SUMMARY REPORTS ON THE PROGRESS OF COVID-19 VACCINATION, FACTORS THAT INFLUENCE VACCINATION & SENTIMENTS OF PUBLIC ABOUT VACCINES

Aditya Nagori

INTRODUCTION

- We will be using four datasets which is about Covid-19 Vaccination drive throughout the world.

- The very first dataset (country_vaccinations.csv) is about country wise daily vaccination drive which was started in different months in respective countries until September 2021. This dataset incudes total vaccinations per date and country, number of people vaccinated, number of people fully vaccinated, daily vaccination, total vaccination percent, people vaccinated percent, daily vaccinations per million, vaccines scheme (the combination of vaccines used by a country).
- Second dataset (country_vaccinations_by_manufacturer.csv) gives insights about the Total number of vaccinations manufactured by various Vaccine companies date wise from January 2021 to September 2021 in each country.
- Third dataset (vaccination_all_tweets.csv) tell us about the attitude of population in view of different vaccination drives according to the tweets related to vaccinations. This dataset was picked up from **tweepy** Python package. This will be used as a text data analysis, which is the main focus of our group.
- Fourth dataset (country_profile_variables.csv) is about the demographics and economy of each country. We will use this to find what influenced these vaccination programmes, and these programmes success.

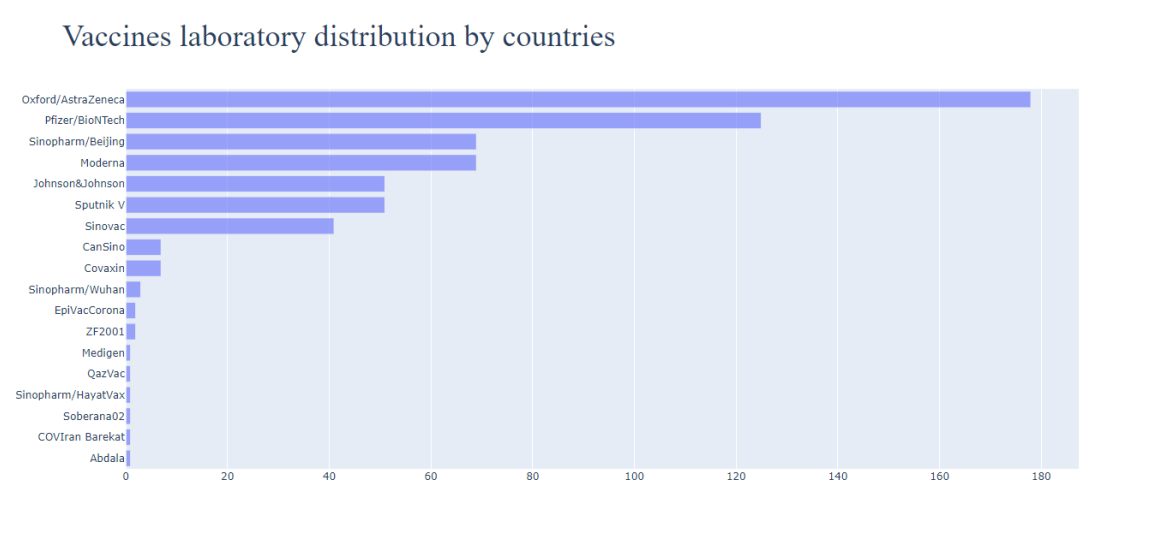
After Explorative Data Analysis we would be able to answer questions like:

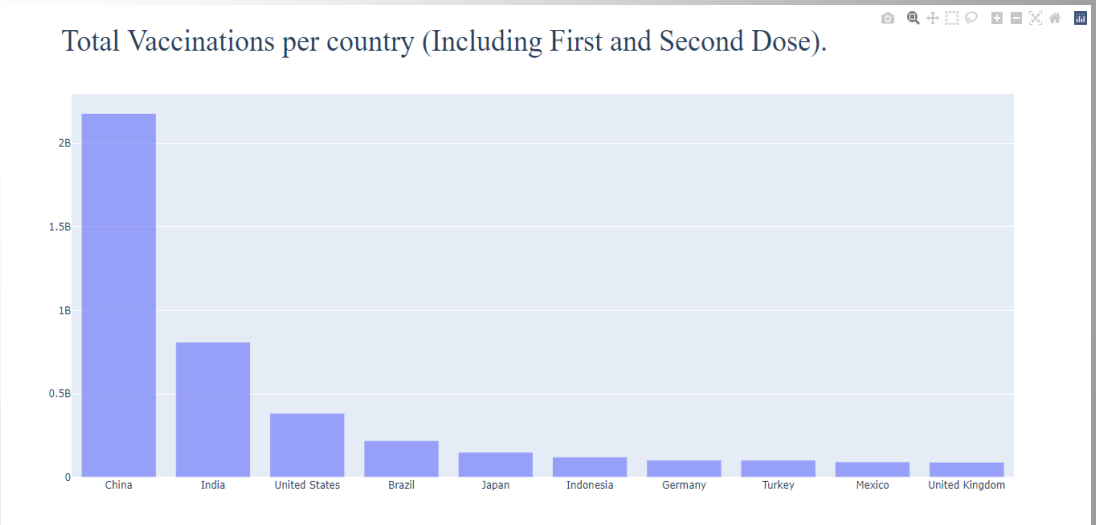
- Which country's vaccine drive was success?
- What percent population of a certain country got vaccinated?
- What vaccination schemes (combination of vaccines) are used and in which countries?
- What manufacturer had produced the most amount of vaccine and sold to what country?
- What was the relationship between vaccination evolution and sentiment toward vaccinations?
- What are the factors that influence vaccinations?
- What country has vaccinated more people?
- What country has immunized the largest percent from its population?
- How progress of vaccination programmes around the World (or in a specific country) is received by the public, as reflected in the tweets about all vaccines.

MAJOR FINDINGS

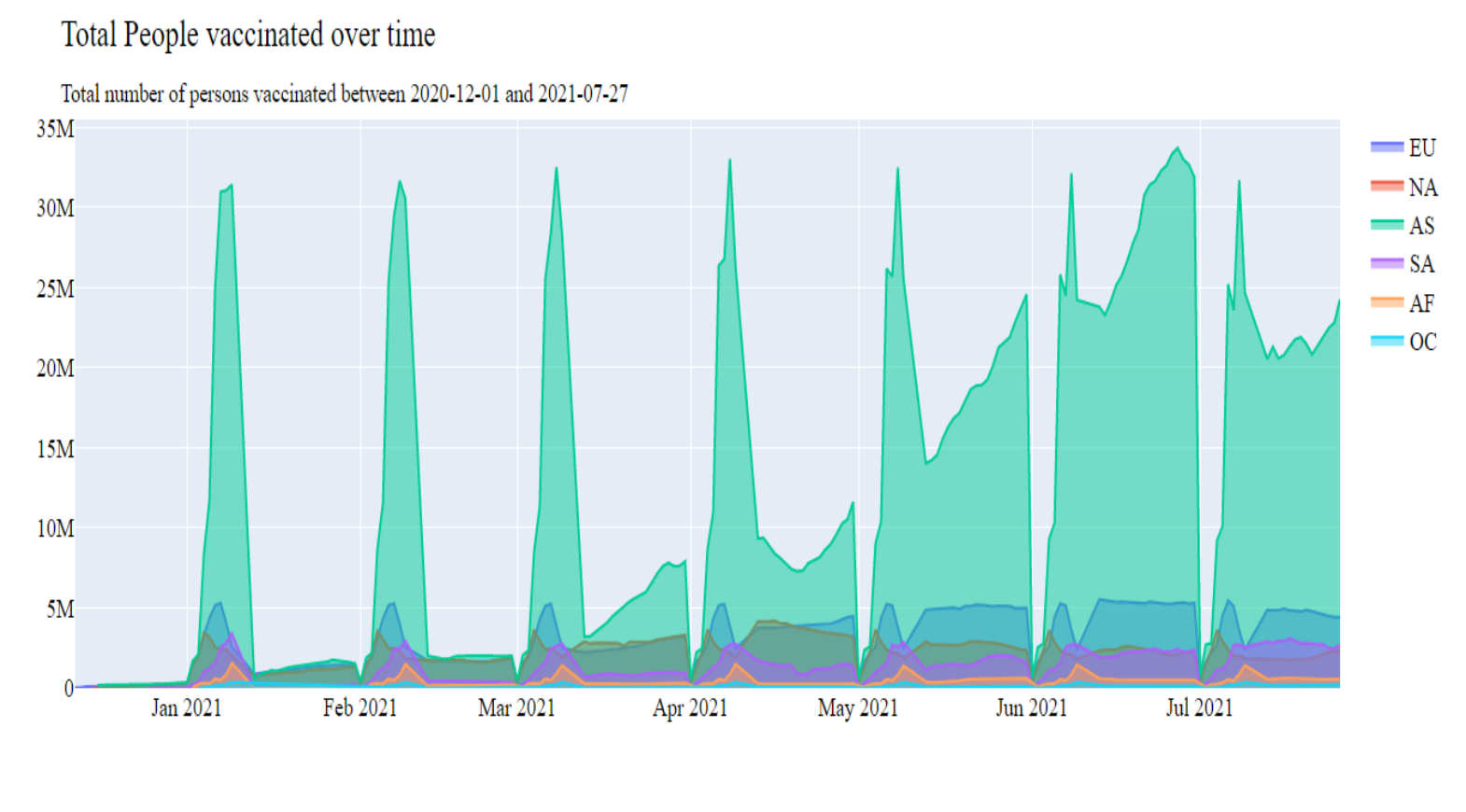
- According to the data and graphs, we can conclude that oxford/ AstraZeneca had manufactured and distributed the most number of vaccines.

- China is the country to have highest number of population vaccinated.



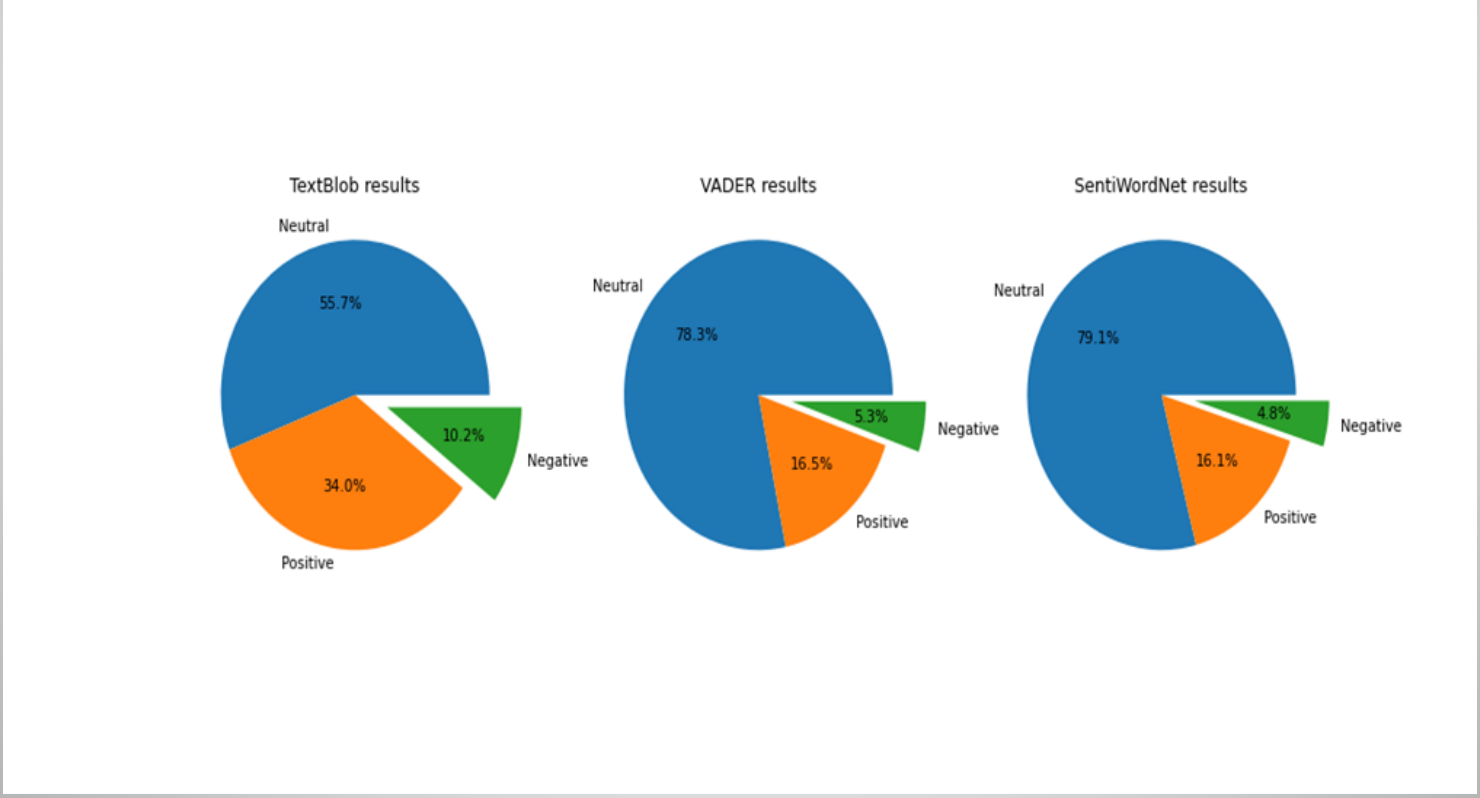


- Asia had the best vaccination drive amongst all other continents.



- According to different sentiment analysis methods mostly the tweets from public were neutral. After that most tweets were positive about vaccination drives.

- We saw that population is inversely proportional to vaccine drive from the regression analysis as the r coefficient is negative between population and daily vaccinations.

