**Project name:** Covid data visualiser

**Aim:**

The project aims to extract covid related data from csv(comma separated values) files and visualise them into graphs. It aims to make analyzing this data a much easier task.

**Modules used:**

* Matplotlib: Used to plot graphs from data
* Pandas: Used to extract data from csv files
* Numpy: Used to create index lists to be plotted
* Random: Used to change the colours of the graph
* Sys: Used to implement the exit functionality of the program

**Functionality of the project:**

The program consists of 10 user defined functions and all of them plot a unique graph with data obtained from two csv files.

The user is presented with a command line styled interface and is asked to enter the number of the graph they wish to see from the given list. After the number is entered, the appropriate function is called and the graph is displayed. The program automatically runs again until “e” or “E” is entered. Invalid input message is shown when the input is a string other than “e/E” or the number is out of bounds.

All the functions use a very similar mechanism for extracting data from the csv files. Data6.csv contains international covid data while data4.csv contains Indian covid data.

Pandas modules is used first to read the desired columns from the csv files and converting them to lists. These values are then fused together into a dictionary (e.g if country and cases columns are read, then dictionary contains country name: cases for that country). This dictionary is then sorted by its values and is again converted back to lists, but now in our desired order.

Numpy module is used to create a list of numbers for data which are not numbers and need to be plotted. After this the pyplot module of the matplotlib library is used to plot graphs using the extracted data. Credit is added at each graph using the text function.

Finally the program displays a thanks message when e/E is pressed and terminates itself. This has been made possible by the exit function of the sys module.

**Functions used:**

Basic functions like print(),input(),zip(),sorted(),append() etc.

* Pandas.read\_csv() – To read csv files
* Numpy.arange() – To create number lists for non plottable data
* Pyplot.bar(),barh() – To plot bar graphs,horizondal bar graphs
* Plt.pie() – to plot pie graphs
* Plt.scatter() – to plot scatter plots
* Random.choice() – To choose a random colour
* Random.shuffle() – to shuffle(change the order) a list
* Plt.xticks() or .yticks() – To plot the numpy list and name them with desired text instead of numbers
* Plt.xlabel,ylabel,title()- for labeling the axes and to give the graph a title
* Plt.text()- to add credit
* Plt.legend()- to add a legend for the graph
* Plt.show() – to display the graph
* Sys.exit()- to exit the program

**Conclusion:**

Thus the program extracts data from csv files using the pandas module, converts them into a desired format and uses the matplotlib module to plot them into graphs.

The program converts raw data from csv files to visually pleasing graphs and makes it much easier to analyse the covid related data.