

PROJEKT SQL

Data o mzdách a cenách potravin a jejich zpracování pomocí SQL

Ing. Renata Kratochvílová

e-mail: renata.venclova@gmail.com

discord: renatakratochvilova

Cíl projektu:

- Cílem projektu je připravit datové podklady, na základě informací o průměrných mzdách v daných odvětvích a průměrných cenách vybraných potravin za roky 2006-2018.
- Připravit dvě tabulky, z nichž první bude obsahovat **porovnání dostupnosti potravin na základě průměrných mezd za určité časové období – primární tabulka**. Druhá tabulka bude obsahovat **informace o HDP, GINI koeficientu a celkové populaci všech evropských států za téže období, tj. 2006-2018 – sekundární tabulka**.
- Připravit podklady pro zodpovězení 5 výzkumných otázek:
 - o *Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají?*
 - o *Kolik je možné si koupit litrů mléka a kilogramů chleba za první a poslední srovnatelné období v dostupných datech cen a mezd?*
 - o *Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší procentuálně meziroční nárůst)?*
 - o *Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (větší než 10 %)?*
 - o *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?*

Využité datové sady:

- **Czechia payroll** – sada obsahuje informace o mzdách v různých odvětvích za několikaleté období (zdroj: Portál otevřených dat ČR)
- **Czechia payroll calculation** – číselník kalkulací v tabulce mezd
- **Czechia payroll industry branch** – číselník odvětví v tabulce mezd
- **Czechia payroll unit** – číselník jednotek hodnot v tabulce mezd
- **Czechia payroll value type** – číselník typů hodnot v tabulce mezd
- **Czechia price** – informace o cenách vybraných potravin za několikaleté období (zdroj: Portál otevřených dat ČR)
- **Czechia price category** – číselník kategorií potravin, které se vyskytují v tabulce cen
- **Countries** – různé informace o zemích na světě, např. hlavní město, měna, národní jídlo, populace apod.
- **Economies** – informace o HDP, GINI koeficientu, daňové zátěži apod. za jednotlivé státy a roky

Primární tabulka

- Seskupuje údaje o průměrných mzdách v jednotlivých odvětvích v období od roku 2006 do roku 2018, na které jsou navázány údaje o průměrných cenách vybraných kategorií.
- Je vytvořena spojením dvou tabulek a to `czechia_payroll` a `czechia_price`.
- Pro zodpovězení výzkumných otázek nám postačí následující informace: rok, název odvětví, průměrná mzda za dané odvětví a daný rok, název komodity, měrná jednotka a průměrná cena komodity vypočtená jako průměr za daný rok.

- Jediným společným jmenovatelem je údaj o roce, na základě něhož jsou tabulky seskupeny.

Předpoklady:

- Průměrná roční mzda je vypočtena jako průměr za čtyři čtvrtletí daného roku v daném odvětví.
- Údaje o průměrné mzdě jsou v přepočtené výši, zohledňují tak délku pracovního poměru. (calculation code = 200)
- Údaje o průměrných cenách jsou vypočteny jako průměr za daný rok.
- Z přehledu je vyloučena položka pod číslem 212101 – Jakostní víno bílé, jelikož údaje o jejích cenách jsou pouze za roky 2015–2018. S položkou tedy není v porovnáních kalkulováno.

