



CLUSTERING ASSIGNMENT

ANKIT BHARDWAJ

Step to procedure in the Assignment

- Problem Statement
- Data Understanding
- Performing Clustering
- Identifying the Countries
- Recommendations and Conclusion

Problem Statement – Identify top countries that are in direst need of aid. Categorization will be based on socio-economic and health factors that determine the development of the country.

Data Understanding

```
In [284]: # Importing the libraries
import pandas as pd
import numpy as np

import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.cluster import KMeans
```

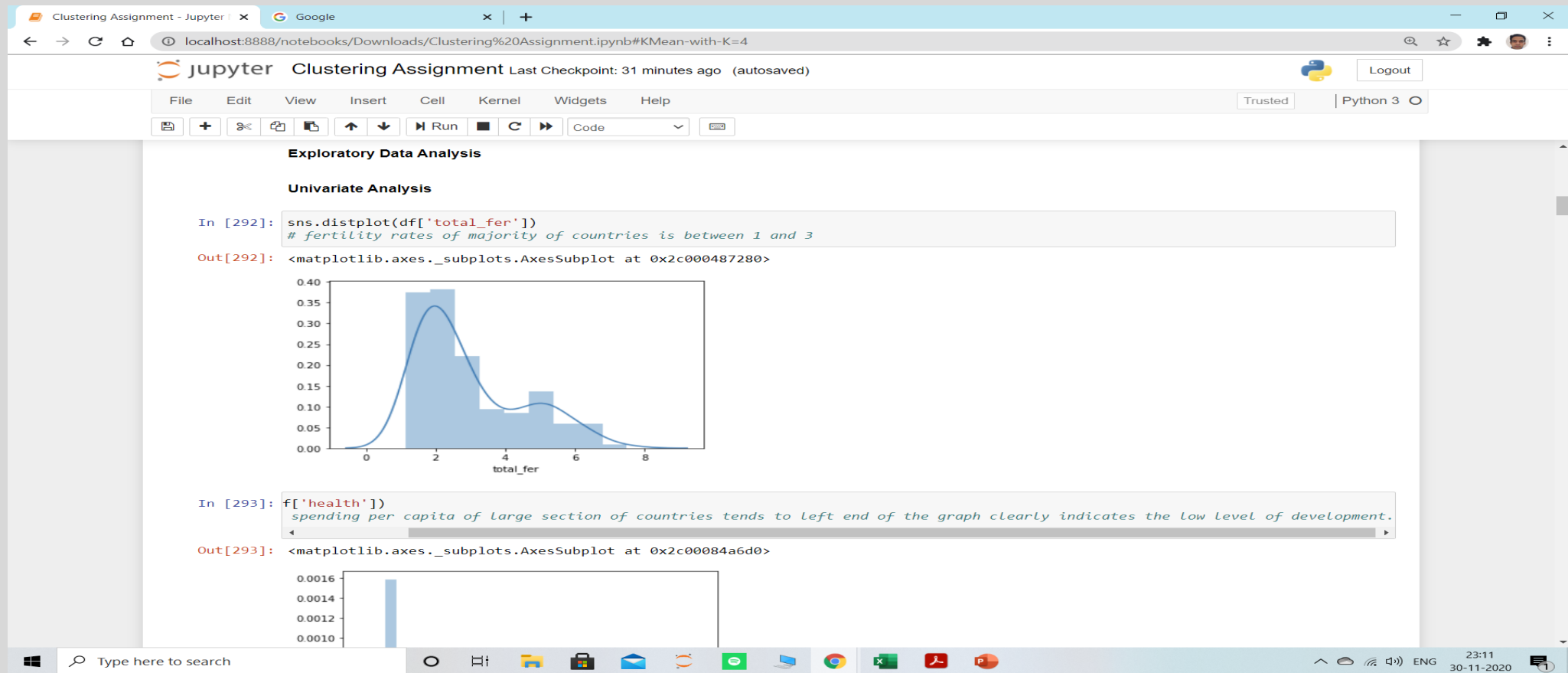
```
In [285]: import warnings
warnings.filterwarnings('ignore')
```

```
In [286]: df = pd.read_csv("Country-data.csv")
df.head()
```

