**PRŮVODNÍ DOKUMENTACE K SQL PROJEKTU**

31. 10. – vytvořen první select k projektu. Na úvod čerpáno z tabulky **covid19\_basic\_differences** a nyní zahájeno zkoumání dat z ostatních tabulek. Z tabulky countries bude potřeba pro vyřešení projektu získat kromě sloupce country také sloupce **population\_density** a **median\_age\_2018**.

Z tabulky economies jsou pro splnění zadání klíčové kromě primárního klíče (country) potřeba také informace o HDP na osobu, které získáme jako celkové **HDP dané země / počet obyvatel** a dále také **giniho koeficient** a **dětská úmrtnost**.

Pro zisk podílů jednotlivých náboženství je potřeba nejdříve převést údaje z tabulky religions na

Pro výpočet průměrné denní teploty je potřeba propojit tabulku **weather** s tabulkou **countries** na základě hlavního města, tak abychom ve výsledném seznamu dostali seřazené průměrné denní teploty podle našich hlavních klíčů **date** a **country**. Ne všechna hlavní města v tabulce countries jsou dostupná i v tabulce **weather**. V tabulce **countries** však některé země mají jiný název než v tabulce **covid19\_basic\_differences**. Bude potřeba postupovat přes propojení iso3 z lookup\_table.

Tabulka **religions** obsahuje větší množství zemí než je celkový počet zemí na světě, protože jako zemi bere např. i ostrovy patřící některým zemím.

Stejně tak jsou zde rozdíly mezi tabulkami Covid 19 basic differences a countries, kdy se v jednotlivých tabulkách některé země jmenují jinak, např. v Covid 19 basic differences se ČR jmenuje Czechia, zatímco v countries je to Czech Republic.

Při napojení tabulek countries a weather jsou viditelné rozdíly, kdy v countries chybí 11 ze 35 hlavních měst uvedených v tabulce weather. Tím pádem data za tyto země budou ve výsledné tabulce chybět. Celkově tady data o počasí po dnech jsou k dispozici jen u 24 zemí z výsledné tabulky.

Roční období jsem zakódoval jako (0=jaro, 1=léto, 2=podzim,3=zima).

Weather: **Truncated incorrect DOUBLE value: '1 °c'**

Pomocí describe, Engeto pokročilé konstrukty – projít tabulky, se kterými pracuji.

**INFORMACE O VÝSTUPNÍCH DATECH**

Součet počtu věřících z tabulky **religions** dle jednotlivých náboženství neodpovídá celkové velikosti populace jednotlivých států dostupné z tabulek ostatních. Proto ve výsledné tabulce procentuální součet věřících dle náboženství u jednotlivých států nedává 100 %.

Tabulka weather obsahuje údaje pouze za 35 hlavních měst a tabulka countries za 235 (zjištěno pomocí count či select distinct), nebylo tedy možné získat údaje o počasí za všechny státy.