SILIGURI INSTITUTE OF TECHNOLOGY
DEPARTMENT OF ELECTRICAL ENGINEERING

Project on :COVID-19 🦠😷 data Analysis

By

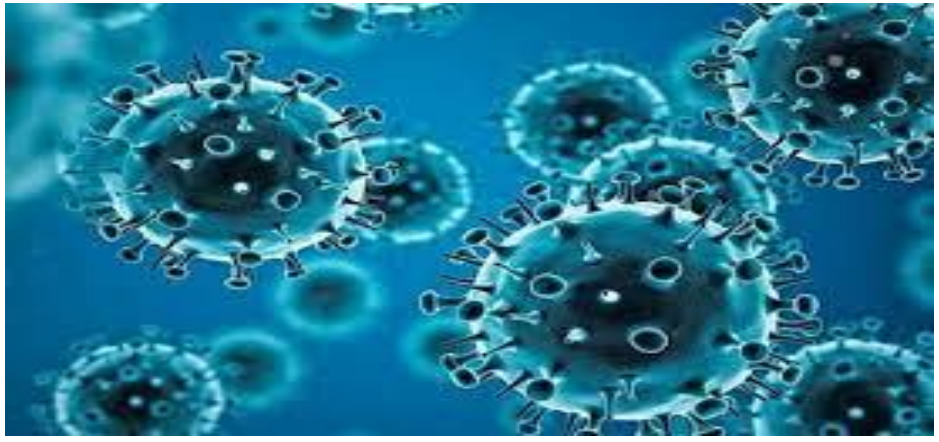
SOURADIP DAS(11901620007)
SAROJ KUMAR GOUDA(11901620006)
ADITYA THAKUR (11901621036)

Using Python 

Under the supervision of

Mr. Ripam Kundu
Sikharthy Infotech Pvt. Ltd.

PROJECT ON



DATA SCIENCE

COVID-19 DATA ANALYSIS

INTRODUCTION

It feels surreal to imagine how the virus began to spread from one person that is patient zero to four million today. It was possible because of the transport system.

To be well prepared for any future pandemic we need to analysis the current covid19 situation, analysis of the data covid19 will help us to understand where we made mistakes and grow to prevent it.

BACKGROUND STUDY

To describe common python funtionality and features used for data science .

To query data frames structures for cleaning and processing.

To understand techniques such as lambdas and manipulating CSV files.

PROBLEMS STATEMENT

A model to predict the data of covid-19 analysis . Several factors affect the spread of covid-19 , we will try to analysis when covid out break is more .

As a result, it will help us to control the spread of covid-19 in affected area.

SOLUTION APPROACH

We will be going through some basic plots available in matplotlib and make it more aesthetically pleasing.

Here are the Visualizations we'll be designing using matplotlib .

Simple Plot

Pie Chart

Initializing Dataset

Importing Dataset on Covid-19 India case time series

```
data = pd.read_csv('case_time_series.csv')
```

SOLUTION APPROACH

case_time_series.csv dataset has 7 column. We will be collecting Daily Confirmed Daily Recovered and Daily Deceased in variables as array.

Y = data.iloc[61:,1].values #Stores Daily Confirmed R = data.iloc[61:,3].values #Stores Daily Recovered D = data.iloc[61:,5].values #Stores Daily DeceasedX = data.iloc[61:,0] #Stores Date‘*Y’* variable stores the ‘Daily Confirmed’ corona virus cases

