

PYTHON PROJECT REPORT

ON

COVID-19 Data Visualisation in India

Course: PYTHON PROGRAMMING (INT 213)



Name: Himangshu Nayak

Registration Number: 12011751

Program: B. Tech CSE (3rd semester)

School: School of Computer Science and Engineering

Name of the University: Lovely Professional University, Punjab, India

Date of submission :20th NOVEMBER 2021

ABSTRACT

As COVID-19 becomes a dangerous pandemic worldwide, there is an urgent need to understand all aspects of it through data visualization. A major challenge we experienced is that, in an open world setting where it is not even clear which datasets are available and useful, generating the right visualizations becomes an extremely tedious process. Traditional data visualization recommendation systems usually assume that the datasets are given, and that the visualizations have a clear objective. We contend that such assumptions do not hold in a COVID-19 setting where one needs to iteratively adjust two moving targets deciding which datasets to use, and generating useful visualizations with the selected datasets. We make it easier to understand all the data in a standard way by using computer language tools.

Contents

- 1) Abstract**
- 2) Introduction**
- 3) Libraries**
- 4) Steps of Visualisation**
- 5) Code Section**
- 6) Conclusion**
- 7) References**

INTRODUCTION

Content: This project is based on visualising the data of Covid-cases in India. Here data is visualised in many different forms from which we can know the complete scenario of the overall cases occurring in India.

Motivation: Visualising Data sources in a correct form is my key aspect. And by looking on some other great projects which are doing very great in their field, by looking their visualising techniques, I too got motivated to make some data visualisation system and that also of the current happening pandemic that is Covid-19.

Idea: The prime idea of this project is to visualise data in the form of graphs of different states in an ordered way, such that user would not get confused to fetch the exact data of all the Covid cases.

LIBRARIES

Pandas: Pandas is the most popular python library that is used for data analysis. We will provide highly optimized performance with back-end source code with the use of Pandas.

Matplotlib: Matplotlib tries to make easy things easy and hard things possible. We will generate plots, histograms, scatterplots, etc., to make our project more appealing and easier to understand.

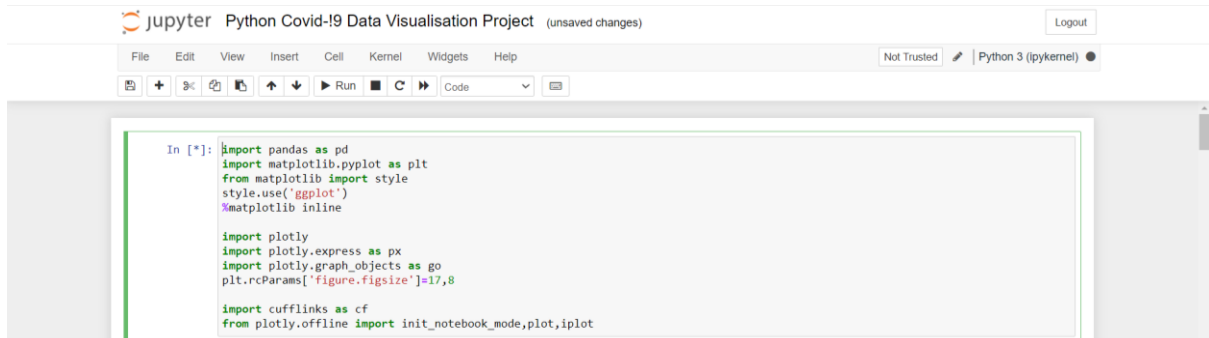
Plotly: Python Plotly Library is an open-source library that can be used for data visualization and understanding data simply and easily. Plotly supports various types of plots like line charts, scatter plots, histograms, cox plots, etc.

Cufflinks: It is another library that connects the Pandas data frame with Plotly enabling users to create visualizations directly from Pandas. The library binds the power of Plotly with the flexibility of Pandas for easy plotting.

STEPS OF VISUALISATION

1. Imported an excel document containing data of Covid Cases of 19th November.
2. Converted it to a csv file in Jupyter Notebook
3. Visualise it in the form of a table.
4. Active cases were calculated and added to the table.
5. Sorted out all the active cases applying some colour in descending order.
6. Perform the graphical visualisation using pandas, matplotlib and plotly libraries.
7. Performed the Object-Oriented Data Visualisation for plotly.

CODE SECTION



```
In [*]: import pandas as pd
import matplotlib.pyplot as plt
from matplotlib import style
style.use('ggplot')
%matplotlib inline

import plotly
import plotly.express as px
import plotly.graph_objects as go
plt.rcParams['figure.figsize']=17,8

import cufflinks as cf
from plotly.offline import init_notebook_mode,plot,iplot
```

- This part will import all the libraries in which we will work upon

It contains only three libraries:

- 1) Pandas
- 2) Matplotlib
- 3) Plotly
- 4) Cufflinks

```
In [130]: df=pd.read_csv(r"C:\Users\Himangshu Nayak\Desktop\Python project\COVID DATA.csv")
```

```
In [131]: df
```

```
Out[131]:
```

	Name of States / UT	Confirmed cases	Cured/Discharged/Migrated	Death	Date
0	Andaman and Nicobar Islands	7675	7541	129	19-11-2021
1	Andhra Pradesh	2070738	2053755	14423	19-11-2021
2	Arunachal Pradesh	55242	54908	280	19-11-2021
3	Assam	614863	605656	6065	19-11-2021
4	Bihar	726178	716476	9663	19-11-2021
5	Chandigarh	65394	64546	820	19-11-2021
6	Chhattisgarh	1006459	992596	13591	19-11-2021
7	Dadar Nagar Haveli	10682	10678	4	19-11-2021
8	Delhi	11140675	1115110	25005	19-11-2021

- These two lines of code will read our csv file and display it in the below. Here on printing df we will get the table of data.

