Assignment 4:

DataBase Design

**ISYS1055**

# **Part D: Data Retrieval and Visualization**

## **Task D.1: The total number of vaccines administered in each observation month and the difference in each of all countries**

### ***Query***

SELECT v1.OM1 AS "Observation Month 1 (OM1)",

v1.CountryName AS "Country Name (CN)",

v1.VOM1 AS "Administered Vaccine OM1 (VOM1)",

v2.OM2 AS "Observation Month 2 (OM2)",

v2.VOM2 AS "Administered Vaccine OM2 (VOM2)",

(v1.VOM1 - v2.VOM2) AS "Difference of Totals (VOM1-VOM2)"

FROM (SELECT CountryName,

strftime('%Y-%m', Date) AS OM1,

SUM(DailyVaccination) AS VOM1

FROM Vaccination

WHERE Date BETWEEN '2022-04-01' AND '2022-04-30'

GROUP BY CountryName, strftime('%Y-%m', Date)) v1

JOIN (SELECT CountryName,

strftime('%Y-%m', Date) AS OM2,

SUM(DailyVaccination) AS VOM2

FROM Vaccination

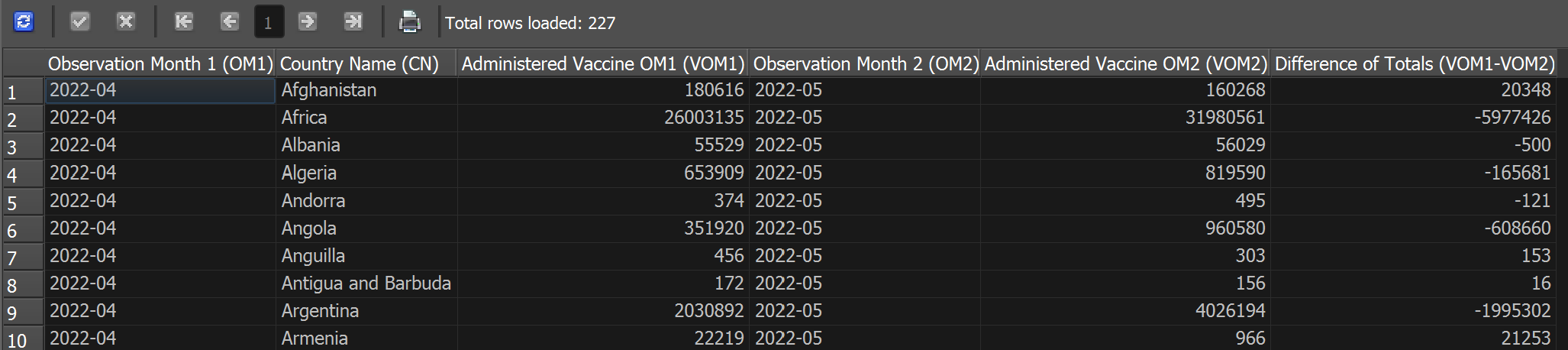
WHERE Date BETWEEN '2022-05-01' AND '2022-05-31'

GROUP BY CountryName, strftime('%Y-%m', Date)) v2

ON v1.CountryName = v2.CountryName

ORDER BY v1.CountryName;

### ***Snapshot***



### ***Visualization***

First, the results are shortened to only four countries including Canada, Denmark, the United States, and Wales that represent the whole data retrieved from the query, to easily maintain the readability for the visualization. Then, since the data for the United States outperforms the figures for other countries, square-root is the mathematical operation that is utilized to transform the data to standardize their scale.

## **Task D.2: Countries with cumulative numbers of doses administered that higher than the average amount**

### ***Query***

SELECT v1.CountryName AS "Country Name",

v1.Month,

v1.CumulativeDoses AS "Cumulative Doses"

FROM (SELECT CountryName,

strftime('%Y-%m', Date) AS Month,

MAX(PeopleVaccinated + PeopleFullyVacinated + TotalBooster) AS CumulativeDoses

FROM Vaccination

GROUP BY CountryName, strftime('%Y-%m', Date)) v1

JOIN (SELECT Month,

AVG(CumulativeDoses) AS AvgCumulativeDoses

FROM (SELECT strftime('%Y-%m', Date) AS Month,

MAX(PeopleVaccinated + PeopleFullyVacinated + TotalBooster) AS CumulativeDoses

FROM Vaccination

GROUP BY CountryName, strftime('%Y-%m', Date))

