

Průvodní zpráva – projekt SQL

Milan Sotona, Discord: misot#3106, milan.sotona@email.cz

Vstupní 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.
8. `countries` – Informace o zemích na světě, například hlavní město, měna, národní jídlo nebo průměrná výška populace.
9. `economies` – HDP, GINI, daňová zátěž atd. pro daný stát a rok.

A) Vytvoření tabulky SQL primary final

`t_milan_sotona_project_SQL_primary_final`

Tabulka obsahuje data mezd a cen potravin za Českou republiku sjednocených na totožné porovnatelné období – společné roky. Dále přidány údaje o Hrubém domácím produktu (HDP) pro zpracování otázky č. 5.

Porovnatelné období: 2006 až 2018 viz níže

Data o mzdách

Tabulka 1 + k ní číselníky v tabulkách 2 až 5, data za období od 2000, do Portálu otevřených dat ČR jsou průběžně doplňována data za čtvrtletí, ke dni zpracování tohoto projektu je poslední uloženo 4. čtvrtletí 2022.

Zdroj dat:

<https://data.gov.cz/datov%C3%A1-sada?iri=https%3A%2F%2Fdata.gov.cz%2Fzdroj%2Fdatov%C3%A1-sady%2F00025593%2F7ddcb833bddeeb84db39004d7e276b87>

Pro vytvoření výstupních dat jsou použity mzdy podle přepočteného stavu zaměstnanců, ne podle fyzického počtu zaměstnanců. Použití přepočteného stavu zaměstnanců je nutné z důvodu srovnatelnosti.

Obecně ke čtvrtletním údajům o mzdách: Z náhledu na data jednotlivých odvětví je zřejmé, že ve 4. čtvrtletí je většinou vyšší průměrná mzda než v 1. - 3. čtvrtletí následujícího roku. To je vysvětlitelné tím, že ve 4. čtvrtletí jsou roční premie (13. plat). Proto při hodnocení podle let je nutné brát pouze roky, které mají data za všechna 4 čtvrtletí, tj. zejména vyloučit poslední rok, pokud nemá údaje za všechna čtvrtletí.

U tohoto projektu je požadavek na porovnatelné období pro data mezd a potravin, takže ve výstupních datech jsou pouze roky 2006 až 2018. Při analýze pouze mezd by mohly být hodnoceny roky 2000 až 2022.

Data o potravinách

Tabulka 6 + číselník v tabulce 7, data jsou za období 2006 až 2018 (ke dni zpracování projektu na Portálu otevřených dat ČR jsou poslední údaje za rok 2018). Od roku 2019 má ČSÚ jinou strukturu zjišťování cen potravin (méně kategorií potravin) – viz web ČSÚ.

Zdroj dat:

<https://data.gov.cz/datov%C3%A1-sada?iri=https%3A%2F%2Fdata.gov.cz%2Fzdroj%2Fdatov%C3%A9-sady%2F00025593%2F3b940c4a1a26578a743d2f650eb6d2f4>

V datech je 27 kategorií potravin, z toho 26 kategorií má údaje za období 2006 až 2018 a kategorie 212101 Jakostní víno bílé má údaje za období 2015 až 2018. Ve zpracování projektu je tato kategorie vyloučena z důvodu nesrovnatelnosti dat za časové rozpětí. V SQL dotazech lze tuto kategorii do výběru zařadit (vypustit WHERE...).

Data o hrubém domácím produktu

Tabulka 8 + 9, vybraný údaje pro období 2006 až 2018 pro státy Evropy.

Zdroj dat: obdrženo v rámci projektu

B) Vytvoření tabulky SQL_secondary_final

t_milan_sotona_project_SQL_secondary_final

Vstupní tabulky:

Countries

Pro každou zemi 1 řádek

Economies

Vybraný údaje pro každou zemi za roky 2006 až 2018.

Údaj GDP (tj. HDP): předpoklad, že se jedná o USD (celosvětové údaje jsou nejčastěji v USD).

Pro zpracování otázky č. 5 jsou použity hodnoty uvedené v dodaných tabulkách. Neurčení měny GDP nemá na zodpovězení otázky č. 5 vliv – pro trend v čase není měna rozhodující.

OTÁZKY

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

Vstupní data obsahují údaje za roky 2006 až 2018 pro 19 odvětví (podle rozdělení Českého statistického úřadu). Rok 2006 se bere jako výchozí, tj. mzdy na úrovni 100 %.

V přehledu jsou použity mzdy pro přepočtený počet zaměstnanců (nikoliv pro fyzický).

V rekapitulační tabulce jsou Kč nárůstu nebo poklesu součty změny průměrné mzdy oproti minulému roku u všech 19 odvětví českého hospodářství. Údaj nemá vazbu na celkovou průměrnou mzdu, je uvedený z důvodu vyjádření velikosti změny mezd oproti minulému roku.

