

## Zadanie projektu

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

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 miezd?
3. Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší percentuální meziroční nárůst)?
4. Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst miezd (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?

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

### Tabulka č.1

Aby bolo pomocou tabuľky 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 tabuľky:

	123 year_sum	abc industry_name	123 average_salary	abc food_name	123 price
1	2,006	Administrativní a podpůrné činnosti	13,414	Banány žluté	27.3
2	2,006	Administrativní a podpůrné činnosti	13,414	Cukr krystalový	21.73
3	2,006	Administrativní a podpůrné činnosti	13,414	Eidamská cihla	110.95
4	2,006	Administrativní a podpůrné činnosti	13,414	Hovězí maso zadní bez kosti	166.34
5	2,006	Administrativní a podpůrné činnosti	13,414	Chléb konzumní kmínový	16.12
6	2,006	Administrativní a podpůrné činnosti	13,414	Jablka konzumní	30.71
7	2,006	Administrativní a podpůrné činnosti	13,414	Jogurt bílý netučný	5.83
8	2,006	Administrativní a podpůrné činnosti	13,414	Kapř živý	69.35
9	2,006	Administrativní a podpůrné činnosti	13,414	Konzumní brambory	12.07
10	2,006	Administrativní a podpůrné činnosti	13,414	Kuřata kuchaň celá	47.47
11	2,006	Administrativní a podpůrné činnosti	13,414	Máslo	104.39
12	2,006	Administrativní a podpůrné činnosti	13,414	Mléko polotučné pasterované	14.44

Pomocou Common Table Expression 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?

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

To znamená, že medzioroční pokles mzdy zaznamenalo v sledovanom období aspoň raz 14 odvetví z 19. Mzdy v odvetví těžba 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
1	Administratívne a podporné činnosti	0
2	Činnosti v oblasti nemovitostí	0
3	Doprava a skladovanie	0
4	Informačné a komunikačné činnosti	0
5	Kultúrne, zábavné a rekreačné činnosti	0
6	Ostatné činnosti	0
7	Peněžníctví a pojišťovnictví	0
8	Profesní, vědecké a technické činnosti	0
9	Stavebnictví	0
10	Těžba a dobývání	1
11	Ubytování, stravování a pohostinství	0
12	Velkoobchod a maloobchod; opravy a údržba motorových vozidel	0
13	Veřejná správa a obrana; povinné sociální zabezpečení	0
14	Výroba a rozvod elektřiny, plynu, tepla a klimatiz. vzduchu	0
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 len odvetvie ťžba a dobývaní. Pokiaľ by sme sledované obdobie natiahli na 6 a viac rokov, potom už stúpajú mzdy vo všetkých odvetviach.

## 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?

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 1262 chlebov a 1409 mliek. V roku 2018 potom 1319 chlebov a 1614 mliek.

	ABC food_name	123 year_sum	123 price	123 salary	123 food_per_salary
1	Chléb konzumní kmínový	2,006	16.12	20,343	1,262
2	Chléb konzumní kmínový	2,018	24.24	31,981	1,319
3	Mléko polotučné pasterované	2,006	14.44	20,343	1,409
4	Mléko polotučné pasterované	2,018	19.82	31,981	1,614

### Otázka č.3: Která kategorie potravin zdražuje najpomaleji (je u ní najnižší percentuální meziroční nárůst)?

Najprv som pomocou Case Expression vyseletoval 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 došlo k najmenšiemu priemernému nárastu (-2.45 %) u cukru.

	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 žlté	27.3	29.32	0.55
4	Vepřová pečeně s kostí	105.18	116.85	0.81
5	Přírodní minerální voda uhlíčitá	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 mezd (větší než 10 %)?

Pomocou self join som vyseletoval 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	6.79	6.52	21,725	20,343	49	46
2	2008 - 2007	8.06	6.12	23,475	21,725	52	49
3	2009 - 2008	3.25	-7.69	24,238	23,475	48	52
4	2010 - 2009	2	2.08	24,722	24,238	49	48
5	2011 - 2010	2.27	4.08	25,284	24,722	51	49
6	2012 - 2011	3.14	5.88	26,077	25,284	54	51
7	2013 - 2012	-1.56	5.56	25,671	26,077	57	54
8	2014 - 2013	2.54	0	26,324	25,671	57	57
9	2015 - 2014	2.42	0	26,960	26,324	57	57
10	2016 - 2015	3.66	-1.75	27,947	26,960	56	57
11	2017 - 2016	6.4	10.71	29,736	27,947	62	56
12	2018 - 2017	7.55	1.61	31,981	29,736	63	62

V záverečnom selekte som v stĺpci real\_price\_growth ukázal rozdiel medzi nárastom cien potravín a nárastom miezd. Výsledkom je tabuľka, ktorá ukazuje, že v období medzi rokmi 2012 a 2013 došlo k výraznému medziročnému nárastu cien potravín oproti mzdám, o 7.12 %. V žiadnom roku však nedošlo k nárastu vyššiemu ako 10%.

	ABC period	123 real_price_growth	123 salary_growth	123 food_price_growth
1	2013 - 2012	7.12	-1.56	5.56
2	2017 - 2016	4.31	6.4	10.71
3	2012 - 2011	2.74	3.14	5.88
4	2011 - 2010	1.81	2.27	4.08
5	2010 - 2009	0.08	2	2.08
6	2007 - 2006	-0.27	6.79	6.52
7	2008 - 2007	-1.94	8.06	6.12
8	2015 - 2014	-2.42	2.42	0
9	2014 - 2013	-2.54	2.54	0
10	2016 - 2015	-5.41	3.66	-1.75
11	2018 - 2017	-5.94	7.55	1.61
12	2009 - 2008	-10.94	3.25	-7.69

## Tabulka 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?**

Základom je skript z otázky č.4, z ktorého som výsledky pre mzdy a ceny spojil s tabuľkou 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 agregáciu 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 6.52 %. Pokiaľ úroveň HDP poklesla, nasledujúci rok mzdy rástli v priemere o 0.99%.

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%	6.52	3.12	5.72
2	gdp growth < 0%	0.99	2.55	-1.83