7	Peněžnictví a pojišťovnictví	1
8	Profesní, vědecké a technické činnosti	1
9	Stavebnictví	1
10	Těžba a dobývání	2
11	Ubytování, stravování a pohostinství	0
12	Velkoobchod a maloobchod; opravy a údržba motorových vozidel	0
13	Veřejná správa a obrana; povinné sociální zabezpečení	1
14	Výroba a rozvod elektřiny, plynu, tepla a klimatiz. vzduchu	3
15	Vzdělávání	0
16	Zásobování vodou; činnosti související s odpady a sanacemi	0
17	Zdravotní a sociální péče	0
18	Zemědělství, lesnictví, rybářství	0
19	Zpracovatelský průmysl	0

V akomkoľvek 5 ročnom období zaznamenalo pokles miezd 6 odvetví zo 17. Pokiaľ by sme sledované obdobie natiahli na 9 a viac rokov, potom už stúpajú mzdy vo všetkých odvetviach.

Otázka č. 2: Kolik je možné si kúpiť litrov mlieka a kilogram chleba za prvú a poslednú srovnateľnú dobu v dostupných dátach cen a miezd?

V tomto prípade stačí spočítať priemernú mzdu pre celé hospodárstvo a vidieť ju cenou danej potraviny v daný rok. V roku 2006 bolo možné z priemernej mzdy kúpiť 1381 chlebov a 1541 mliek. V roku 2018 potom 1541 chlebov a 1710 mliek.

	ABC food_name	123 year_sum	123 price	123 salary	123 food_per_salary
1	Chléb konzumní kmínový	2,006	16.12	22,259	1,381
2	Chléb konzumní kmínový	2,018	24.24	33,902	1,399
3	Mléko polotučné pasterované	2,006	14.44	22,259	1,541
4	Mléko polotučné pasterované	2,018	19.82	33,902	1,710

Otázka č.3: Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší percentuální meziroční nárůst)?

Najprv som pomocou Case Expresion vyselekoval ceny každej potraviny pre roky 2006 a 2018 a tieto som v ďalšom selekte porovnal pomocou vzorca na výpočet percentuálneho rozdielu.

Medzi rokmi 2006 a 2018 najmenej zdražili banány (0.55 %). Najvýraznejšie zlacnil cukor (-2.45 %).

	ABC food_name	123 price_2006	123 price_2018	123 avg_percentual_growth
1	Cukr krystalový	21.73	15.75	-2.45
2	Rajská jablka červená kulatá	57.83	44.49	-2
3	Banány žluté	27.3	29.32	0.55
4	Vepřová pečeně s kostí	105.18	116.85	0.81
5	Přírodní minerální voda uhličitá	7.69	8.64	0.9
6	Pečivo pšeničné bílé	38.6	43.84	0.98
7	Jablka konzumní	30.71	36.17	1.27
8	Šunkový salám	116.77	144.88	1.67
9	Konzumní brambory	12.07	15.08	1.73

Otázka č. 4: Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst miezd (větší než 10 %)?

Pomocou self join som vyselekoval priemerné mzdy v hospodárstve a ceny potravín tak, aby bol v jednom riadku zaznamenaný medziročný rozdiel v mzdách aj cenách potravín. Následne som urobil ďalší selekt, kde som spočítal medziročnú zmenu pre mzdy a ceny potravín pre každé dvojročné obdobie. Tieto selekty som uložil do CTE.

	ABC period	123 salary_growth	123 food_price_growth	123 current_salary	123 prev_salary	123 current_food_price	123 prev_food_price
1	2007 - 2006	5.57	6.52	23,498	22,259	49	46
2	2008 - 2007	8.26	6.12	25,440	23,498	52	49
3	2009 - 2008	4.42	-7.69	26,565	25,440	48	52
4	2010 - 2009	0.09	2.08	26,588	26,565	49	48
5	2011 - 2010	1.97	4.08	27,111	26,588	51	49
6	2012 - 2011	5.45	5.88	28,589	27,111	54	51
7	2013 - 2012	-4.46	5.56	27,314	28,589	57	54
8	2014 - 2013	1.76	0	27,794	27,314	57	57
9	2015 - 2014	2.93	0	28,608	27,794	57	57
10	2016 - 2015	3.26	-1.75	29,540	28,608	56	57
11	2017 - 2016	7.96	10.71	31,891	29,540	62	56
12	2018 - 2017	6.31	1.61	33,902	31,891	63	62