Závěr:

Z přehledu je zřejmé, že při prvním ekonomickém oslabení s minimem v roce 2009 nedošlo k významnějšímu poklesu mezd. Podstatná změna nastala až při druhém oslabení ekonomiky s minimem v roce 2013.

rok	nárůst	pokles	Kč nárůst	Kč pokles	Kč celkem
2007	19	0	27 659	0	27 659
2008	19	0	33 067	0	33 067
2009	15	4	15 476	-1 253	14 223
2010	16	3	9 799	-673	9 126
2011	15	4	12 130	-782	11 348
2012	19	0	14 585	0	14 585
2013	8	11	2 618	-10 587	-7 969
2014	18	1	13 164	-185	12 979
2015	18	1	14 069	-641	13 428
2016	18	1	19 482	-183	19 299
2017	19	0	33 917	0	33 917
2018	19	0	44 938	0	44 938

viz dotaz_1.sql

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

potravina	první rok	množství		množství posl. rok	změna %
		první rok	poslední rok		
Chléb konzumní kmínový 1 kg	2006	1 313	2018	1 365	3,98
Mléko polotučné pasterované 1 l	2006	1 466	2018	1 670	13,90

viz dotaz_2.sql

Ze zadání nevyplývá požadavek na rozdělení podle jednotlivých odvětví.

V případě upřesnění požadavku je připraven rozbor podle jednotlivých odvětví (dotaz_2.sql).

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

V datech od Českého statistického úřadu za období 2006 až 2018 je 27 položek potravin. Jakostní víno bílé má údaje pouze za roky 2015 až 2018. Proto je vyloučeno ze zpracování z důvodu nesrovnatelnosti období. Celkem je tedy ve výstupní tabulce 26 položek, které jsou seřazeny od nejnižší po nejvyšší změnu ceny.

potravina	cena 2006	cena 2018	% změna
Cukr krystalový 1 kg	21,73	15,75	-27,52
Rajská jablka červená kulatá 1 kg	57,83	44,49	-23,07
Banány žluté 1 kg	27,31	29,32	7,36
Vepřová pečeně s kostí 1 kg	105,18	116,85	11,10
Přírodní minerální voda uhlíčitá 1 l	7,69	8,65	12,48
Pečivo pšeničné bílé 1 kg	38,60	43,84	13,58
Jablka konzumní 1 kg	30,71	36,18	17,81
Šunkový salám 1 kg	116,77	144,88	24,07
Konzumní brambory 1 kg	12,08	15,08	24,83
Eidamská cihla 1 kg	110,95	142,45	28,39
Hovězí maso zadní bez kosti 1 kg	166,34	223,26	34,22
Kapr živý 1 kg	69,35	93,46	34,77
Mléko polotučné pasterované 1 l	14,44	19,82	37,26
Pivo výčepní, světlé, lahvé 0,5 l	8,45	11,81	39,76
Rostlinný roztíratelný tuk 1 kg	69,45	99,40	43,12
Kuřata kuchařská celá 1 kg	47,47	69,31	46,01
Pomeranče 1 kg	24,73	36,50	47,59
Chléb konzumní kmínový 1 kg	16,12	24,24	50,37
Pšeničná mouka hladká 1 kg	7,41	11,44	54,39
Mrkev 1 kg	14,41	22,45	55,79
Jogurt bílý netučný 150 g	5,83	9,17	57,29
Vejce slepičí čerstvá 10 ks	23,49	38,39	63,43
Rýže loupaná dlouhozrná 1 kg	21,29	36,18	69,94
Papriky 1 kg	35,31	60,47	71,25
Těstoviny vaječné 1 kg	26,10	47,88	83,45
Máslo 1 kg	104,39	207,08	98,37

viz dotaz_3.sql

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

Tuto otázku by bylo potřeba přesněji definovat.

Pokud bereme rozdíl mezi procentní změnou u cen potravin a mezd nedošlo ke kumulativnímu rozdílu nad 10 %. Největší nárůst cen potravin vůči změně mezd byl v roce 2013, kdy ceny vzrostly o 5,1 %, ale mzdy klesly o 1,6 %, tj. celkový rozdíl byl 6,7 %.

Agregované údaje potvrzují zjištění u ostatních otázek: první pokles ekonomiky s minimem v roce 2009 nezapůsobil na pokles mezd. Teprve druhý pokles s minimem v roce 2013 měl vliv na pokles mezd.

