

Zadání projektu

Úvod do projektu

Na vašem analytickém oddělení nezávislé společnosti, která se zabývá životní úrovní občanů, jste se dohodli, že se pokusíte odpovědět na pár definovaných výzkumných otázek, které adresují **dostupnost základních potravin široké veřejnosti**. Kolegové již vydefinovali základní otázky, na které se pokusí odpovědět a poskytnout tuto informaci tiskovému oddělení. Toto oddělení bude výsledky prezentovat na následující konferenci zaměřené na tuto oblast.

Potřebuji k tomu **od vás připravit robustní datové podklady**, ve kterých bude možné vidět **porovnání dostupnosti potravin na základě průměrných příjmů za určité časové období**.

Jako dodatečný materiál připravte i tabulku s HDP, GINI koeficientem a populací **dalších evropských států** ve stejném období, jako primární přehled pro ČR.

Datové sady, které je možné použít pro získání vhodného datového podkladu

Primární tabulky:

1. [czechia_payroll](#) – Informace o mzdách v různých odvětvích za několikaleté období. Datová sada pochází z Portálu otevřených dat ČR.
2. [czechia_payroll_calculation](#) – Číselník kalkulací v tabulce mezd.
3. [czechia_payroll_industry_branch](#) – Číselník odvětví v tabulce mezd.
4. [czechia_payroll_unit](#) – Číselník jednotek hodnot v tabulce mezd.
5. [czechia_payroll_value_type](#) – Číselník typů hodnot v tabulce mezd.
6. [czechia_price](#) – Informace o cenách vybraných potravin za několikaleté období. Datová sada pochází z Portálu otevřených dat ČR.

7. `czechia_price_category` – Číselník kategorií potravin, které se vyskytují v našem přehledu.

Číselníky sdílených informací o ČR:

1. `czechia_region` – Číselník krajů České republiky dle normy CZ-NUTS 2.
2. `czechia_district` – Číselník okresů České republiky dle normy LAU.

Dodatečné tabulky:

1. `countries` – Všechné informace o zemích na světě, například hlavní město, měna, národní jídlo nebo průměrná výška populace.
2. `economies` – HDP, GINI, daňová zátěž, atd. pro daný stát a rok.

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?

Výstup projektu

Pomozte kolegům s daným úkolem. Výstupem by měly být dvě tabulky v databázi, ze kterých se požadovaná data dají získat. Tabulky pojmenujte `t_{jmeno}_{prijmeni}_project_SQL_primary_final` (pro data mezd a cen potravin za Českou republiku sjednocených na totožné porovnatelné období – společné roky) a `t_{jmeno}_{prijmeni}_project_SQL_secondary_final` (pro dodatečná data o dalších evropských státech).

Dále připravte sadu SQL, které z vámi připravených tabulek získají datový podklad k odpovězení na vytyčené výzkumné otázky. Pozor, otázky/hypotézy mohou vaše výstupy podporovat i vyvracet! Záleží na tom, co říkají data.

Na svém GitHub účtu vytvořte repozitář (může být soukromý), kam uložíte všechny informace k projektu – hlavně SQL skript generující výslednou tabulku, popis mezivýsledků (průvodní listinu) a informace o výstupních datech (například kde chybí hodnoty apod.).

Zpracování finálních tabulek

Primární tabulka **t_petr_melicharik_project_SQL_primary_final** byla vytvořena spojením dvou tabulek: **t_avg_pay** (průměrné mzdy v různých odvětvích za určité období) a **t_avg_price** (průměrné ceny různých potravin v určitém období).

Výsledkem je tabulka, která zobrazuje průměrné mzdy dle odvětví a průměrné ceny potravin v období let 2006 – 2018.