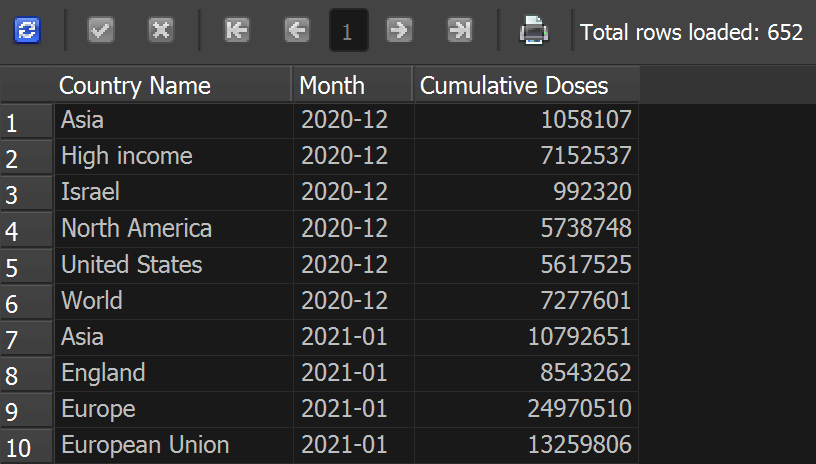
GROUP BY Month) v2

ON v1.Month = v2.Month

WHERE v1.CumulativeDoses > v2.AvgCumulativeDoses

ORDER BY v1.Month;

### ***Snapshot***



### ***Visualization***

## **Task D.3: Vaccine types that are administered by several countries**

### ***Query***

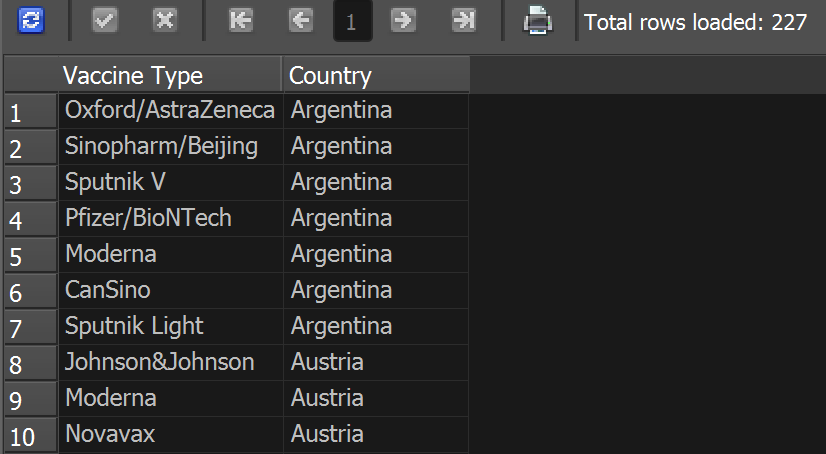
SELECT DISTINCT

VaccineName AS "Vaccine Type",

CountryName AS "Country"

FROM Vaccines;

### ***Snapshot***



### ***Visualization***

## **Task D.4: Total of vaccines administered retrieved from specific data source**

### ***Query***

SELECT v.CountryName AS "Country Name",

l.SourceURL AS "Source Name (URL)",

SUM(v.DailyVaccination) AS "Total Administered Vaccines"