Ze zpracování je vyloučena položka Jakostní víno bílé, data jsou k dispozici jen za roky 2015-2018.

rok	průměr cen	% změny cen	průměr mezd	% změny mezd	% změn ceny a mzdy
2006	45,52		21 165		
2007	48,59	6,7	22 621	6,9	-0,2
2008	51,60	6,2	24 361	7,7	-1,5
2009	48,29	-6,4	25 110	3,1	-9,5
2010	49,23	1,9	25 590	1,9	-0,0
2011	50,88	3,4	26 188	2,3	1,1
2012	54,30	6,7	26 955	2,9	3,8
2013	57,07	5,1	26 536	-1,6	6,7
2014	57,49	0,7	27 219	2,6	-1,9
2015	55,82	-2,9	27 926	2,6	-5,5
2016	55,03	-1,4	28 941	3,6	-5,0
2017	60,60	10,1	30 726	6,2	3,9
2018	61,86	2,1	33 092	7,7	-5,6

viz dotaz_4.sql

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?

Ze zpracovaných dat lze odvodit následující:

V období 2006 až 2018 došlo ke dvěma propadům HDP. První s minimem v 2009 s následujícím mírným oživením a druhý propad v 2013 s následujícím růstem až do 2018.

Mzdy

Růst mezd má mnohem větší setrvačnost než změna u HDP. V roce 2009 byl největší propad HDP, ale změna mezd zdaleka nereagovala poklesem, který by měl úměrnou vazbu na pokles HDP. Teprve při druhém poklesu s minimem v roce 2013 došlo k poklesu mezd. Uvedené je zřejmé z grafu 1.

Nad rámec dotazu je velmi významné se podívat na kumulativní vývoj růstu mezd v porovnání s růstem HDP. Je zcela evidentní, že od roku 2007 roste disproporce mezi růstem mezd a růstem HDP a to negativním směrem – růst mezd je čím dál rychlejší. Viz graf 2.

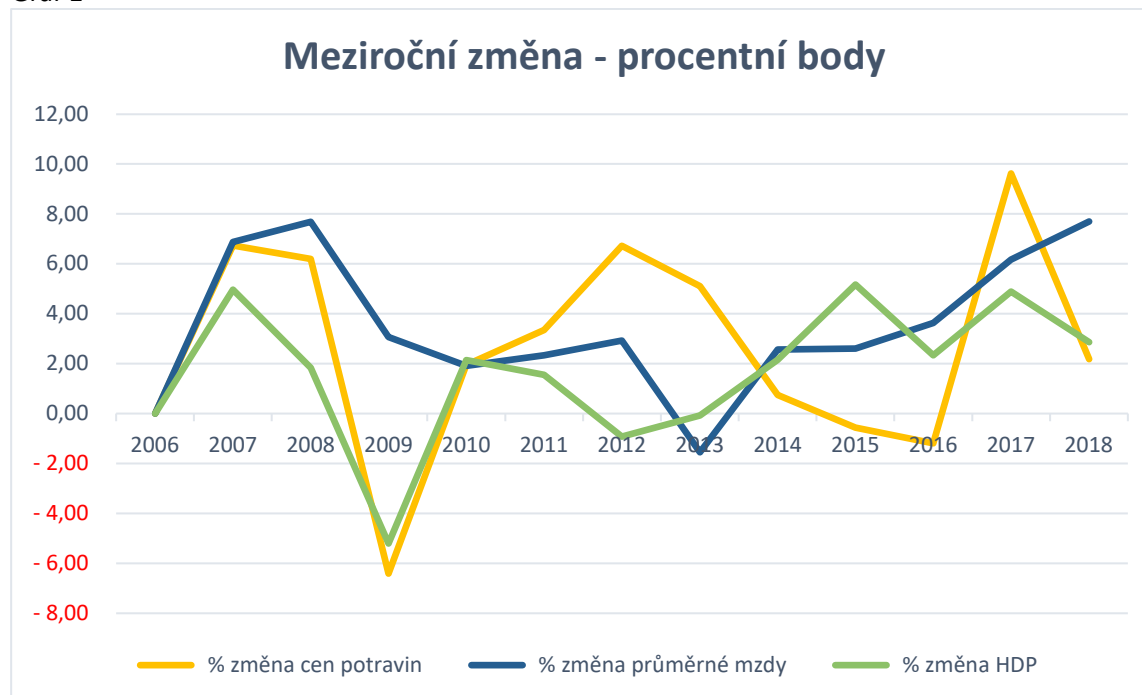
Tj. loď se sice nepotápí, ale orchestr by neměl hrát tak vesele.

