COVID19 ANALYSIS - REPORT

Introduction

We have been provided data regarding the COVID cases in the UK (including the number of deaths, cases, recovered, and hospitalised cases) have been provided along with information regarding the vaccination rates in the UK (including the number of vaccinated, first and second dose cases) within each province in the UK. We have also been provided with data of tweets relating to the #coronavirus hashtag.

Approach:

Our approach is to use Python to analyse the data provided to us.
We upload UK COVID cases data, UK Vaccination cases Data and Twitter Tweets data to python. We clean the data by identifying and removing null values and duplicates. The UK COVID and UK Vaccination cases data is merged to perform better and more wholistic data analysis.

On the Dataframe that is generated, we visualise certain data in the form of table to derive the following information:

- Province/State having the highest number of individuals with a first dose but not a second dose
- Province/State having the highest percentage of individuals with a first dose but not a second dose.
- Change in number of vaccinations taken over time

We further derive graphs based on the data for the purpose of providing an easy, quick and simple understanding of data in the form of bar and line graphs.
We created bar graphs to depict the following information:

- Area wise comparison of vaccination doses.
- Monthly deaths across regions.

- Region wise recovery numbers over time.
- Monthly hospitalisations across regions.
- Location wise comparison of Total Vaccinations

From these visualisations, we can identify those areas with the most first dose and second dose vaccinations, most total vaccination doses, most deaths, most recoveries, and most hospitalisations.

We further analyse the twitter data to analyse to identify the most trending hashtags and understand the most discussed topics among the citizens.

FINDINGS AND CONCLUSIONS

a. Finding:

From our initial analysis of the UK COVID Cases Dataframe, we have deduced the following findings:

- The UK average number of deaths greatly exceeds the number of recoveries and hospitalisations. The average number of recoveries is very low when comparing to the total number of cases and deaths.
- When looking at the Gibraltar province, the average number of deaths is much lower than average number of recoveries and hospitalisations.

a. Conclusions:

From the above findings, we can conclude the following:

- When looking at the UK, the British citizens were severely affected by the COVID19 pandemic and resulted in many deaths in comparison to recoveries and hospitalisations.
- However, the province of Gibraltar managed to have lower death rate in comparison to the rest of the country. Hence, we must understand what Gibraltar did correctly to ensure this.

b. Finding:

From our initial analysis of the merged Dataframe containing UK COVID Case data and UK Vaccination Case data, we have deduced the following findings:

- Vaccinations began in the month of January 2021.
- The Province of Gibraltar has the highest number of individuals who have received a first dose but not a second dose with 264745 individuals.
- The province of Turks and Caicos Islands have the highest percent of individuals who have received a first dose but not a second dose with a percentage of 2.30709%.
- The number of individuals who received the first dose was highest during February 2021 post which there has been decrease in the number of first dose vaccinations which increases only in the months of May and July 2021.
- The number of individuals who received the second dose had been rising till it reached its highest during May 2021 post which there has been decrease in the number of second dose vaccinations.

b. <u>Conclusion:</u>

From the above findings we can conclude the following:

- Gibraltar had the highest number of individuals that received at least one vaccination dose. Hence, we can ensure that there is a correlation between the low deaths rates and high vaccination numbers.
- There was high awareness around vaccine efficacy in the initial months from the release of the vaccine. However, from May 2021, there has be a gradual decrease in those taking both vaccine doses except for the month of June 2021 where there is an increase in first dose vaccinations.

c. <u>Finding:</u>

From our visualisations of the merged Dataframe, we deduced the following findings:

- Number of Deaths were highest in the provinces of 'Others'. A large number of all deaths were from the province 'Others'. As the number is so large, we have excluded it from our visualisation as it skews the data.
- Number of Deaths were highest in Gibraltar and the Channel Islands. The number of deaths seemed to hit its peak during the months of August and September. The number of deaths dropped extremely quickly from September to October.
- The overall number of deaths has greatly decreased over the last month of October.
- Hospitalisations peaked in the moth of January 2021 and has since been decreasing with a small spike in months of June to September 2021.

c. <u>Conclusion:</u>

From the above findings we can conclude the following:

- In Comparison to most provinces excluding 'Others', Gibraltar and the Channel Islands the most deaths.
- In comparison to most provinces, the number of cases within Gibraltar and Channel Islands were highest excluding the province 'Others'.
- As the number of deaths has greatly decreased, this could be viewed as a result of consistent vaccinations over time. Hospitalisation
- There is a trend with the number of cases, deaths, hospitalisations, and recoveries. As the number of cases rose around January 2021, there was more awareness around COVID19. This resulted in a large number of people taking the vaccine as soon as it was available. However, the overall vaccine numbers dropped as well soon after the release.
- The region 'Others' has had the greatest number of cases and the most number of deaths. However, in terms of vaccinations, the region 'Others' is the 2nd lowest.

d. <u>Finding:</u>

From our visualisations of the merged Dataframe, we deduced the following findings:

- On observation of the top 10 hashtags within our twitter data, we can see that 8 out of the top 10 tweets were about the COVID19 pandemic.

d. Conclusion:

From the above finding, we can conclude:

- During the period, the COVID19 was one of the most discussed topics among the twitter users. Hence, COVID19 was widely discussed during this period.

RECOMMENDATIONS:

From the analysis of UK COVID case data, UK Vaccination Case Data and Tweets analysis, the following recommendations can be provided:

- Even as the rate of Covid Cases, Deaths and hospitalisation are reducing. The province 'Others' has suffered the most with the most number of cases and most number of deaths. Hence, it is important that the government direct its efforts towards marketing the vaccine to this population.
- Government directs efforts towards requesting residents to take both doses of the vaccine.
- While awareness is still high on twitter, the government can make efforts on twitter to advertise and raise awareness on efficacy of the vaccine.