RMIT UNIVERSITY

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

a) Query

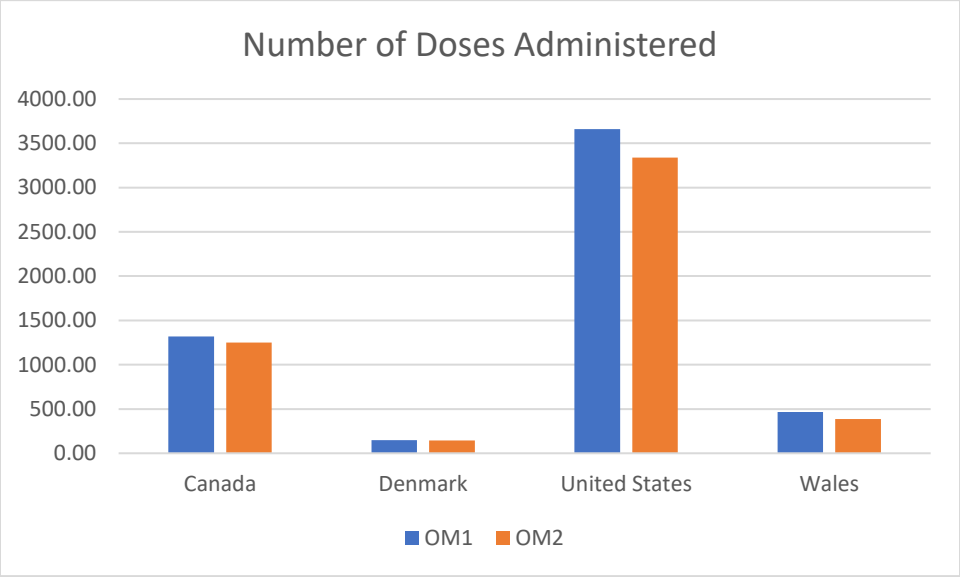
```
SELECT v1.OM1 AS "Observation Month 1 (OM1)",
   v1.CountryName AS "Country Name (CN)",
   v1.VOM1 AS "Administered Vaccine OM1 (VOM1)",
   v2.OM2 AS "Observation Month 2 (OM2)",
   v2.VOM2 AS "Administered Vaccine OM2 (VOM2)",
   (v1.VOM1 - v2.VOM2) AS "Difference of Totals (VOM1-VOM2)"
FROM (SELECT 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;
```

b) Snapshot

Total rows loaded: 227

	Observation Month 1 (OM1)	Country Name (CN)	Administered Vaccine OM1 (VOM1)	Observation Month 2 (OM2)	Administered Vaccine OM2 (VOM2)	Difference of Totals (VOM1-VOM2)
1	2022-04	Afghanistan	180616	2022-05	160268	20348
2	2022-04	Africa	26003135	2022-05	31980561	-5977426
3	2022-04	Albania	55529	2022-05	56029	-500
4	2022-04	Algeria	653909	2022-05	819590	-165681
5	2022-04	Andorra	374	2022-05	495	-121
6	2022-04	Angola	351920	2022-05	960580	-608660
7	2022-04	Anguilla	456	2022-05	303	153
8	2022-04	Antigua and Barbuda	172	2022-05	156	16
9	2022-04	Argentina	2030892	2022-05	4026194	-1995302
10	2022-04	Armenia	22219	2022-05	966	21253