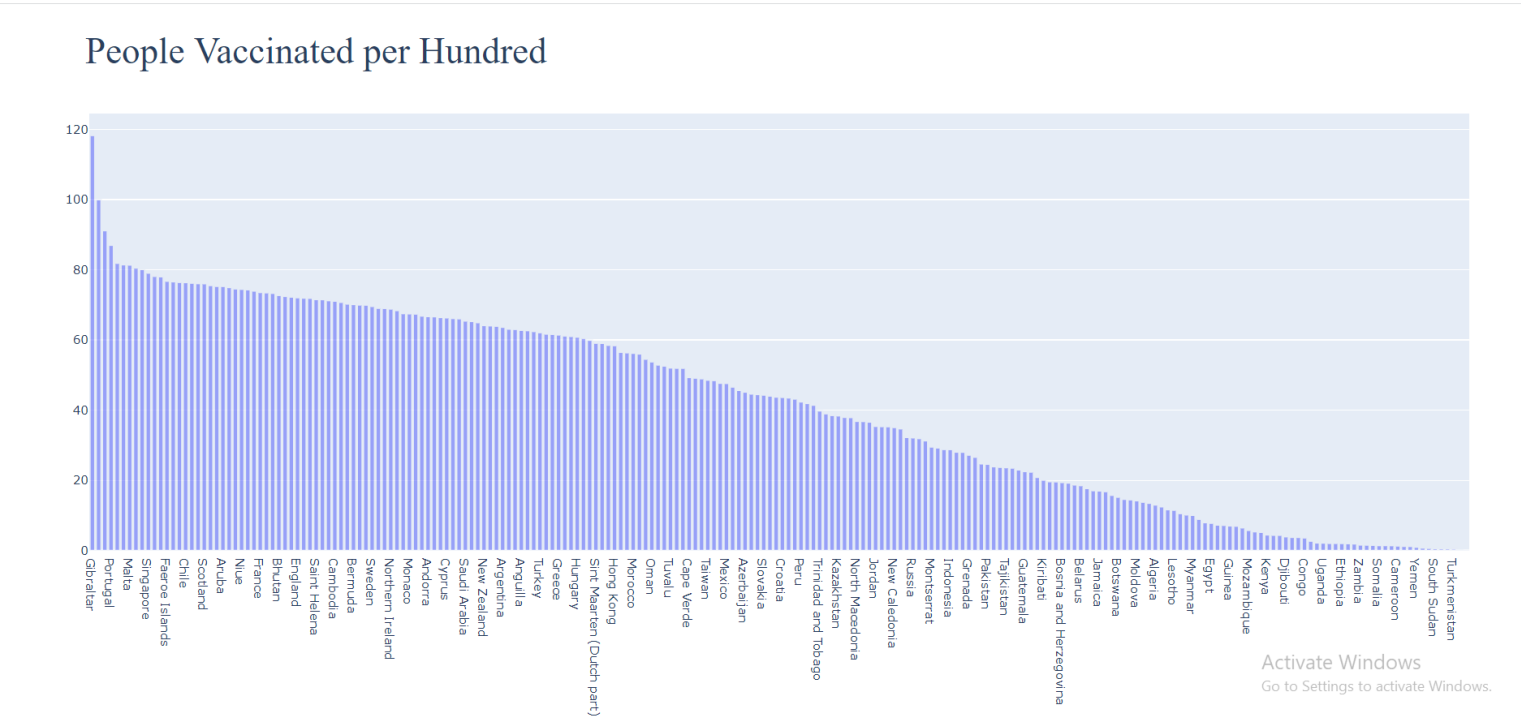


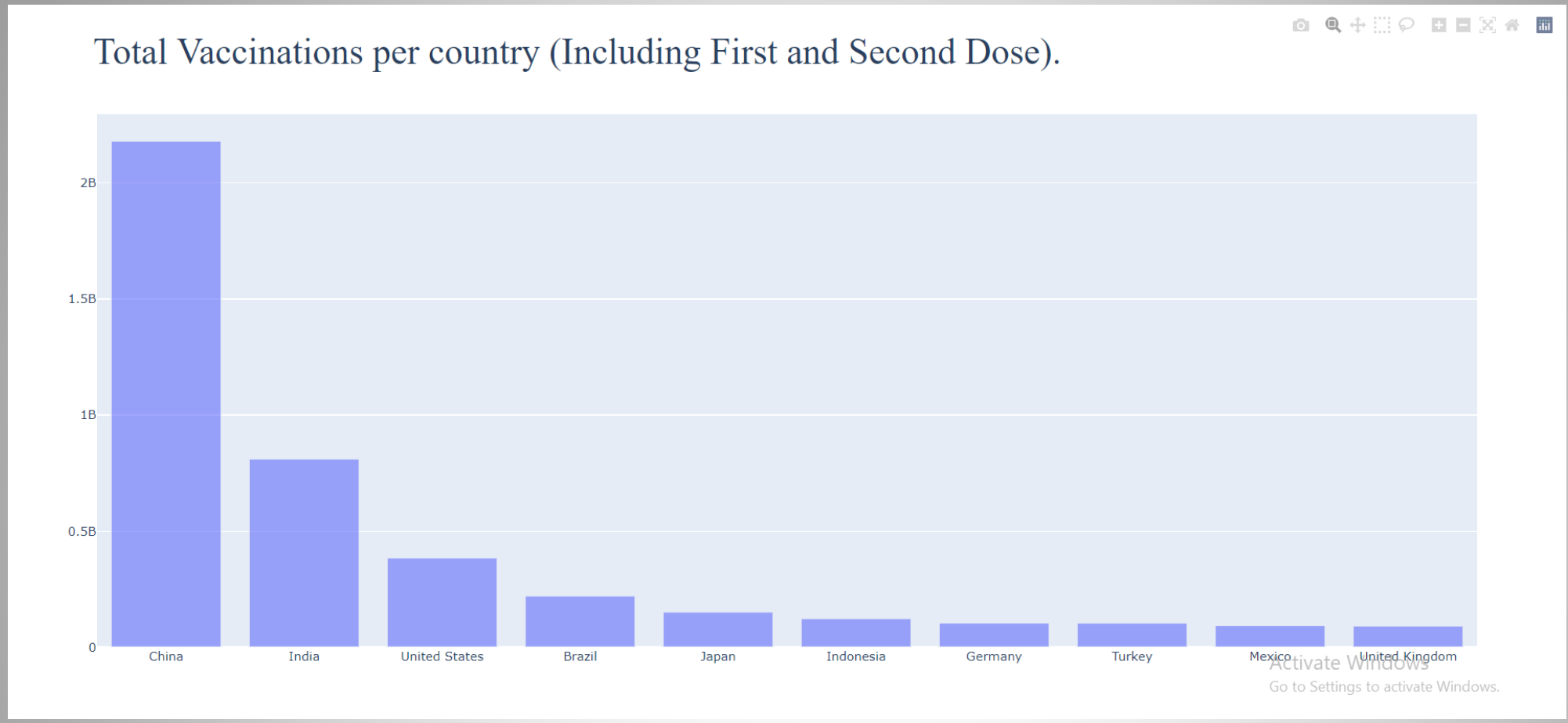
Analytical Overview:

- Excel data was cleaned before analysing such as duplicate rows and non-essential column was removed from the databases.
- For all graphs and data visualisation, we used different methods to showcase our knowledge of python.
- All major findings and recommendations are based on EDA which is explained in Documentation Page.

OUTPUT GRAPHS DESCRIPTION

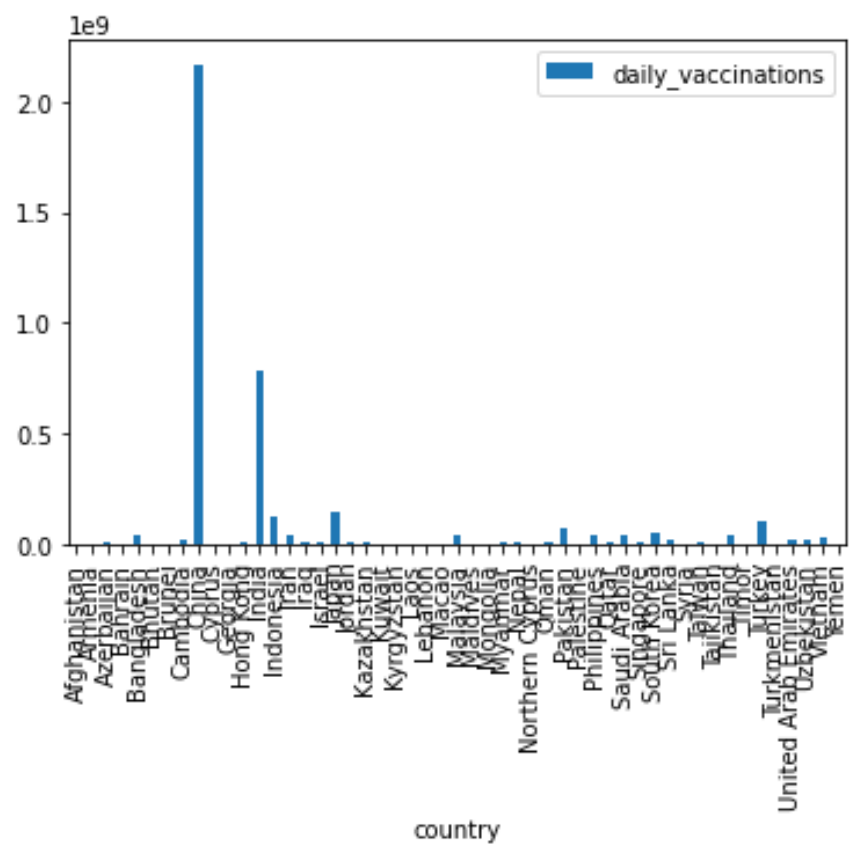
- China shows the highest rate of vaccinations of both the doses followed by India and United States. On the other hand, Indonesia, Germany, Mexico and Turkey stand on almost same number of vaccinations.

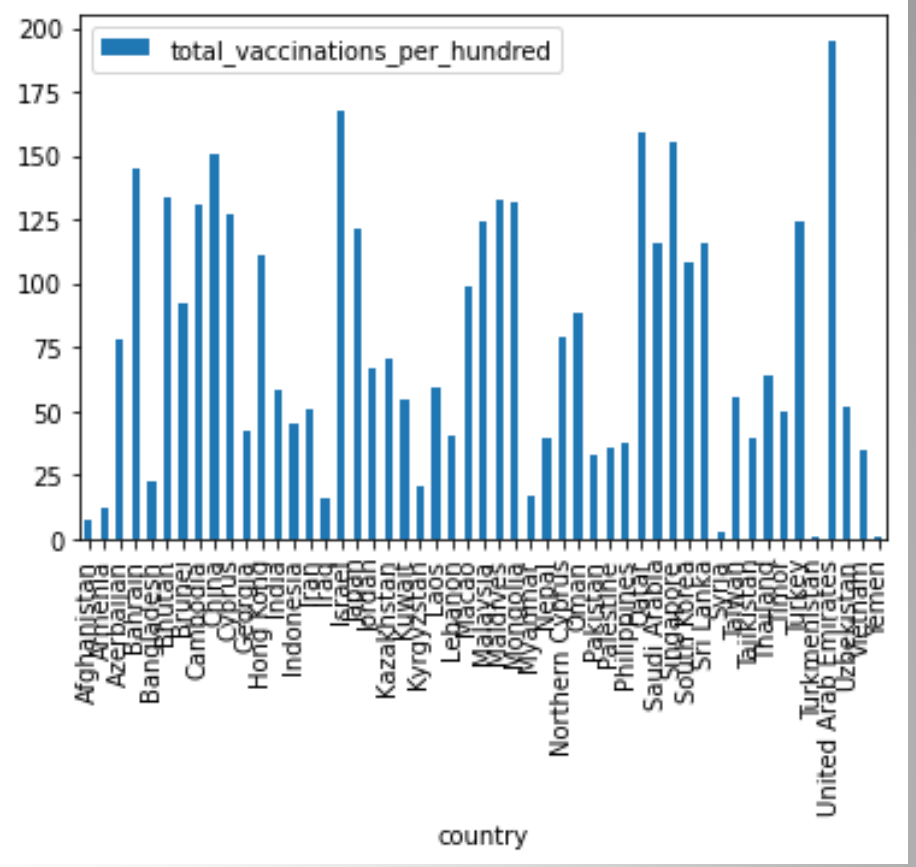




From the above bar plot we can see that most of the vaccinations is done only in China from December of 2020 till October 2021. India stands in second place in that list. As the population of both the countries is very huge, the number of people vaccinated are more.

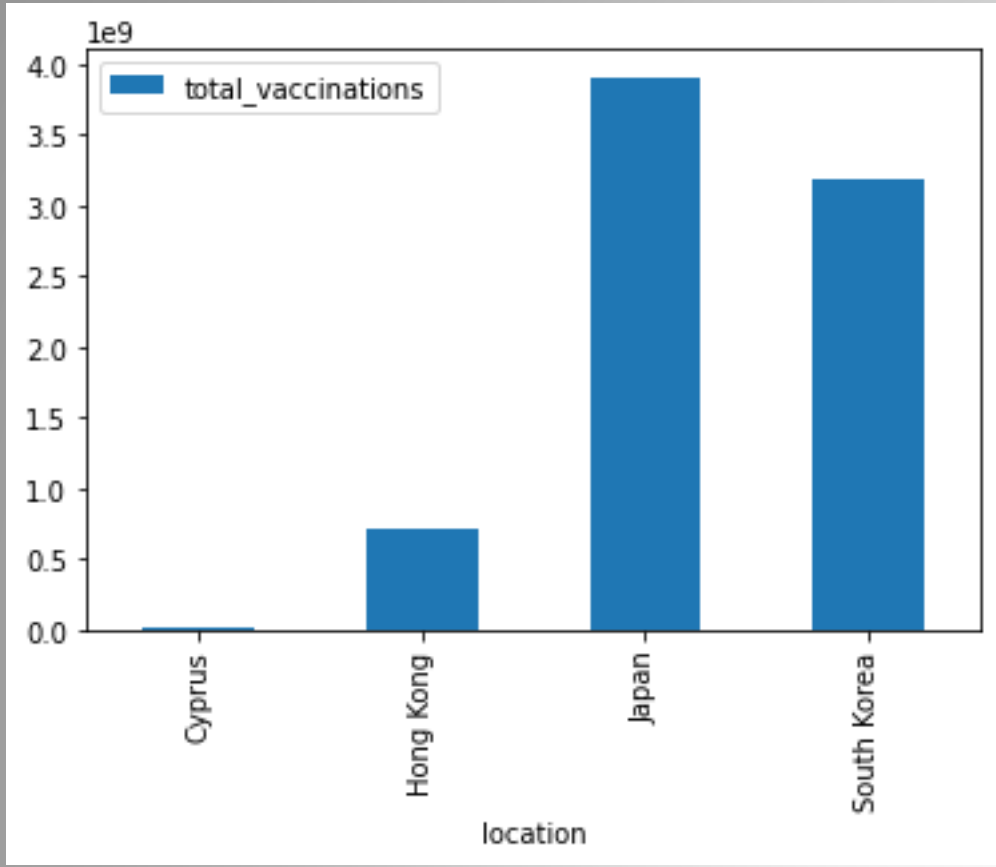
ASIA



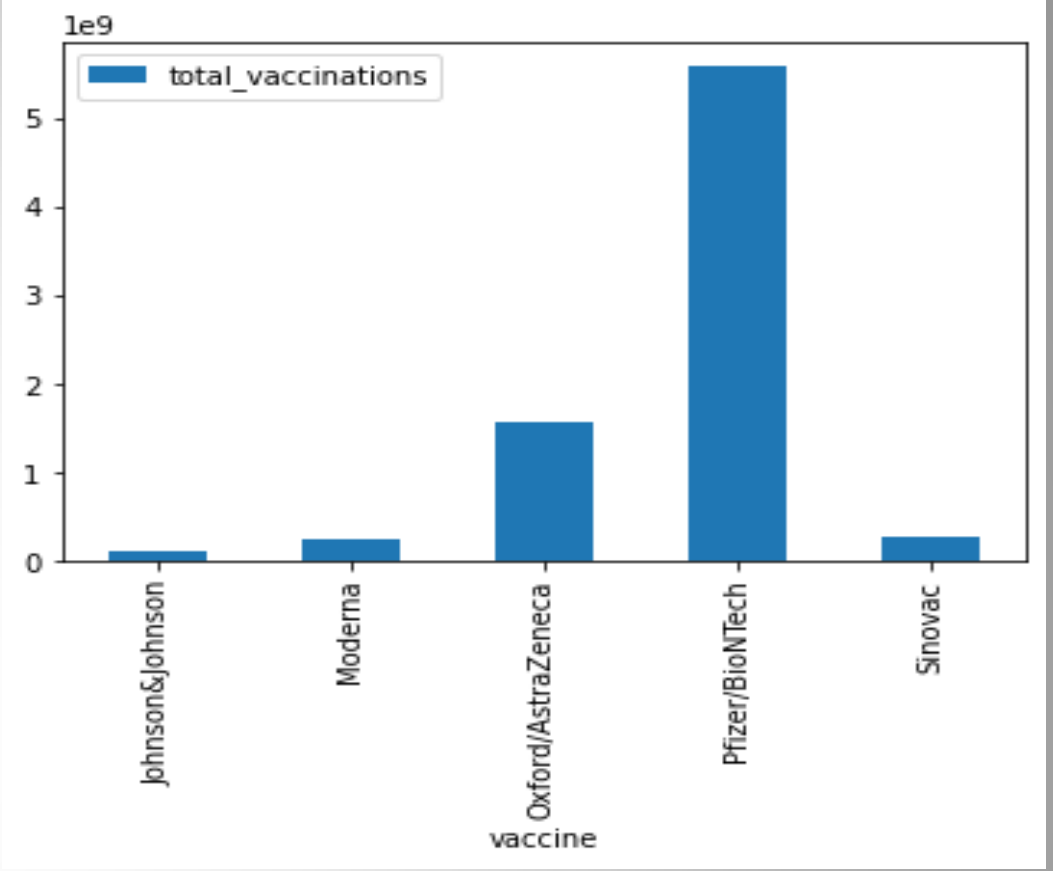


- From the above bar plot we can see that most of the vaccinations is done only in China from December of 2020 till October 2021. India stands in second place in that list. As the population of both the countries is very huge, the number of people vaccinated are more.

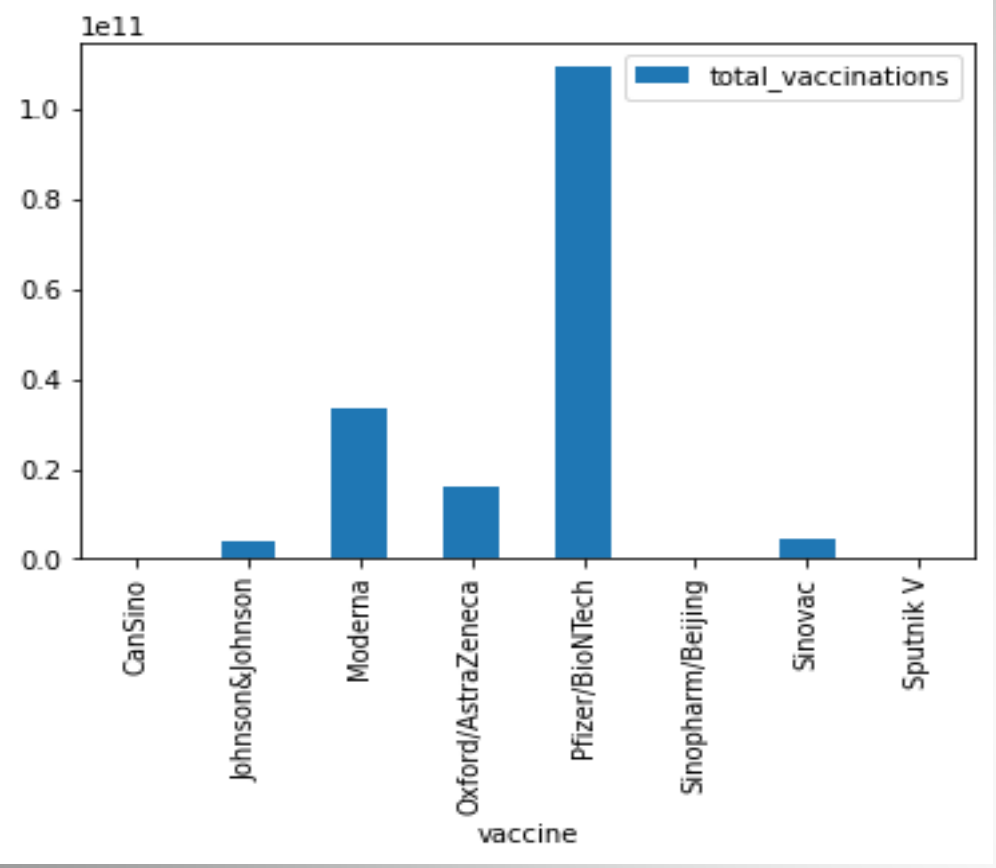
- With a higher population Asia depicts mixed figures of number of vaccinations. Arab Emirates gets the greatest number of vaccinations. Next comes Israel, Azerbaijan, China, Arabia, Singapore, Bahrain, Bhutan, Cyprus and Japan. Syria, Turkmenistan and Yemen have marked the least count of vaccinations along with Afghanistan, Armenia, Myanmar, and Iraq. Average numbers are depicted by Malaysia, Maldives, Vietnam, Uzbekistan, Cambodia, Georgia, Lebanon, Kyrgyzstan, Jordan, Nepal, Northern Cyprus, Timor, Kazakhstan, and Tajikistan.