FROM Vaccination v

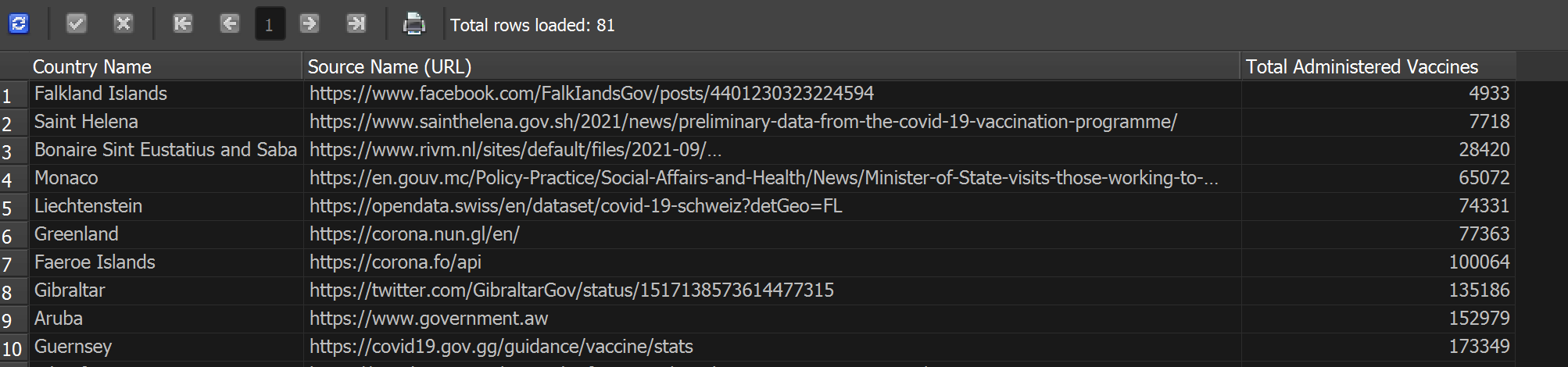
JOIN Location l

ON v.CountryName = l.CountryName

GROUP BY SourceURL

ORDER BY SUM(v.DailyVaccination);

### ***Snapshot***



### ***Visualization***

## **Task D.5: Number of people fully vaccinated in the United States, Canada, Denmark, and Wales**

### ***Query***

SELECT Month AS "Date Range (Months)",

MAX(CASE WHEN CountryName = 'United States' THEN TotalFullyVaccinated ELSE 0 END) AS "United States",

MAX(CASE WHEN CountryName = 'Wales' THEN TotalFullyVaccinated ELSE 0 END) AS "Wales",

MAX(CASE WHEN CountryName = 'Canada' THEN TotalFullyVaccinated ELSE 0 END) AS "Canada",

MAX(CASE WHEN CountryName = 'Denmark' THEN TotalFullyVaccinated ELSE 0 END) AS "Denmark"

FROM (SELECT strftime('%Y-%m', Date) AS Month,

CountryName,

MAX(CASE WHEN PeopleFullyVacinated = '' THEN 0 ELSE PeopleFullyVacinated END) AS TotalFullyVaccinated

FROM Vaccination

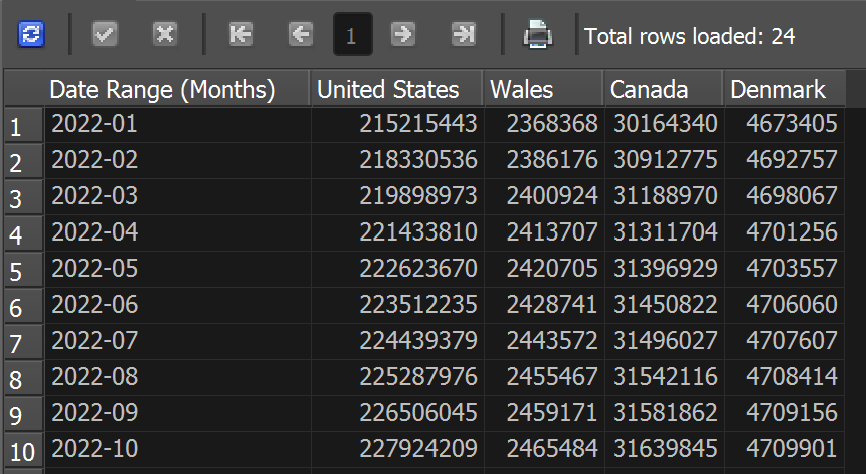
GROUP BY strftime('%Y-%m', Date), CountryName

HAVING CountryName IN ('United States', 'Wales', 'Canada', 'Denmark')

AND strftime('%Y', Date) IN ('2022', '2023'))

GROUP BY Month;

### ***Snapshot***



### ***Visualization***

The data of the total number of people fully vaccinated with the second dose for each country are modified to determine the speed of administering vaccines to their population. By taking the amount for the following month minus the figure for the previous month, the difference between each month will form a line to see how fluctuating the vaccine administration is. Therefore, the graph shows each nation’s performance and its changes. Moreover, the graph displays data only from 20222 as coming to the following year, several months failed to record any values of vaccine administration. Since the data for the United States is too large comparing to Wales, Canada, or Denmark, the transformation is done for a better scale that ensure the readability of the chart.