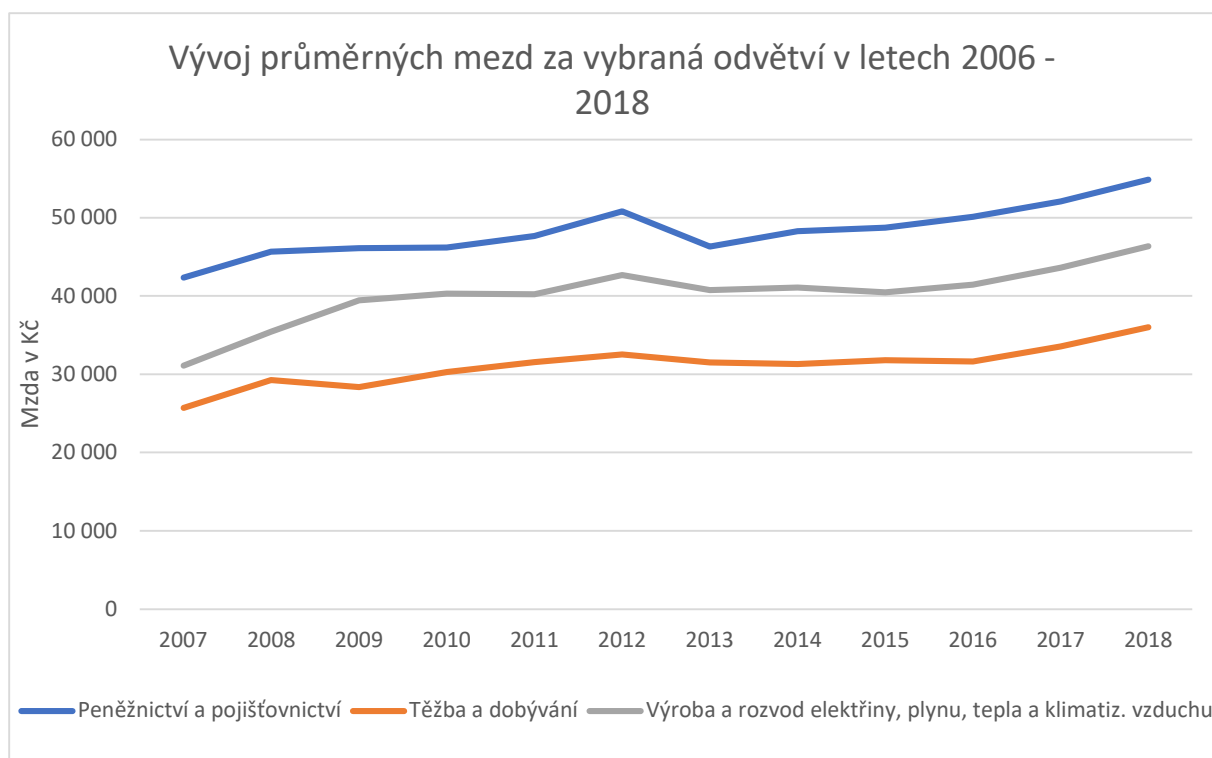
Sekundární tabulka

- Seskupuje údaje o zemích Evropy a vybraných ekonomických ukazatelích za roky 2006-2018.
- Vychází z tabulky countries, která obsahuje informace o tom, zda se země nachází v Evropě či nikoli a tabulky economies, kde jsou uvedeny vybrané ekonomické ukazatele za dané roky.

Výzkumné otázky:

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

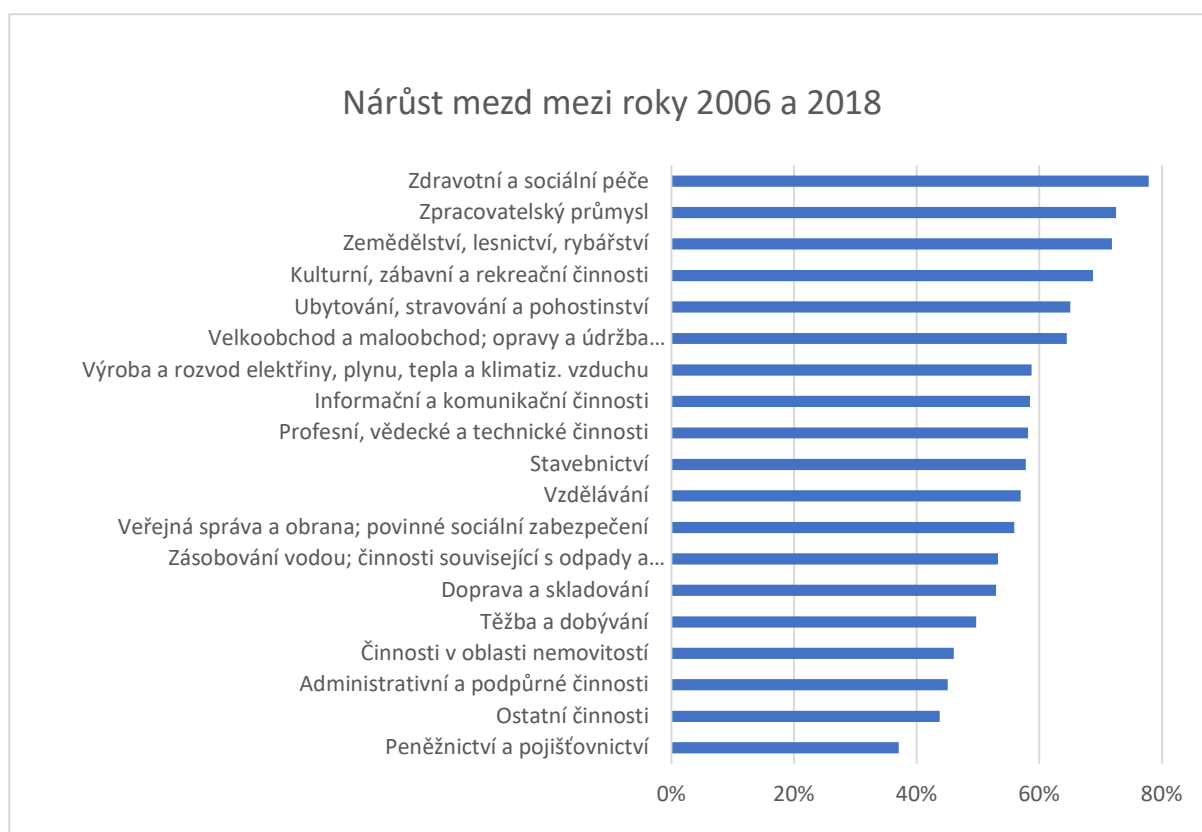
- Podklad v SQL vychází z primární tabulky, kde jsme si vyfiltrovali průměrnou mzdu dle odvětví a vypočetli roční průměr z kvartálních hodnot. K těmto údajům jsme dohráli informaci o průměrné mzdě za předchozí rok. Srovnáváme data za období 2006-2018.
- Jak je vidno z výsledné tabulky v SQL nebo níže v grafu, v průběhu let, došlo k výraznému poklesu průměrné mzdy **v roce 2013 v Peněžnictví a pojišťovnictví (-8,83 %), Výrobě a rozvodu elektřiny, plynu, tepla a klimatiz. vzduchu (-4,44 %), Těžbě a dobývání (-3,24 %).**
- Z uvedených dat by se mohlo zdát, že v tomto roce došlo k poklesu průměrné mzdy. Podíváme-li se však podrobněji na výsledný graf, lze z něho vyčíst, že „snížení“ průměrné mzdy předcházelo **navýšení** průměrné mzdy v předchozím období. Nejedná se tak o postupný pokles mzdy. Tento jev lze vysvětlit změnou, ke které došlo v rámci výpočtu daně z příjmu a zavedení tzv. solidární daně v roce 2013. Lze předpokládat, že z důvodu zavedení solidární daně byly vyplaceny roční odměny o rok dříve, za účelem daňové úspory, což uměle navýšilo průměrnou mzdu v roce 2012 a snížilo v roce 2013.
- V souhrnu lze konstatovat, že v průběhu let má vývoj průměrné mzdy spíše stoupající tendenci. Jeho trend je ovlivněn více faktory jako je např. výše HDP, inflace, cenová hladina apod.