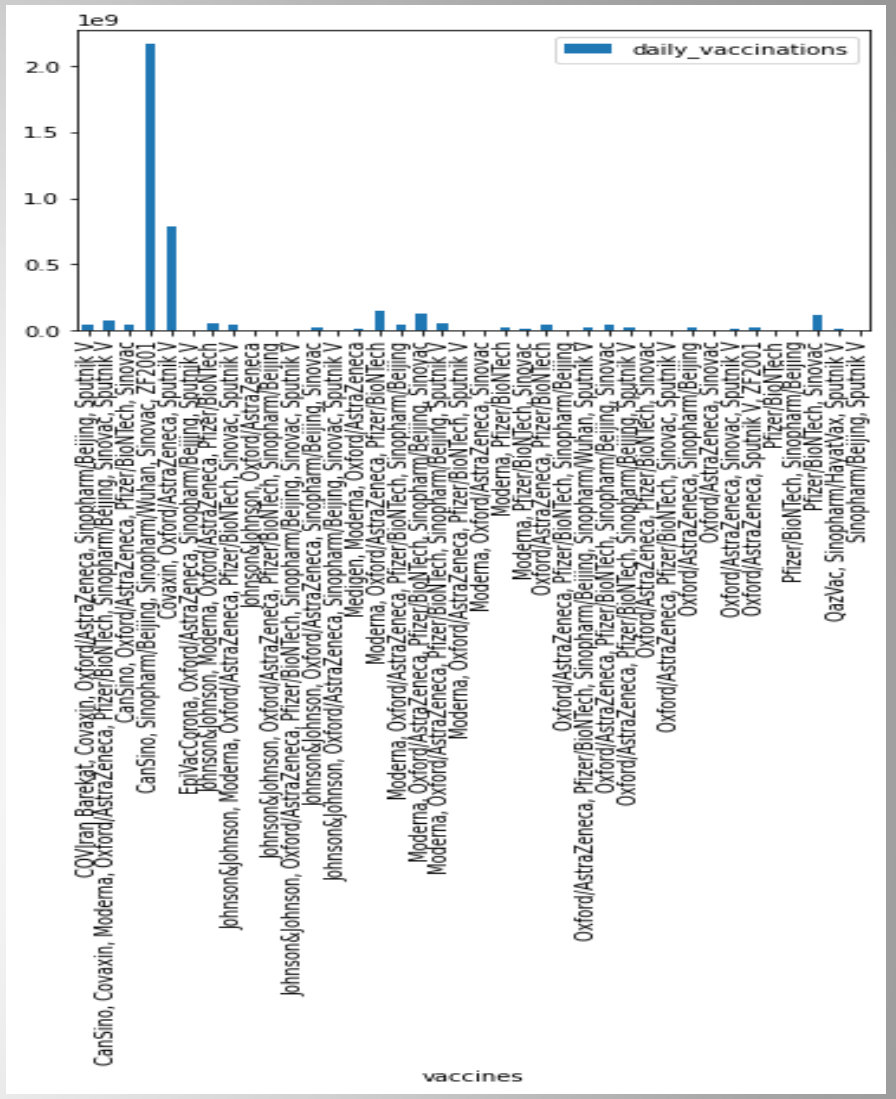
- The above graph shows which Asian continent country received max doses from manufacture.



- From the above bar plot we can see that almost 5.8 Billion Vaccines produced by the company Pfizer/BioNTech were manufactured by the Countries in Asia. A Decent amount of vaccines produced by the company Oxford/AstraZeneca were manufactured and a very few vaccines developed by the companies namely Johnson&Johnson, Moderna and Sinova were manufactured.

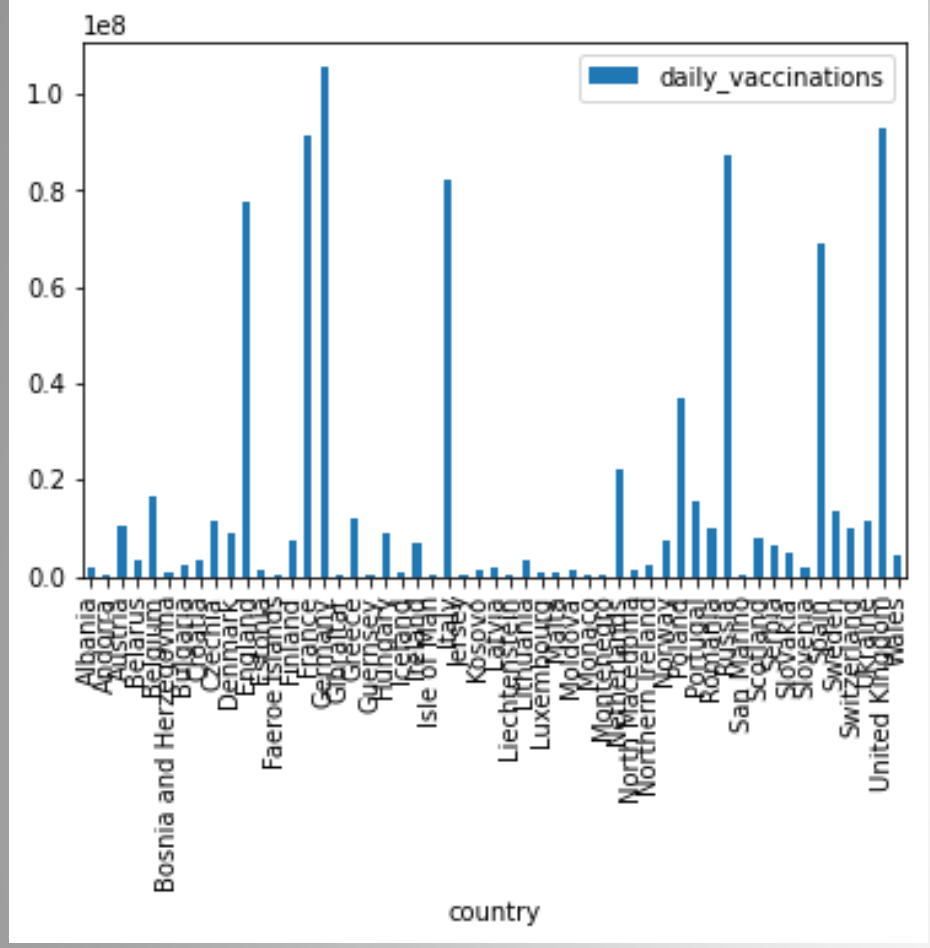


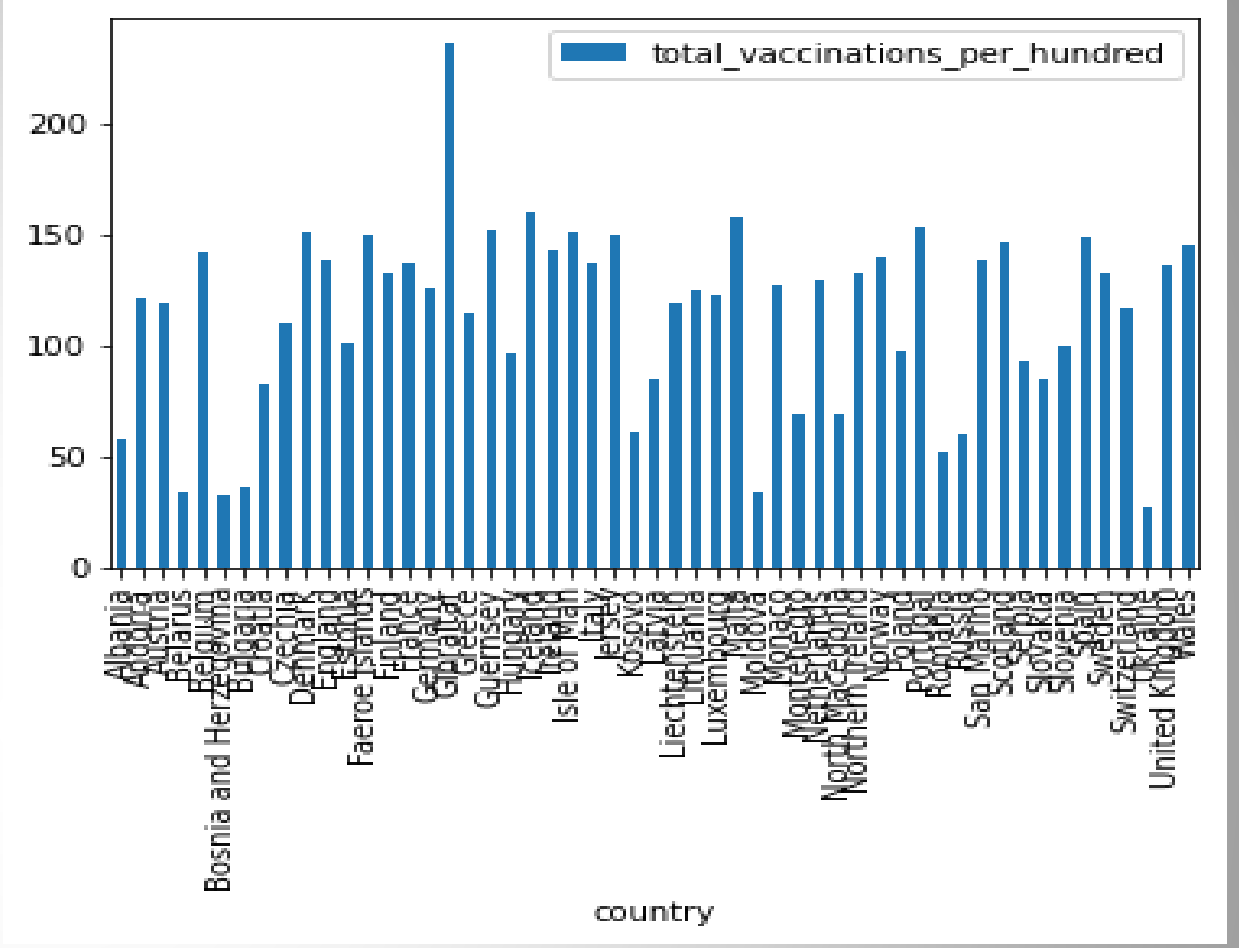
- Most of the Vaccinations were manufactured by Pfizer/BioNTech. Companies like Sputnik V, Sinopharm/Beijing and CanSino has not manufactured any vaccines and few of the vaccines were supplied by the Companies namely Moderna, Oxford/AstraZeneca, Sinovac and Johnson&Johnson.



- The above graph shows which country uses which vaccine (Data vaccination country wise) according to the countries in Asia.

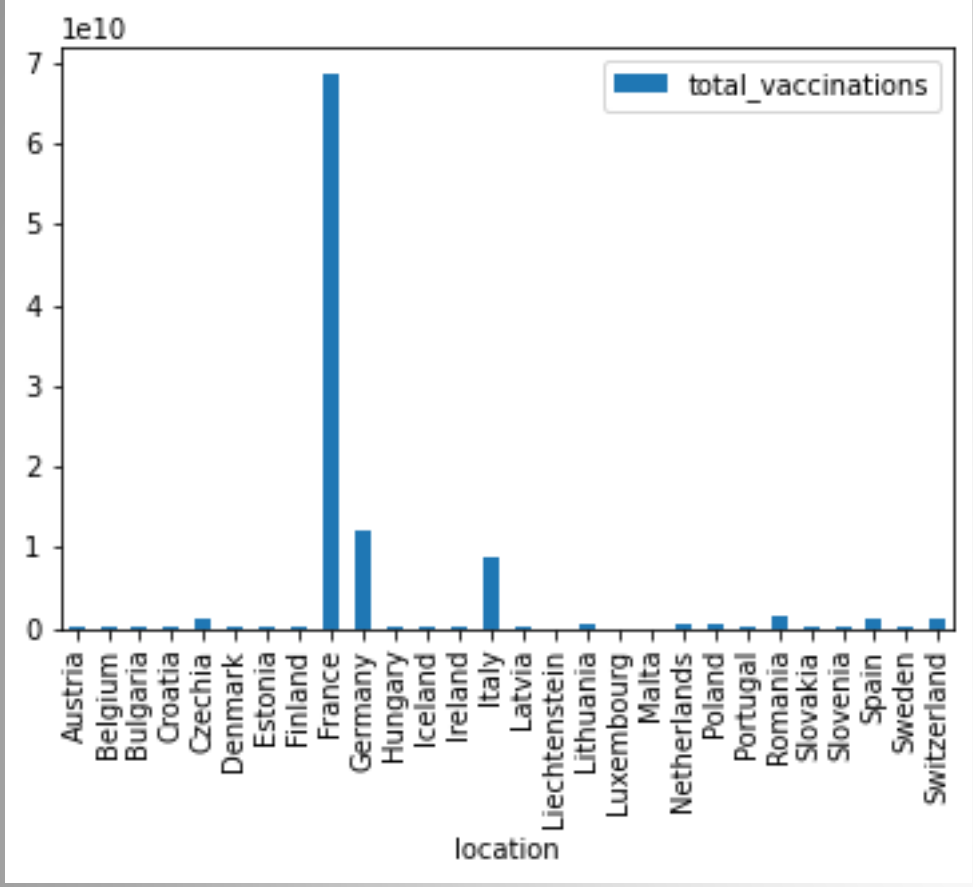
EUROPE

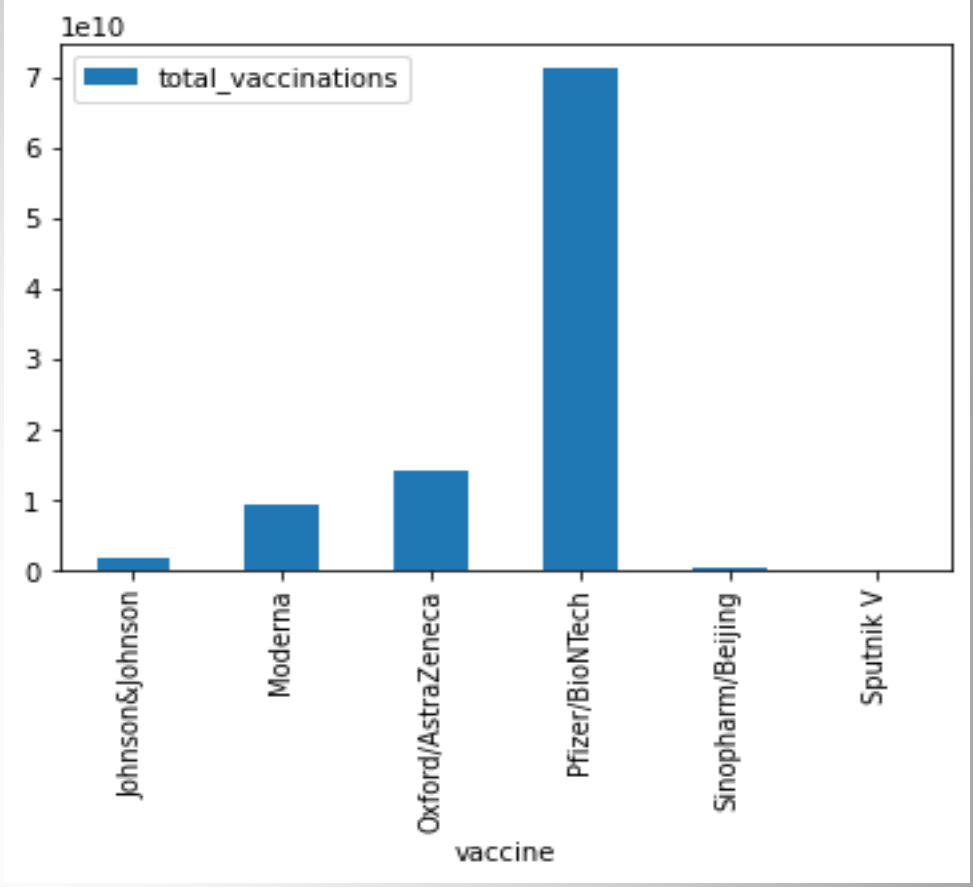




- The above graph shows which European continent country received max doses from manufacture. In Europe, we can see that Germany stands first as more number of Daily Vaccinations are done in Germany. Followed by France , United Kingdom, Russia , Italy, England and Spain. We can also notice that only few vaccinations were done in some countries like Albania, Andora, Bulgaria, Faeroe Islands, Guernsey, Iceland, Isle of Man, Kosovo, Malta , Moldova, Monaco and San Marino.

