

---

## Zadání 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.

Potřebují 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.

### 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 percentuální 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 zodpovězení 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.).

## Odpovědi na výzkumné otázky:

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

ODPOVĚĎ: Pro zodpovězení otázky jsem se rozhodla porovnávat nejstarší, nejnovější, a průměrně stará data, tedy z let 2006, 2012 a 2018. Provedla jsem porovnání růstu mezd mezi lety:

2012 a 2018

2006 a 2012

2006 a 2018

Název_odvětví	123 Průměrná_mzda_2006	123 Průměrná_mzda_2012	123 Průměrná_mzda_2018	abc Změna mzdy 2012-2018	abc Změna mzdy 2006-2012	abc Změna mzdy 2006-2018
Administrativní a podpůrné činnosti	13,414	15,935	20,165.5	+	+	+
Činnosti v oblasti nemovitostí	17,550.75	21,438.5	27,059.5	+	+	+
Doprava a skladování	18,993.75	22,998.5	29,135.75	+	+	+
Informační a komunikační činnosti	34,942	45,689.5	55,474.75	+	+	+
Kulturní, zábavní a rekreační činnosti	15,477.5	19,550.5	26,761.75	+	+	+
Ostatní činnosti	14,955	17,853	21,249.5	+	+	+
Peněžnictví a pojišťovnictví	39,353.25	49,706.5	53,339.5	+	+	+
Profesní, vědecké a technické činnosti	23,105.75	31,141.25	37,162.75	+	+	+
Stavebnictví	17,643.75	22,571.75	27,821.75	+	+	+
Těžba a dobývání	24,017.5	32,487	35,942.75	+	+	+
Ubytování, stravování a pohostinství	11,105.75	12,705.75	18,270.25	+	+	+
Velkoobchod a maloobchod; opravy	17,507.25	22,441	28,833.75	+	+	+
Veřejná správa a obrana; povinné soc	22,852.75	26,285.25	35,714.5	+	+	+
Výroba a rozvod elektřiny, plynu, tepl	28,477	42,107.25	45,982.5	+	+	+
Vzdělávání	18,197.25	21,912.5	28,047.5	+	+	+
Zásobování vodou; činnosti souvisejíc	18,240.75	23,261.75	28,272	+	+	+
Zdravotní a sociální péče	17,973.5	23,621.5	31,629.5	+	+	+
Zemědělství, lesnictví, rybářství	14,421	19,537	25,115.25	+	+	+
Zpracovatelský průmysl	18,276.75	24,221.75	31,646.25	+	+	+

Pokud ve sledovaném rozmezí mzda rostla, nechala jsem si zobrazit +, v opačném případě potom -. Na základě výsledků moje odpověď zní: **Mzdy rostou ve všech odvětvích.**

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

ODPOVĚĎ:

abc Produkt	123 Rok	123 Průměrná_mzda	123 Cena	123 Zakoupitelné_množství
Chléb konzumní kmínový	2,006	20,342.38	16.12	1,262
Chléb konzumní kmínový	2,018	31,980.26	24.24	1,319
Mléko polotučné pasterované	2,006	20,342.38	14.44	1,409
Mléko polotučné pasterované	2,018	31,980.26	19.82	1,614

Porovnatelné období je v letech 2006 až 2018. **Konzumního kmínového chleba bylo možné s průměrnou mzdou zakoupit 1262 bochníků v roce 2006 a 1319 bochníků o 12 let později. V případě polotučného mléka bychom si v roce 2006 mohli zakoupit 1409 litrů, ale v roce 2018 již 1614 litrů.**

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

ABC Produkt	123 Maximální_rozdíl_procenta	123 Průměrná_roční_změna_procenta
Cukr krystalový	44	-1.92
Rajská jablka červená kulatá	55.2	-0.75
Banány žluté	25.2	0.83
Vepřová pečeně s kostí	16	0.99
Přírodní minerální voda uhlíč	9.5	1.02
Šunkový salám	10	1.85
Jablka konzumní	38.6	2.02
Pečivo pšeničné bílé	57.5	2.22
Hovězí maso zadní bez kosti	12.7	2.53
Kapr živý	16.3	2.58
Pivo výčepní, světlé, lahvové	8.2	2.85
Eidamská cihla	49.3	2.92
Mléko polotučné pasterované	27.7	2.98
Rostlinný roztíratelný tuk	26.1	3.22
Kuřata kuchaňská celá	18.4	3.38
Pomeranče	30.5	3.6
Jogurt bílý netučný	15.2	3.96
Chléb konzumní kmínový	35.8	3.97
Konzumní brambory	83.8	4.19
Rýže loupaná dlouhozrná	43	5
Mrkev	68.4	5.23
Pšeničná mouka hladká	66.1	5.25
Těstoviny vaječné	15.9	5.26
Vejce slepičí čerstvá	69.9	5.55
Máslo	49.6	6.68
Papriky	111.8	7.29