	year	industry_code	industry_name	average_pay	product_code	product_name	price_value	price_unit	average_price
1	2006	N	Administrativní a podpůrné činnosti	13,929	116,103	Banány žluté	1	kg	27.305
2	2006	N	Administrativní a podpůrné činnosti	13,929	118,101	Čukr krystalový	1	kg	21.727
3	2006	N	Administrativní a podpůrné činnosti	13,929	114,501	Eidamská cihla	1	kg	110.947
4	2006	N	Administrativní a podpůrné činnosti	13,929	112,101	Hovězí maso zadní bez kosti	1	kg	166.337
5	2006	N	Administrativní a podpůrné činnosti	13,929	111,301	Chléb konzumní krmínový	1	kg	16.124
6	2006	N	Administrativní a podpůrné činnosti	13,929	116,104	Jablka konzumní	1	kg	30.715
7	2006	N	Administrativní a podpůrné činnosti	13,929	114,401	Jogurt bílý netučný	150	g	5.829
8	2006	N	Administrativní a podpůrné činnosti	13,929	2,000,001	Kapři živý	1	kg	69.35
9	2006	N	Administrativní a podpůrné činnosti	13,929	117,401	Konzumní brambory	1	kg	12.075
10	2006	N	Administrativní a podpůrné činnosti	13,929	112,401	Kuřata kuchařská celá	1	kg	47.474
11	2006	N	Administrativní a podpůrné činnosti	13,929	115,101	Máslo	1	kg	104.388
12	2006	N	Administrativní a podpůrné činnosti	13,929	114,201	Mléko polotučné pasteurované	1	l	14.438
13	2006	N	Administrativní a podpůrné činnosti	13,929	117,106	Míkev	1	kg	14.409
14	2006	N	Administrativní a podpůrné činnosti	13,929	117,103	Papriky	1	kg	35.31
15	2006	N	Administrativní a podpůrné činnosti	13,929	111,303	Pecivo pšeničné bílé	1	kg	38.596
16	2006	N	Administrativní a podpůrné činnosti	13,929	213,201	Pivo výčepní, světlé, lahvé	0.5	l	8.45
17	2006	N	Administrativní a podpůrné činnosti	13,929	116,101	Pomeranče	1	kg	24.733
18	2006	N	Administrativní a podpůrné činnosti	13,929	122,102	Přírodní minerální voda uhlíkatá	1	l	7.688

Sekundární tabulka zobrazuje výši HDP, populace a gini koeficient v letech 2006 – 2018 v evropských státech.

economies 1 ✕					
SELECT e.country, e.year, e.GDP, e.population, e.gini FROM ¹ Enter a SQL expression to filter res					
	country	year	GDP	population	gini
100	Croatia	2,014	57,909,810,141.88	4,238,389	32.1
101	Croatia	2,015	59,318,112,197.572	4,203,604	31.1
102	Croatia	2,016	61,393,139,836.765	4,174,349	30.9
103	Croatia	2,017	63,504,583,355.145	4,124,531	30.4
104	Croatia	2,018	65,287,544,197.869	4,087,843	29.7
105	Czech Republic	2,006	197,470,142,753.551	10,238,905	26.7
106	Czech Republic	2,007	208,469,898,850.69	10,298,828	26
107	Czech Republic	2,008	214,070,259,127.502	10,384,603	26.3
108	Czech Republic	2,009	204,100,298,391.036	10,443,936	26.2
109	Czech Republic	2,010	209,069,940,963.177	10,474,410	26.6
110	Czech Republic	2,011	212,750,323,790.759	10,496,088	26.4
111	Czech Republic	2,012	211,080,224,602.858	10,510,785	26.1
112	Czech Republic	2,013	210,983,331,025.624	10,514,272	26.5
113	Czech Republic	2,014	215,755,991,069.112	10,525,347	25.9
114	Czech Republic	2,015	227,381,745,549.048	10,546,059	25.9
115	Czech Republic	2,016	233,151,067,380.883	10,566,332	25.4
116	Czech Republic	2,017	245,202,002,265.920	10,594,428	24.9

Výzkumné otázky

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

Vytvořením porovnávací tabulky t_avg_pay_comp získáme průměrné mzdy v odvětvích v letech 2006 a 2018, ze kterých vidíme, že mzda roste ve všech odvětvích.