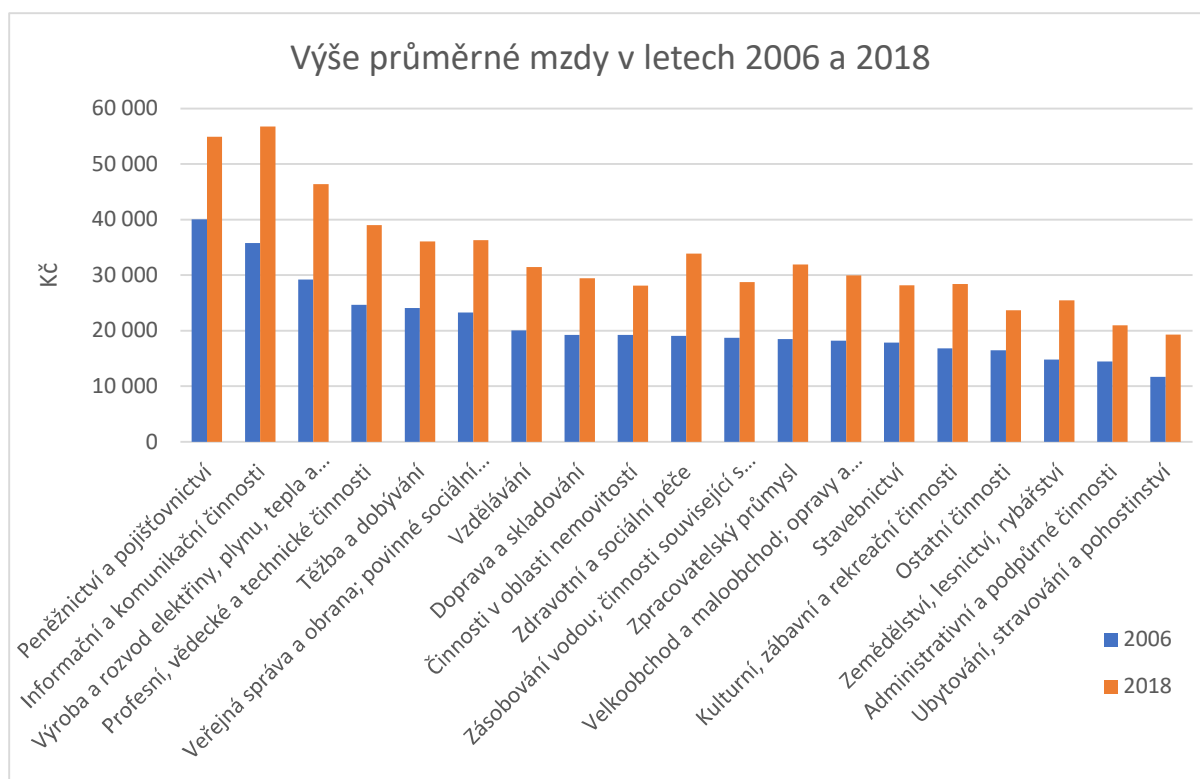


2. Kolik je možné si koupit litrů mléka a kilogramů chleba za první a poslední srovnatelné období v dostupných datech cen a mezd?

- Podklad pro zodpovězení otázky vytvoříme z primární tabulky, kde si vyfiltrujeme dvě komodity (Chléb konzumní kmínový a Mléko polotučné pasterované) za roky 2006 a 2018. Vypočítáme průměrnou cenu za daný rok a přihrajeme informaci o průměrné mzdě. Ta je vypočtena jako průměr za všechna odvětví v daném roce. Pak jednoduše vypočteme počet kusů, které je možné za průměrnou mzdu koupit.
- Ze získané tabulky lze vyčíst, že v roce 2006 jsme si mohli za průměrnou mzdu koupit 1313 ks chleba nebo 1466 ks mléka. V roce 2018 jsme si mohli koupit 1365 ks chleba (o 52 ks více) nebo 1670 ks mléka (o 204 ks více). Vliv na kupované množství má nejen zvýšení průměrné mzdy cca o 56 % ale také nárůst průměrné ceny dané komodity. Chleba zdražil v průměru o cca 50 % a mléko o 37 %.
- Pokud bychom se chtěli podívat na nějaké anomálie, spustíme si druhý příkaz v daném SQL souboru, kde jsou uvedeny a vypočteny rozdíly v kupovaném množství dle jednotlivých odvětví.
 - o Největší pokles v počtu kupovaného množství je v odvětví **Peněžnictví a pojišťovnictví, Ostatních činnostech, Administrativní podpůr. činnosti a Činnosti v oblasti nemovitostí**. Např. osoba pracující v Peněžnictví a pojišťovnictví si mohla v roce 2018 koupit o 219 ks chleba méně než v roce 2006. Tato změna nebyla způsobena pouze nárůstem ceny chleba (cca o 50 %), ale i nejnižším nárůstem mezd mezi sledovanými roky, jak je vidět i na grafu níže. V tomto odvětví došlo meziročně k nejnižšímu nárůstu o 37 %. Další odvětví s nejnižším meziročním přírůstkem mzdy jsou také výše zmíněná odvětví.