Jupyter Python Covid-19 Data Visualisation Project Last Checkpoint: 3 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

In [178]: `df['Active_cases']=df['Confirmed cases']-(df['Cured/Discharged/Migrated']+df['Death'])`

In [179]: `df`

Out[179]:

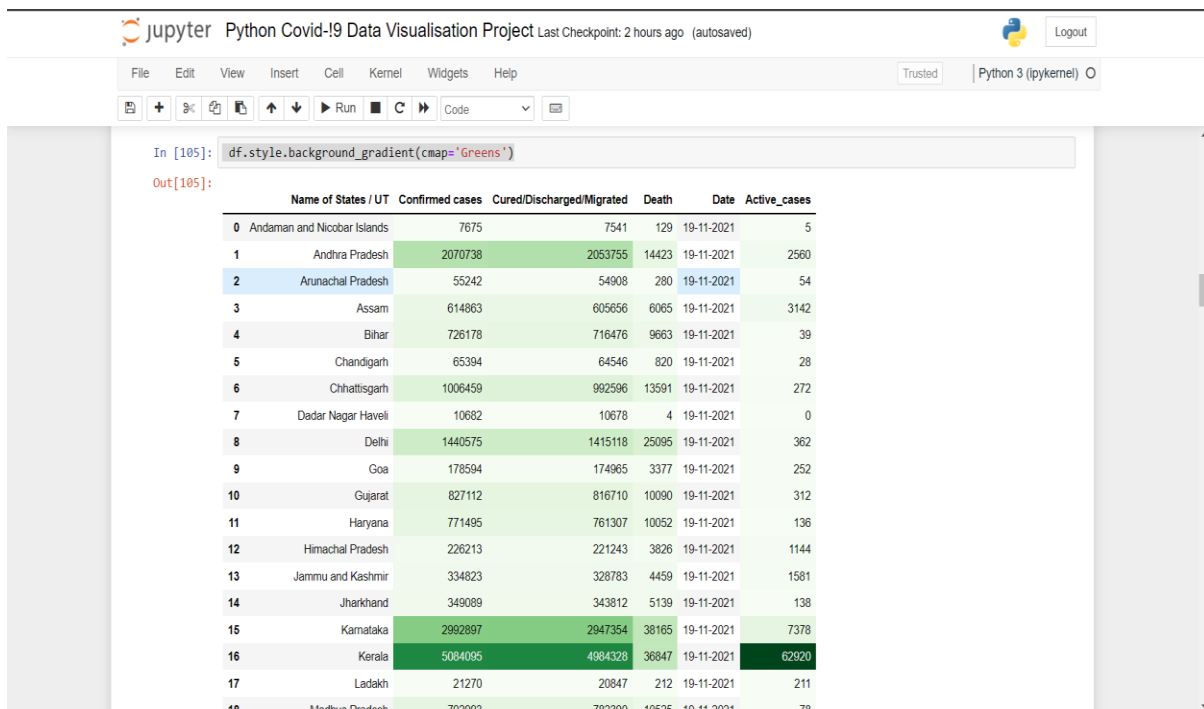
	Name of States / UT	Confirmed cases	Cured/Discharged/Migrated	Death	Date	Active_cases
0	Andaman and Nicobar Islands	7675	7541	129	19-11-2021	5
1	Andhra Pradesh	2070738	2053755	14423	19-11-2021	2560
2	Arunachal Pradesh	55242	54908	280	19-11-2021	54
3	Assam	614863	605656	6065	19-11-2021	3142
4	Bihar	726178	716476	9663	19-11-2021	39
5	Chandigarh	65394	64546	820	19-11-2021	28
6	Chhattisgarh	1006459	992596	13591	19-11-2021	272
7	Dadar Nagar Haveli	10682	10678	4	19-11-2021	0
8	Delhi	1440575	1415118	25095	19-11-2021	362
9	Goa	178594	174965	3377	19-11-2021	252
10	Gujarat	827112	816710	10090	19-11-2021	312
11	Haryana	771495	761307	10052	19-11-2021	136
12	Himachal Pradesh	226213	221243	3826	19-11-2021	1144
13	Jammu and Kashmir	334823	328783	4459	19-11-2021	1581
14	Jharkhand	349089	343812	5139	19-11-2021	138
15	Karnataka	2992897	2947354	38165	19-11-2021	7378
16	Kerala	5084095	4984328	36847	19-11-2021	62920

- These two lines of code will create a new column and add data to it by performing some operations with second, third and fourth column.

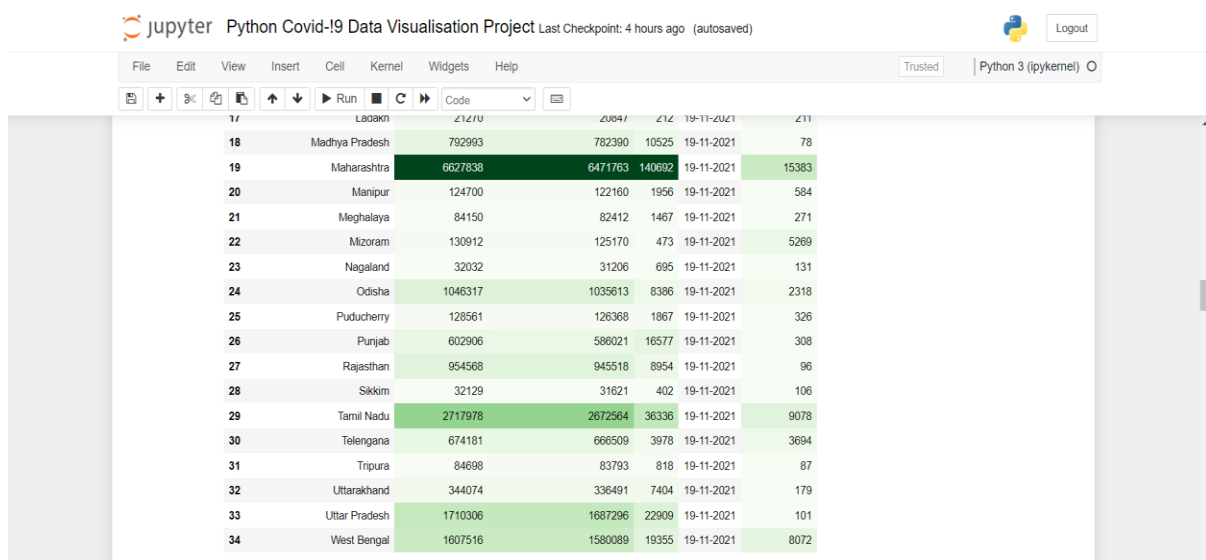
Jupyter Python Covid-19 Data Visualisation Project Last Checkpoint: 4 hours ago (autosaved) Logout