	industry_name	123 year1	123 avg_pay_2006	123 year2	123 avg_pay_2018
1	Administrativní a podpůrné činnosti	2,006	13,929	2,018	20,560
2	Činnosti v oblasti nemovitostí	2,006	18,396	2,018	27,584
3	Doprava a skladování	2,006	19,125	2,018	29,298
4	Informační a komunikační činnosti	2,006	35,368	2,018	56,101
5	Kulturní, zábavní a rekreační činnosti	2,006	16,152	2,018	27,580
6	Ostatní činnosti	2,006	15,719	2,018	22,473
7	Peněžnictví a pojišťovnictví	2,006	39,690	2,018	54,111
8	Profesní, vědecké a technické činnosti	2,006	23,875	2,018	38,074
9	Stavebnictví	2,006	17,747	2,018	27,994
10	Těžba a dobývání	2,006	24,042	2,018	35,991
11	Ubytování, stravování a pohostinství	2,006	11,390	2,018	18,770
12	Velkoobchod a maloobchod; opravy a údržba motorových vozidel	2,006	17,865	2,018	29,405
13	Veřejná správa a obrana; povinné sociální zabezpečení	2,006	23,069	2,018	36,014
14	Výroba a rozvod elektřiny, plynu, tepla a klimatiz. vzduchu	2,006	28,844	2,018	46,179
15	Vzdělávání	2,006	19,114	2,018	29,745
16	Zásobování vodou; činnosti související s odpady a sanacemi	2,006	18,490	2,018	28,498
17	Zdravotní a sociální péče	2,006	18,508	2,018	32,746
18	Zemědělství, lesnictví, rybářství	2,006	14,619	2,018	25,291
19	Zpracovatelský průmysl	2,006	18,379	2,018	31,768

Kontrolou pomocí příkazu CASE, kdy vytvoříme nový sloupec s hodnotami 1 – růst mzdy nebo 0 – pokles mzdy dojdeme ke stejnému výsledku.

	industry_name	123 year1	123 avg_pay_2006	123 year2	123 avg_pay_2018	123 pay_raise
1	Administrativní a podpůrné činnosti	2,006	13,929	2,018	20,560	1
2	Činnosti v oblasti nemovitostí	2,006	18,396	2,018	27,584	1
3	Doprava a skladování	2,006	19,125	2,018	29,298	1
4	Informační a komunikační činnosti	2,006	35,368	2,018	56,101	1
5	Kulturní, zábavní a rekreační činnosti	2,006	16,152	2,018	27,580	1
6	Ostatní činnosti	2,006	15,719	2,018	22,473	1
7	Peněžnictví a pojišťovnictví	2,006	39,690	2,018	54,111	1
8	Profesní, vědecké a technické činnosti	2,006	23,875	2,018	38,074	1
9	Stavebnictví	2,006	17,747	2,018	27,994	1
10	Těžba a dobývání	2,006	24,042	2,018	35,991	1
11	Ubytování, stravování a pohostinství	2,006	11,390	2,018	18,770	1
12	Velkoobchod a maloobchod; opravy a údržba motorových vozidel	2,006	17,865	2,018	29,405	1
13	Veřejná správa a obrana; povinné sociální zabezpečení	2,006	23,069	2,018	36,014	1
14	Výroba a rozvod elektřiny, plynu, tepla a klimatiz. vzduchu	2,006	28,844	2,018	46,179	1
15	Vzdělávání	2,006	19,114	2,018	29,745	1
16	Zásobování vodou; činnosti související s odpady a sanacemi	2,006	18,490	2,018	28,498	1
17	Zdravotní a sociální péče	2,006	18,508	2,018	32,746	1
18	Zemědělství, lesnictví, rybářství	2,006	14,619	2,018	25,291	1
19	Zpracovatelský průmysl	2,006	18,379	2,018	31,768	1

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

V roce 2006 si za průměrnou mzdu je možné koupit 1289 ks chleba a 1441 litrů mléka.

V roce 2018 si za průměrnou mzdu je možné koupit 1344 ks chleba a 1643 litrů mléka.

year	avg_pay_in_year	product_name	avg_product_price_in_year	quantity_to_buy
2,006	20,754	Chléb konzumní kmínový	16.1	1,289
2,006	20,754	Mléko polotučné pasterované	14.4	1,441
2,018	32,536	Chléb konzumní kmínový	24.2	1,344
2,018	32,536	Mléko polotučné pasterované	19.8	1,643

3. Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší procentuálně meziroční nárůst)?

Nejpomaleji zdražují banány žluté, které od roku 2006 do roku 2018 zdražily o 7,4 %.

