

Project_SQL

1. Zadání projektu

Na našem analytickém oddělení firmy, která se zabývá životní úrovní občanů, jsme se dohodli, že se pokusíme 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.

Na konferenci bude potřeba nachystat datové podklady. Z nich bude patrné porovnání dostupnosti potravin na základě průměrných příjmů za určité časové období.

Dále připravujeme dodatečné podklady, jako je tabulka s HDP, GINI koeficientem a populací dalších evropských států ve stejném období jako primární přehled pro ČR.

Datové podklady obsahují dvě tabulky a odpovědi na 5 otázek.

2. První část projektu – 2 tabulky

V tomto projektu byly vytvořeny dvě tabulky.

První tabulka (`t_{Anezka}_{Kinclova}_project_SQL_primary_final`) shrnuje data z ČR o cenách potravin a mezd v letech 2006 – 2018.

Druhá tabulka (`t_{Anezka}_{Kinclova}_project_SQL_secondary_final`) je ze stejného časového období (2006 – 2018). V této tabulce najdete informace o evropských zemích, které jsou rozděleny dle geopolitických částí:

- Central Europe and the Baltics
- Euro area
- Europe & Central Asia
- Europe & Central Asia (excluding high income)
- Europe & Central Asia (IDA & IBRD countries)
- European Union

V této tabulce jsou informace o HDP, populaci a koeficientu gini v jednotlivých částech Evropy.

Tyto dvě tabulky posloužily jako podklad pro získání detailnějších informací o ČR i Evropě.

3. Druhá část projektu – odpovědi na otázky

V další části projektu jsme se zabývali těmito otázkami:

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?

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

K dopovězení první otázky (Rostou v průběhu let mzdy ve všech odvětvích, nebo v některých klesají?) nám poslouží tabulka č. 1

year	name	avg_salary	prev_avg_salary	difference_salary
2009	Těžba a dobývání	28 200 Kč	29 300 Kč	-1 100 Kč
2009	Zemědělství, lesnictví, rybářství	17 400 Kč	17 500 Kč	-100 Kč
2009	Ubytování, stravování a pohostinství	12 100 Kč	12 200 Kč	-100 Kč
2010	Veřejná správa a obrana; povinné sociální zabezpečení	26 700 Kč	26 800 Kč	-100 Kč
2010	Profesní, vědecké a technické činnosti	30 800 Kč	31 000 Kč	-200 Kč
2010	Vzdělávání	22 000 Kč	22 400 Kč	-400 Kč
2011	Veřejná správa a obrana; povinné sociální zabezpečení	26 100 Kč	26 700 Kč	-600 Kč
2011	Ubytování, stravování a pohostinství	12 800 Kč	12 900 Kč	-100 Kč
2013	Kulturní, zábavní a rekreační činnosti	19 900 Kč	20 200 Kč	-300 Kč
2013	Velkoobchod a maloobchod; opravy a údržba motorových vozidel	22 700 Kč	22 900 Kč	-200 Kč
2013	Těžba a dobývání	31 600 Kč	32 500 Kč	-900 Kč
2013	Profesní, vědecké a technické činnosti	31 000 Kč	32 000 Kč	-1 000 Kč
2013	Administrativní a podpůrné činnosti	16 400 Kč	16 500 Kč	-100 Kč
2013	Zásobování vodou; činnosti související s odpady a sanacemi	23 400 Kč	23 500 Kč	-100 Kč
2013	Informační a komunikační činnosti	45 700 Kč	46 200 Kč	-500 Kč
2013	Výroba a rozvod elektřiny, plynu, tepla a klimatiz. vzduchu	40 500 Kč	42 400 Kč	-1 900 Kč
2013	Peněžnictví a pojišťovnictví	45 800 Kč	50 300 Kč	-4 500 Kč
2013	Činnosti v oblasti nemovitostí	21 600 Kč	22 000 Kč	-400 Kč
2013	Stavebnictví	22 200 Kč	22 700 Kč	-500 Kč
2014	Těžba a dobývání	31 300 Kč	31 600 Kč	-300 Kč
2015	Výroba a rozvod elektřiny, plynu, tepla a klimatiz. vzduchu	40 500 Kč	41 000 Kč	-500 Kč
2016	Těžba a dobývání	31 500 Kč	31 700 Kč	-200 Kč

Tabulka č.1 Klesající mzdy v letech 2006 – 2018

V této tabulce můžeme vidět, ve kterých odvětvích klesly ceny mezd v letech 2006 – 2018. Nejhuře na tom bylo odvětví Peněžnictví a pojišťovnictví v roce 2013, kdy rozdíl mezi roky 2012 a 2013 je 4500.

