

## Tvorba tabulek

Při tvorbě tabulek jsem vycházela z výzkumných otázek a údajů, které budu potřebovat pro jejich vypracování. Primární tabulka uvádí ceny jednotlivých produktů, mzdy v jednotlivých odvětvích v jednotlivých letech a kód, podle kterého poznáme, že se jedná o pracující osoby (sloupec payroll\_code). Výchází pro mě byly údaje z tabulek czechia\_price a czechia\_payroll, které jsem propojila pomocí LEFT JOIN tak, aby byly zachovány i sloupce s nulovými hodnotami tabulek. Napojila jsem na ně i další tabulky, které obsahovaly údaje potřebné pro vypracování odpovědí na zadané otázky. Při propojování tabulek jsem vycházela zejména z ERD diagramů, kde jsou dobře vidět jednotlivé vazby mezi tabulkami. Tabulky czechia\_price a czechia\_payroll jsem propojila pomocí sloupce s rokem, kde jsem pomocí funkce YEAR musela upravit sloupec s datem z tabulky czechia\_price na roky tak, aby se údaje obou tabulek shodovali svým zápisem.

Při tvorbě druhé tabulky, která měla obsahovat údaje o jednotlivých zemích jsem se rozhodla propojit tabulky countries s tabulkou economies, protože při vypracování páté otázky budu potřebovat znát HDP, které je v tabulce economies uvedeno. Tabulky jsem opět propojila pomocí jediného společného sloupce, kterým je v tabulce countries rok získání nezávislosti v tabulce economies máme jednotlivé roky pro účely ekonomických dat také uvedené.

### 1. otázka

Při zodpovídání první otázky jsem vycházela z primární tabulky, kterou jsem vytvořila. Využila jsem podmínku CASE, kde jsem zadala dvě podmínky, na základě kterých zjistím, jestli mzdy rostou, klesají nebo stagnují, a přidala jsem nový sloupec payroll\_changes, ze kterého je možné tento údaj vyčíst.

Vzhledem k tomu, že jsem potřebovala pracovat pouze s daty v konkrétních odvětvích, vyfiltrovala jsem si v dotazu pomocí klauzule WHERE pouze údaje ze sloupce industry\_branch, kde je odvětví uvedeno. Klauzuli GROUP BY jsem pak využila k seskupení dat podle jednotlivých odvětví.

Výsledkem dotazu je tabulka, kde vidíme ze sloupce payroll\_changes, že mzdy v průběhu jednotlivých let v jednotlivých odvětvích mzdy stagnují.

### 2. otázka

U zodpovídání druhého dotazu jsem postupovala tak, že jsem chtěla vytvořit tabulku, ve které budu mít čtyři řádky. Chtěla jsem v každém řádku mít uvedenou minimální a maximální hodnotu, kterou si při daných cenách mléka a chleba lze za danou mzdu pořídit. Vytvořila jsem si čtyři dílčí dotazy, které jsem pak spojila pomocí klauzule UNION ALL. Použití mi přišlo vhodnější než výběr s použitím UNION, protože při výběru potraviny dojde k duplicitám v daném sloupci.

Při SELECTu jsem u jednotlivých sloupců s roky vždy vybírala minimální nebo maximální datum ve všech čtyřech dílčích příkazech. Do SELECTu sloupců jsem mimo datum zařadila i cenu potraviny, požadované potraviny a průměrnou hroubou mzdu u zaměstnanců. Přidala jsem vždy také sloupec, který mzdu vydělil průměrnou cenou požadované potraviny, abych získala informace o množství, které si lze za danou mzdu pořídit.

V každém dílčím dotazu jsem vždy zařadila dotaz WHERE, kde jsem vybrala požadovanou potravinu pomocí operátoru LIKE ze sloupce food\_category, abych dostala požadované hodnoty pouze pro chléb a mléko.

Výsledkem dotazu je, že za průměrnou hrubou mzdou koupíme v roce 2006 tj. v prvním srovnatelném období 924, 5 kg chleba a 1 126 l mléka, stejně tak i v roce 2018 tj. v posledním srovnatelném období.

### 3. otázka

Při řešení třetí otázky jsem potřebovala zjistit procentuální nárůst cen za potraviny vzhledem k cenám roku pro každou kategorii potraviny, tudíž jsem v SELECTu vybrala potravinu, její cenu, pomocí funkce YEAR jsem si upravila sloupec date\_from tak, aby zobrazoval pouze roky a pomocí funkce ROUND pro zaokrouhlení na 2 desetinná místa jsem si vypočítala procentuální nárůst jako rozdíl ceny současného a minulého roku vydělený cenou minulého roku násobeného 100. Následně jsem navázala tabulku samu na sebe, a pomocí SUBSELECTu jsem se rozhodla dopočítat a vytvořit sloupec s daty za minulý rok. Tabulku jsem provázela tedy prostřednictvím roku, kde zjistím maximální cenu potraviny za předchozí rok pro každou kategorii potraviny.

Nejpomaleji zdražuje hovězí maso zadní bez kosti.

### 4. otázka

Pro vyřešení této otázky jsem vycházela z primární tabulky. V SELECTu jsem si přidala dva sloupce, které počítají průměrné ceny potravin a mezd.

Vycházím z průměrných mezd zaměstnanců, které jsou určeny mzdovým kódem v klauzuli WHERE. Do klauzule jsme také přidala podmínku, že musí být ve sloupci se mzdou mzda vyplněna. Data jsem seskupila podle roku a kategorie potravin. Na závěr jsem pomocí HAVING přidala do dotazu podmínku, kde vypočítám z poměru průměrných cen a průměrných mezd, překračují 110 %.

Dotazem jsme žádná data nenašli, tudíž takový rok neexistuje.

### 5. otázka

Při řešení tohoto úkolu jsem spojovala primární a sekundární tabulku, které jsem si vytvořila, a kde byla uvedena výše HDP v jednotlivých letech.

Obě tabulky jsem propojila na základě sloupce s roky.

Klauzulí WHERE jsem si stanovila podmínku pro výběr dat, kde rozdíl mezi cenou jídla a mzdou přesahuje 10%, což znamená že buď je cena jídla o více než 10% vyšší než mzda nebo naopak.

Výška HDP na změny ve mzdách a cenách nemá vliv.