Nejvíce ve stejném období zdražilo máslo – o 98,4 %.

Dvě položky jsou dokonce levnější – cukr krystalový o 27,5 % a rajská jablka červená kulatá o 23,1 %.

	name	yearly_price	avarage_price	yearly_price	avarage_price	percentual_change
1	Cukr krystalový	2,006	21.727	2,018	15.75	-27.5
2	Rajská jablka červená kulatá	2,006	57.83	2,018	44.488	-23.1
3	Banány žluté	2,006	27.305	2,018	29.321	7.4
4	Vepřová pečeně s kostí	2,006	105.178	2,018	116.853	11.1
5	Přírodní minerální voda uhlíčitá	2,006	7.688	2,018	8.645	12.4
6	Pečivo pšeničné bílé	2,006	38.596	2,018	43.835	13.6
7	Jablka konzumní	2,006	30.715	2,018	36.175	17.8
8	Šunkový salám	2,006	116.774	2,018	144.876	24.1
9	Konzumní brambory	2,006	12.075	2,018	15.079	24.9
10	Eidamská cihla	2,006	110.947	2,018	142.454	28.4
11	Hovězí maso zadní bez kosti	2,006	166.337	2,018	223.259	34.2
12	Kapr živý	2,006	69.35	2,018	93.459	34.8
13	Mléko polotučné pasterované	2,006	14.438	2,018	19.817	37.3
14	Pivo výčepní, světlé, lahvé	2,006	8.45	2,018	11.809	39.8
15	Rostlinný roztíratelný tuk	2,006	69.452	2,018	99.398	43.1
16	Kuřata kuchařská celá	2,006	47.474	2,018	69.31	46
17	Pomeranče	2,006	24.733	2,018	36.501	47.6
18	Chléb konzumní kmínový	2,006	16.124	2,018	24.239	50.3
19	Pšeničná mouka hladká	2,006	7.408	2,018	11.435	54.4
20	Mrkev	2,006	14.409	2,018	22.453	55.8
21	Jogurt bílý netučný	2,006	5.829	2,018	9.171	57.3
22	Vejce slepičí čerstvá	2,006	23.495	2,018	38.39	63.4
23	Rýže loupáná dlouhozrnná	2,006	21.29	2,018	36.175	69.9
24	Papriky	2,006	35.31	2,018	60.468	71.2
25	Těstoviny vaječné	2,006	26.096	2,018	47.874	83.5
26	Máslo	2,006	104.388	2,018	207.084	98.4

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

Nejdříve jsme si připravil pomocnou tabulku, ve které se nachází průměrné mzdy a průměrné ceny potravin v ČR za období 2006 – 2018.

payroll_year	avarage_payment	avarage_product_price
2,006	20,754	45.52
2,007	22,173	48.59
2,008	23,918	51.6
2,009	24,674	48.29
2,010	25,156	49.23
2,011	25,736	50.88
2,012	26,516	54.3
2,013	26,103	57.07
2,014	26,771	57.49
2,015	27,443	57.18
2,016	28,444	56.49
2,017	30,231	61.93
2,018	32,536	63.27

Tuto tabulku jsem napojil na sebe samou, abych měl v řádku srovnání dvou let.

payroll_year	avarage_payment	avarage_product_price	payroll_year	avarage_payment	avarage_product_price
2,006	20,754	45.52	2,007	22,173	48.59
2,007	22,173	48.59	2,008	23,918	51.6
2,008	23,918	51.6	2,009	24,674	48.29
2,009	24,674	48.29	2,010	25,156	49.23
2,010	25,156	49.23	2,011	25,736	50.88
2,011	25,736	50.88	2,012	26,516	54.3
2,012	26,516	54.3	2,013	26,103	57.07
2,013	26,103	57.07	2,014	26,771	57.49
2,014	26,771	57.49	2,015	27,443	57.18
2,015	27,443	57.18	2,016	28,444	56.49
2,016	28,444	56.49	2,017	30,231	61.93
2,017	30,231	61.93	2,018	32,536	63.27

Po potřebných výpočet lze dojít k závěru, že v žádném roce nebyl meziroční nárůst cen potravin větší o více jak 10 procent než meziroční nárůst mezd.

yearly_pay_change	yearly_price_change
6.8	6.7
7.9	6.2
3.2	-6.4
2	1.9
2.3	3.4
3	6.7
-1.6	5.1
2.6	0.7
2.5	-0.5
3.6	-1.2
6.3	9.6
7.6	2.2

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?