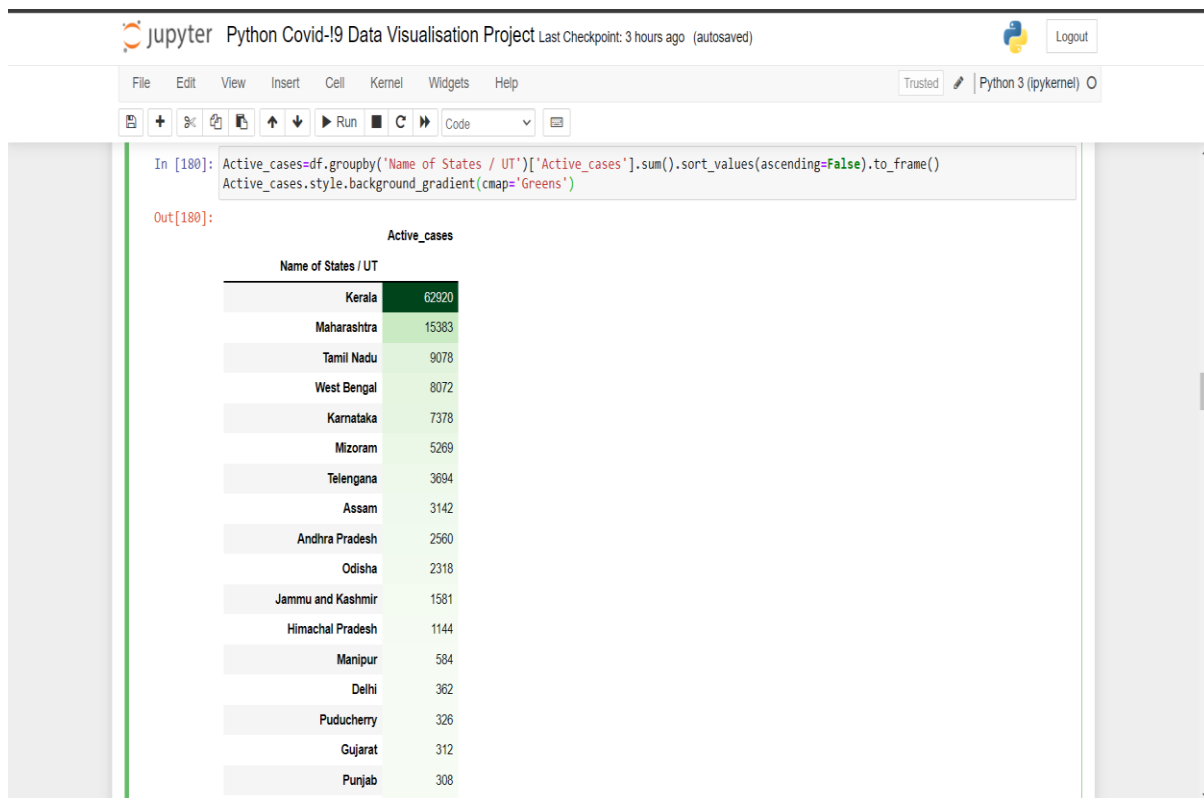
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

18	Madhya Pradesh	792993	782390	10525	19-11-2021	78
19	Maharashtra	6627838	6471763	140692	19-11-2021	15383
20	Manipur	124700	122160	1956	19-11-2021	584
21	Meghalaya	84150	82412	1467	19-11-2021	271
22	Mizoram	130912	125170	473	19-11-2021	5269
23	Nagaland	32032	31206	695	19-11-2021	131
24	Odisha	1046317	1035613	8386	19-11-2021	2318
25	Puducherry	128561	126368	1867	19-11-2021	326
26	Punjab	602906	586021	16577	19-11-2021	308
27	Rajasthan	954568	945518	8954	19-11-2021	96
28	Sikkim	32129	31621	402	19-11-2021	106
29	Tamil Nadu	2717978	2672564	36336	19-11-2021	9078
30	Telangana	674181	666509	3978	19-11-2021	3694
31	Tripura	84698	83793	818	19-11-2021	87
32	Uttarakhand	344074	336491	7404	19-11-2021	179
33	Uttar Pradesh	1710306	1687296	22909	19-11-2021	101
34	West Bengal	1607516	1580089	19355	19-11-2021	8072

