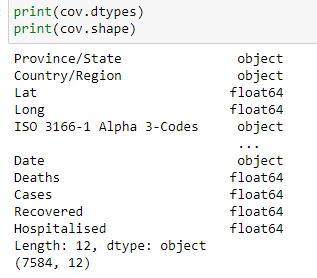
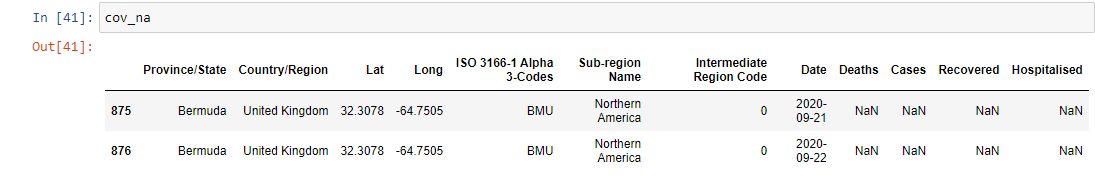
**Data analytics using Python and UK Covid/vaccine data**

After loading the data into Python and creating data frames (cov and vac), I found that the, ‘covid\_19\_uk\_cases.csv’ file contained 7584 rows and 12 columns. Two rows had NA’s or missing data, both belonging to the ‘Province/State’ Bermuda.



*Figure 1: Determining the Data types in the DataFrame and number of rows and columns*

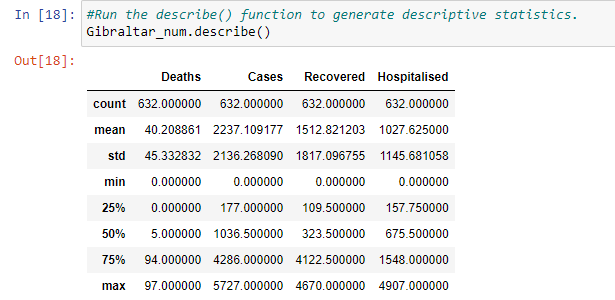


*Figure 2: Determining the number of missing values.*

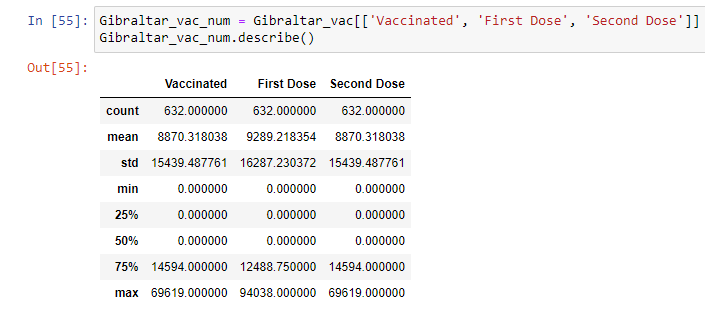
The ‘covid\_19\_uk\_vaccinations.csv’ file contained 7584 rows and 11 columns. There were no NA’s.

The 2 DataFrames do not have a default index. After sorting the cov DataFrame by ‘Date’ (*cov.sort\_values('Date')*), I could see that the number of covid cases increases from the beginning (22/01/2020) all the way to the end (14/10/2021), however, it fluctuates and so it would be nice to be able to visualise this data with a line chart to see the changes over time. A similar observation can be made with the vaccination numbers. The recovered data stops from August 2021; is this an error? On their own, the descriptive statistics will not be very meaningful; what would help is to compare the results with other State’s in the DataFrames and see how they compare. We can then try to interpret, based on the mean, for example, do states that have a higher mean vaccination lead to fewer hospitalisations or deaths?

The numerical data for Gibraltar seem to suggest the numbers are inflated. For example, the sum of the ‘Vaccinated’ is 5,606,041; Gibraltar’s population in 2020 was only 33,691, so how can the total vaccinated exceed this number? The sum for ‘Cases’, ‘Recovered’ and ‘Hospitalised’ also exceed the population of Gibraltar (1,413,853; 956,103 and 649,459). Where is the data from and is it reliable?