Potraviny

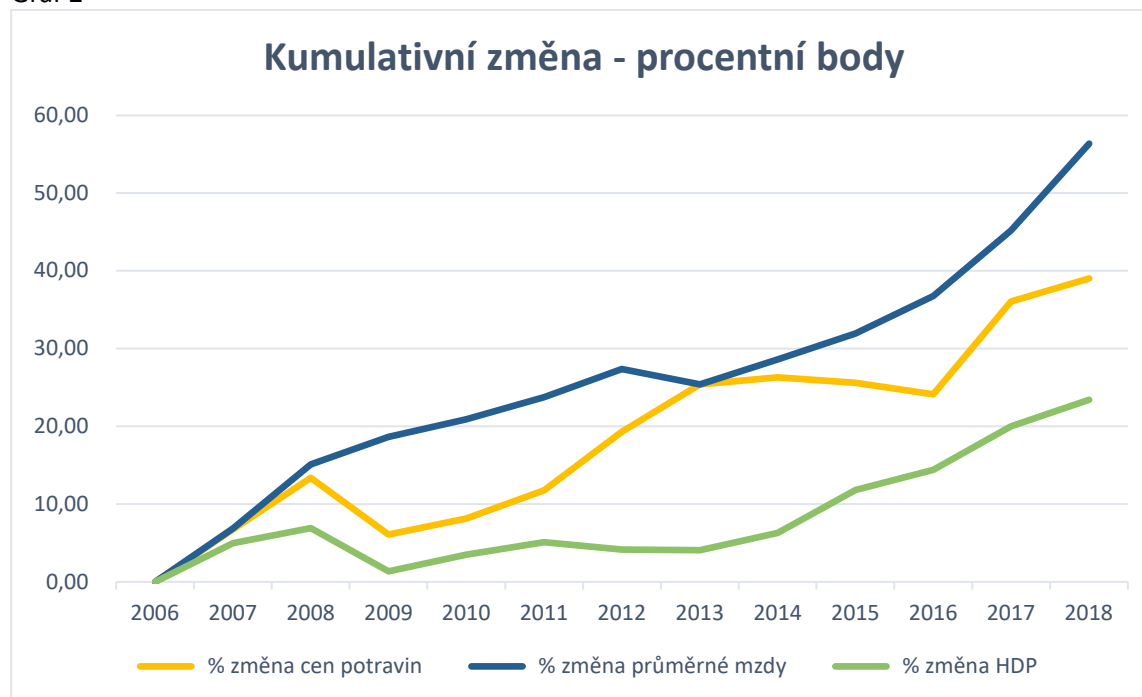
Při prvním poklesu do minima v roce 2009 změna ceny potravin kopírovala změnu HDP. V roce 2010 byl předpoklad konce hospodářské recese a očekával se opětovný růst ekonomiky. Cena potravin rostla až do roku 2012, kdy zřejmý pokles HDP ovlivnil i změnu cenu potravin, které spadly do svého minima (záporné změny) v letech 2015 a 2016 (graf 1).

Obecně lze konstatovat, že změna HDP má rychlejší vliv na změnu ceny potravin než na mzdy. U mezd je zřejmé, že se těžko snižují i při poklesu HDP.

Graf 1



Graf 2



Vstupní data pro grafy:

rok	průměrné ceny potravin Kč	% změna cen potravin	% kumul. změny cen potravin	průměrná mzda Kč	% změna průměrné mzdy	% kumul. změny mzdy	HDP na obyvatele USD	% změna HDP	kumul. % změna HDP
2006	45,52	0,00	0,00	21 165	0,00	0,00	19 286	0,00	0,00
2007	48,59	6,74	6,74	22 621	6,88	6,88	20 242	4,96	4,96
2008	51,60	6,19	13,36	24 361	7,69	15,10	20 614	1,84	6,89
2009	48,29	-6,41	6,09	25 110	3,07	18,64	19 542	-5,20	1,33
2010	49,23	1,95	8,15	25 590	1,91	20,91	19 960	2,14	3,49
2011	50,88	3,35	11,78	26 188	2,34	23,73	20 269	1,55	5,10
2012	54,30	6,72	19,29	26 955	2,93	27,36	20 082	-0,92	4,13
2013	57,07	5,10	25,37	26 536	-1,55	25,38	20 066	-0,08	4,04
2014	57,49	0,74	26,30	27 219	2,57	28,60	20 499	2,16	6,29
2015	57,17	-0,56	25,59	27 926	2,60	31,94	21 561	5,18	11,80
2016	56,49	-1,19	24,10	28 941	3,63	36,74	22 065	2,34	14,41
2017	61,93	9,63	36,05	30 726	6,17	45,17	23 144	4,89	20,00
2018	63,28	2,18	39,02	33 092	7,70	56,35	23 805	2,86	23,43

viz dotaz_5.sql

HDP na obyvatele: hodnota HDP(GDP) vydělená počtem obyvatel (population) v tabulce economies. U údaje GDP není zřejmé, o jakou měnu se jedná, při přepočtu na USD i EUR při průměrném ročním kurzu těchto měn podle České národní banky neodpovídá HDP na údaje od Českého statistického úřadu. Pro potřeby % změn toto nemá vliv. V případě potřeby by bylo nutné dohledat, jak byla hodnota GDP určena a v jaké je měně. Předpoklad: HDP je v USD.