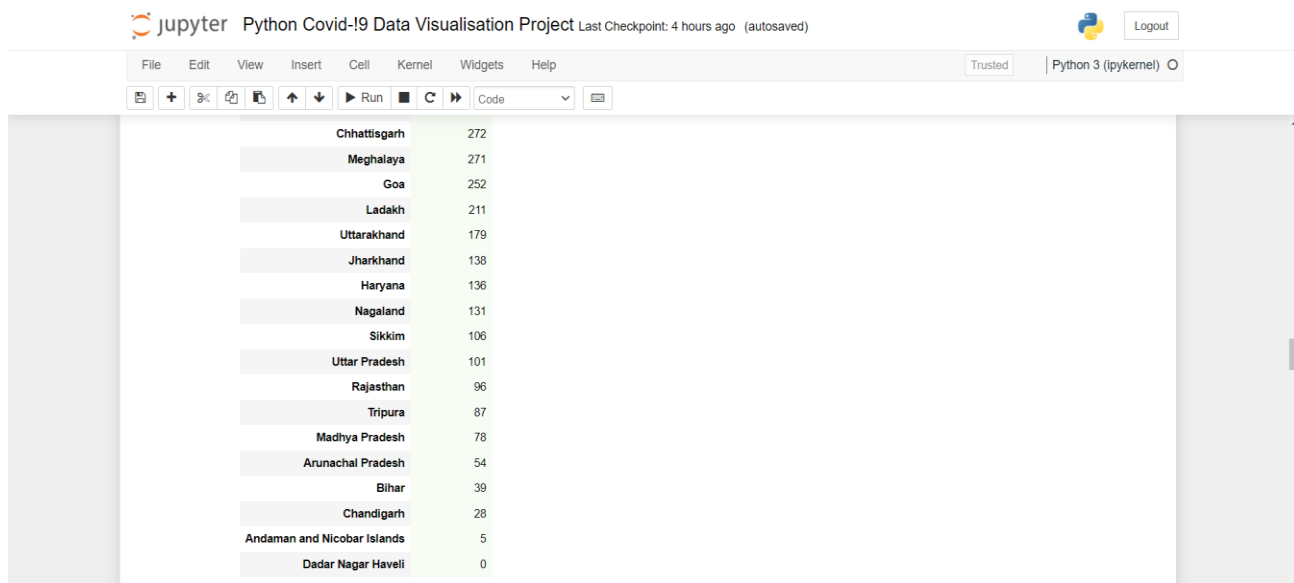


- This line of code will present the data in a perfect order. We can easily distinguish data from the table. In this code we had given an inbuilt feature in which we can present our data very easily and in a good way.
- Here we used green colour to categorized it. Dark coloured portion represent highest impact and faded one is low.





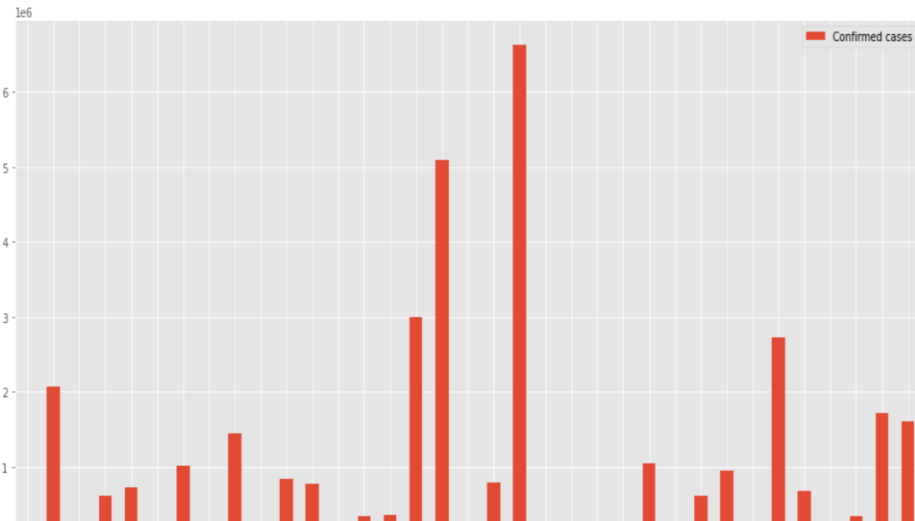
- This code gives the command to show the data of a column in a sorted way.
- We used inbuilt feature in it to display the data in a categorized way.
- Here green colour is been used to categorized.



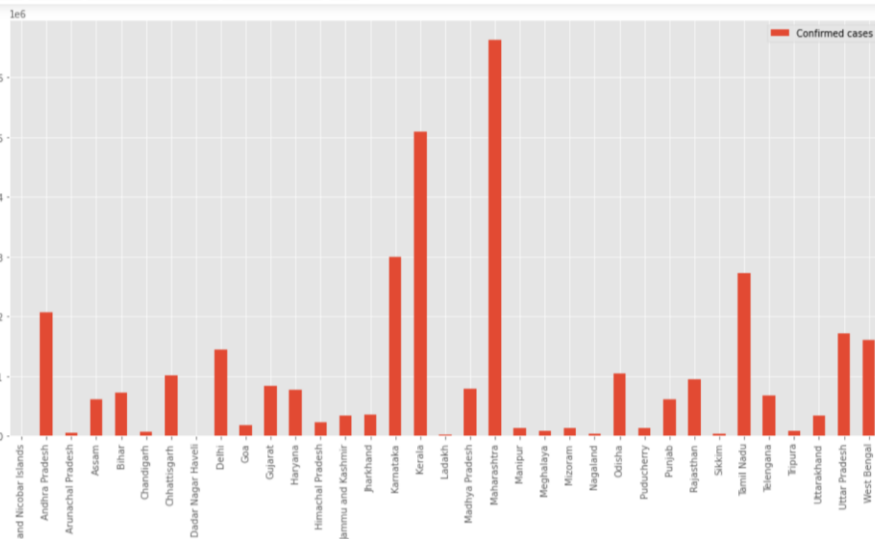
Graphical Representation (Pandas Visualisation)