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

Druhá věc, která nás zajímala bylo, kolik je možné si koupit litrů mléka a kilogramů chleba za první a poslední srovnatelné období v dostupných datech cen a mezd? Odpověď na tuto otázku se nám zobrazí v tabulce číslo 2.

year	food	value	unit	price_food	avg_salary	quantity_food
2006	Chléb konzumní kmínový	1,0	kg	16,94 Kč	20 754 Kč	1225 ks
2018	Chléb konzumní kmínový	1,0	kg	24,31 Kč	32 536 Kč	1338 ks
2006	Mléko polotučné pasterované	1,0	l	14,04 Kč	20 754 Kč	1478 ks
2018	Mléko polotučné pasterované	1,0	l	19,15 Kč	32 536 Kč	1699 ks

Tabulka č.2 Počet litrů mléka a kilogramů chleba v letech 2006 – 2018

Z druhé tabulky vidíme, že v roce 2018 bychom koupili více litrů mléka i více kilogramů chleba. Hodnoty jsou počítány tak, že v jednotlivém roce nakupují pouze chleba nebo pouze mléko.

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

Ve třetí otázce - Která kategorie potravin zdražuje nejpomaleji (je u ní nejnižší procentuální meziroční nárůst)? – jsme dosáhly těchto výsledků, které jsou shrnuty v tabulce číslo 3.

food	year_food	value	unit	price	before_price	prc_year_food
Cukr krystalový	2009	1.0	kg	19,49 Kč	19,49 Kč	0,00%
Chléb konzumní kmínový	2018	1.0	kg	24,31 Kč	24,31 Kč	0,00%
Pěčivo pšeničné bílé	2012	1.0	kg	45,78 Kč	45,78 Kč	0,00%
Pěčivo pšeničné bílé	2018	1.0	kg	45,15 Kč	45,15 Kč	0,00%

Tabulka č.3 Potraviny, které zdražují nejpomaleji v letech 2006 – 2018

Jak je z tabulky č. 3 patrné, tak nejpomaleji se zdražoval cukr krystal, chléb bílý konzumní a pečivo pšeničné bílé. Rozdíl cen je vždy brán z daného roku, který je uveden v tabulce a z roku předešlého. Tyto tři suroviny se ve dvou letech po sobě vůbec nezdražovaly.

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

Předposlední dotaz směřuje k porovnání nárůstu cen potravin a mezd - Existuje rok, ve kterém byl meziroční nárůst cen potravin výrazně vyšší než růst mezd (větší než 10 %)? Výsledky jsou prezentovány v tabulce č. 4

year	average_salary	prc_avg_salary	price	prc_price	differ_prc	price_vs_salary
2006	20 754 Kč		45 Kč			
2007	22 173 Kč	6,84%	47 Kč	6,34%	0,50%	Růst cen v normě
2008	23 918 Kč	7,87%	50 Kč	6,41%	1,46%	Růst cen v normě
2009	24 674 Kč	3,16%	47 Kč	-6,80%	9,96%	Růst cen v normě
2010	25 156 Kč	1,95%	48 Kč	1,77%	0,18%	Růst cen v normě
2011	25 736 Kč	2,30%	50 Kč	3,50%	-1,20%	Růst cen v normě
2012	26 516 Kč	3,03%	53 Kč	6,92%	-3,89%	Růst cen v normě
2013	26 103 Kč	-1,56%	56 Kč	5,55%	-7,11%	Růst cen v normě
2014	26 771 Kč	2,56%	56 Kč	0,89%	1,67%	Růst cen v normě
2015	27 443 Kč	2,51%	56 Kč	-0,56%	3,07%	Růst cen v normě
2016	28 444 Kč	3,65%	55 Kč	-1,12%	4,77%	Růst cen v normě
2017	30 231 Kč	6,28%	61 Kč	9,98%	-3,70%	Růst cen v normě
2018	32 536 Kč	7,62%	62 Kč	1,84%	5,78%	Růst cen v normě

Tabulka č.4 Srovnání růstu cen mezd a potravin v letech 2006 – 2018

V žádném roce nebyl rozdíl cen mezd a potravin překročen o deset a více procent. Vždy byl růst cen mezd a potravin v normě. Zelené čísla ve sloupci differ_prc nám určují, že růst cen mezd byl větší, jak potravin a červená čísla ve stejném sloupci nám udávají vyšší růst cen potravin.