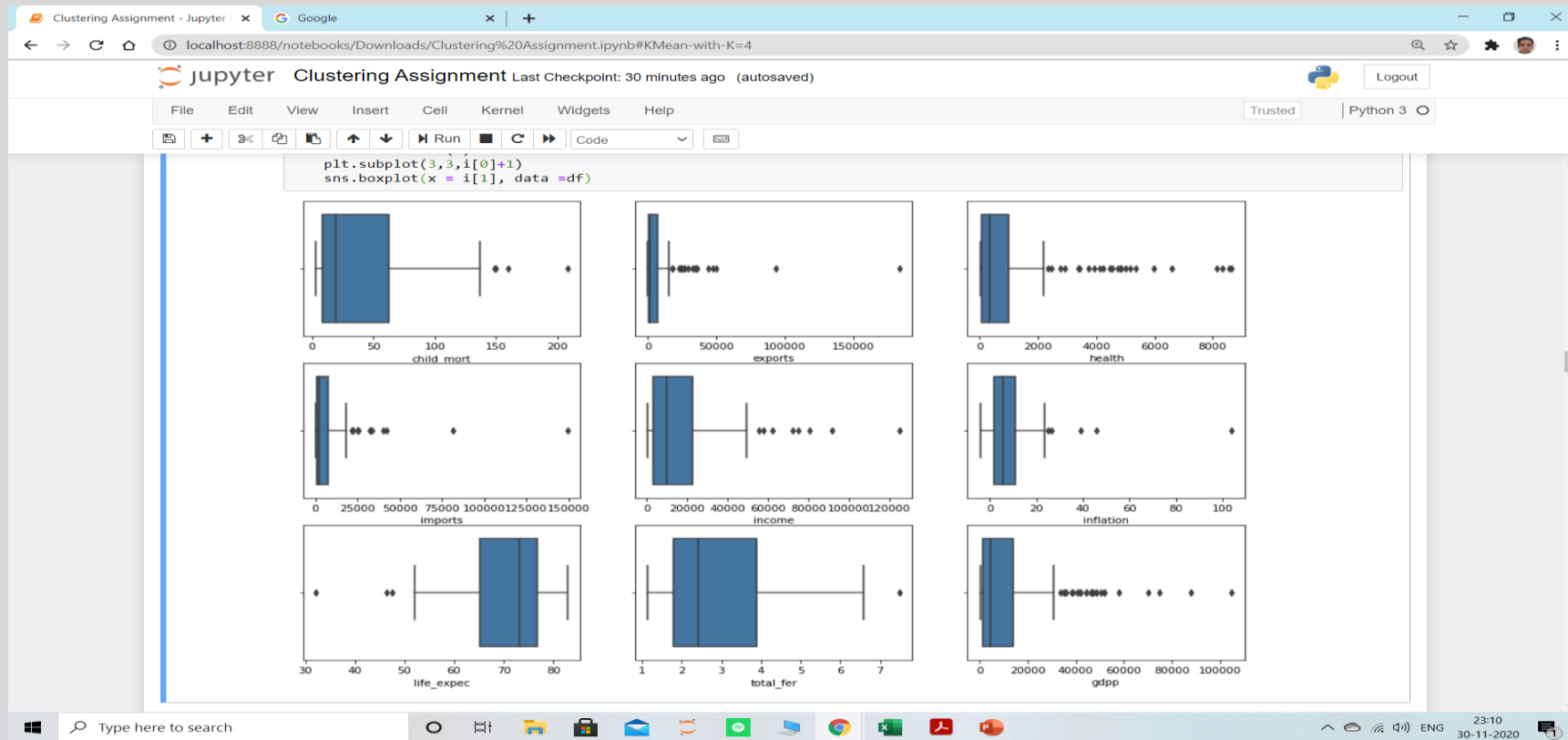
Out[286]:

	country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp
0	Afghanistan	90.2	10.0	7.58	44.9	1610	9.44	56.2	5.82	553
1	Albania	16.6	28.0	6.55	48.6	9930	4.49	76.3	1.65	4090
2	Algeria	27.3	38.4	4.17	31.4	12900	16.10	76.5	2.89	4460
3	Angola	119.0	62.3	2.85	42.9	5900	22.40	60.1	6.16	3530
4	Antigua and Barbuda	10.3	45.5	6.03	58.9	19100	1.44	76.8	2.13	12200