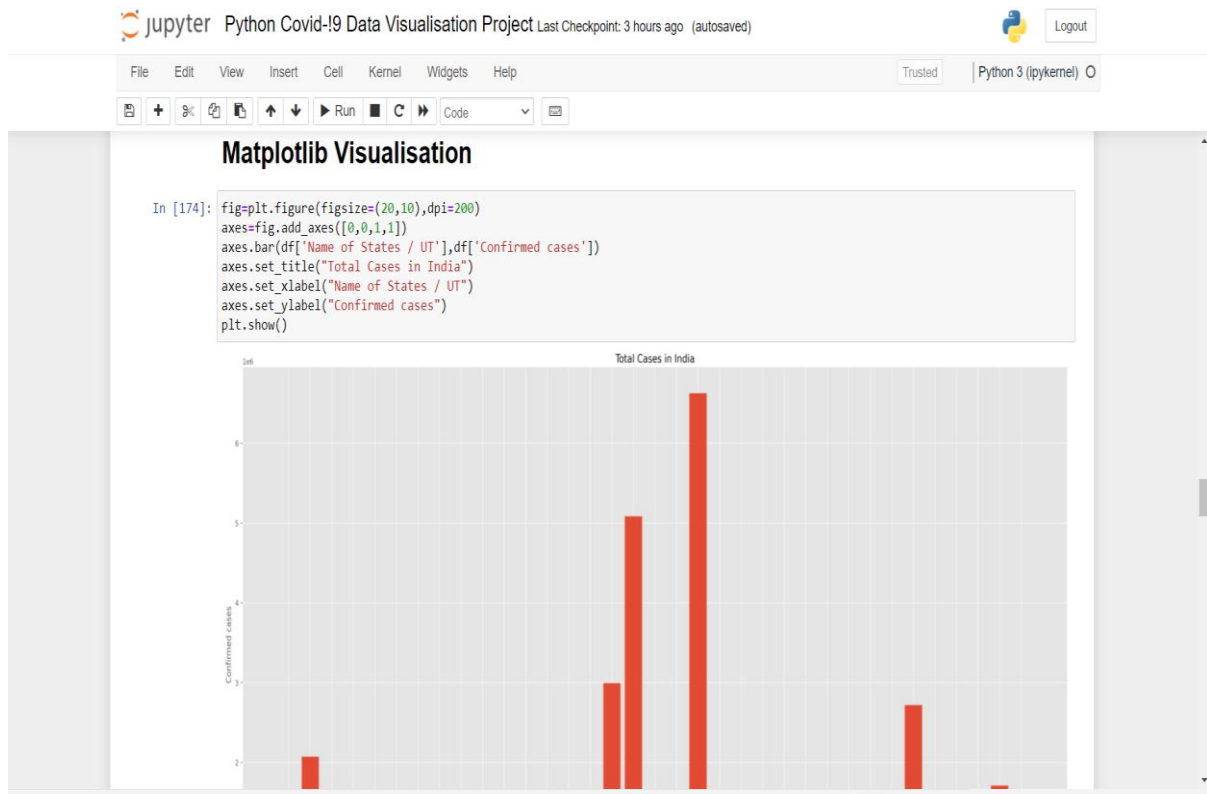
```
In [181]: df.plot(kind='bar',x='Name of States / UT',y='Confirmed cases')
```

```
Out[181]: <AxesSubplot: xlabel='Name of States / UT'>
```