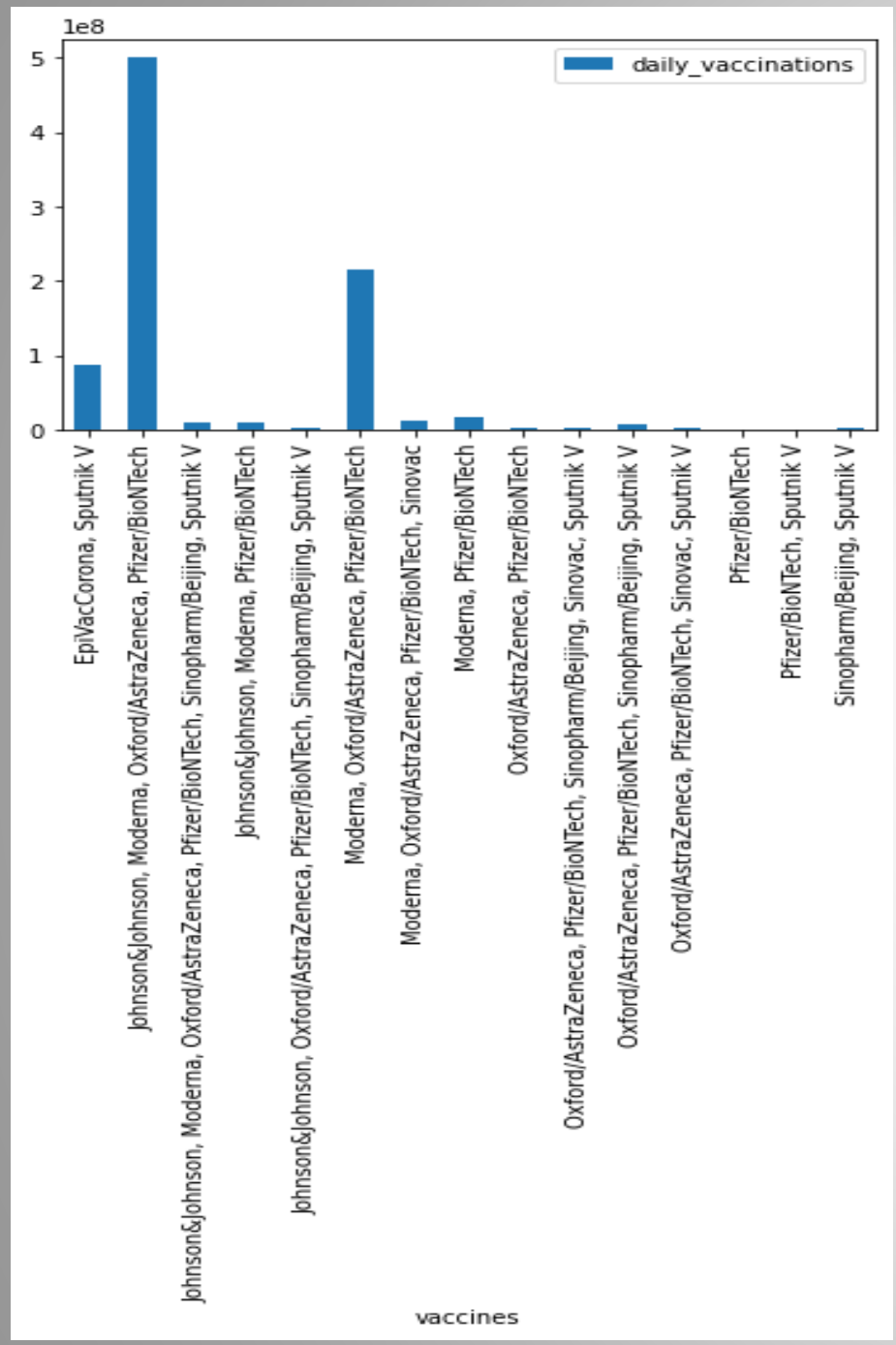
- From the above statistics it is clear that Europe has a smaller number of countries with lower count on vaccinations. On that note Moldova, Belarus, Bulgaria, Croatia and Ukraine stand on a lower rate of vaccinations comparatively. Gibraltar has marked the highest count for vaccinations. Belgium, Sweden, Malta, Iceland, Faroe Islands, Portugal, Russia, Wales, Scotland and United Kingdom have had a fair number of vaccinations along with Andorra, Netherlands, Norway, Austria, Poland and Czechia.





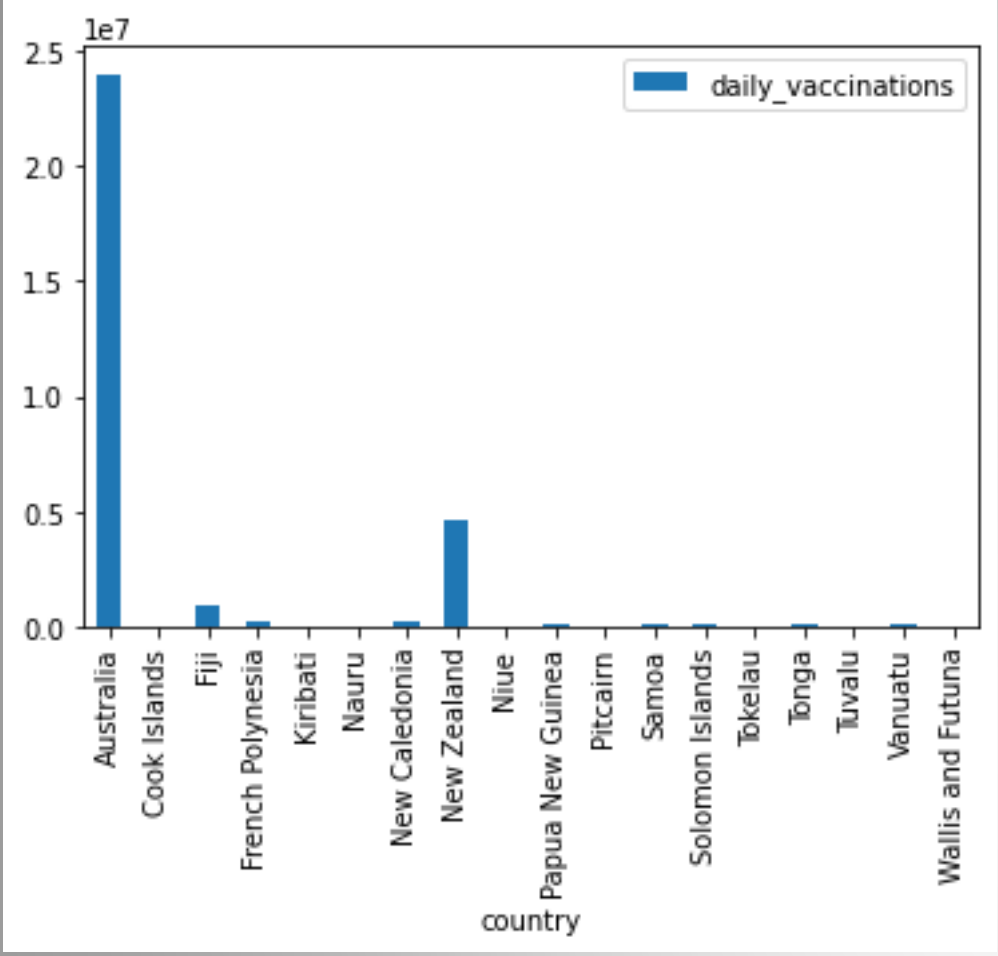
- When coming to the total vaccinations, France stands first with more number of vaccinations done followed by Germany and then Italy. All the remaining countries namely Austria, Belgium, Bulgaria, Croatia, Czechia, Denmark, Estonia, Finland, Hungary, Iceland, Ireland, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and Switzerland were almost approximately equal and the vaccinations done are also very less.

- Most of the vaccines that were taken by the people living in the European Countries is supplied by the Company Pfizer/BioNTech. Almost 70% of the people has taken the vaccine produced by Pfizer/BioNTech. Very few of them has taken the shot of the vaccine developed by the companies Oxford/AstraZeneca, Moderna, Johnson&Johnson and almost 1% of the total vaccines were developed by Sinopharm/Beijing.

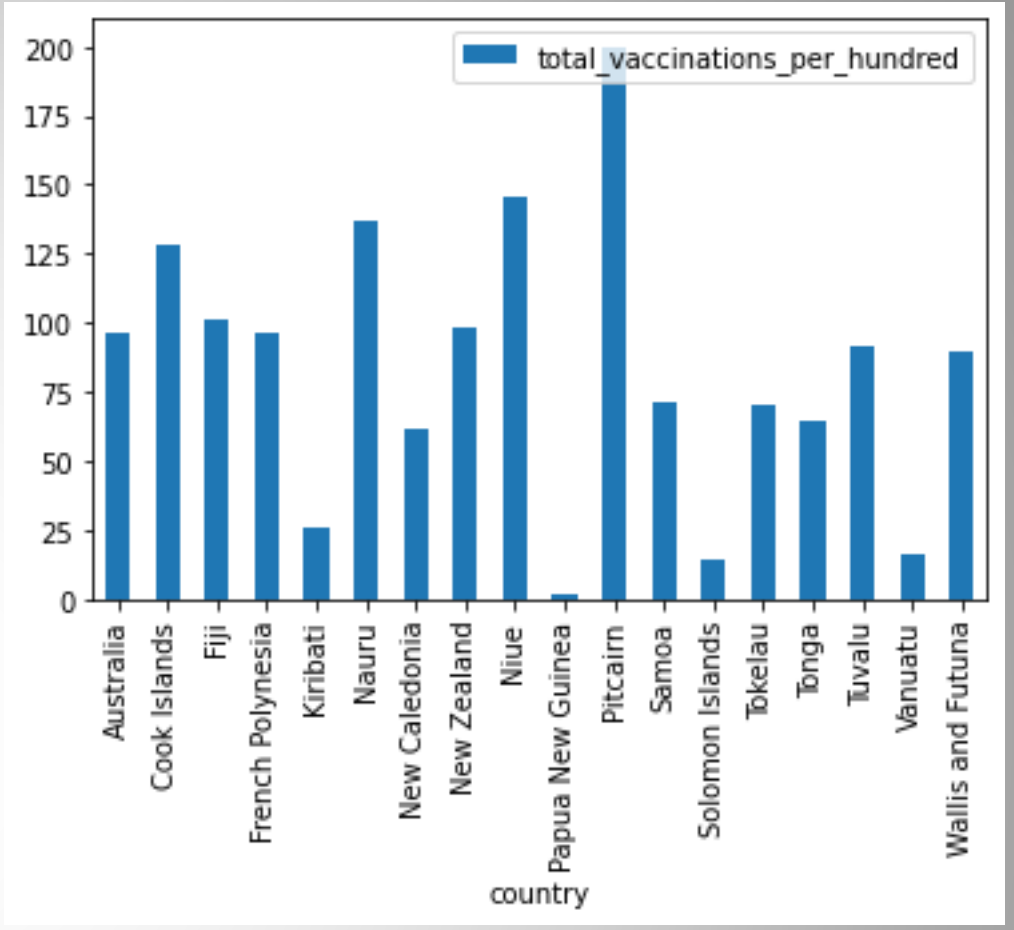


- The graph shows which country uses which vaccine (Data vaccination country wise) according to the countries in Europe.

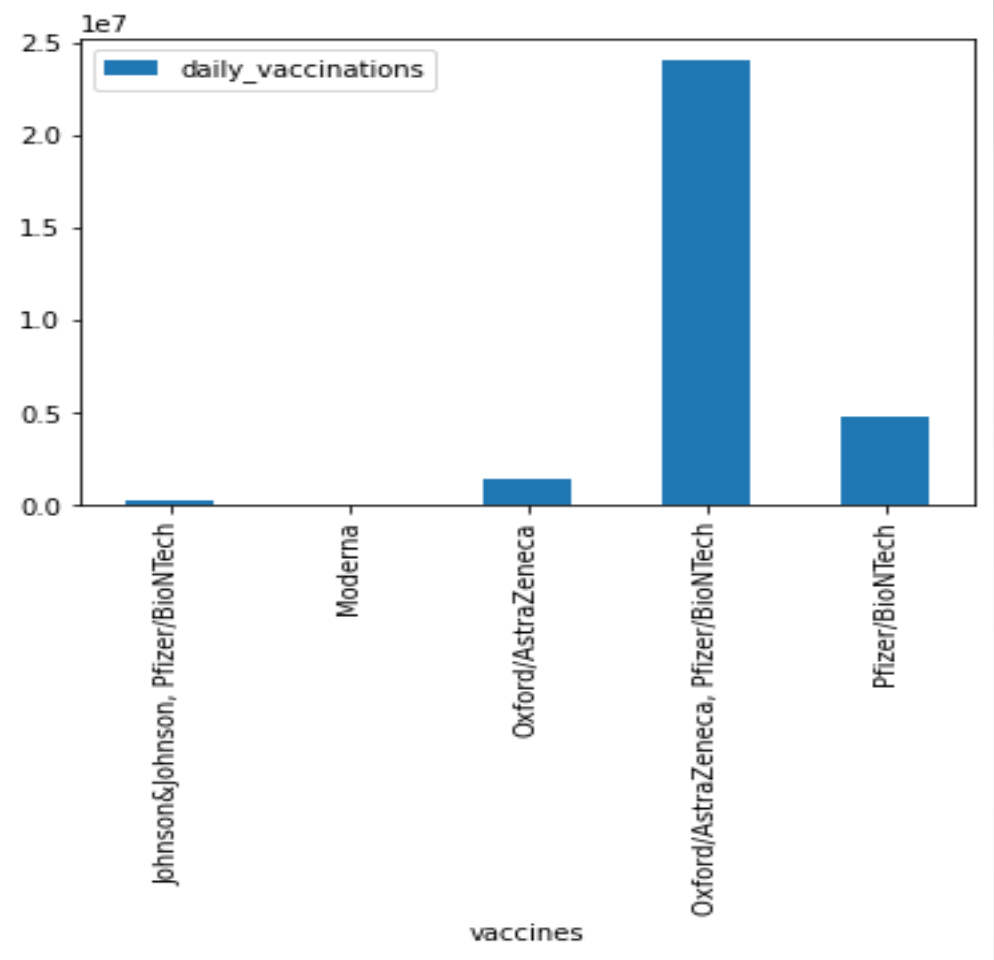
OCEANIA



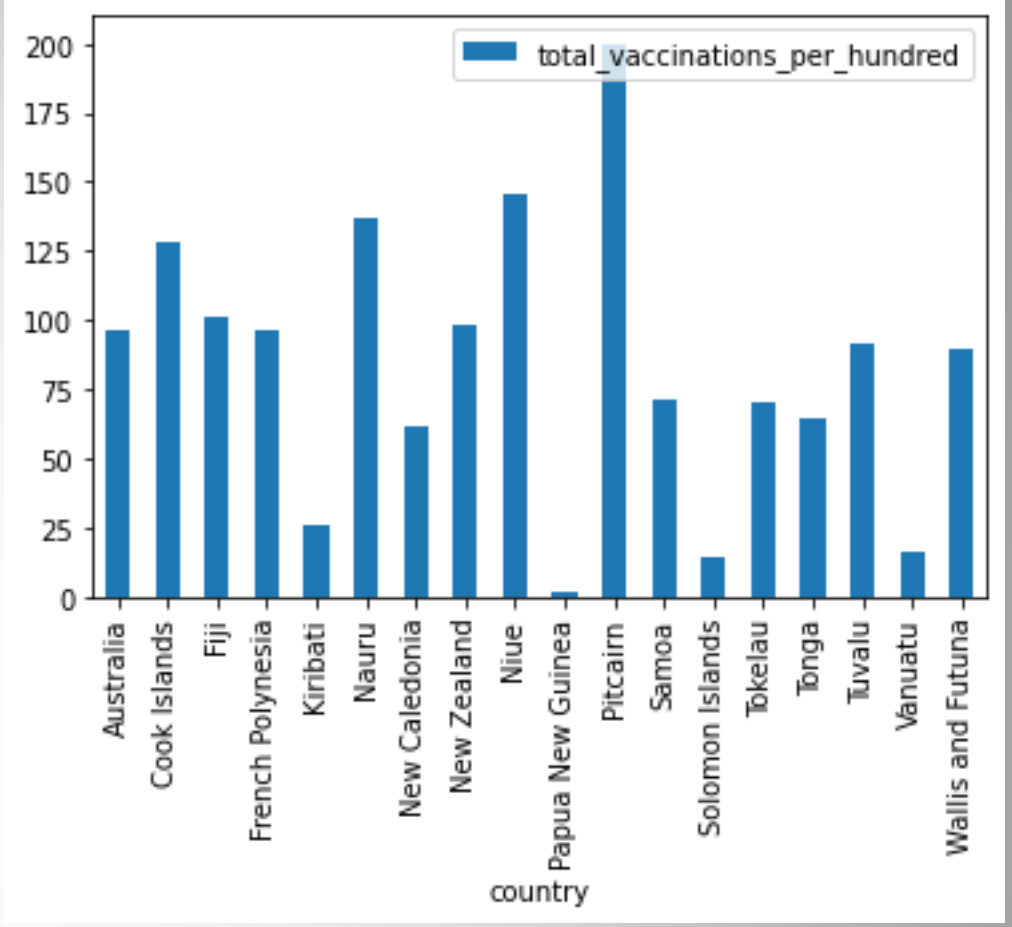
- When comes to the continent Oceania, most of the daily vaccinations were done at Australia followed by New Zealand. Countries like French Polynesia, New Caledonia, Papua New Guinea, Samoa, Solomon Islands, Tonga and Vanuatu might have just started their daily vaccinations.