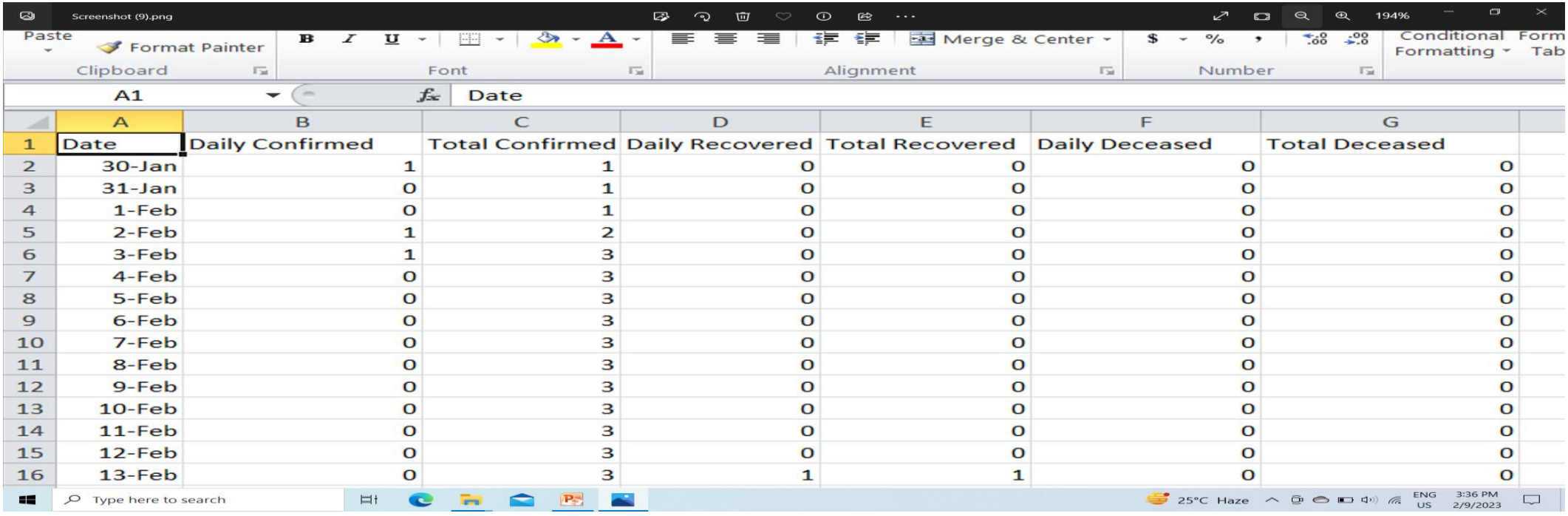
‘*R’* variable stores the ‘Daily Recovered’ corona virus cases

‘D*’* variable stores the ‘Daily Deceased’ corona virus cases

And ‘X’ variable stores the ‘Date’ column

DATASET

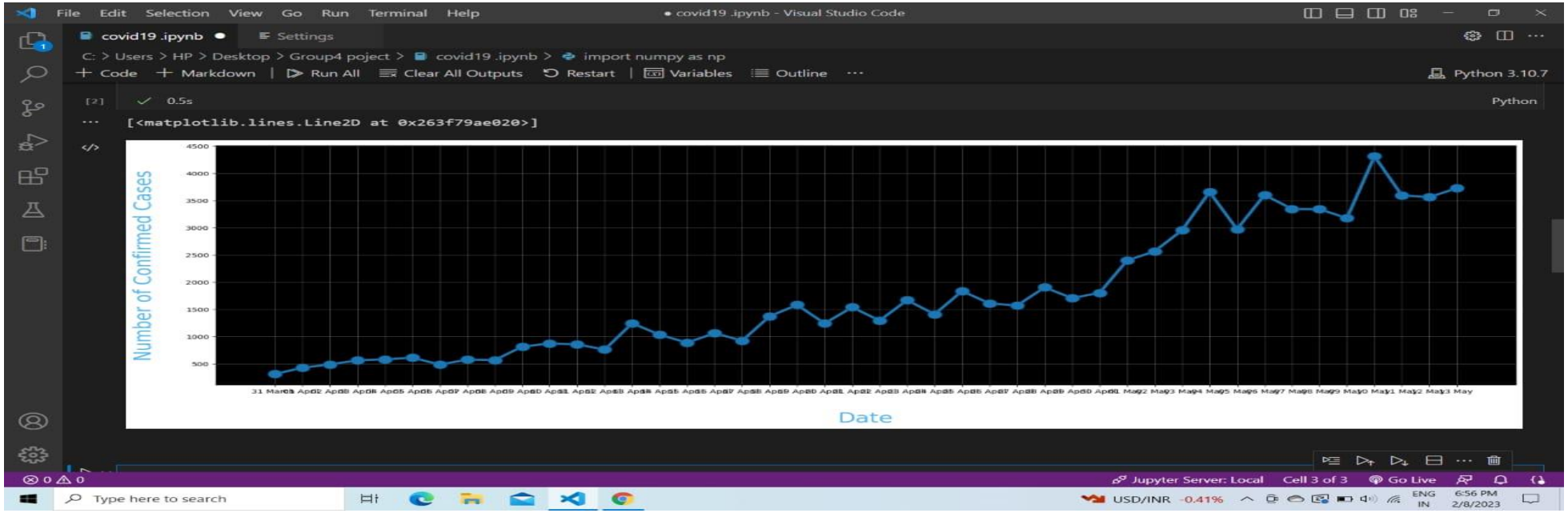
- **The data we have used in this project was given by our instructor . The dataset consists of 106 rows and 7 columns with no null values. Column data consist of daily confirmed cases ,total confirmed cases ,daily recovered cases , daily decreasing case, total recovered cases .**