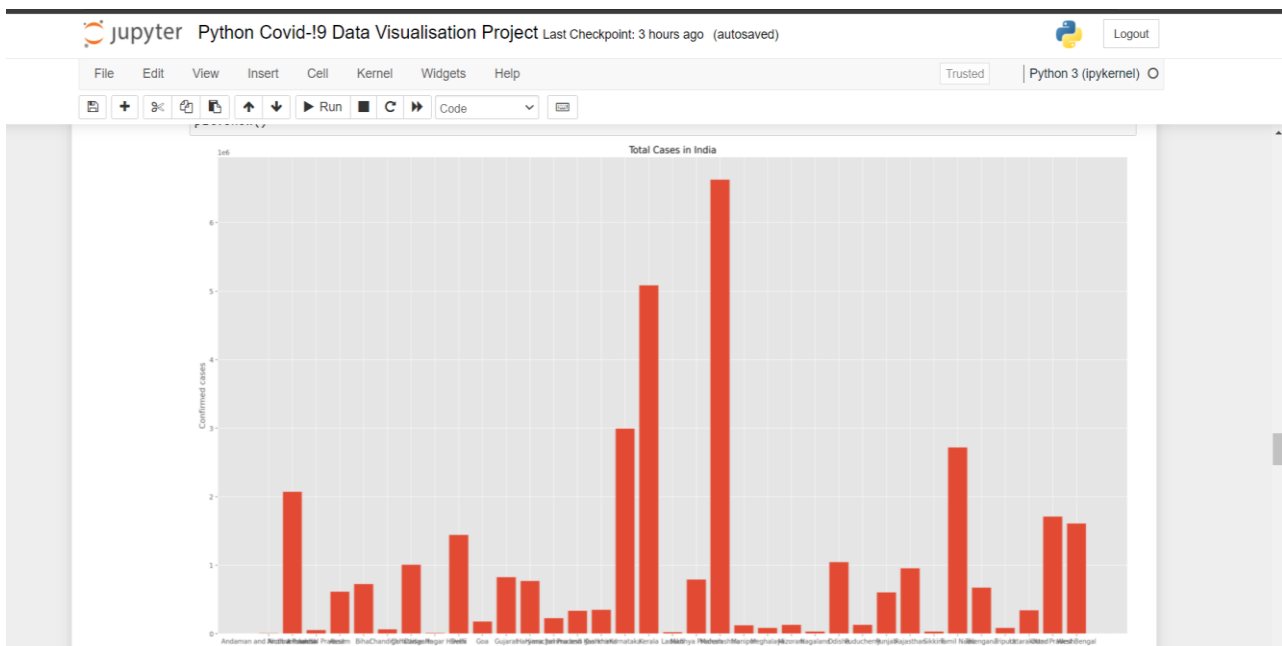


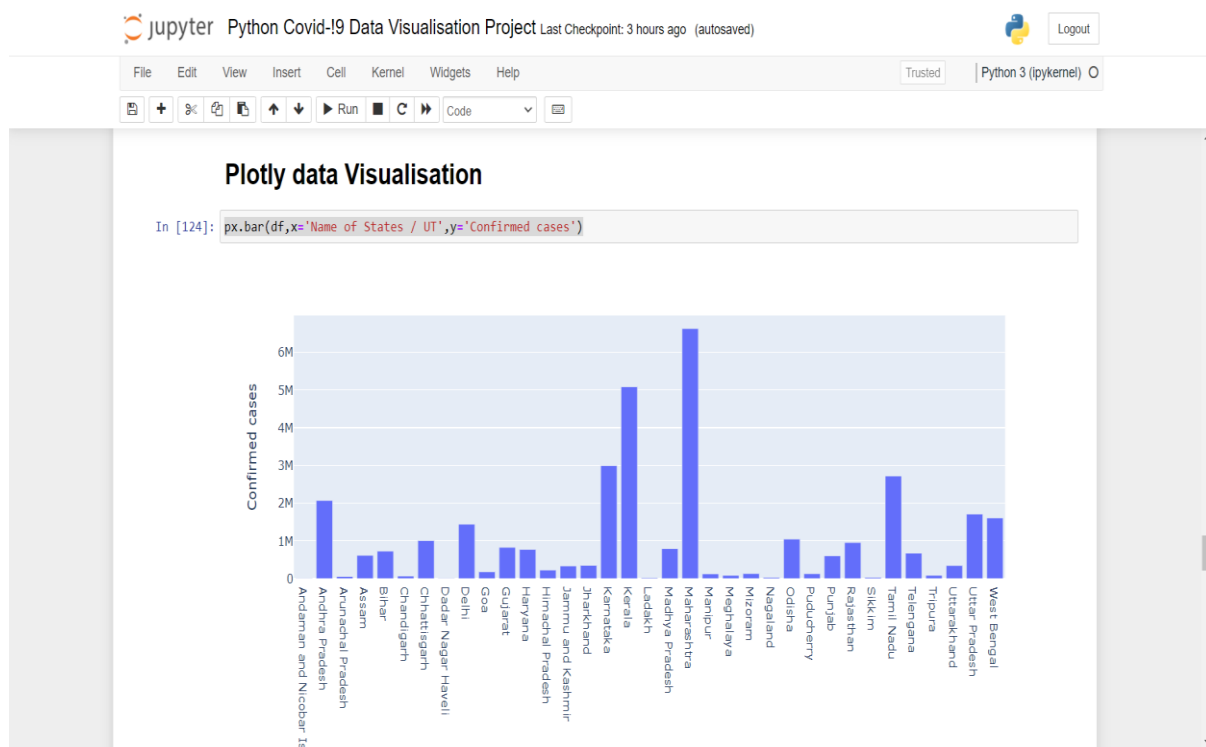
- Here comes the graphical representation of data in which we can plot the data in a graphical manner.
- Here we are using pandas library to plot it.
- Here we are using two axis, X-axis and Y-Axis to plot.