- Pitcairn takes up the most number of total vaccinations as per the figures denoted above. After Pitcairn Cook islands, Nauru, Niue and Tuvalu have a greater number comparatively, followed by Australia, Fiji, French Polynesia and Wallis and Futuna. The lowest number of figures are depicted by Papa New Guinea followed by Solomon Islands, Vanuatu and Kiribati.

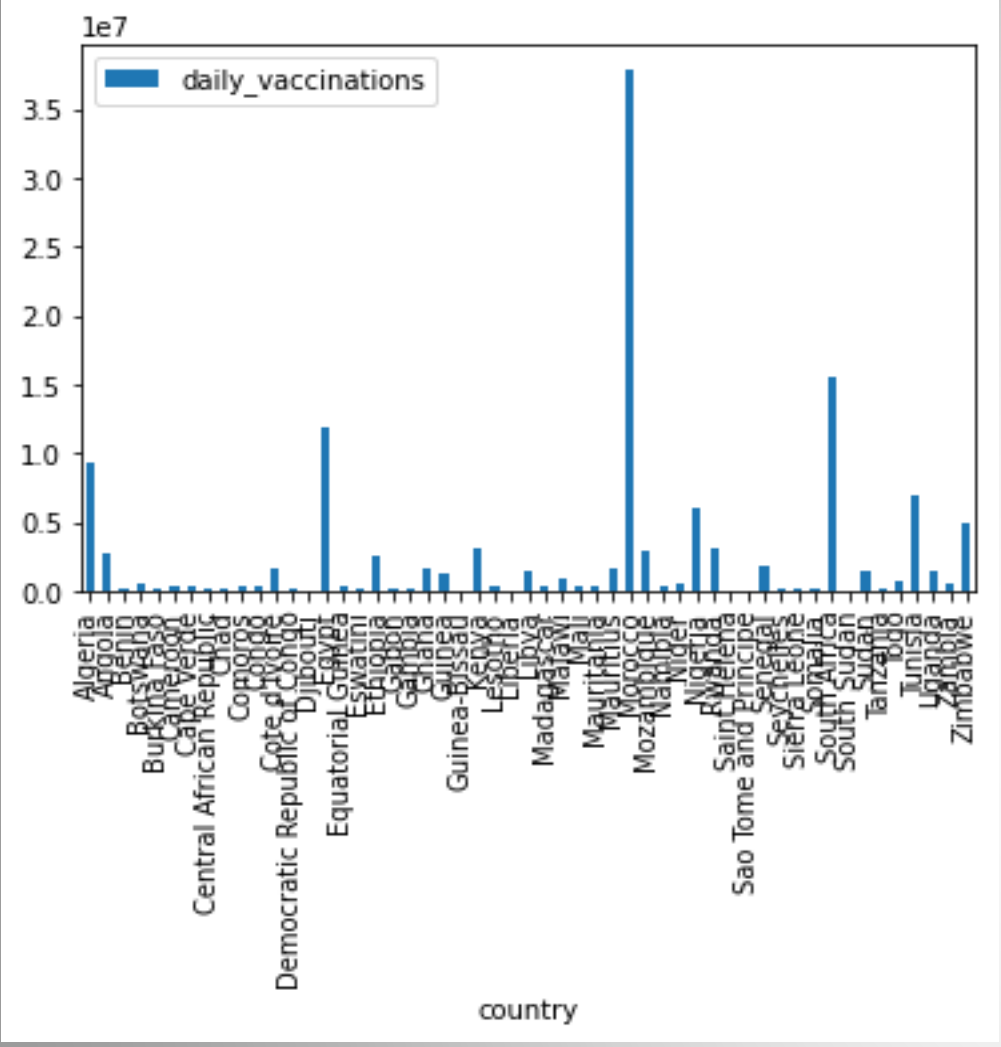


- When coming to the daily Vaccinations, most of the Vaccinations were produced by the Companies Oxford/AstraZeneca, Pfizer/BioNTech. No vaccines were produced by the company Moderna.

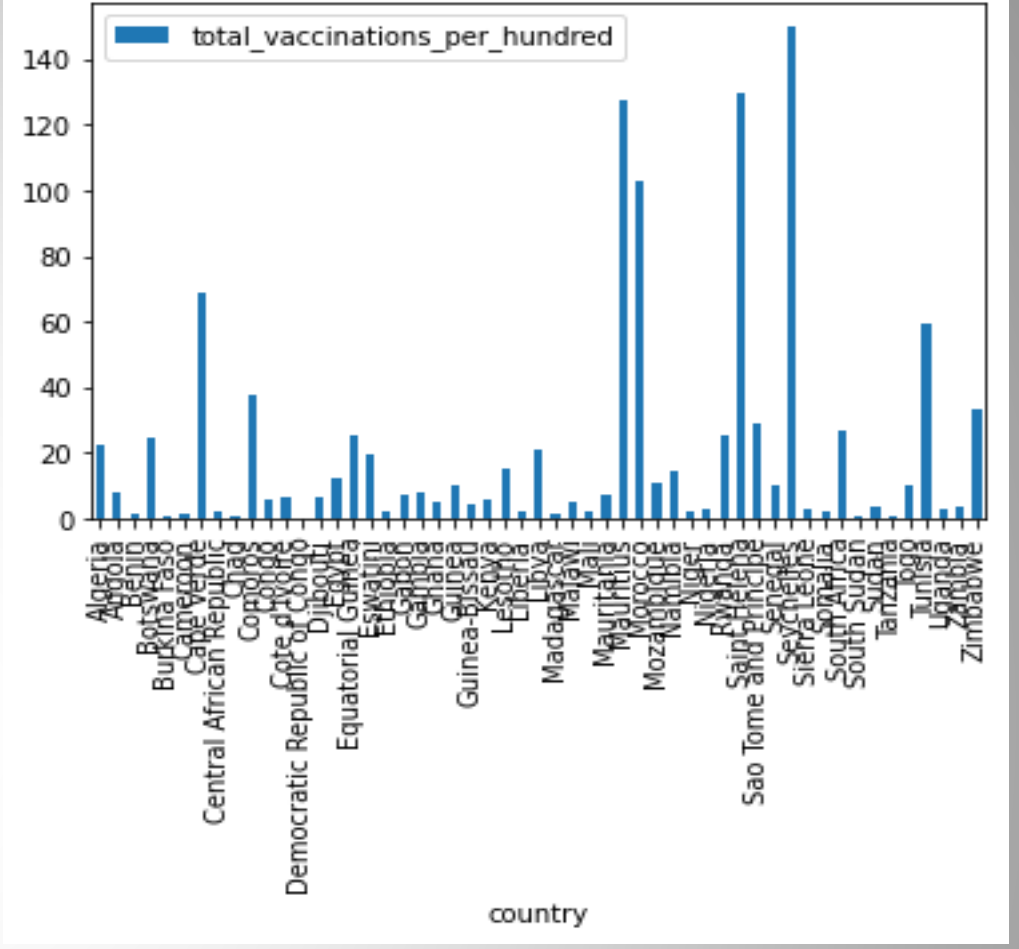


- Of the total Vaccinations done from the continent Oceania, we can see that most vaccinations are done by the country Pitcairn followed by Niue, Nauru, Cook Islands, Fiji, Australia, French Polynesia, New Zealand and least number of vaccinations were taken by the people living in the country Papua New Guinea, Solomon Islands, Vanuatu and Kiribati.

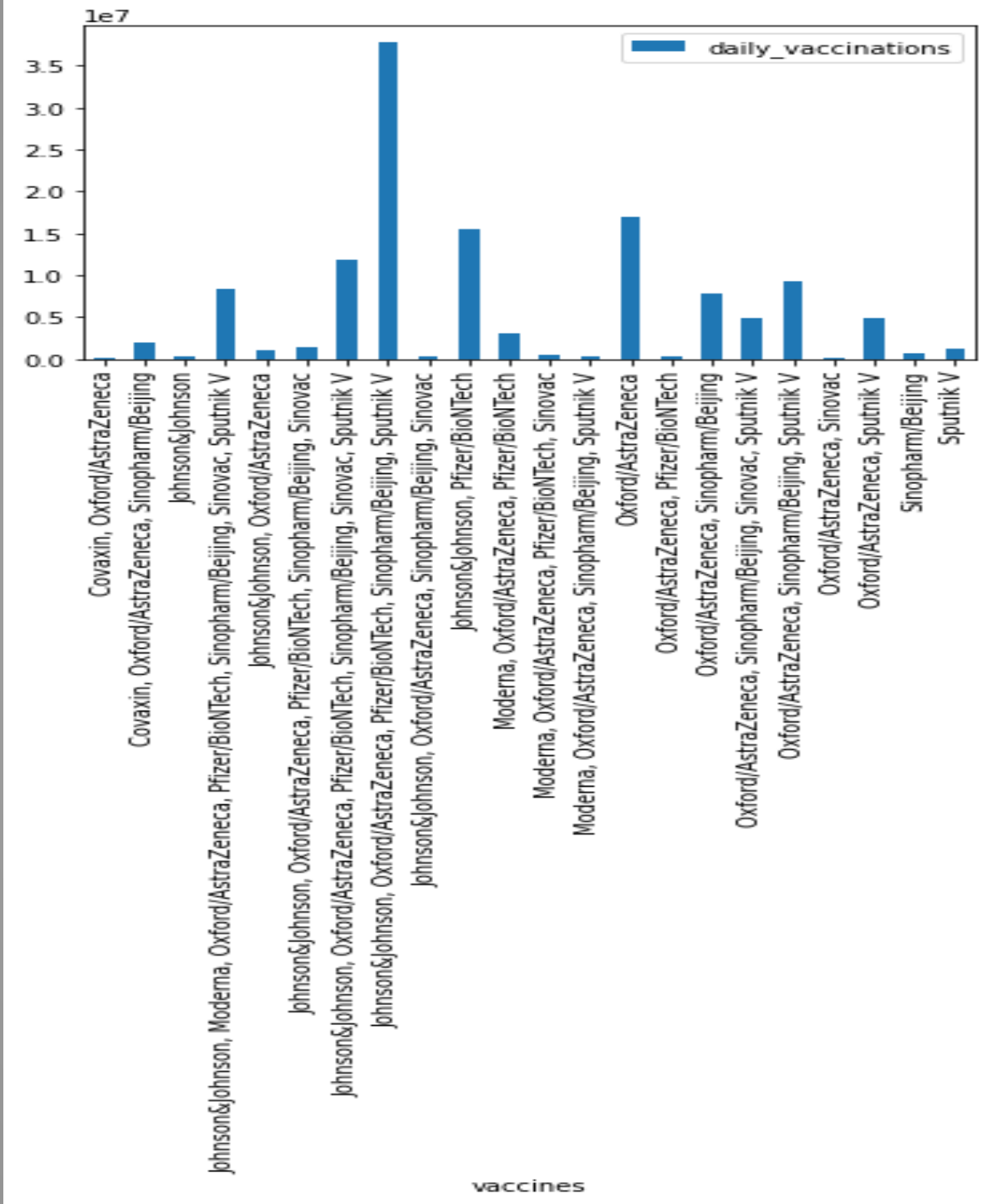
AFRICA



- The above graph shows daily vaccinations in African countries.

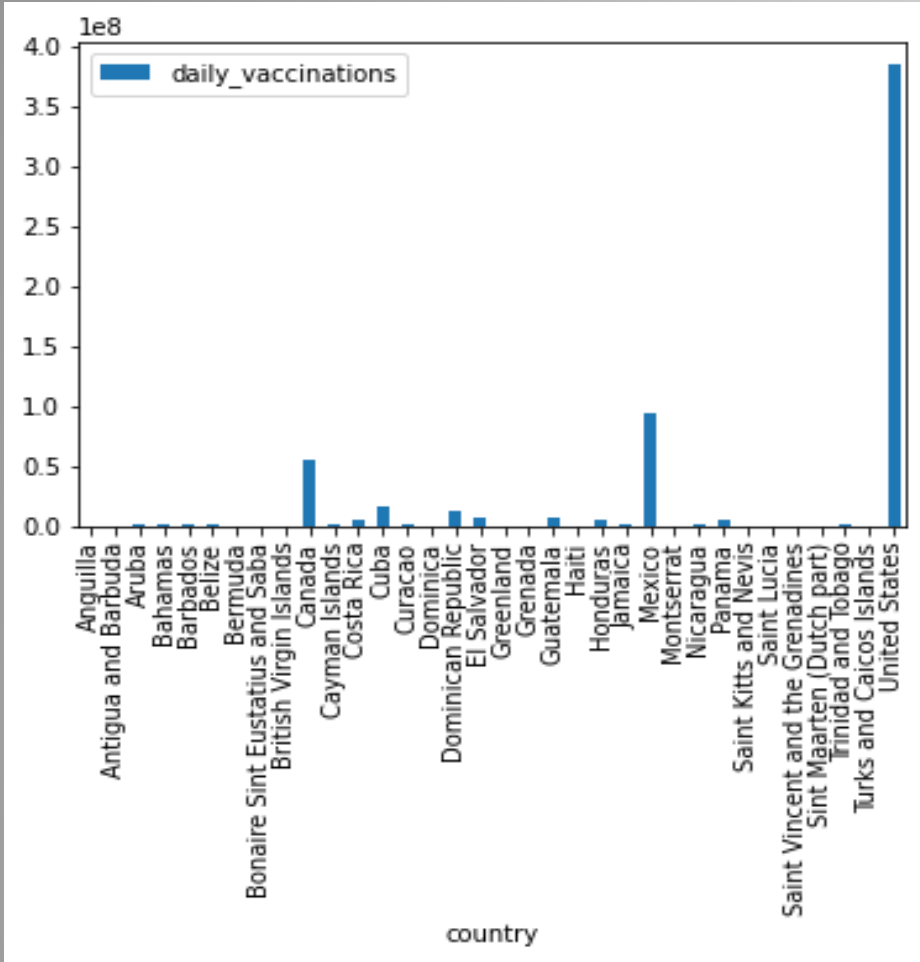


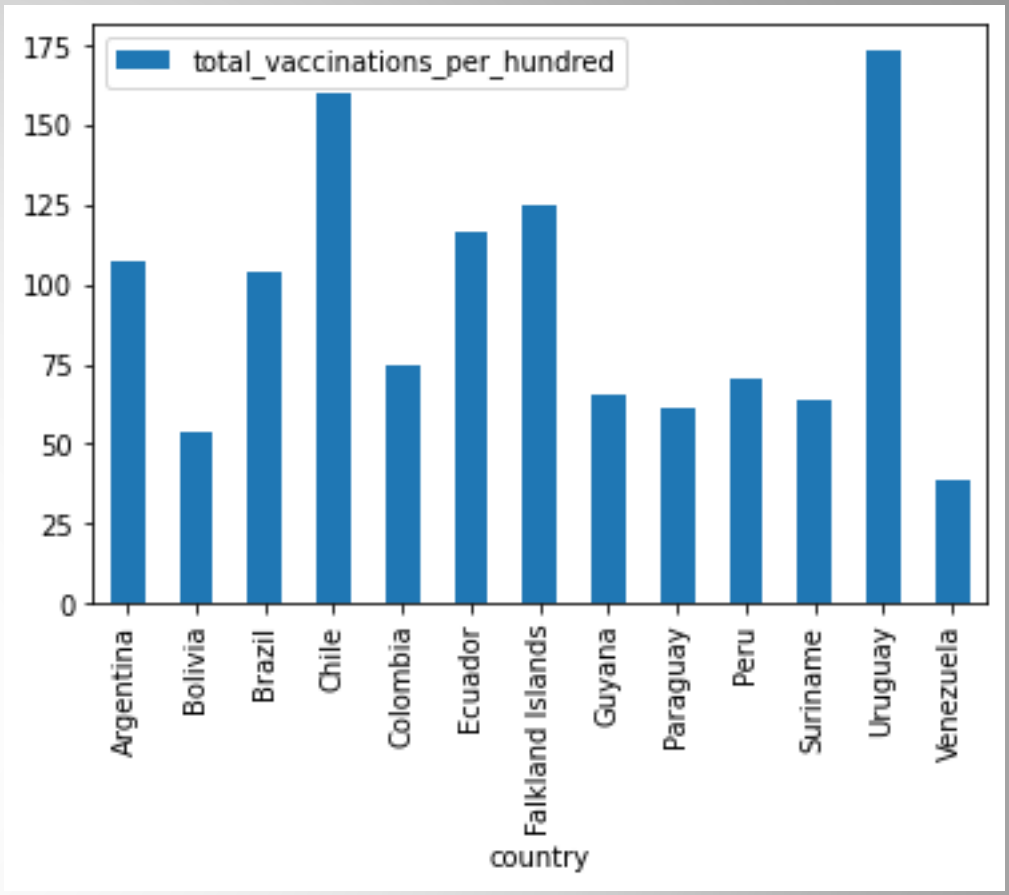
- From the above picture of total number of vaccinations done, there are countries with more number of vaccinatons and countries with less number of vaccinations. The countries with lowest rate of vaccinations are Benin, Cameroon, Central African Republic, Comoros, Ethopia, Liberia, Madgascar,Mali, Niger and Nigeria as well. Seychelles takes up the maximum number of vaccinations whereas Saint Helena, Mauritus and Morocco have average rate of vaccinations.