	A	B	C	D	E	F	G
1	Date	Daily Confirmed	Total Confirmed	Daily Recovered	Total Recovered	Daily Deceased	Total Deceased
2	30-Jan	1	1	0	0	0	0
3	31-Jan	0	1	0	0	0	0
4	1-Feb	0	1	0	0	0	0
5	2-Feb	1	2	0	0	0	0
6	3-Feb	1	3	0	0	0	0
7	4-Feb	0	3	0	0	0	0
8	5-Feb	0	3	0	0	0	0
9	6-Feb	0	3	0	0	0	0
10	7-Feb	0	3	0	0	0	0
11	8-Feb	0	3	0	0	0	0
12	9-Feb	0	3	0	0	0	0
13	10-Feb	0	3	0	0	0	0
14	11-Feb	0	3	0	0	0	0
15	12-Feb	0	3	0	0	0	0
16	13-Feb	0	3	1	1	0	0

OBSERVATION

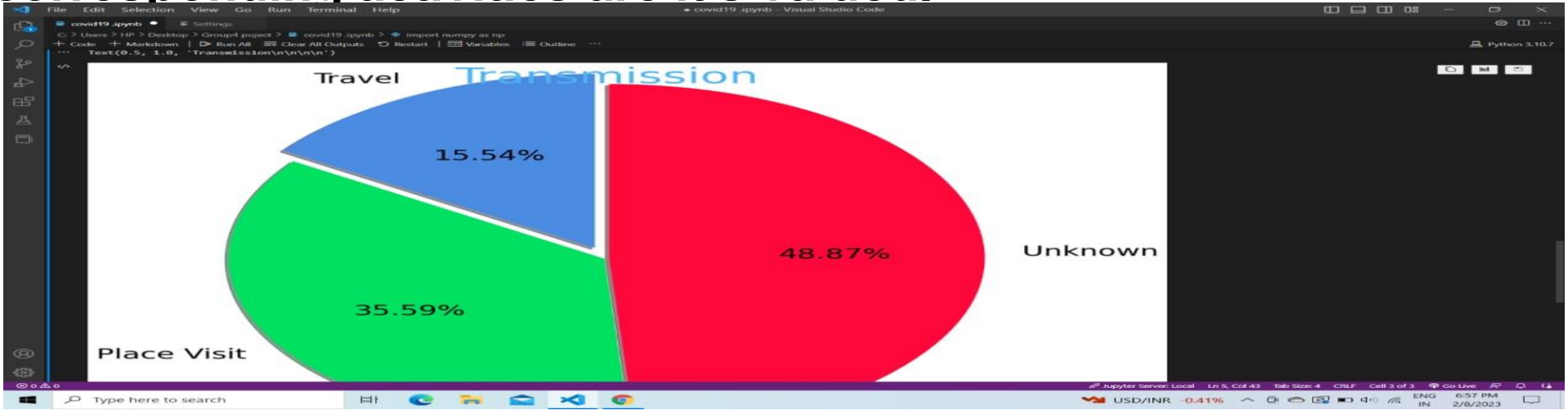
We get a Simple Plot on the execution of above code which looks like this where X-axis has Dates and Y-axis has Number of Confirmed cases.



OBSERVATION

We'll be plotting the Transmission Pie Chart to understand the how the virus is spreading based on Travel, Place Visit and Unknown reason.

slices = [62, 142, 195] activities = ['Travel', 'Place Visit', 'Unknown'] So we have created list slices based on which our Pie Chart will be divided and the corresponding activities are it's valued.



CONCLUSION

At first ,there were many reasons why the

model didn’t work, it was showing many error than we analyse the code and debug

it.

we divided the work among our team member.

In this project , we have developed a python Machine Learning project to analyse the number of covid19 cases.

REFERENCE

Wikipedia

Kaggle

Books on Machine learning

THANK YOU