Covid Vaccination Trend Analysis
CINS 370
Dr. Jaime Raigoza
Project - 1
Team - 2
Team members: *Ray Perez, Alexes Gonzalez, Feroz khan*

Lately, Coronavirus is the World number's one worry. This virus has infected and killed millions of people all over the world. Every country is trying its best to reduce the number of cases from one day to the next. As this problem arises, data have been playing a key resource for measuring trends like the impact of the Covid-19 pandemic. Our team will be using these related datasets to cultivate an analysis of covid related trends.

Our domain will be covid rates by country and the progress of vaccination by country. As well as analyzing the population by each country thus tracking the overall population by each country revealing trends of covid cases for each country and how well they are handling vaccination processes. By looking at datasets it will reveal the trends of vaccination. The entities that we will be using to support our queries are country population, Covid cases by Month/Year, total Covid cases, and Vaccination progress by country. Our overall goal will be to model and populate information about Covid trends within the given domain above.

5 queries :
1. **Providing graphical visualization of the progress between countries of their vaccination rates.**
2. **The change over time of active covid cases and as well as change over time of vaccination progress by country.**
3. **Is there a relationship between death rates and percentage of vaccination. If higher vaccination would affect the percentage of death rates.**
4. **How does population density affect covid cases? Does this hinder vaccination procedures?**
5. **How does the vaccination progress affect the total number of cases by continent/region? Are there any wildcards that affect the overall progression?**

Dataset 1- Country <u>Population density</u>

- This dataset is the information of countries' population densities through yearly change including birth rates. In Regards to the overall land mass as well as the percentage of migrants that make up the population.
  Contains 236 rows

	Country (or dependency)	Population (2020)	Yearly Change	Net Change	Density (P/KmÂ²)	Land Area (KmÂ²)	Migrants (net)	Fert. Rate	Med. Age	Urban Pop %	World Share
1											
2	China	1440297825	0.39%	5540090	153	9388211	-348399	1.7	38	61%	18.47%
3	India	1382345085	0.99%	13586631	464	2973190	-532687	2.2	28	35%	17.70%
4	United States	331341050	0.59%	1937734	36	9147420	954806	1.8	38	83%	4.25%
5	Indonesia	274021604	1.07%	2898047	151	1811570	-98955	2.3	30	56%	3.51%
6	Pakistan	221612785	2.00%	4327022	287	770880	-233379	3.6	23	35%	2.83%
7	Brazil	212821986	0.72%	1509890	25	8358140	21200	1.7	33	88%	2.73%
8	Nigeria	206984347	2.58%	5175990	226	910770	-60000	5.4	18	52%	2.64%
9	Bangladesh	164972348	1.01%	1643222	1265	130170	-369501	2.1	28	39%	2.11%
10	Russia	145945524	0.04%	62206	9	16376870	182456	1.8	40	74%	1.87%
11	Mexico	129166028	1.06%	1357224	66	1943950	-60000	2.1	29	84%	1.65%
12	Japan	126407422	-0.30%	-383840	347	364555	71560	1.4	48	92%	1.62%
13	Ethiopia	115434444	2.57%	2884858	115	1000000	30000	4.3	19	21%	1.47%
14	Philippines	109830324	1.35%	1464463	368	298170	-67152	2.6	26	47%	1.41%
15	Egypt	102659126	1.94%	1946331	103	995450	-38033	3.3	25	43%	1.31%
16	Vietnam	97490013	0.91%	876473	314	310070	-80000	2.1	32	38%	1.25%
17	DR Congo	90003954	3.19%	2770836	40	2267050	23861	6	17	46%	1.15%
18	Turkey	84495243	1.09%	909452	110	769630	283922	2.1	32	76%	1.08%
19	Iran	84176929	1.30%	1079043	52	1628550	-55000	2.2	32	76%	1.08%
20	Germany	83830972	0.32%	266897	240	348560	543822	1.6	46	76%	1.07%

Dataset 2 - Covid cases by <u>Month/Year</u>

- This dataset reveals the trend of covid cases and the connection between time intervals. Revealing the trends of covid cases as we progress through the years with comparisons to the death and case rates.
  This dataset contains 53,000 rows.