V záverečnom selekte som v stĺpci real_price_growth ukázal rozdiel medzi nárastom cien potravín nárastom miezd. Výsledkom je tabuľka, ktorá ukazuje, že v období 2012-2013 bol medziročný nárast cien potravín o 10.02 % vyšší ako nárast miezd (ceny potravín narástly o 5.56 % a mzdy poklesly o 4.46 %).

	ABC period	123 real_price_growth	123 salary_growth	123 food_price_growth
1	2013 - 2012	10.02	-4.46	5.56
2	2017 - 2016	2.75	7.96	10.71
3	2011 - 2010	2.11	1.97	4.08
4	2010 - 2009	1.99	0.09	2.08
5	2007 - 2006	0.95	5.57	6.52
6	2012 - 2011	0.43	5.45	5.88
7	2014 - 2013	-1.76	1.76	0
8	2008 - 2007	-2.14	8.26	6.12
9	2015 - 2014	-2.93	2.93	0
10	2018 - 2017	-4.7	6.31	1.61
11	2016 - 2015	-5.01	3.26	-1.75
12	2009 - 2008	-12.11	4.42	-7.69

Tabuľka č. 2

Aby bolo možné odpovedať na otázku č. 5, je potrebné vytvoriť ďalšiu tabuľku s údajmi o raste HDP pre ČR medzi rokmi 2005 a 2017 (o rok pozadu oproti výsledkom pre mzdy a ceny). Najprv som urobil self join tabuľky economies skrz stĺpce year (a year +1) a následne pomocou vzorca spočítal medziročný rast HDP.

	ABC country	123 year	123 gdp_growth
1	Czech Republic	2,017	5.17
2	Czech Republic	2,016	2.54
3	Czech Republic	2,015	5.39
4	Czech Republic	2,014	2.26
5	Czech Republic	2,013	-0.05
6	Czech Republic	2,012	-0.79
7	Czech Republic	2,011	1.76
8	Czech Republic	2,010	2.43
9	Czech Republic	2,009	-4.66
10	Czech Republic	2,008	2.69
11	Czech Republic	2,007	5.57
12	Czech Republic	2,006	6.77
13	Czech Republic	2,005	6.6

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?

Základem je skript z otázky č.4, z ktorého som výsledky pre mzdy a ceny spojil s tabulkou č.2 (pre HDP) a pomocou selektu som dostal do jedného riadku údaje o raste cien a miezd z jedného roku a o raste HDP z predchádzajúceho roku. Nakoniec som použil agregáčnú funkcie AVG a operátor UNION.

Výška HDP má vplyv na výšku miezd v nasledujúcom roku. Pokiaľ bol rast HDP vyšší ako 5%, nasledujúci rok rástli mzdy v priemere o 5.72 %. Pokiaľ úroveň HDP poklesla, nasledujúci rok mzdy poklesli v priemere o 0.87%.

Výška HDP nemá významný vplyv na vývoj cien potravín v nasledujúcom roku. Pokiaľ bol rast HDP vyšší ako 5%, nasledujúci rok rástli ceny v priemere o 3.12 %. Pokiaľ úroveň HDP poklesla, nasledujúci rok ceny rástli v priemere o 2.55%. Rozdiel v raste je cien nie je veľký a úroveň rastu HDP je zrejme len jeden z mnohých faktorov, ktoré majú vplyv na rast cien potravín.

	abc gdp	123 average_salary_growth	123 average_food_price_growth	123 average_gdp_growth
1	gdp growth > 5%	5.85	3.12	5.72
2	gdp growth < 0%	-0.87	2.55	-1.83