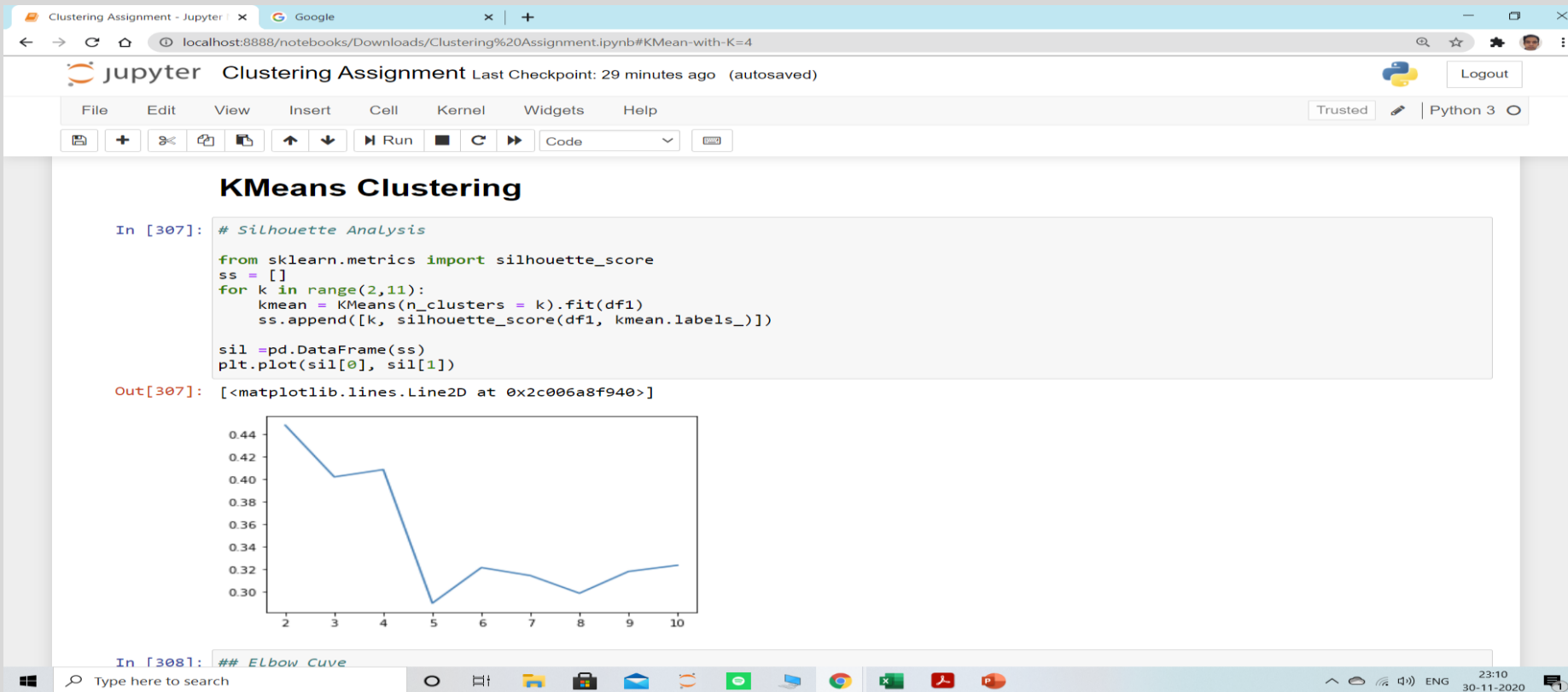
Data Understanding – Various steps will be involved such as Exploratory Data Analysis including uni and bivariate analysis