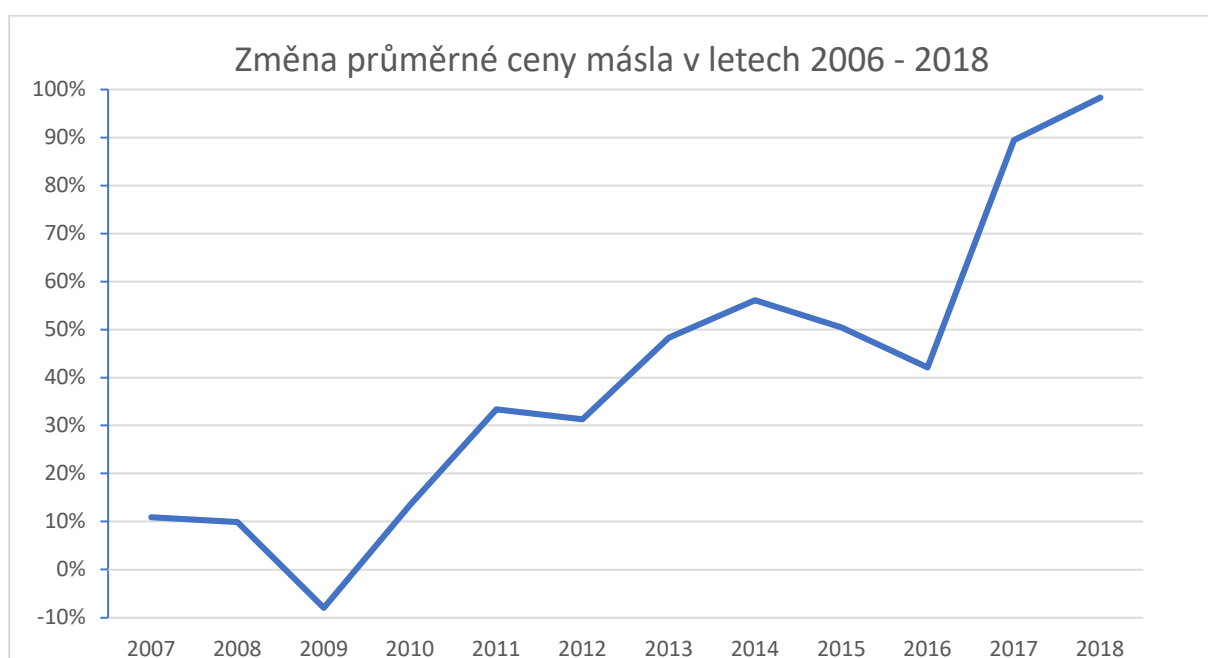
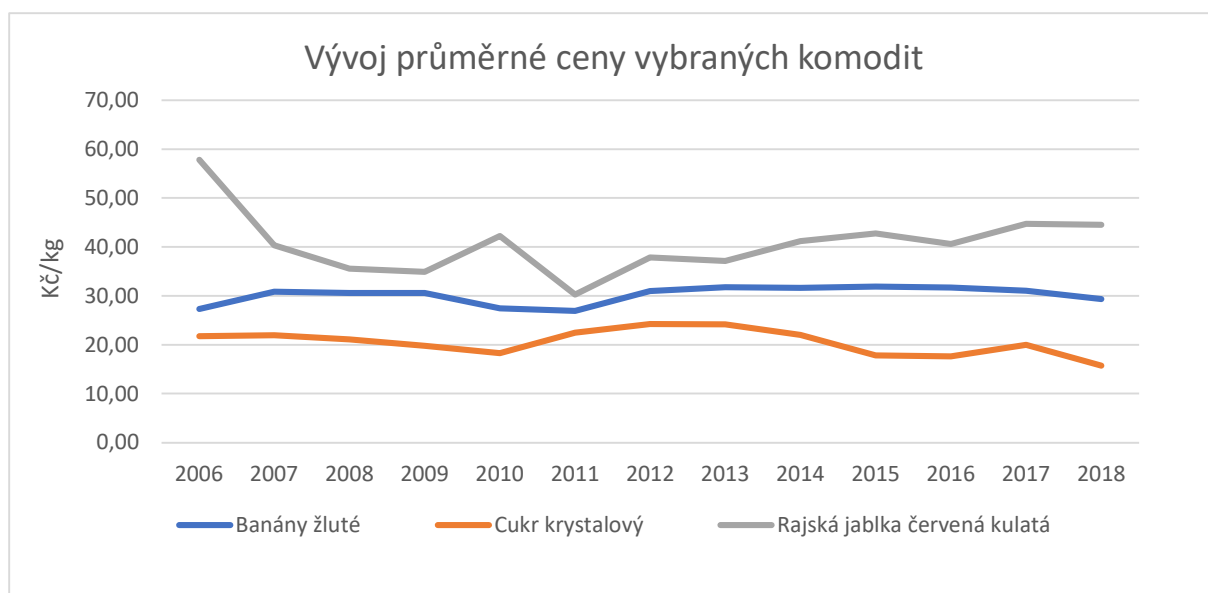
- Naopak největší nárůst nakupovaného množství – tentokrát mléka – je v odvětví **Zdravotní a sociální péče, Informačních a komunikačních činnostech, Zpracovatelském průmyslu**. Vysvětlení je obdobné jako v předešlém bodě. K nárůstu došlo především z důvodu největšího meziročního nárůstu mezd v těchto odvětvích. Největší nárůst kupovaného mléka o 390 ks je ve Zdravotní a sociální péči. Za zmínku stojí také odvětví **Informační a komunikační činnosti**, kde došlo k druhému největšímu nárůstu nakupovaného množství mléka o 383 ks. Jak je vidět v grafu, v nárůstu mezd je toto odvětví až na 8. místě. Vysoký nárůst množství je však také ovlivněn nominální výší mzdy, za kterou je možné komoditu nakoupit. Toto odvětví patří k odvětvím s nejvyšší nominální mzdou viz druhý graf Výše průměrné mzdy v letech 2006 a 2018.