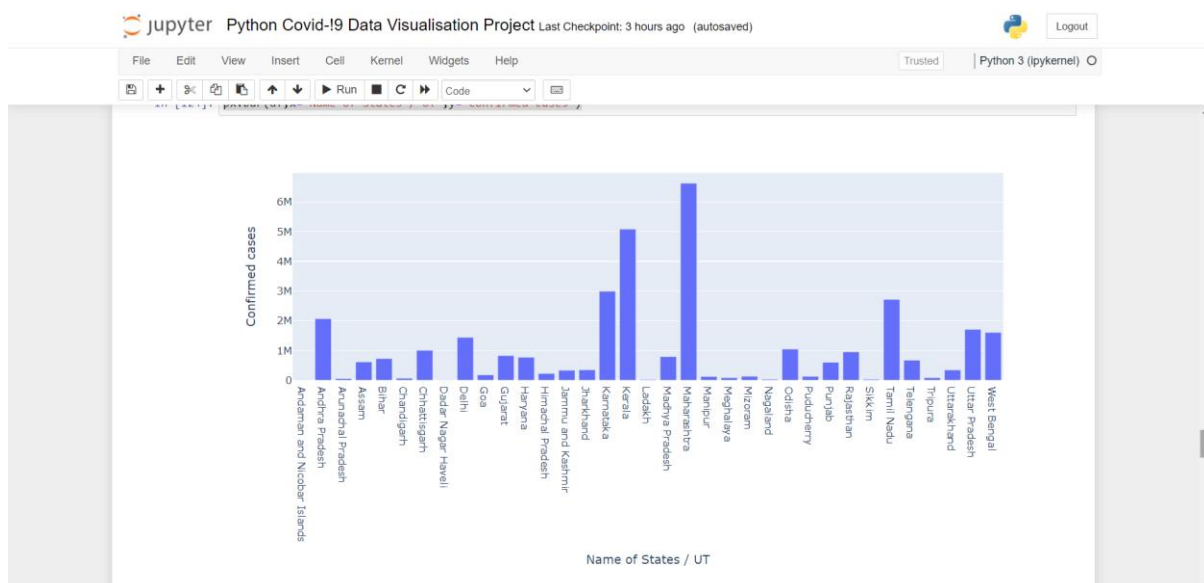


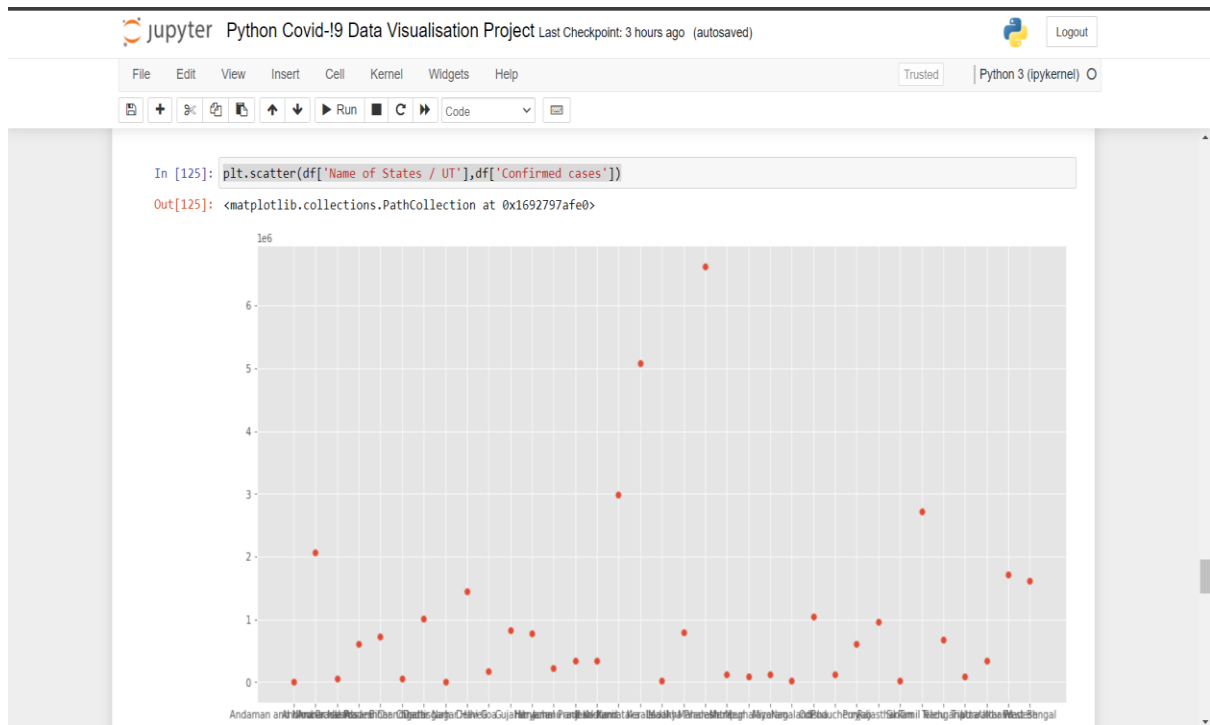
- Here is the Matplotlib Visualisation in which we had plotted the graph in an ordered manner.
- In the above line of codes, we have assigned title to the graph, set X-axis and Y-axis with their column names and displayed it with the function supported by the library.