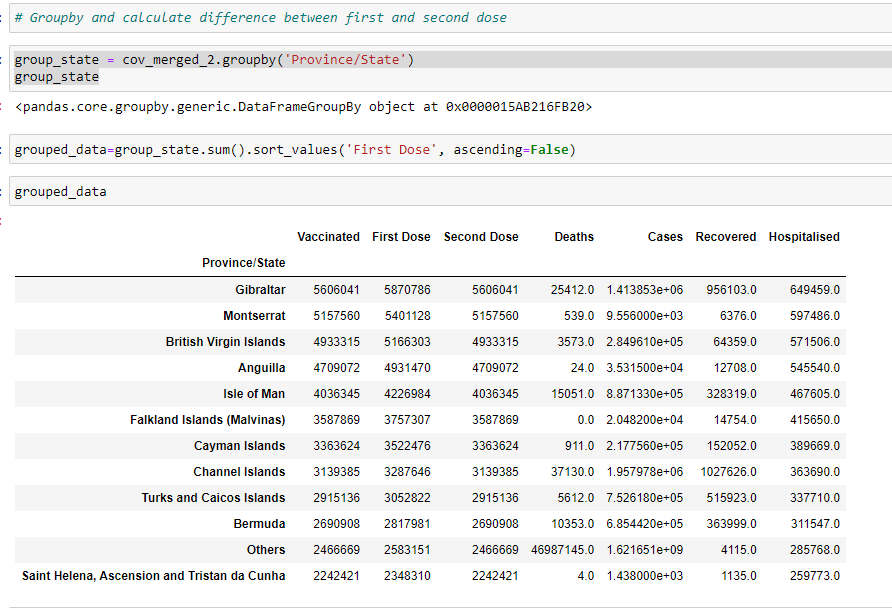


*Figure 3: Running the descriptive statistics on Gibraltar\_num DataFrame*

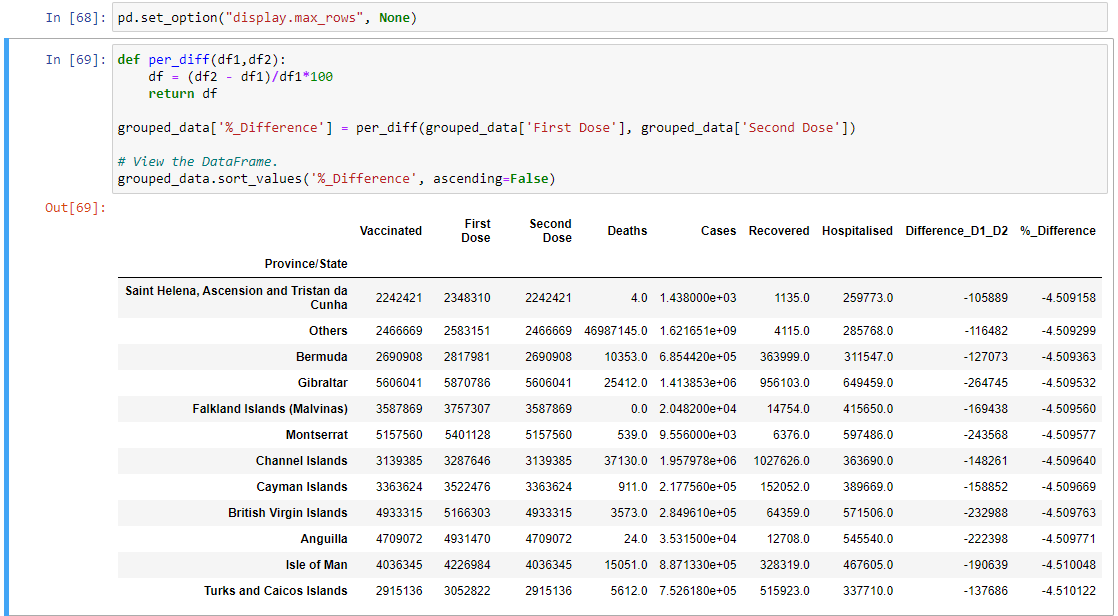


*Figure 4: Running the descriptive statistics on Gibraltar\_vac\_num DataFrame*

The State that has received the highest number of First Dose is Gibraltar, although the State that has the highest percentage who have received the first dose but not the second dose is, ‘Saint Helena, Ascension and Tristan da Cunha’.



*Figure 5: Grouped data to show highest number of First Dose*



*Figure 6: Grouped data to show highest percentage to not receive Second Dose after receiving First Dose*

The number of individuals who have received the vaccine has increased in time as we can see that the number vaccinated in 2020 was zero and the sum that received the First Dose in 2021 is 46,966,364.

