/\* Note: The data was obtained from https://ourworldindata.org/coronavirus on May 2, 2022 and was originally 1 Excel file.

It was then split into 2 files in order to showcase joins,

\*/

-- Inspecting the data

SELECT \*

FROM `directed-reef-321809.covid\_case\_study.covid\_deaths05022022` dea

JOIN `directed-reef-321809.covid\_case\_study.covid\_vaccinations05022022` vac

ON dea.location = vac.location

AND dea.date = vac.date

/\* I noticed that the continent column was blank and the location column had a continent.

So looking a bit closer... \*/

SELECT DISTINCT(dea.location)

FROM `directed-reef-321809.covid\_case\_study.covid\_deaths05022022` dea

JOIN `directed-reef-321809.covid\_case\_study.covid\_vaccinations05022022` vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent IS NULL

-- Comparing this to the location column where the continent column is not null

SELECT DISTINCT(dea.location)

FROM `directed-reef-321809.covid\_case\_study.covid\_deaths05022022` dea

JOIN `directed-reef-321809.covid\_case\_study.covid\_vaccinations05022022` vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent IS NOT NULL

/\* The dataset already has pre-aggregated data by continent, income, and even the world as a whole,

and they're categorized under the location column.

\*/

/\* Here, I selected data I started to analyze and visualize.

I'm keeping both continent and location to manipulate in tableau \*/

SELECT dea.continent,

dea.location,

dea.date,

dea.population,

dea.new\_cases,

dea.new\_deaths,

vac.new\_vaccinations,

vac.people\_fully\_vaccinated

FROM `directed-reef-321809.covid\_case\_study.covid\_deaths05022022` dea

JOIN `directed-reef-321809.covid\_case\_study.covid\_vaccinations05022022` vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent IS NOT NULL

ORDER BY dea.location, dea.date

/\* Aggregated regional data but ended up barely using this for visualizing.

Used window functions to see if there were any notable differences between SUM(new cases/deaths) and MAX(total cases/deaths)

without grouping the variables and to be able to showcase them in my SQL code, too \*/

WITH regional\_data AS (

SELECT location,

population,

date,

new\_cases,

total\_cases,

new\_deaths,

total\_deaths,

new\_cases\_smoothed,

new\_deaths\_smoothed

FROM `directed-reef-321809.covid\_case\_study.covid\_deaths05022022`

WHERE continent IS NULL and location IN ('Africa', 'Asia', 'Europe', 'North America', 'Oceania', 'South America'))

SELECT \*,

SUM(new\_cases) OVER (PARTITION BY location ORDER BY date) AS continental\_case\_count,

SUM(new\_deaths) OVER (PARTITION BY location ORDER BY date) AS continental\_death\_count

FROM regional\_data

ORDER BY date DESC, location

/\* Global data using the pre-aggregated data.

I also compute for the global fully vaccinated rate. Since the first variable is a total(sum), it acts as a rolling rate \*/

SELECT dea.location,

dea.date,

dea.population,

dea.new\_cases,

dea.total\_cases,

dea.new\_deaths,

dea.total\_deaths,

vac.new\_vaccinations,

vac.people\_fully\_vaccinated,

vac.total\_boosters,

ROUND(vac.people\_fully\_vaccinated/dea.population\*100, 2) AS fully\_vaccinated\_rate

FROM `directed-reef-321809.covid\_case\_study.covid\_deaths05022022` dea

JOIN `directed-reef-321809.covid\_case\_study.covid\_vaccinations05022022` vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent IS NULL AND dea.location = 'World'

ORDER BY dea.location, dea.date

/\* Looking at the number of regular and ICU hospital patients to check the severity of the Omicron variant

relative to other variants (variant dates visualized on tableau). Only ended up using sum\_icu.

Used the CAST() function to convert weekly admissions data into integers as they were stored as strings \*/

SELECT continent,

date,

SUM(icu\_patients) AS sum\_icu,

SUM(hosp\_patients) AS sum\_hosp,

SUM(CAST(weekly\_icu\_admissions AS int64)) AS weekly\_icu,

SUM(CAST(weekly\_hosp\_admissions AS int64)) AS weekly\_hosp

FROM `directed-reef-321809.covid\_case\_study.covid\_deaths05022022`

WHERE continent IS NOT NULL

GROUP BY continent, date

ORDER BY continent, date DESC

-- Tableau: https://public.tableau.com/app/profile/ajcdl#!/