- The above graph shows which country uses which vaccine (Data vaccination country wise) according to the countries in Africa.

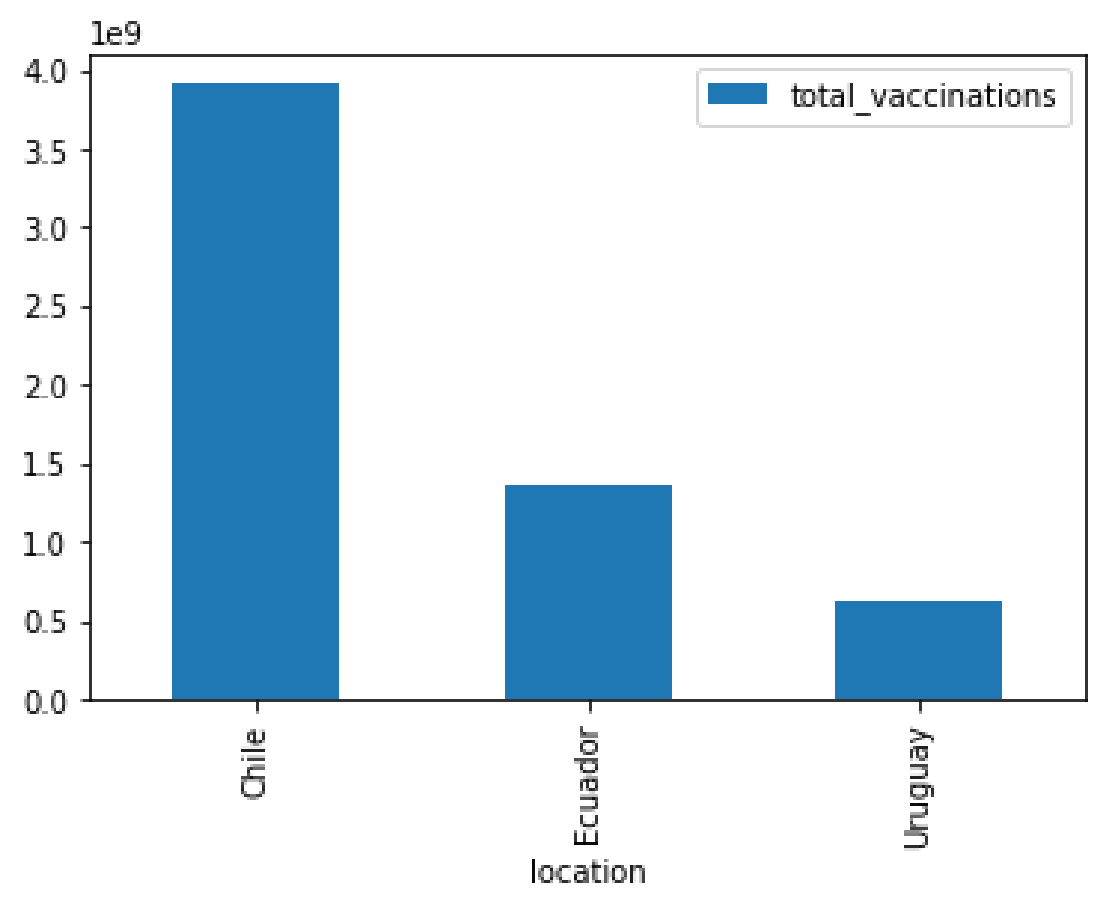
SOUTH AMERICA

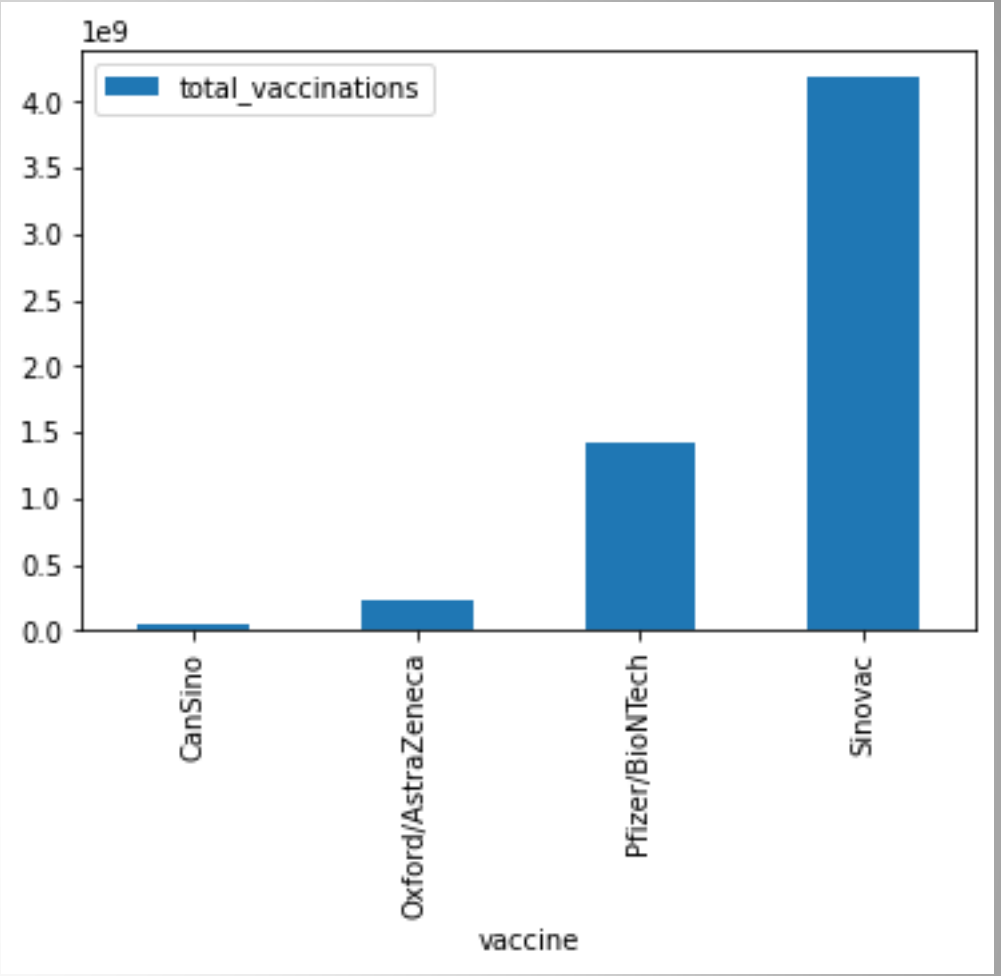




- When you observe the daily vaccination status of the countries present in South America Continent, we can see most of the countries has very less progress in their Daily vaccines. Only few countries like United States, Mexico and Canada has some progress.

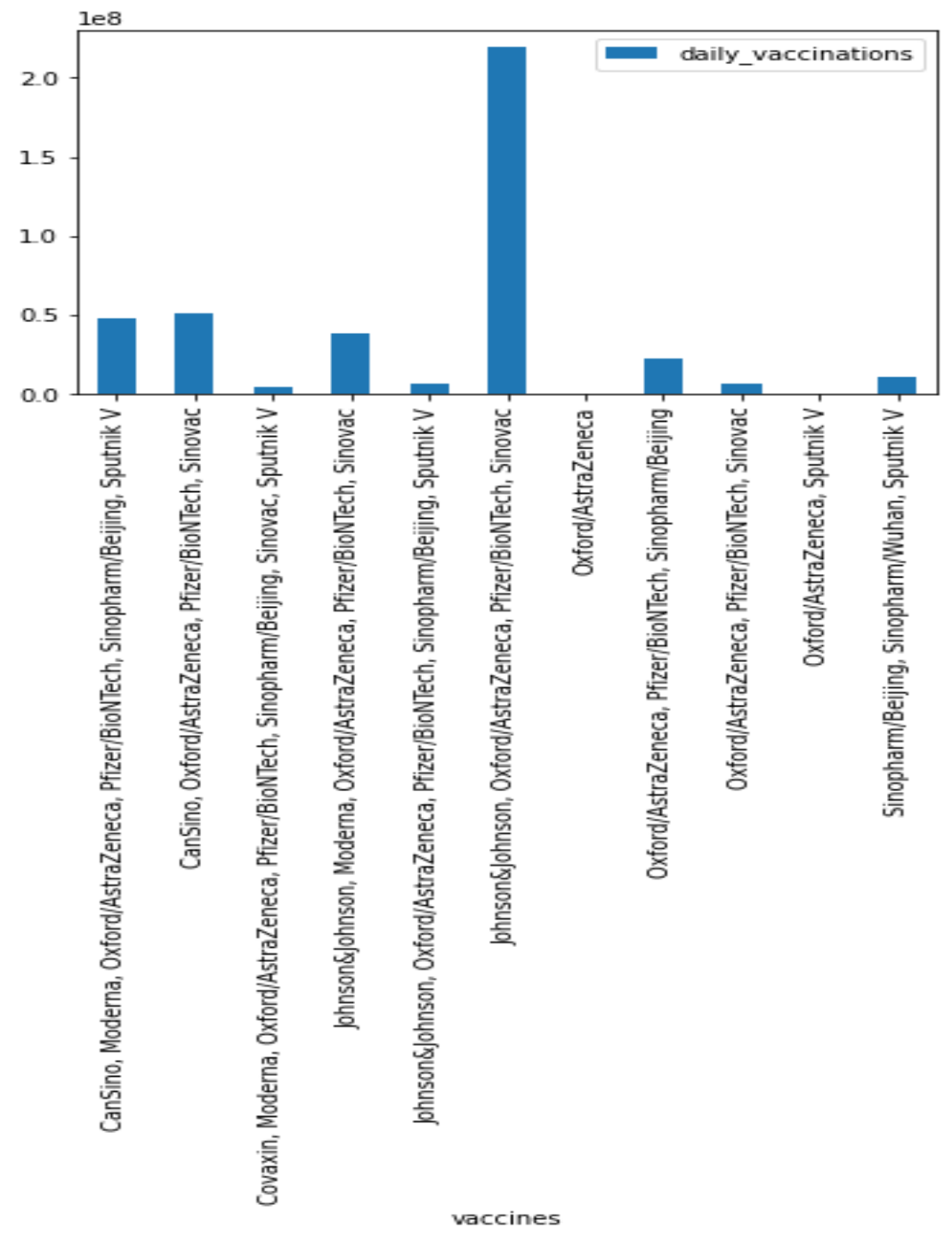
- South America has a real fair set of vaccinations with almost no country bearing an extremely lower number of vaccinations. Among the above, Uruguay gets the highest number of vaccinations followed by Chile, Argentina, Ecuador, Falkland Islands. The countries with a mean value of number of vaccinations include Bolivia, Guyana, Columbia, Suriname and Venezuela.





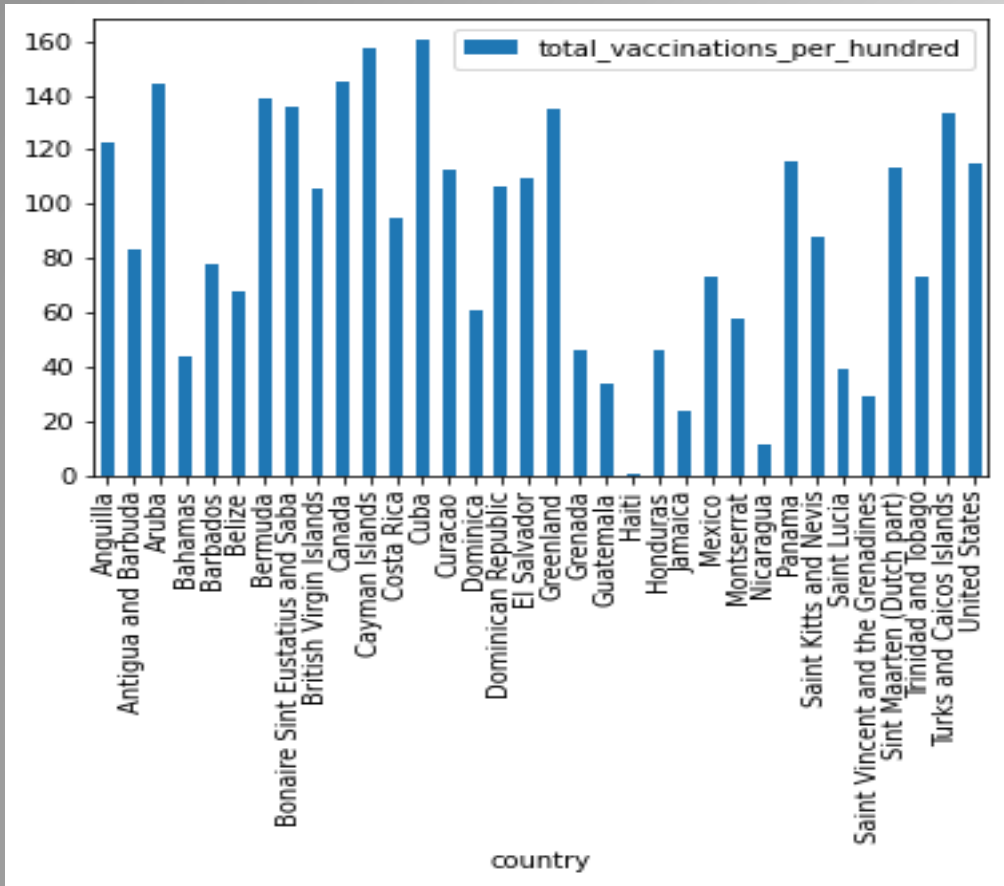
- The above graph shows which SOUTH AMERICAN continent country received max doses from manufacture.

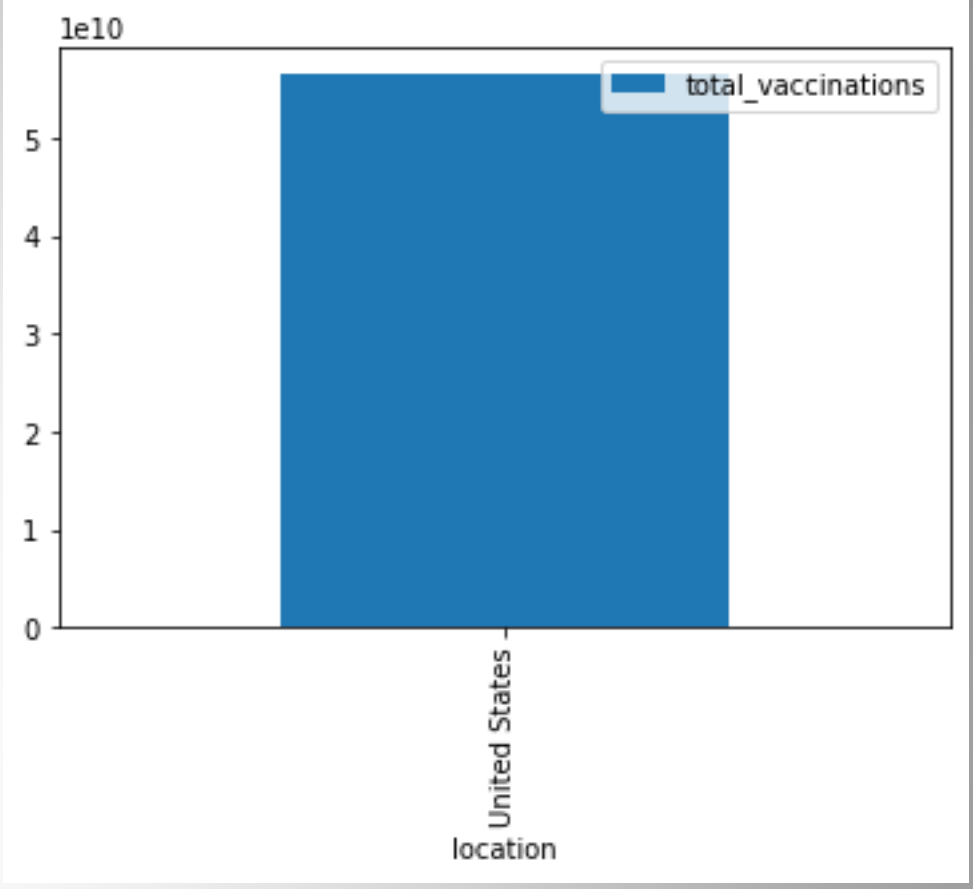
- Sinovac has produced the most number of Vaccinations followed by Pfizer/BioNTech, Oxford/AstraZeneca and CanSino.



- The above graph shows which country uses which vaccine (Data vaccination country wise) according to the countries in South America.