	Date.Day	Date.Month	Date.Year	Data.Cases	Data.Deaths	Location.Country	Data.Population	Location.Continent	Data.Rate
1									
2	5	11	2020	27	0	China	1433783692	Asia	0.02754948
3	5	11	2020	50210	704	India	1366417756	Asia	48.09217365
4	5	11	2020	102507	1102	United_States_of_America	329064917	America	349.5371097
5	5	11	2020	3356	113	Indonesia	270625567	Asia	17.96652125
6	5	11	2020	1302	26	Pakistan	216565317	Asia	6.18520093
7	5	11	2020	23976	610	Brazil	211049519	America	138.0022098
8	5	11	2020	155	4	Nigeria	200963603	Africa	0.82651782
9	5	11	2020	1517	21	Bangladesh	163046173	Asia	12.90002679
10	5	11	2020	19768	389	Russia	145872260	Europe	168.722278
11	5	11	2020	5225	635	Mexico	127575529	America	59.6282066
12	5	11	2020	938	8	Japan	126860299	Asia	7.34193445
13	5	11	2020	379	9	Ethiopia	112078727	Africa	6.03415133
14	5	11	2020	976	49	Philippines	108116622	Asia	23.95006385
15	5	11	2020	207	13	Egypt	100388076	Africa	2.43654436
16	5	11	2020	4	0	Vietnam	96462108	Asia	0.06427394
17	5	11	2020	23	0	Democratic_Republic_of_the_Congo	86790568	Africa	0.44244439
18	5	11	2020	2391	77	Turkey	82003882	Asia	37.90430312
19	5	11	2020	8452	419	Iran	82913893	Asia	121.6659794
20	5	11	2020	19990	118	Germany	83019213	Europe	247.5740164

Dataset 3- Covid Cases by Country

- This dataset includes information about Covid cases by country. It holds the total number of cases and total death caused by Covid-19. Each row represents each country worldwide. It has 210 rows and 6 columns. Each attribute holds information regarding the country or Covid.

	Country/Region	Continent	Population	TotalCases	TotalDeaths	WHO Region
2	USA	North America	331198130	5032179	162804	Americas
3	Brazil	South America	212710692	2917562	98644	Americas
4	India	Asia	1381344997	2025409	41638	South-EastAsia
5	Russia	Europe	145940924	871894	14606	Europe
6	South Africa	Africa	59381566	538184	9604	Africa
7	Mexico	North America	129066160	462690	50517	Americas
8	Peru	South America	33016319	455409	20424	Americas
9	Chile	South America	19132514	366671	9889	Americas
10	Colombia	South America	50936262	357710	11939	Americas
11	Spain	Europe	46756648	354530	28500	Europe
12	Iran	Asia	84097623	320117	17976	EasternMediterranean
13	UK	Europe	67922029	308134	46413	Europe
14	Saudi Arabia	Asia	34865919	284226	3055	EasternMediterranean
15	Pakistan	Asia	221295851	281863	6035	EasternMediterranean
16	Bangladesh	Asia	164851401	249651	3306	South-EastAsia
17	Italy	Europe	60452568	249204	35187	Europe
18	Turkey	Asia	84428331	237265	5798	Europe
19	Argentina	South America	45236884	228195	4251	Americas
20	Germany	Europe	83811260	215210	9252	Europe

Dataset 4 - Vaccination progress by Continents

- This dataset contains the overall progression of vaccination progress by continents. Showcasing data of doses administered and the percentage of Fully vaccinated population. This dataset contains 219 rows of data.

	Continent	Doses Administered	Doses per 1000	Fully Vaccinated Population (%)	Vaccine being used in a country
1	Continent	Doses Administered	Doses per 1000	Fully Vaccinated Population (%)	Vaccine being used in a country
2	World	10102904922	1330.3	54.79	
3	North America	3000532000	2154.4	87.92	CanSino, Sinopharm/Beijing, Sinopharm/Wuhan, Sinovac, ZF2001
4	South America	1664786038	1230.8	52.49	Covaxin, Oxford/AstraZeneca, Sputnik V
5	Asia	539337296	1648.5	64.51	Johnson&Johnson, Moderna, Pfizer/BioNTech
6	Europe	362503209	1730.6	71.58	Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac
7	Africa	312709071	1168.3	44.44	Moderna, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing, Sinovac
8	South America	205744165	1626.1	78.87	Moderna, Oxford/AstraZeneca, Pfizer/BioNTech
9	Australia/Oceania	181280001	1897.4	45.61	Moderna, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing, Sputnik V
10	Asia	174085175	820.3	38.05	CanSino, Covaxin, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing, Sinovac, Sputnik V

Reference:

CORGIS Datasets Project (corgis-edu.github.io)

https://www.kaggle.com/imdevskp/corona-virus-report?select=worldometer_data.csv

https://www.kaggle.com/gpreda/covid-world-vaccination-progress

https://www.kaggle.com/imdevskp/corona-virus-report?select=worldometer_data.csv