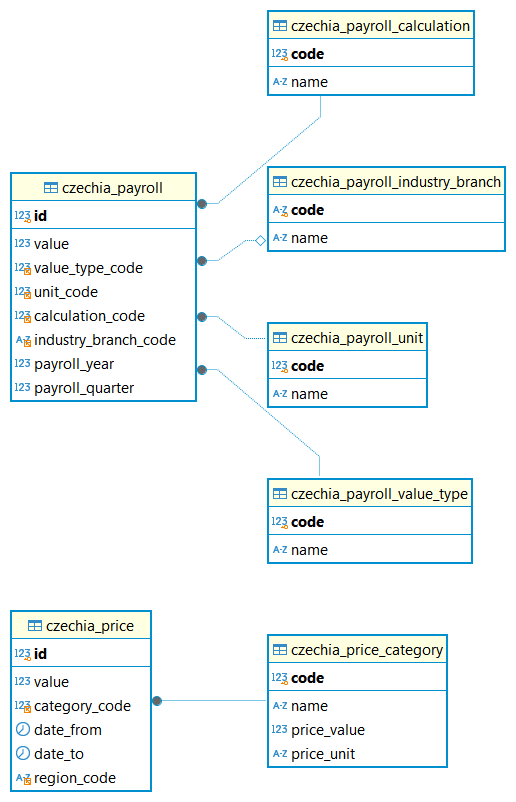
Průvodní listina k SQL projektu

Tento dokument slouží jako průvodní listina k projektu z datové akademie Engeto.

Nejprve vytvoříme souhrnnou tabulku (t\_robert\_zunt\_project\_SQL\_primary\_final), kterou použijeme jako datový zdroj k dalším úkolům.



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

Záleží v jakém horizontu na tuto otázku chceme odpovědět. Pokud bude porovnávat trend mezd (increase/decrease/no change) mezi průměrnou roční mzdou v jednotlivých letech a pro každé odvětví, tak dojdeme k zjištění, že v některých odvětvích došlo i k meziročnímu poklesu. Pro ukázání těchto dat lze použít první část skriptu První\_dotaz.sql

Pokud se na otázku pokusíme odpovědět v horizontu porovnatelných dat (tabulka t\_robert\_zunt\_project\_SQL\_primary\_final obsahuje data pro období 2006-2018), tak můžeme vidět, že všechny odvětví měly nárůst průmětné roční mzdy. Pro tento pohled slouží druhá část skriptu První\_dotaz.sql.

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

*Pro odpověď slouží k skript druhy\_dotaz.sql*

Během prvotní agregace dat jsem zjistil, že porovnatelné období pro které jsou dostupné data o mzdách i potravinách je 2006-2018.

|  |  |  |
| --- | --- | --- |
| food\_name | quantity\_for\_2006 | Quantity\_for\_2018 |
| Chléb konzumní kmínový | 1284 | 1342 |
| Mléko polotučné pasterované | 1436 | 1641 |

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

*Pro odpověď slouží skript treti\_dotaz.sql*

Pro zjištění, která kategorie potravin zdražuje nejpomaleji jsem nejprve vypočítal meziroční rozdíl ceny pomocí LAG a dopočetl procenta. Pro odpověď na otázku jsem výsledky seřadil podle procentuálního rozdílu. V datech lze vidět, že existují potraviny, které v některých letech dokonce i zlevňovaly. Na otázku tedy nelze odpověď, že nějaká kategorie " meziročně zdražuje nejpomaleji. Pro zobrazení pouze těchto potravin a let, které zlevňovaly slouží druhá část skriptu.

Datový set neobsahuje informace o ceně Jakostního bílého vína pro celé období 2006-2018 ale pouze 2015-2018.

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

*Pro odpověď slouží skript ctvrty\_dotaz.sql*

Vytvořil jsem přehled zobrazující procentuální rozdíl průměrné ceny všech potravin pro sledované roky. Z dat lze vidět, že v žádném sledovaném roce nedošlo k nárůstu přes 10%.

|  |  |
| --- | --- |
| Year | Percentage difference |
| 2007 | 6.78 |
| 2008 | 6.21 |
| 2009 | -6.41 |
| 2010 | 1.93 |
| 2011 | 3.34 |
| 2012 | 6.74 |
| 2013 | 5.1 |
| 2014 | 0.72 |
| 2015 | -0.54 |
| 2016 | -1.18 |
| 2017 | 9.62 |
| 2018 | 2.18 |

## 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 odpověď slouží skript paty\_dotaz.sql*

Z dat nemůžeme jednoznačně říct, že změna HDP má vliv na změnu potravin a mezd. Např. v roce 2007 můžeme vidět nárůst ve všech 3 kategoriích, ale v roce 2015 už tento vzorec neplatí.