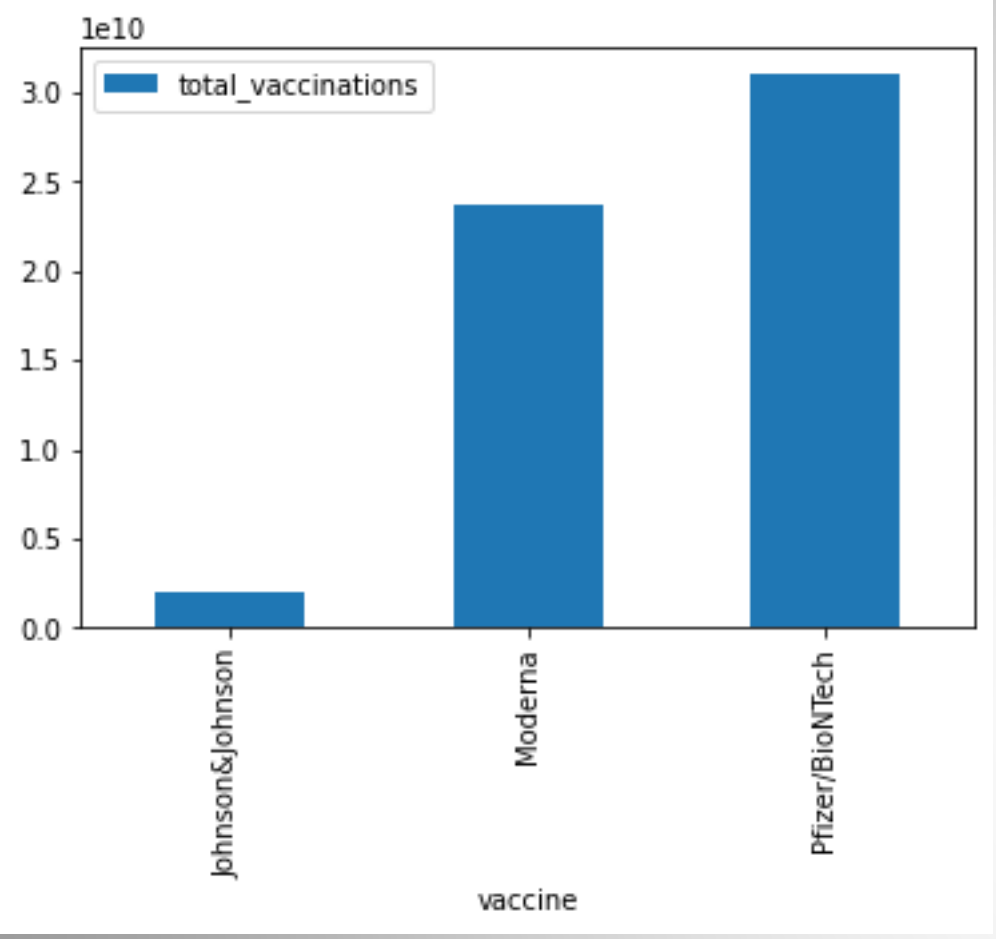
NORTH AMERICA



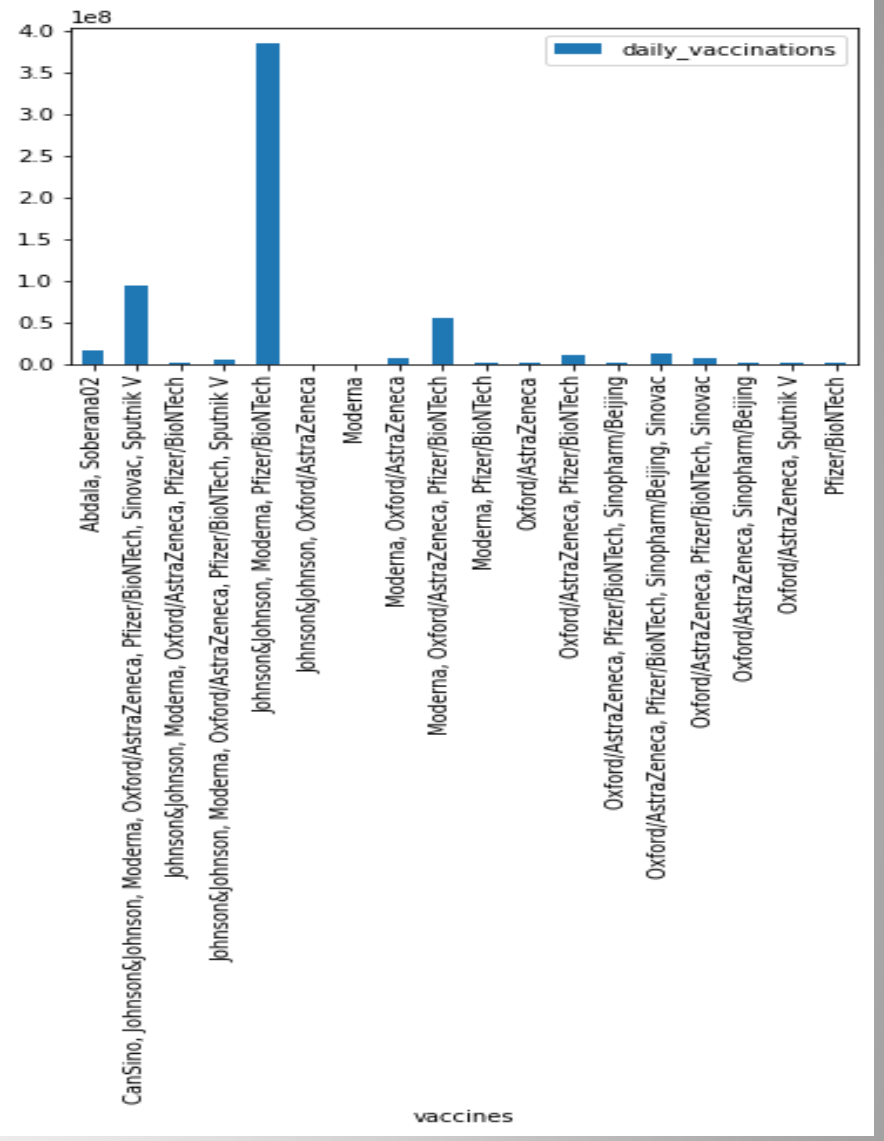


- In the North American figures of vaccinations, Cuba stands the highest in the total number of vaccinations followed by Aruba, Cayman Islands, Panama, St. Maarten, Turks and Caico Islands, United States, Anguilla and Canada. Among the average count of vaccinations Bahamas, Dominica, Honduras, St. Lucia, Trinidad, and Tobago mark their number. Haiti followed by Nicaragua take up the least count of vaccinations.

- The above graph shows which North American continent country received max doses from manufacture.

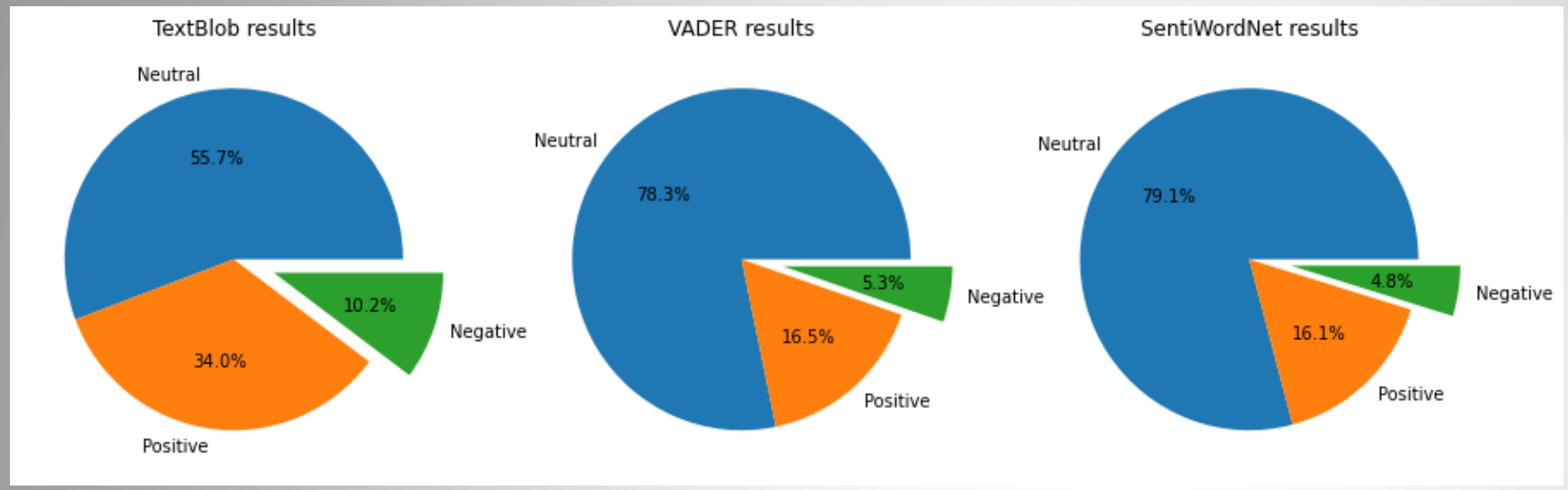


- Of all the Vaccines produced, Pfizer/BioNTech company produced the most number of vaccines. Moderna also produced about 40% of the total vaccines. The least amount of vaccines were produced by Johnson&Johnson.



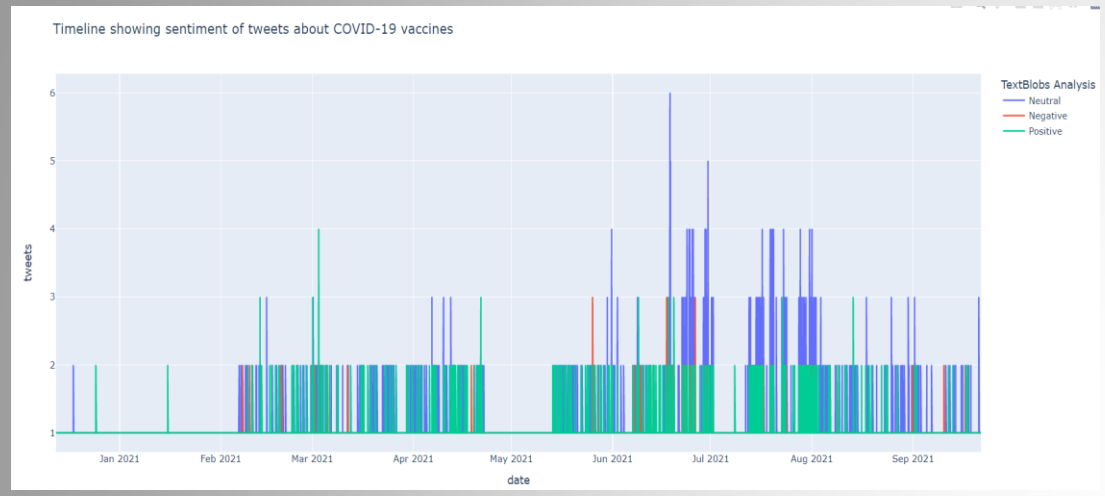
- The above graph shows which country uses which vaccine (Data vaccination country wise) according to the countries in North America.

OUTPUT OF ALL 3 SENTIMENT ANALYSIS USING PIE CHART:

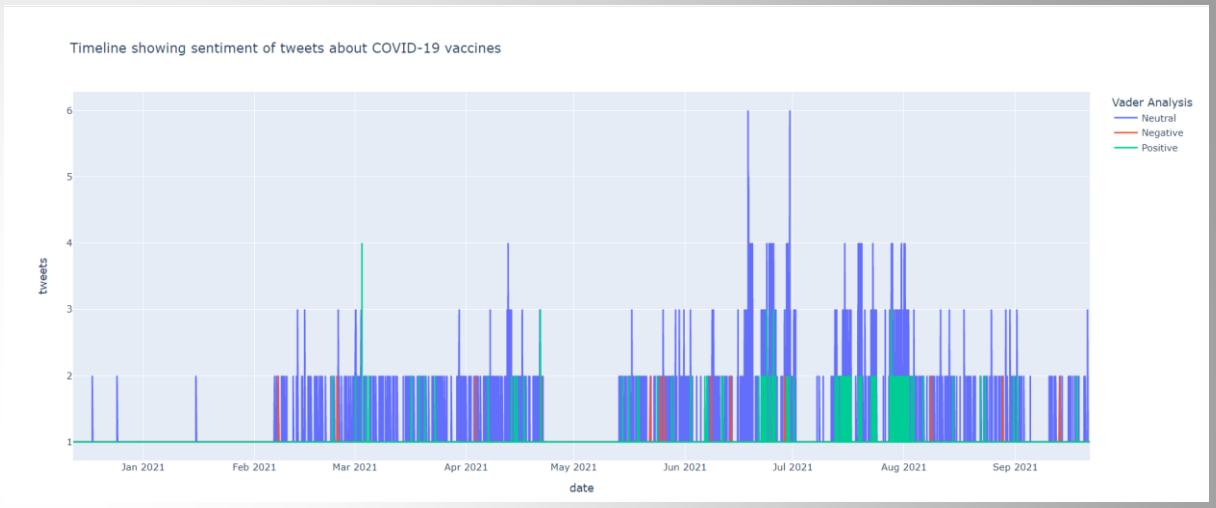


- Timeline of sentiments of tweets about vaccines using TextBlob's analysis.

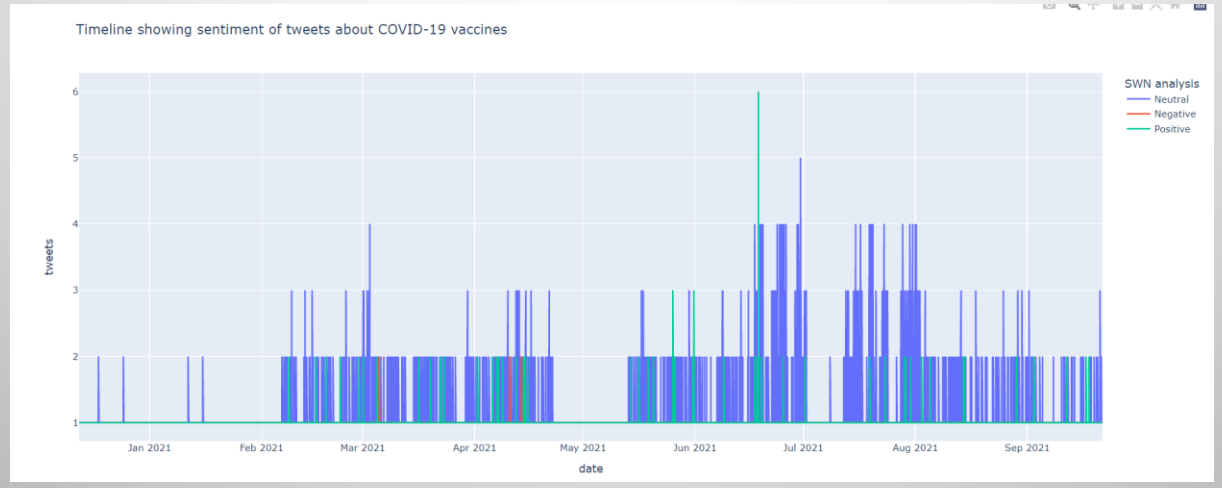
Timeline of sentiments of tweets about vaccines using TextBlob's analysis :



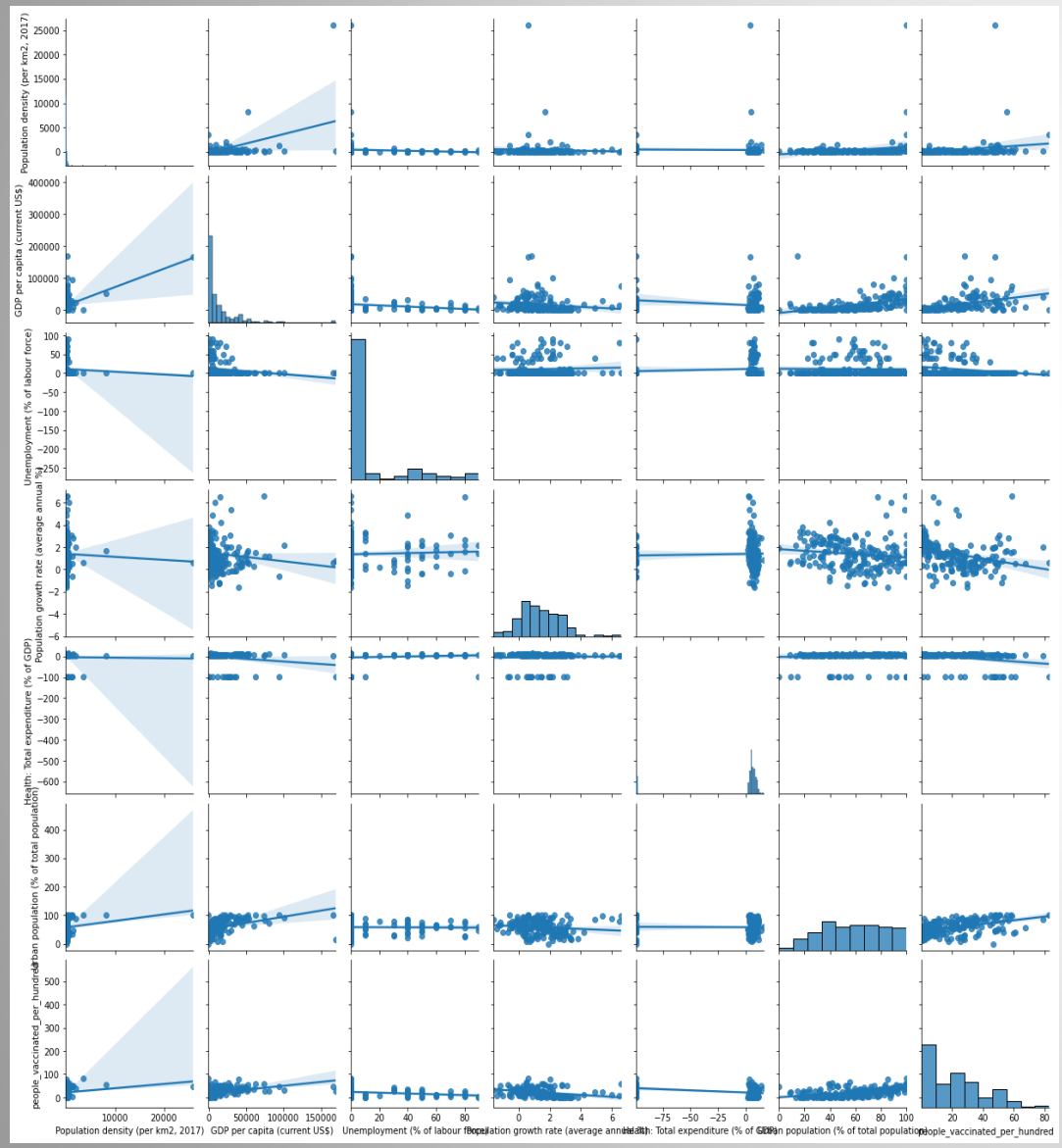
Timeline of sentiments of tweets about vaccines using VADER analysis :