ODPOVĚĎ: Otázka je položena tak, že se nás ptá konkrétně na potraviny, které zdražují. Správně bychom tedy měli vyjmout ty potraviny, které naopak zlevňují (tedy cukr a rajská jablka). **Odpověď by pak byly banány žluté, které zdražují v průměru o 0.83 % ročně.**

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

123 Rok	123 Průměr_cen	123 Průměr_mezd	123 Rozdíl_mezd_a_cen
2,013	6	-0.78	6.78
2,012	7.48	2.82	4.66
2,011	4.85	2.15	2.7
2,007	9.26	6.91	2.35
2,008	8.91	7.42	1.49
2,017	7.24	6.72	0.52
2,010	1.52	2.18	-0.67
2,014	-0.63	2.5	-3.13
2,015	-0.68	2.77	-3.45
2,018	2.37	7.67	-5.3
2,016	-1.56	3.96	-5.52
2,009	-6.58	3.11	-9.69

ODPOVĚĎ: **Ne, neexistuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd.** Nejvyšší rozdíl mezi meziročním růstem cen vůči mzdám byl v roce 2013, kdy činil 6.78 %

**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?**

Pro účely zodpovězení této otázky jsem použila již dříve vypočítané meziroční změny ve mzdách a cenách potravin, a přidala jsem sloupec s meziročním rozdílem v HDP. Data jsem porovnávala pouze pro Českou Republiku, protože k ostatním evropským státům nemám k dispozici data o mzdách ani cenách potravin.

<sup>123</sup> Rok	<sup>ABC</sup> Země	<sup>123</sup> Meziroční_rozdíl_HDP	<sup>123</sup> Meziroční_procentuální_změna_mezd	<sup>123</sup> Meziroční_procentuální_změna_cen
2,007	Czech Republic	[NULL]	5.2	9.26
2,008	Czech Republic	2.69	2.4	8.91
2,009	Czech Republic	-4.66	4.1	-6.58
2,010	Czech Republic	2.43	1.2	1.52
2,011	Czech Republic	1.76	1.9	4.85
2,012	Czech Republic	-0.79	2.7	7.48
2,013	Czech Republic	-0.05	0.6	6
2,014	Czech Republic	2.26	1	-0.63
2,015	Czech Republic	5.39	2.3	-0.68
2,016	Czech Republic	2.54	5.9	-1.56
2,017	Czech Republic	5.17	7.5	7.24
2,018	Czech Republic	3.2	6.9	2.37

Jak vidíme, meziroční změny HDP se v nadcházejícím roce vždy projevily na růstu mezd. V případě propadu HDP v následujícím roce mzdy rostly oproti předcházejícím letům pomaleji. Naopak v případě rychle rostoucího HDP (například v roce 2015) můžeme v následujícím roce pozorovat veliký nárůst mezd, a to o 5.9 %. Naproti tomu vztah mezi HDP a cenami potravin již tak zřejmý není. V některých letech by se mohlo zdát, že rostoucí HDP má za důsledek zlevňování (roky 2014 až 2016), nicméně toto nám vyvrací hned rok 2017, kdy je nárůst HDP sice značný, ale ceny vyrostly o více než 7 %. Zde bych tedy žádnou vazbu na základě těchto dat nehledala.

ODPOVĚĎ: **Výška HDP má vliv na nárůst mezd v následujících letech, zatímco vliv na ceny sledovaných produktů pozorován nebyl.**

---

## *Background projektu – s jakými tabulkami se pracovalo?*

---

### PRIMÁRNÍ TABULKY

<b>czechia_payroll</b>	Informace o mzdách v různých odvětvích za několikaleté období. Datová sada pochází z Portálu otevřených dat ČR.
<b>czechia_payroll_calculation</b>	Číselník kalkulací v tabulce mezd.
<b>czechia_payroll_industry_branch</b>	Číselník odvětví v tabulce mezd.
<b>czechia_payroll_unit</b>	Číselník jednotek hodnot v tabulce mezd.
<b>czechia_payroll_value_type</b>	Číselník typů hodnot v tabulce mezd.
<b>czechia_price</b>	Informace o cenách vybraných potravin za několikaleté období. Datová sada pochází z Portálu otevřených dat ČR.
<b>czechia_price_category</b>	Číselník kategorií potravin, které se vyskytují v našem přehledu.

### ČÍSELNÍKY SDÍLENÝCH INFORMACÍ O ČR

<b>czechia_region</b>	Číselník krajů České republiky dle normy CZ-NUTS 2.
<b>czechia_district</b>	Číselník okresů České republiky dle normy LAU.

### DODATEČNÉ TABULKY

<b>countries</b>	Všemožné informace o zemích na světě, například hlavní město, měna, národní jídlo nebo průměrná výška populace.
<b>economies</b>	HDP, GINI, daňová zátěž, atd. pro daný stát a rok.