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

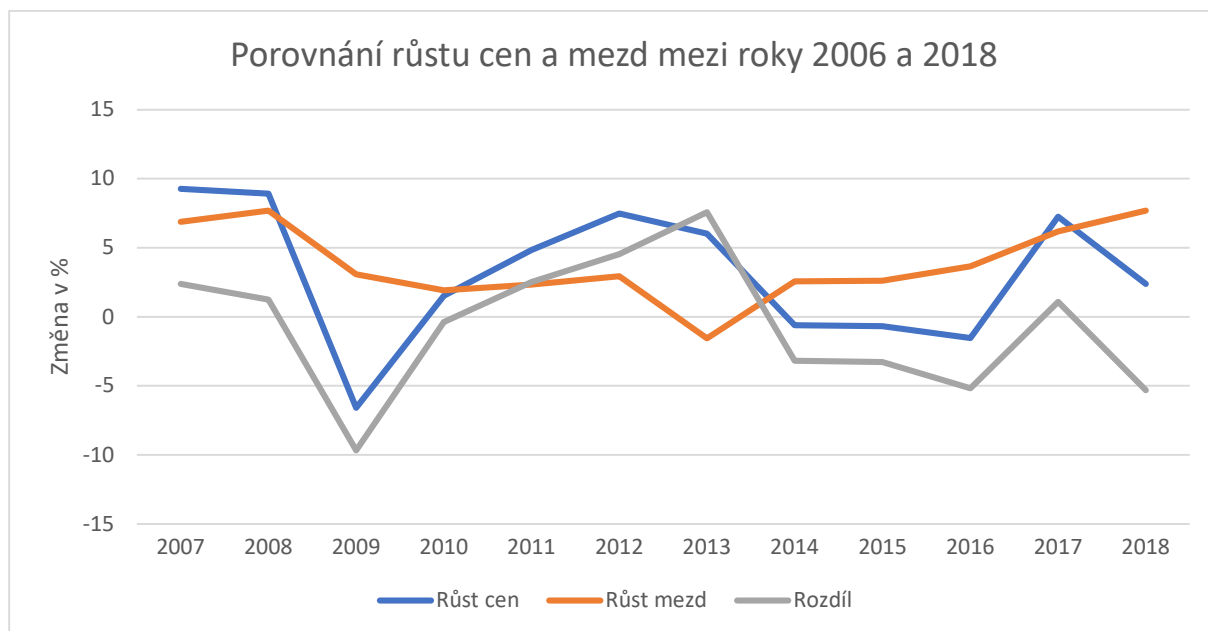
- I v případě této otázky vycházíme z primární tabulky. Z té jsou ve výsledné tabulce vyexportovány jednotlivé komodity s údaji o průměrných cenách v roce 2006 a 2018. Pro zjednodušení výpočtu budeme počítat meziroční nárůst ceny pouze jako poměr mezi roky 2006 a 2018, nikoli jako průměr meziročních přírůstků v jednotlivých letech. Výsledná data jsou totiž srovnatelná, budeme tedy pracovat pouze se zjednodušeným porovnáním.
- V porovnání je kalkulováno s nominálními cenami, nikoli však s reálnými. Ceny tak nezohledňují přepočty na současnou reálnou hodnotu což může ovlivnit výsledky.
- Z výsledné tabulky je zřejmé, že k nejmenšímu nárůstu cen dochází u komodit **Cukr krystalový, Rajská jablka červená kulatá, Banány žluté**. Z grafu níže můžeme vyčíst, že cena **Cukru krystalového** v průběhu let dokonce klesá.
- Naopak komodita s nejvyšším nárůstem ceny je **Máslo**. Zde došlo v průběhu sledovaného období k nárůstu téměř o 100 %.