c) 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

a) Query

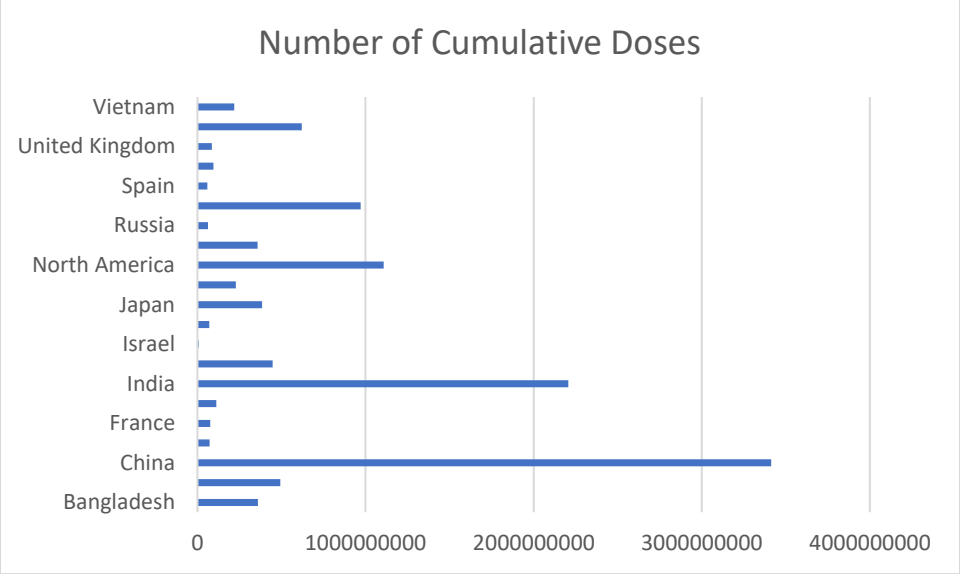
```
SELECT v1.CountryName AS "Country Name",
   v1.Month,
   v1.CumulativeDoses AS "Cumulative Doses"
FROM (SELECT CountryName,
      strftime('%Y-%m', Date) AS Month,
      MAX(PeopleVaccinated + PeopleFullyVacinated + TotalBooster) AS CumulativeDoses
   FROM Vaccination
   GROUP BY CountryName, strftime('%Y-%m', 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;
```

b) Snapshot

Total rows loaded: 652

	Country Name	Month	Cumulative Doses
1	Asia	2020-12	1058107
2	High income	2020-12	7152537
3	Israel	2020-12	992320
4	North America	2020-12	5738748
5	United States	2020-12	5617525
6	World	2020-12	7277601
7	Asia	2021-01	10792651
8	England	2021-01	8543262
9	Europe	2021-01	24970510
10	European Union	2021-01	13259806

c) Visualization



RMIT UNIVERSITY

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

a) Query

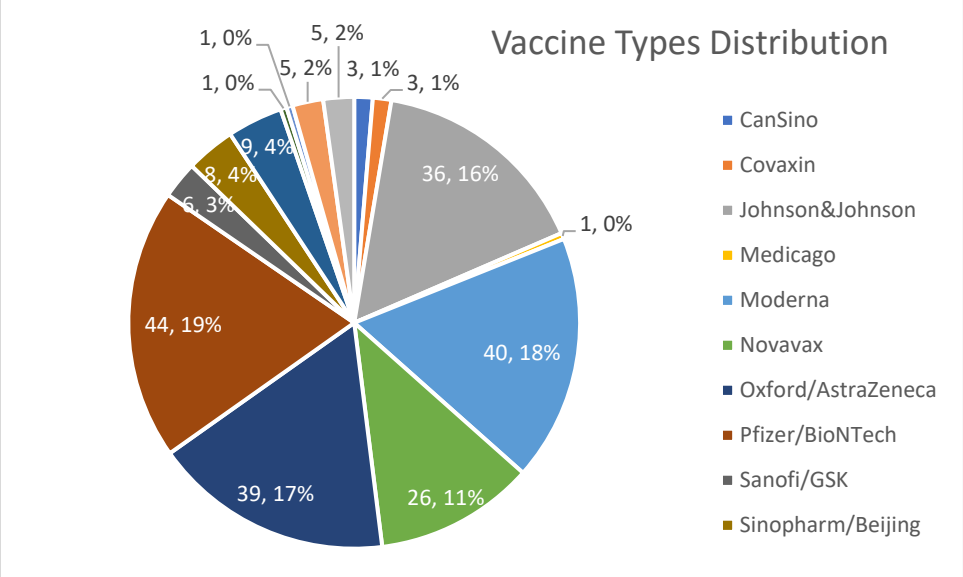
```
SELECT DISTINCT
   VaccineName AS "Vaccine Type",
   CountryName AS "Country"
FROM Vaccines;
```

