

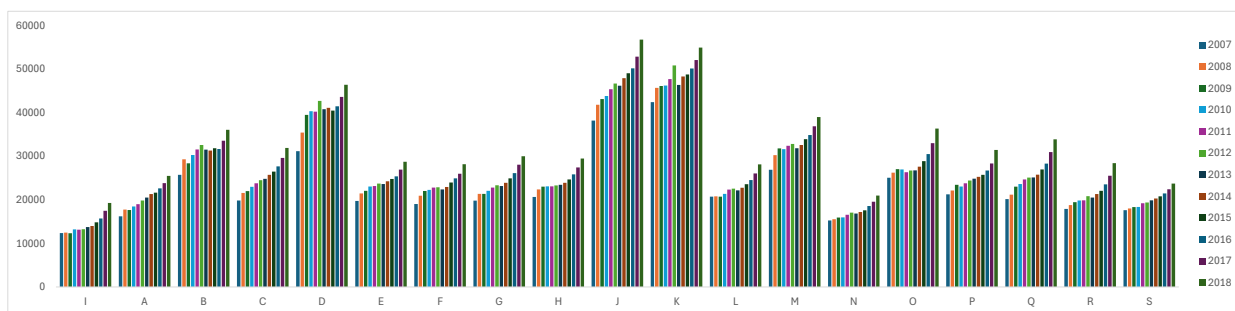
SQL PROJEKT

Abychom mohli odpovědět na výzkumné otázky, je důležité připravit si správná data do primární tabulky. Postupovala jsem tak, že jsem si nejprve z dat dostupných v tabulce *czechia_payroll* vzala potřebné sloupce a připojila k nim pomocí JOIN sloupce z tabulky *czechia_price*. Napojení jsem provedla pomocí roku/data, které je dostupné v obou tabulkách a následně provedla zpřesnění (vyfiltrování), která data potřebuji přidat - jedná se o *value_type_code* 5958, který označuje položky Průměrná hrubá mzda na zaměstnance a *calculation_code* roven 200 jako přepočtený. Dále jsem přidala *czechia_price_category*, aby měla v tabulce kromě kódu potravin i jejich název. V posledním kroku jsem přidala klauzuli GROUP BY, do které jsem uvedla všechny sloupce, které nejsou obsaženy v agregačních funkcích. Tímto jsem si připravila kompletní data, ze kterých budu schopna zodpovědět následující dotazy. Kromě primární tabulky jsem připravila data také pro ostatní evropské státy, kde jsem vybrala data z tabulky *countries* a doplnila ji sloupce z tabulky *economies*. Napojení jsem provedla pomocí klauzule JOIN a klíčem byly názvy států, které jsou dostupné v obou tabulkách. Podmínkou v klauzuli *where* bylo, že data jsou pro evropský kontinent a hodnota GDP je neprázdná.

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

K zodpovězení otázky jsem použila CTE - klauzuli WITH. Z primární tabulky jsem vybrala potřebné sloupce jako kód odvětví, rok, průměrnou mzdu a mzdu z předchozího roku, ke které jsem došla přes window funkci *lag*. Dále jsem si z této dočasné tabulky vybrala potřebné sloupce do výsledkové tabulky, kde mám navíc spočítaný průměrný procentuální meziroční přírůstek mezd.

Dle dat v tabulce, kterou jsme získali pomocí SQL dotazu i dle grafu přiloženého níže můžeme tvrdit, že mzdy rostou ve všech sledovaných odvětvích.



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

Nejprve jsem zjišťovala MIN a MAX v dostupných datech, došla jsem k závěru, že budu porovnávat data z let 2006 a 2018. V tabulce jsem rovněž omezila data pouze pro kódy pro chléb a mléko. Veškerá data jsou dostupná z primární tabulky, takže k výsledku bylo snadné se dostat.

V roce 2006 jsme si za průměrnou mzdu 14 818 Kč mohli koupit 919 kg chleba a 1026 litrů mléka.

Oproti tomu v roce 2018 jsme si za průměrnou mzdu 25 467 Kč mohli koupit 1051 kg chleba a 1285 litrů mléka. Vzrostla tedy průměrná mzda i kupní síla.