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

- V tomto případě budeme v SQL vytvářet VIEW, jelikož výsledná data budeme využívat i při zodpovězení otázky č. 5.
- Výsledná tabulka vychází opět z tabulky primární, kde vypočítáme průměrnou změnu cen a průměrnou změnu mezd v daném roce.
- Abychom zajistili porovnatelnost dat o cenách je průměrný růst cen počítán jako průměr z průměrného růstu jednotlivých komodit.

- Jak je vidět z výsledné tabulky i z grafu níže, z námi dostupných dat vyplývá, že neexistuje rok, ve kterém by byl růst cen potravin větší než 10 % oproti růstu mezd. Můžeme však konstatovat, že největší rozdíl mezi cenami potravin a mzdami je v roce 2013 a to 7,6 %. Propad průměrných mezd v roce 2013 jde vysvětlit důvodem z první výzkumné otázky, tj. předčasná výplata ročních odměn již v roce 2012.



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?

- Podklad v SQL si vytvoříme tak, že si ze sekundární tabulky vyfiltrujeme údaje o HDP za Českou republiku za období 2006–2018. K těmto údajům si přihrajeme data z pohledu vytvořeného ve čtvrtém úkolu (v_price_payroll_growth). Díky tomuto spojení získáme informace o vývoji HDP, cenách a mzdách v daném období.
- Podíváme-li se na řešení otázky z makroekonomického pohledu, můžeme konstatovat, že v případě, že roste HDP zákonitě musí růst i mzdy a růst cenová hladina. Toto však nelze potvrdit ani vyvrátit z námi dostupných dat.
- Jak je vidět z grafu níže v roce 2010 došlo k výraznějšímu růstu HDP a to ovlivnilo růst cen, ve mzdách však k žádnému výraznému navýšení nedošlo ani v roce 2010 ani v roce 2011. K dalšímu výraznějšímu růstu HDP došlo v roce 2015 (5,39 %), kdy cenová hladina klesala a mzdy mírně vzrostly. V roce 2017 došlo opět k růstu HDP (5,17 %) a tentokrát došlo i k růstu cen (7,24 %) a růstu mezd (6,17 %). Z tohoto zjištění bohužel nemůžeme potvrdit, že pokud výrazněji roste HDP, tak ceny potravin i mzdy taktéž rostou ať už ve stejném nebo následujícím roce.