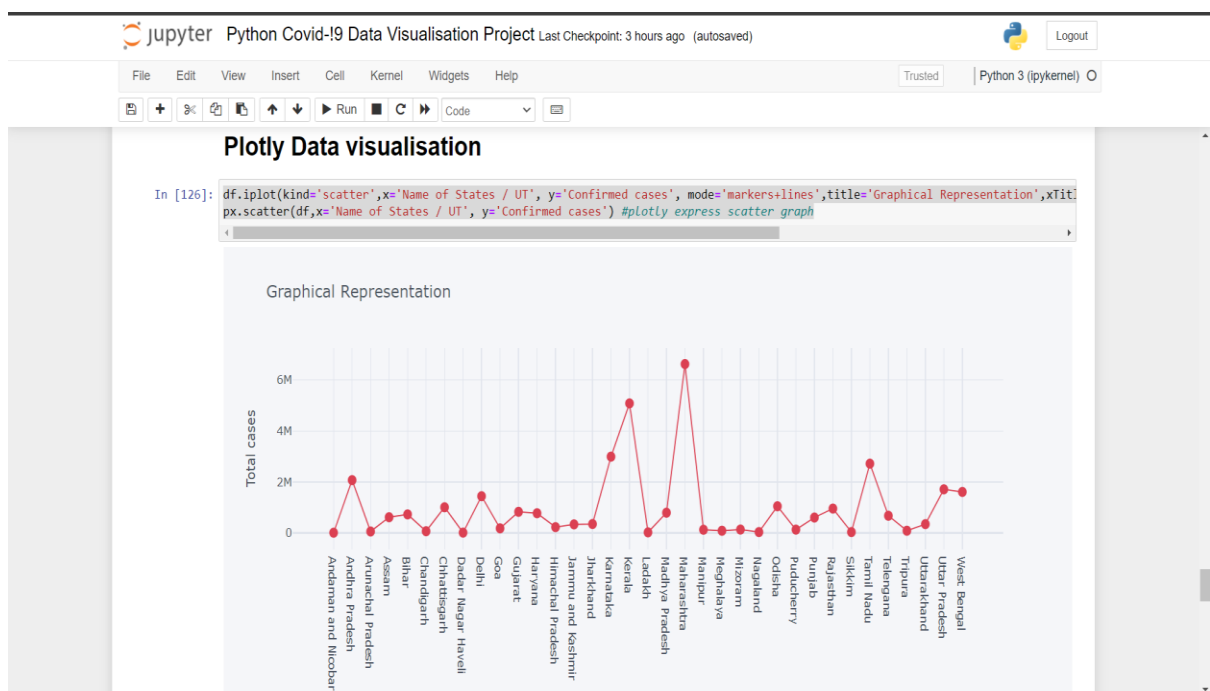


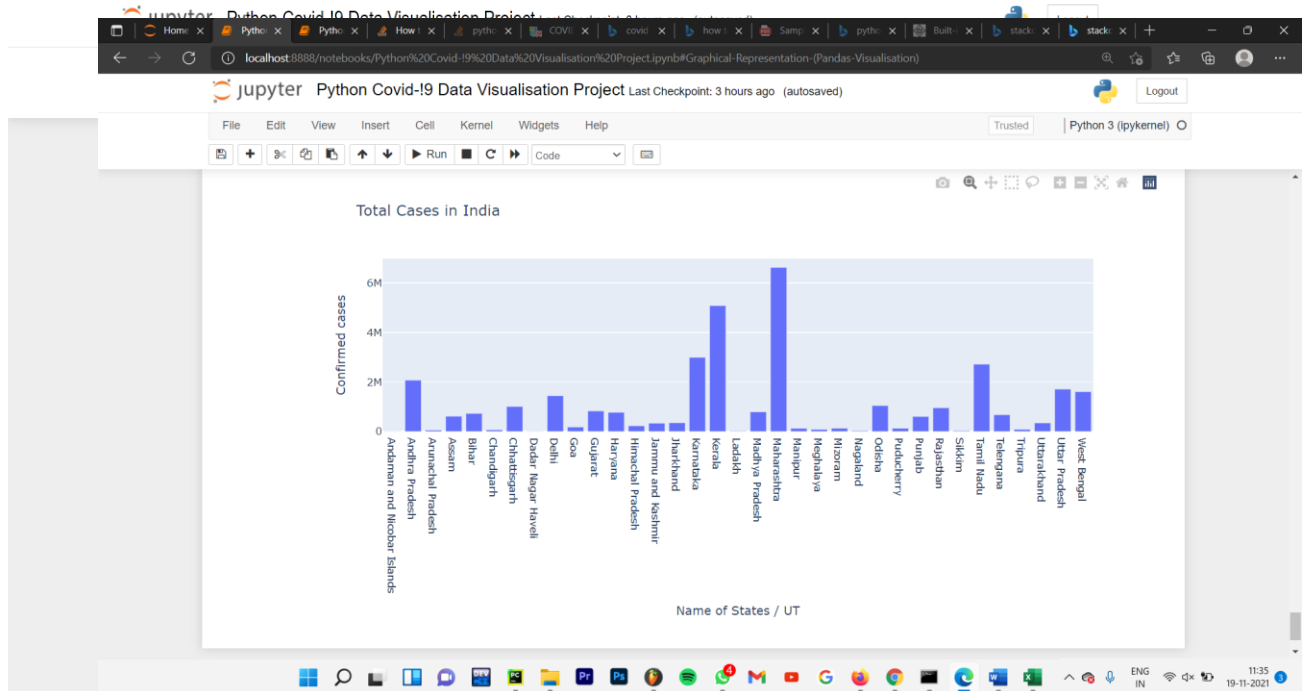
- This is Plotly Visualisation
- Here we put the data in an ordered form, it is similar to that of Matplotlib Visualisation.
- Here the contents of a particular column are displayed in a proper way.
- Visualisation is much clear than that of Matplotlib Visualisation.





- This is scattered Visualisation of data, using some inbuilt features we can plot it.





- This is Plotly Data along with Express Scatter graph
- Above graph gives more clear-cut Data Visualisation than the express Scatter one
- The presence of marker lines makes the visualisation more perfect to fetch data.

```

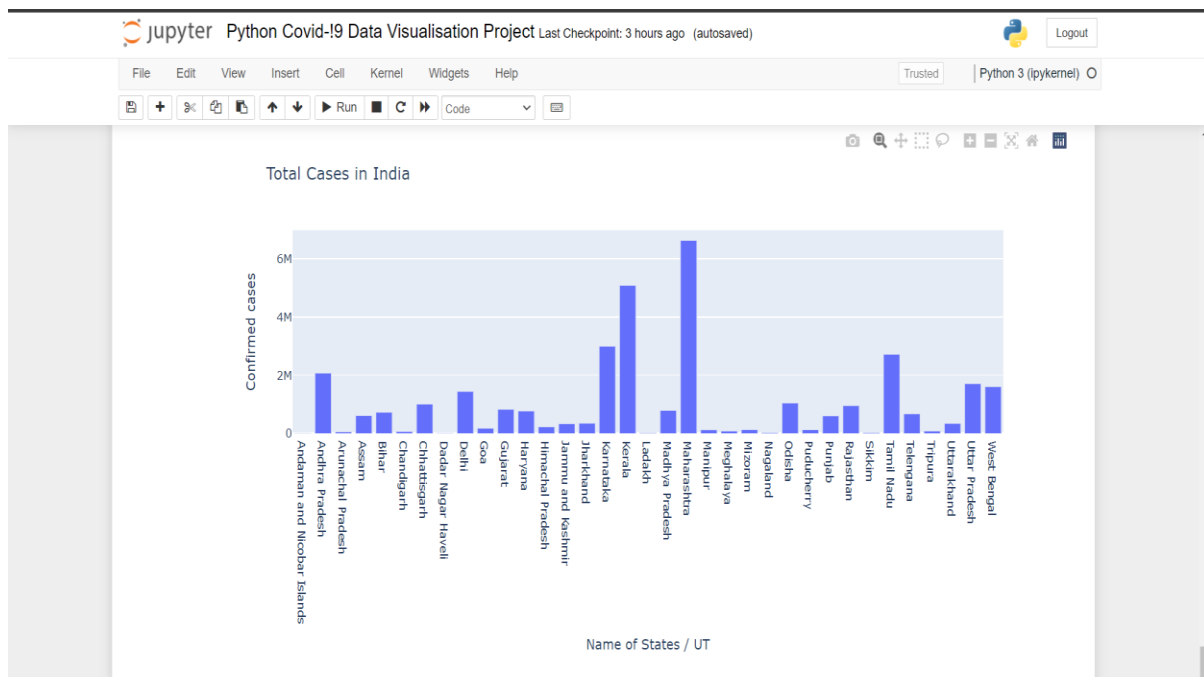
Jupyter Python Covid-19 Data Visualisation Project Last Checkpoint: 3 hours ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel) O

In [127]: fig=go.Figure()
fig.add_trace(go.Bar(x=df['Name of States / UT'], y=df['Confirmed cases']))
fig.update_layout(title='Total Cases in India',xaxis=dict(title='Name of States / UT'),yaxis=dict(title='Confirmed cases'))

```

- Here is the Object-Oriented Method for Plotly
- Here we are creating plots for two columns, we can also perform it for other columns also.



- We can also create for matplotlib also by using Object oriented Method.
- We can set here the title, X-Axis and Y-Axis and also have control over the scale.
- It gives a perfect data visualisation overall.

CONCLUSION

In this project report, I visualise all the data in different forms using various types of methods. One can easily go through this data easily. Nowadays data are not easy to handle and manage it properly and the most important is that to display it in a systematically order. Here I have done all the possible methods which I have learnt till yet and implemented it in my project work. Hope this data visualisation work will be far enough to fetch data in a proper way. One will not get confused in reading the data and can go through it easily.

REFERENCES:

- <https://mygov.in>
- <https://stackoverflow.com>
- YouTube
- <https://www.geeksforgeeks.org>
- Hacker Rank etc.