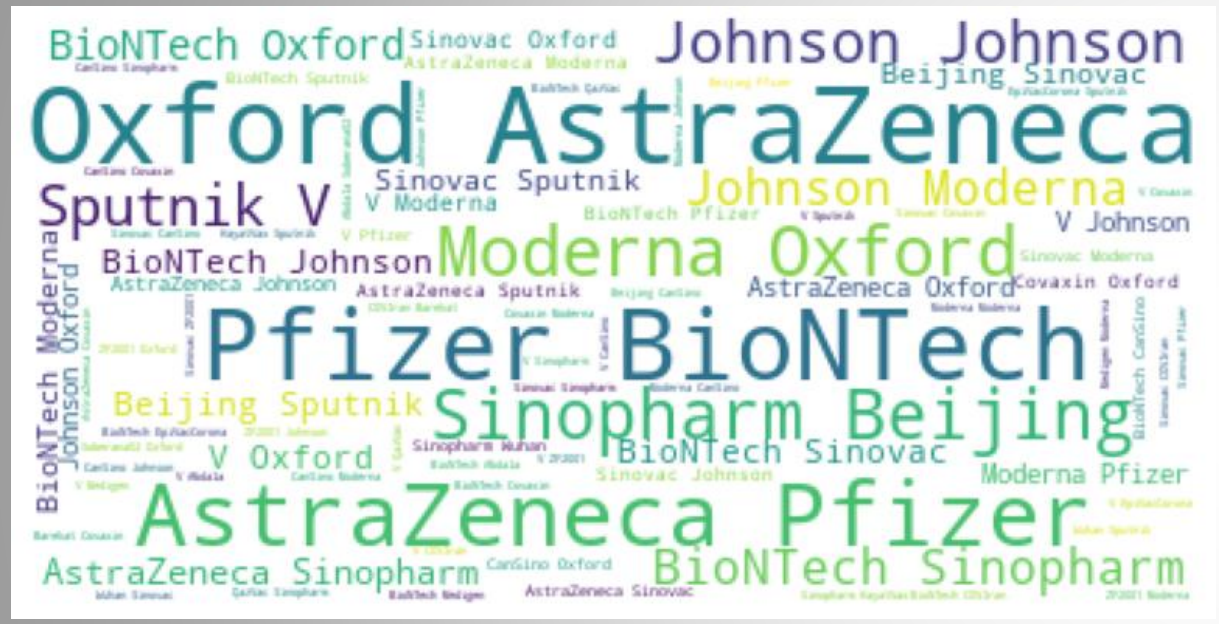
Timeline of sentiments of tweets about vaccines using SentiWordNet analysis :



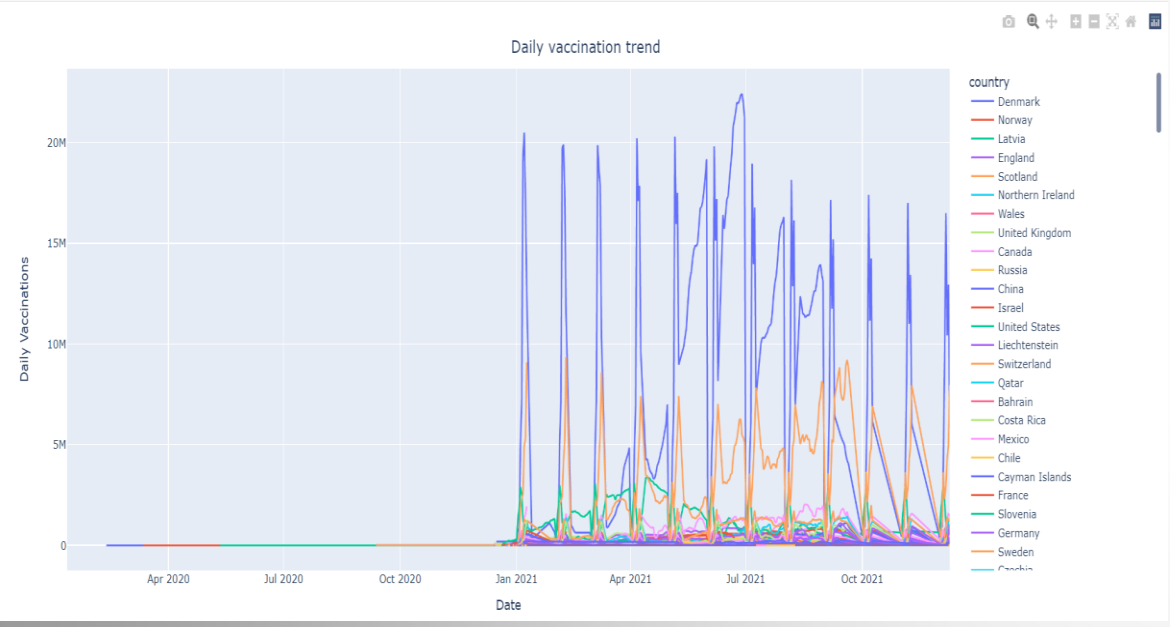
DEMOGRAPHICS AND VACCINATION CORRELATIONS:



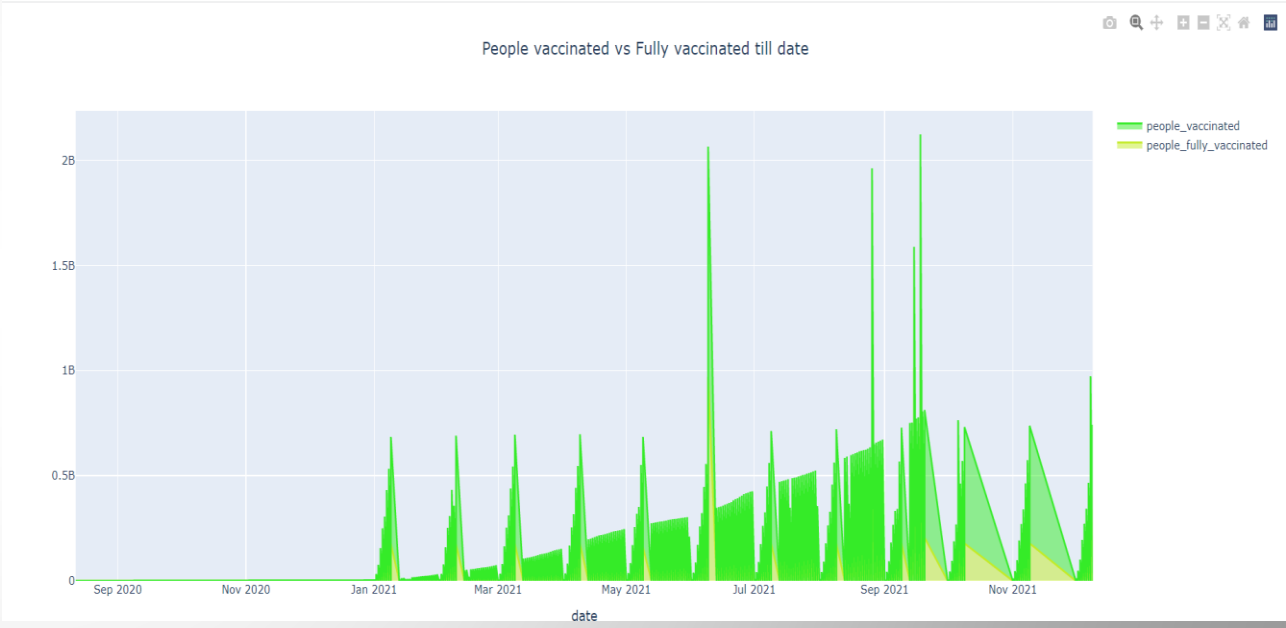
Word Cloud for Vaccine names in dataCV :



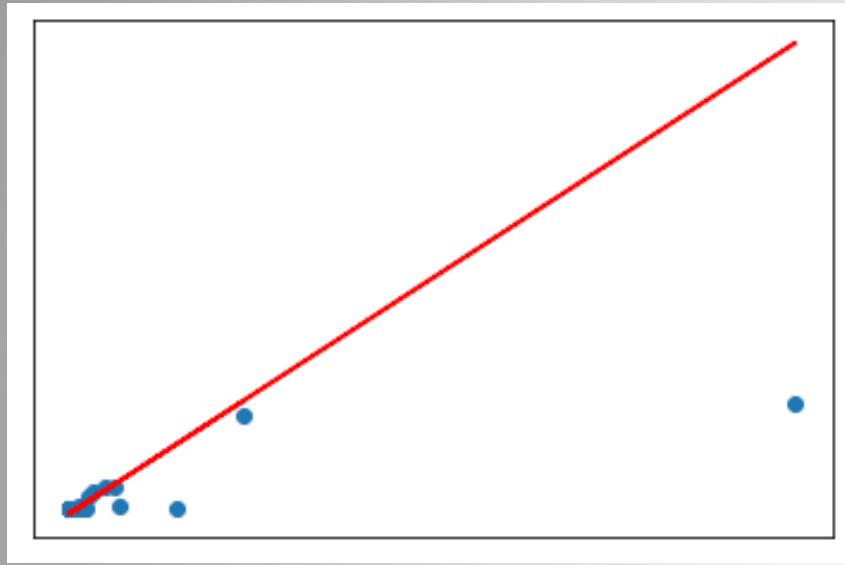
Daily vaccination timeline:



People vaccinated vs fully vaccinated till date :

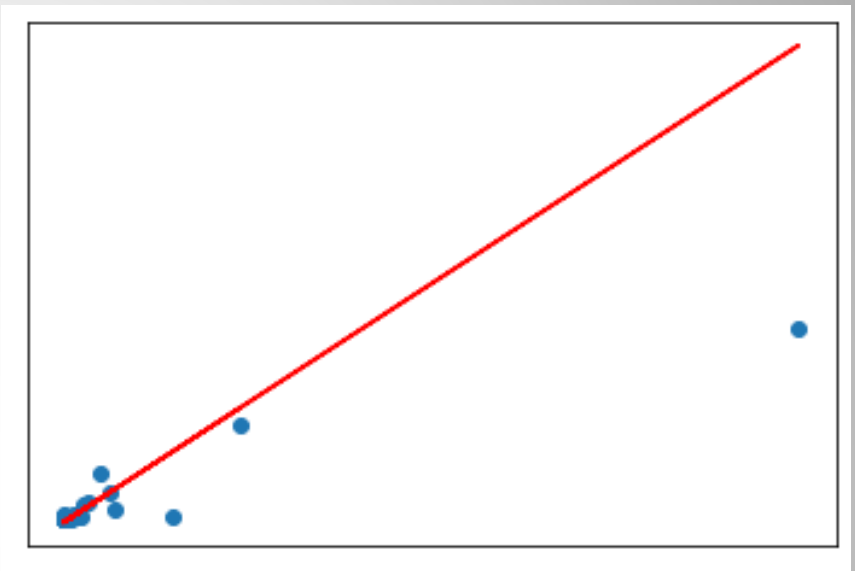


Regression Analysis between Population & people fully vaccinated:



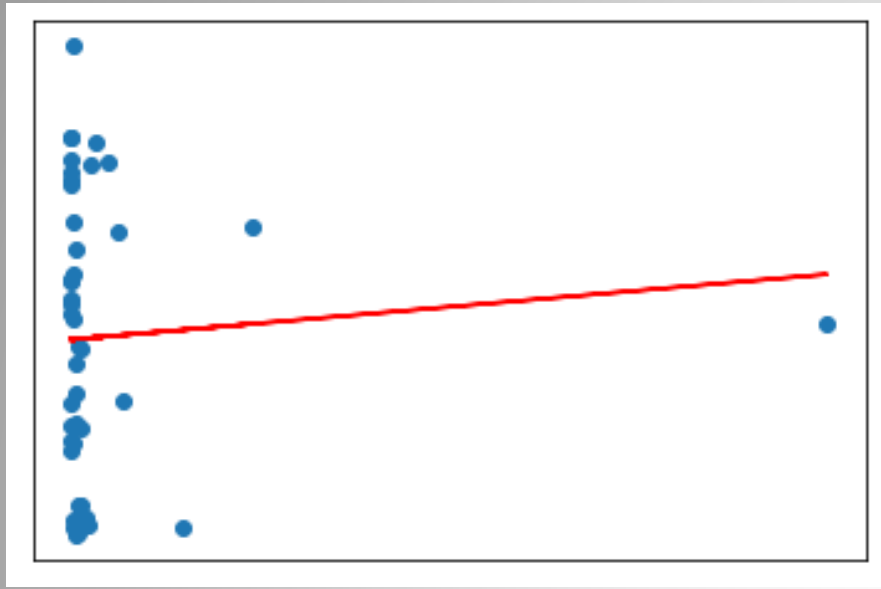
- Here R square can have a negative value when the model selected does not follow the trend of the data, therefore leading to a worse fit than the horizontal line. It is usually the case when there are constraints on either the intercept or the slope of the linear regression line.
- This is the worse model to predict something. As this is the worse fit.

Regression Analysis between Population & daily vaccinations:



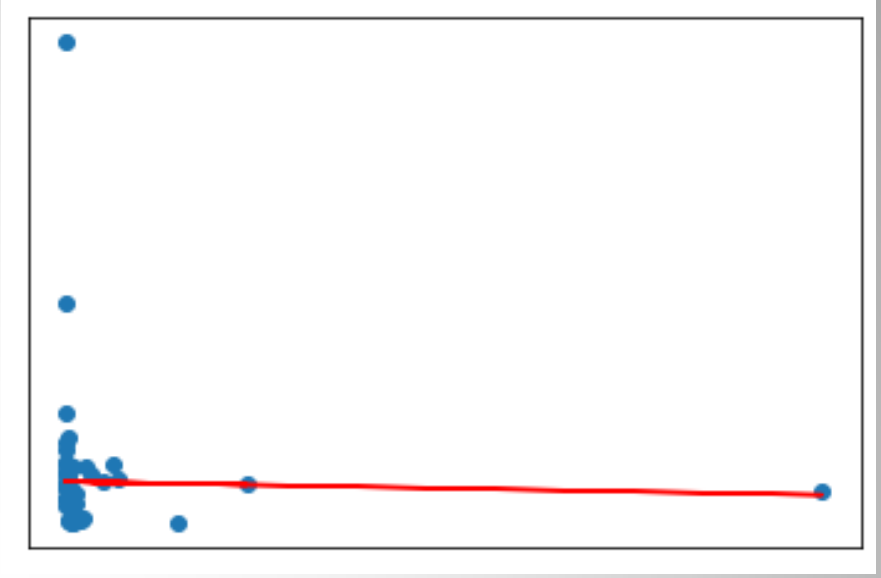
- Here R square can have a negative value when the model selected does not follow the trend of the data, therefore leading to a worse fit than the horizontal line. It is usually the case when there are constraints on either the intercept or the slope of the linear regression line.
- This is the worse model to predict something. As this is the worse fit.

Regression Analysis between Population & people vaccinated per hundred:



- Here R square can have a negative value when the model selected does not follow the trend of the data, therefore leading to a worse fit than the horizontal line. It is usually the case when there are constraints on either the intercept or the slope of the linear regression line.
- This is the worse model to predict something. As this is the worse fit.

Regression Analysis between Population & daily vaccinations per million:



Here the r2 is negative as well as regression coefficient which suggests there is an inverse relationship between the population and daily vaccinations per million, suggesting more the population less people are vaccinated per million.

Also here R square can have a negative value when the model selected does not follow the trend of the data, therefore leading to a worse fit than the horizontal line. It is usually the case when there are constraints on either the intercept or the slope of the linear regression line.

This is the worse model to predict something. As this is the worse fit.