Potřebná tabulka vznikla spojením tabulky s průměry mezd a cen a sekundární tabulky s informacemi o výši HDP za potřebné období pro Českou republiku. Poslední 3 sloupce jsou potřebné výpočty o meziročních změnách HPD, mezd, cenách produktů.

year	GDP	average_payment	average_product_price	year	GDP	average_payment	average_product_price	gdp_change	pay_change	product_change
2006	197,470,142,753.551	20,754	45.52	2007	208,469,898,850.69	22,173	48.59	5.57	6.84	6.74
2007	208,469,898,850.69	22,173	48.59	2008	214,070,259,127.502	23,918	51.6	2.69	7.87	6.19
2008	214,070,259,127.502	23,918	51.6	2009	204,100,298,391.036	24,674	48.29	-4.66	3.16	-6.41
2009	204,100,298,391.036	24,674	48.29	2010	209,069,940,963.177	25,156	49.23	2.43	1.95	1.95
2010	209,069,940,963.177	25,156	49.23	2011	212,750,323,790.759	25,736	50.88	1.76	2.31	3.35
2011	212,750,323,790.759	25,736	50.88	2012	211,080,224,602.858	26,516	54.3	-0.79	3.03	6.72
2012	211,080,224,602.858	26,516	54.3	2013	210,983,331,025.624	26,103	57.07	-0.05	-1.56	5.1
2013	210,983,331,025.624	26,103	57.07	2014	215,755,991,069.112	26,771	57.49	2.26	2.56	0.74
2014	215,755,991,069.112	26,771	57.49	2015	227,381,745,549.048	27,443	57.18	5.39	2.51	-0.54
2015	227,381,745,549.048	27,443	57.18	2016	233,151,067,380.883	28,444	56.49	2.54	3.65	-1.21
2016	233,151,067,380.883	28,444	56.49	2017	245,202,003,265.939	30,231	61.93	5.17	6.28	9.63
2017	245,202,003,265.939	30,231	61.93	2018	253,045,172,103.95	32,536	63.27	3.2	7.62	2.16

gdp_change	pay_change	product_change
5.57	6.84	6.74
2.69	7.87	6.19
-4.66	3.16	-6.41
2.43	1.95	1.95
1.76	2.31	3.35
-0.79	3.03	6.72
-0.05	-1.56	5.1
2.26	2.56	0.74
5.39	2.51	-0.54
2.54	3.65	-1.21
5.17	6.28	9.63
3.2	7.62	2.16

Pro zodpovězení výzkumné otázky jsem vybral roky 2007, 2015 a 2017, kdy došlo k procentuálnímu nejvyššímu růstu HDP (> 5 %).

V roce 2007 došlo k růstu HDP o 5,57 %. Ve stejném roce se mzdy navýšily o 6,84% a ceny produktů o 6,74 %. V následujícím roce 2018 mzdy stouply o 7,87 % a ceny produktů o 6,19%.

V roce 2015 došlo k růstu HDP o 5,39%. Ve stejném roce se výše mezd zvýšila o 2,51%, ceny produktů se snížily o 0,54 procent. V následujícím

roce 2016 se mzdy zvýšily o 3,65%, ceny produktů opět klesly a to o 1,21%.

V roce 2017 došlo k růstu HDP o 5,17 %. Mzdy se v tomto roce zvýšily o 6,28 %, ceny potravin o 9,63%. V následujícím roce 2018 výše mezd nadále rostla, konkrétně o 7,62 %, ceny potravin vzrostly také, avšak nešlo již o tak razantní růst (2,16%).

Závěr:

Při výraznějším růstu HDP docházelo v témže nebo následujícím roce také k vyššímu růstu průměrné mzdy, avšak průměrné ceny potravin dokonce v některých letech při růstu HDP klesaly. Vyšší růst HDP tedy ovlivnil i růst průměrné mzdy, na průměrnou cenu potravin však vliv neměl.