b) Snapshot

Total rows loaded: 227

	Vaccine Type	Country
1	Oxford/AstraZeneca	Argentina
2	Sinopharm/Beijing	Argentina
3	Sputnik V	Argentina
4	Pfizer/BioNTech	Argentina
5	Moderna	Argentina
6	CanSino	Argentina
7	Sputnik Light	Argentina
8	Johnson&Johnson	Austria
9	Moderna	Austria
10	Novavax	Austria

c) Visualization



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

a) Query

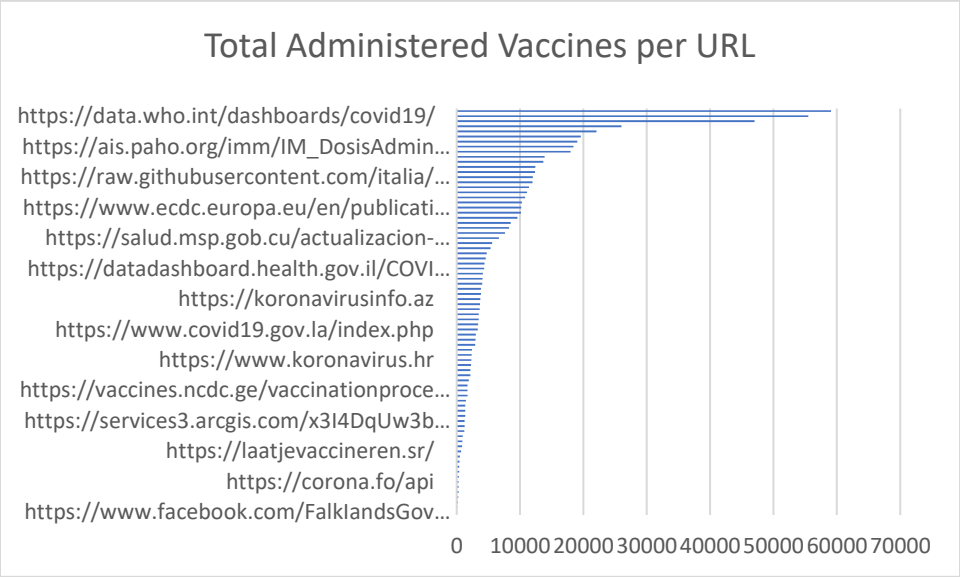
```
SELECT v.CountryName AS "Country Name",
   l.SourceURL AS "Source Name (URL)",
   SUM(v.DailyVaccination) AS "Total Administered Vaccines"
FROM Vaccination v
   JOIN Location l
   ON v.CountryName = l.CountryName
GROUP BY SourceURL
ORDER BY SUM(v.DailyVaccination);
```

b) Snapshot

Total rows loaded: 81

	Country Name	Source Name (URL)	Total Administered Vaccines
1	Falkland Islands	https://www.facebook.com/FalkIandsGov/posts/4401230323224594	4933
2	Saint Helena	https://www.sainthelena.gov.sh/2021/news/preliminary-data-from-the-covid-19-vaccination-programme/	7718
3	Bonaire Sint Eustatius and Saba	https://www.rivm.nl/sites/default/files/2021-09/...	28420
4	Monaco	https://en.gouv.mc/Policy-Practice/Social-Affairs-and-Health/News/Minister-of-State-visits-those-working-to-...	65072
5	Liechtenstein	https://opendata.swiss/en/dataset/covid-19-schweiz?detGeo=FL	74331
6	Greenland	https://corona.nun.gl/en/	77363
7	Faeroe Islands	https://corona.fo/api	100064
8	Gibraltar	https://twitter.com/GibraltarGov/status/1517138573614477315	135186
9	Aruba	https://www.government.aw	152979
10	Guernsey	https://covid19.gov.gg/guidance/vaccine/stats	173349