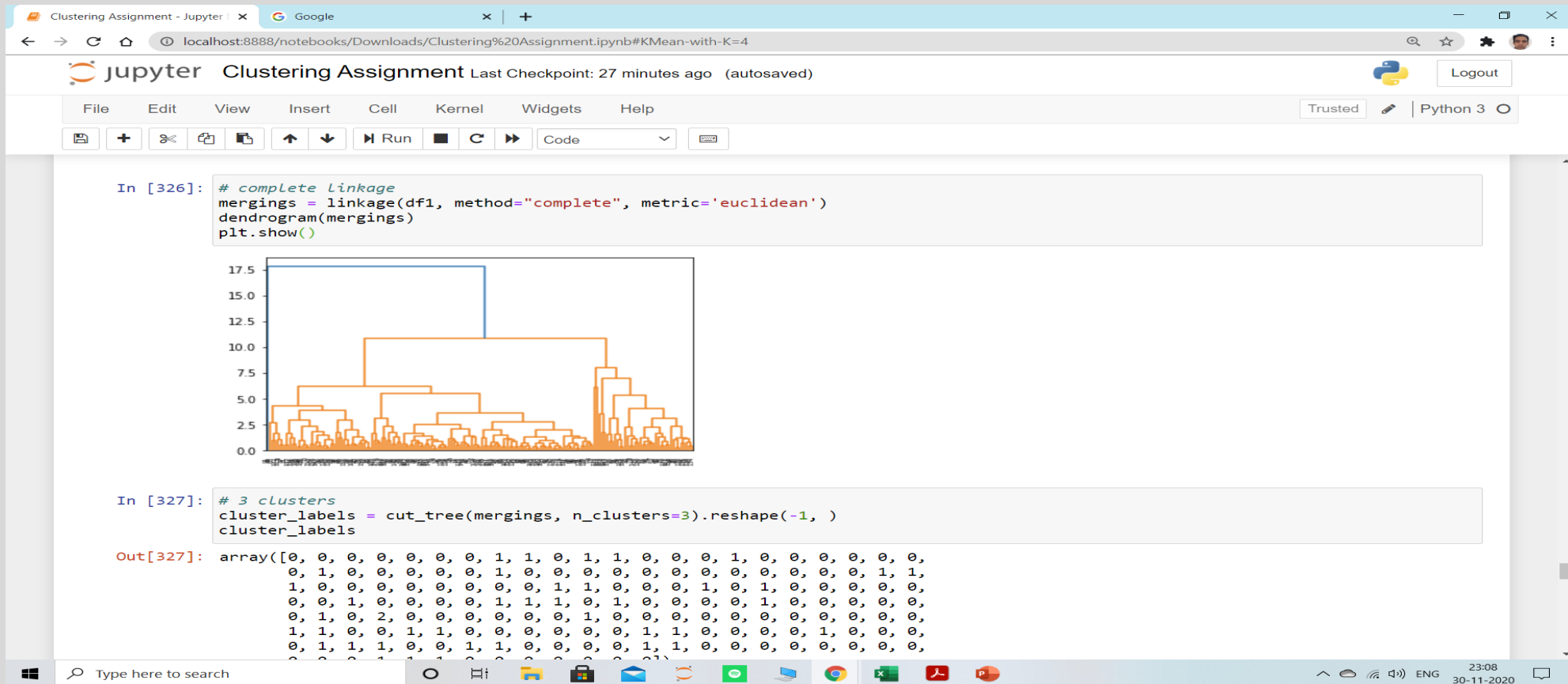
Data Understanding – EDA will be followed by outlier analysis



Performing Clustering – Both kind will be used K-Means and Hierarchical Clustering



Both kind of Hierarchical Clustering will be used single and complete linkage.



Identifying the top five countries in dire need for aid

Clustering Assignment - Jupyter | Google

localhost:8888/notebooks/Downloads/Clustering%20Assignment.ipynb#KMean-with-K=4

jupyter Clustering Assignment Last Checkpoint: 28 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

25 Burkina Faso 116.0 110.400 38.7550 170.200 1430.0 6.810 57.9 5.87 575.0 0 0

```
In [344]: c1.sort_values(by = ['gdpp', 'child_mort', 'income'], ascending = [True, False, True]).head()
```

Out[344]:

	country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	cluster_id	cluster_labels
112	Niger	123.00	77.25600	18.438976	170.8680	837.68	2.55	58.8	7.49	364.32	0	0
37	Congo, Dem. Rep.	116.00	137.27400	26.419400	165.6640	837.68	20.80	57.5	6.54	364.32	0	0
26	Burundi	93.60	53.48304	26.796000	124.4188	837.68	12.30	57.7	6.26	364.32	0	0
88	Liberia	89.30	62.45700	38.586000	302.8020	837.68	5.47	60.8	5.02	364.32	0	0
132	Sierra Leone	145.16	67.03200	52.269000	137.6550	1220.00	17.20	55.0	5.20	399.00	0	0

The top Five countries eligible to recieve financial aid are :-

- Niger
- Democratic Republic of Congo
- Burundi
- Liberia
- Sierra Leone

Type here to search

23:08 30-11-2020

Recommendations and Conclusion

- The top five countries were selected on the basis of three key variables that is GDPP (GDP Per Capita), Child Mortality rate and Income of the country.
- The socio-economic and health factors determine are key indicators of a country to receive any kind of financial aid.

THE END