3.5. 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?

Poslední dotaz se týkal i informací z druhé tabulky (t_{Anezka}_{Kinclova}_project_SQL_secondary_final), ve které se srovnávalo HDP evropských zemí a změny ve mzdách a cenách potravin. Výsledky jsou prezentovány v tabulce č. 5. Hodnoty avg_GDP je průměr HDP geopolitického celku s názvem European Union – teda Evropská unie – ve které se ČR nachází. Proto jsem zvolila právě tento celek.

year	avg_GDP	prc_GDP	avg_salary	prc_avg_salary	price	prc_price	GDP	salary	price
2006	14 299 677 831 554 Kč		20 754 Kč		45 Kč				
2007	14 748 189 125 973 Kč	3,14%	22 173 Kč	6,84%	47 Kč	6,35%	roste	roste	roste
2008	14 843 586 065 671 Kč	0,65%	23 918 Kč	7,87%	50 Kč	6,41%	roste	roste	roste
2009	14 205 575 522 837 Kč	-4,30%	24 674 Kč	3,16%	47 Kč	-6,80%	klesá	roste	klesá
2010	14 518 958 244 622 Kč	2,21%	25 156 Kč	1,95%	48 Kč	1,77%	roste	roste	roste
2011	14 781 071 268 002 Kč	1,81%	25 736 Kč	2,30%	50 Kč	3,50%	roste	roste	roste
2012	14 669 122 338 810 Kč	-0,76%	26 516 Kč	3,03%	53 Kč	6,92%	klesá	roste	roste
2013	14 661 016 876 632 Kč	-0,06%	26 103 Kč	-1,56%	56 Kč	5,55%	klesá	klesá	roste
2014	14 890 776 587 454 Kč	1,57%	26 771 Kč	2,56%	56 Kč	0,89%	roste	roste	roste
2015	15 233 927 171 587 Kč	2,30%	27 443 Kč	2,51%	56 Kč	-0,56%	roste	roste	klesá
2016	15 539 223 069 142 Kč	2,00%	28 444 Kč	3,65%	55 Kč	-1,12%	roste	roste	klesá
2017	15 973 737 289 265 Kč	2,80%	30 231 Kč	6,28%	61 Kč	9,98%	roste	roste	roste
2018	16 310 490 150 108 Kč	2,11%	32 536 Kč	7,62%	62 Kč	1,94%	roste	roste	roste

Tabulka č.5 Vliv HDP evropských zemí na růst cen potravin a mezd v letech 2006 – 2018

Z tabulky vyplývá, že růst nebo pokles HDP evropských zemí nemá vliv na růst či pokles cen potravin nebo mezd. Nezávisle na sobě se ve sloupcích GDP, salary a price střídá „roste“ a „klesá“. Z toho usuzuji, že jsou hodnoty na sobě nezávislé.

4. Závěr

Datové sady k tomuto projektu byly kvantitou dat dobré. Jen v datové sadě pro druhou tabulku (t_{Anezka}_{Kinclova}_project_SQL_secondary_final) chybí gini koeficient. Je potřeba tento koeficient do sady doplnit. V mém projektu to zase až tak velký problém nebyl. Nepotřebovalo jsem jej pro další části projektu.

Ještě byl menší problém s roky v datové sadě `czechia_price`, která je vedená od roku 2006. A protože jsem měla porovnávat roky 2006 až 2018, tak k porovnání roku 2006 nám chybí (díky `czechia_price`) rok 2005. Proto v některých částech tabulek jsou prázdné buňky. Popř. mohu srovnávat od roku 2007, kde už data budou všechny.