c) Visualization

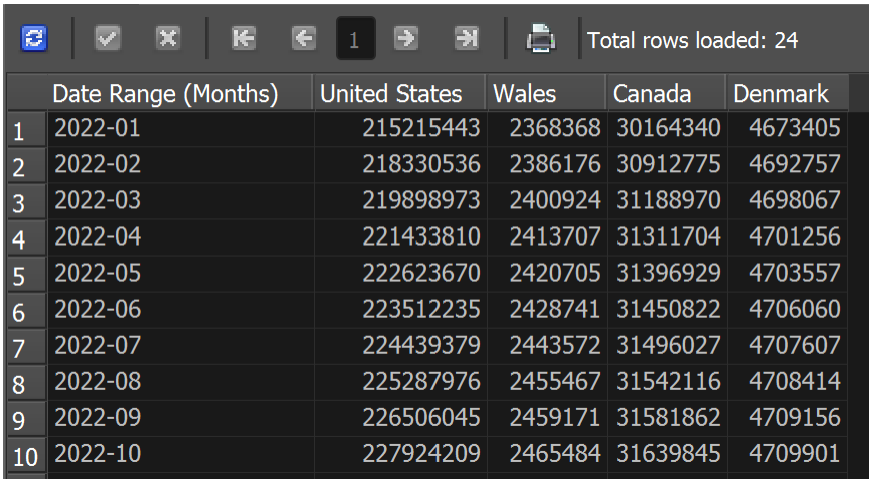


Task D.5: Number of people fully vaccinated in the United States, Canada, Denmark, and Wales

a) Query

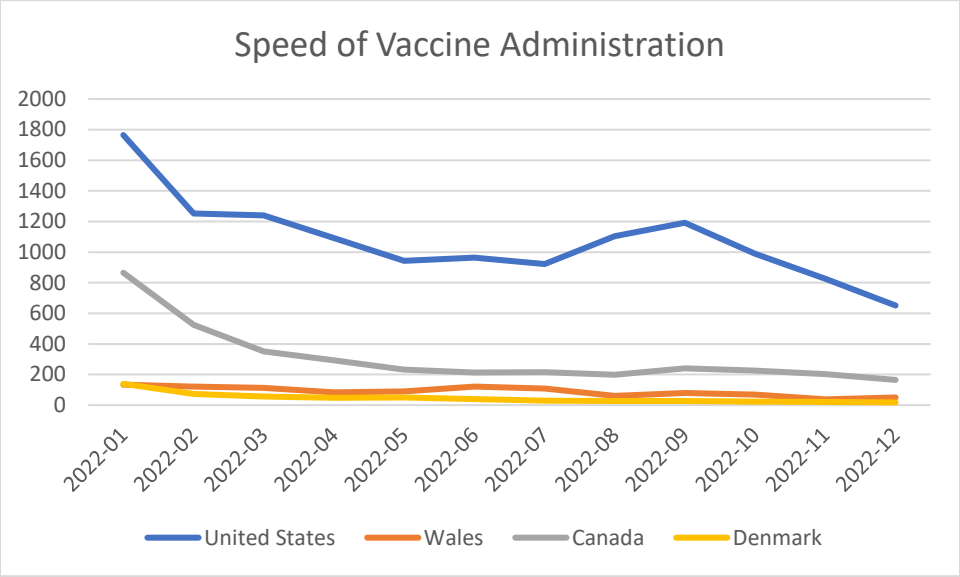
```
SELECT Month AS "Date Range (Months)",
   MAX(CASE WHEN CountryName = 'United States' THEN TotalFullyVaccinated ELSE 0 END) AS "United States",
   MAX(CASE WHEN CountryName = 'Wales' THEN TotalFullyVaccinated ELSE 0 END) AS "Wales",
   MAX(CASE WHEN CountryName = 'Canada' THEN TotalFullyVaccinated 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;
```

b) Snapshot



	Date Range (Months)	United States	Wales	Canada	Denmark
1	2022-01	215215443	2368368	30164340	4673405
2	2022-02	218330536	2386176	30912775	4692757
3	2022-03	219898973	2400924	31188970	4698067
4	2022-04	221433810	2413707	31311704	4701256
5	2022-05	222623670	2420705	31396929	4703557
6	2022-06	223512235	2428741	31450822	4706060
7	2022-07	224439379	2443572	31496027	4707607
8	2022-08	225287976	2455467	31542116	4708414
9	2022-09	226506045	2459171	31581862	4709156
10	2022-10	227924209	2465484	31639845	4709901

c) 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.