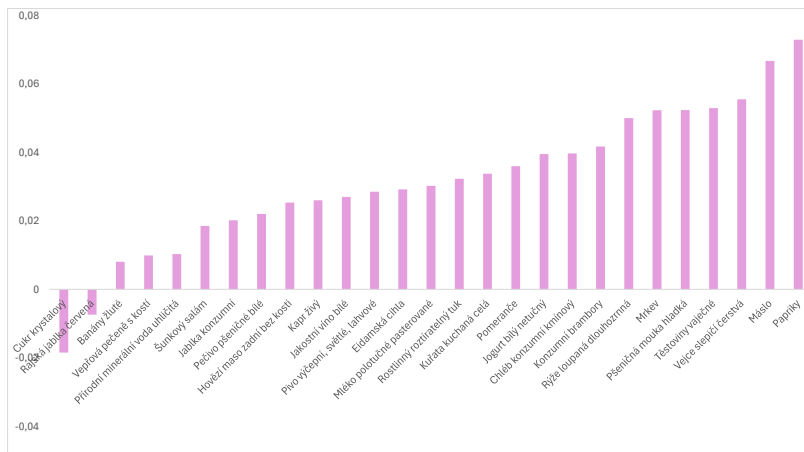
food_category	actual_yr	avg_price	average_wage	avg_amount
Chléb konzumní kminový	2006	16,12	14 818 Kč	919
Mléko polotučné pasterované	2006	14,44	14 818 Kč	1 026
Chléb konzumní kminový	2018	24,24	25 467 Kč	1 051
Mléko polotučné pasterované	2018	19,82	25 467 Kč	1 285

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

Opět jsem použila CTE, kdy jsem si nejprve připravila dočasnou tabulku *growth_prices*, kde jsem vybrala sloupce kategorie potravin, průměrné ceny a pomocí *lag* funkce taky průměrnou cenu potravin v předchozím roce. Ve výsledné tabulce je pak zobrazena kategorie potravin a průměrný meziroční přírůstek ceny potravin.

Při seřazení podle tohoto přírůstku zjišťujeme, že nejpomalejší meziroční nárůst cen je u cukru, rajčat a banánů. Kompletní pohled nám dává graf i tabulka níže.

Nejpomaleji zdražuje:	
1. Cukr krystalový	15. Rostlinný rozstříratelný tuk
2. Rajská jablka červená	16. Kuřata kuchařská celá
3. Banány žluté	17. Pomeranče
4. Vepřová pečeně s kostí	18. Jogurt bílý netučný
5. Přírodní minerální voda	19. Chléb konzumní kminový
6. Šunkový salám	20. Konzumní brambory
7. Jablka konzumní	21. Rýže loupáaná dlouhozrná
8. Pečivo pšeničné bílé	22. Mrkev
9. Hovězí maso zadní bk	23. Pšeničná mouka hladká
10. Kapr živý	24. Těstoviny vaječné
11. Jakostní víno bílé	25. Vejce slepičí čerstvá
12. Pivo výčepní, světlé	26. Máslo
13. Eidamská cihla	27. Papriky
14. Mléko polotučné paster.	



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

Opět jsem využila CTE, kdy jsem si připravila tabulky týkající se cen a mezd, k údajům z předchozího roku jsem došla opět přes po užití window funkce lag. Ve výsledné tabulce jsem vyselektovala aktuální rok a průměrný meziroční nárůst cen potravin a mezd.

Ze závěrů v přiložené tabulce vyplývá, že takový rok naštěstí neexistuje, průměrný meziroční nárůst cen potravin je mnohem pomalejší než průměrný meziroční nárůst mezd.

actual_yr	price_growth_AVG_perc	payroll_growth_AVG_perc
2007	0,49	9,25
2008	0,47	9,74
2009	-0,35	-0,67
2010	0,08	4,55
2011	0,26	2,92
2012	0,39	4,45
2013	0,32	3,47
2014	-0,03	3,80
2015	-0,04	1,62
2016	-0,07	4,51
2017	0,37	5,23
2018	0,13	6,99

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?

V poslední tabulce jsem postupovala podobně, respektive jsem použila i data z předchozí otázky, pomocí CTE jsem si připravila tři dočasné tabulky týkající se GDP, růstu cen a mezd. Tabulku GDP jsem napojila na sekundární tabulku, která údaje o GDP obsahuje a omezila jsem ji podmínkou na data týkající se ČR po roce 2006. Průměrná data za předchozí rok jsem opět zobrazila pomocí funkce lag. Ve finální tabulce, která je rovněž uvedena níže jsem pak nechala zobrazit aktuální rok a meziroční procentuální přírůstky u GDP, cen potravin i mezd.

Z výsledků mého zkoumání usuzuji, že výše průměrných mezd koreluje s výší HDP, v případě cen potravin ale tato závislost není tak silná a jejich pohyb tak nejspíš závisí i na dalších faktorech. U mezd by se dalo říct, že kopírují křivku HDP a v ten samý rok se projevuje jejich změna.

actual_yr	GDP_growth_AVG_perc	price_growth_AVG_perc	payroll_growth_AVG_perc
2007	5,57	0,49	9,25
2008	2,69	0,47	9,74
2009	-4,66	-0,35	-0,67
2010	2,43	0,08	4,55
2011	1,76	0,26	2,92
2012	-0,79	0,39	4,45
2013	-0,05	0,32	3,47
2014	2,26	-0,03	3,80
2015	5,39	-0,04	1,62
2016	2,54	-0,07	4,51
2017	5,17	0,37	5,23
2018	3,20	0,13	6,99

