

# LOCHITA GUPTA

## OASIS INFOBYTE - TASK 2

## UNEMPLOYMENT ANALYSIS

### OBJECTIVE :-

TO ANALYSIS THE UNEMPLOYMENT RATE DURING COVID 19

### Import Libraries

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
```

### Read data

```
In [2]: data = pd.read_csv('Unemployment_Rate_upto_11_2020.csv')
print(data.head())
```

	Region	Date	Frequency	Estimated Unemployment Rate (%) \
0	Andhra Pradesh	31-01-2020	M	5.48
1	Andhra Pradesh	29-02-2020	M	5.83
2	Andhra Pradesh	31-03-2020	M	5.79
3	Andhra Pradesh	30-04-2020	M	20.51
4	Andhra Pradesh	31-05-2020	M	17.43

	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1 \
0	16635535	41.02	South
1	16545652	40.90	South
2	15881197	39.18	South
3	11336911	33.10	South
4	12988845	36.46	South

	longitude	latitude
0	15.9129	79.74
1	15.9129	79.74
2	15.9129	79.74
3	15.9129	79.74
4	15.9129	79.74

In [3]: data.info

```
Out[3]: <bound method DataFrame.info of
Unemployment Rate (%) \
0    Andhra Pradesh    31-01-2020    M    5.48
1    Andhra Pradesh    29-02-2020    M    5.83
2    Andhra Pradesh    31-03-2020    M    5.79
3    Andhra Pradesh    30-04-2020    M    20.51
4    Andhra Pradesh    31-05-2020    M    17.43
..    ...    ...    ...    ...
262    West Bengal    30-06-2020    M    7.29
263    West Bengal    31-07-2020    M    6.83
264    West Bengal    31-08-2020    M    14.87
265    West Bengal    30-09-2020    M    9.35
266    West Bengal    31-10-2020    M    9.98

    Estimated Employed    Estimated Labour Participation Rate (%) Region.1 \
0    16635535    41.02    South
1    16545652    40.90    South
2    15881197    39.18    South
3    11336911    33.10    South
4    12988845    36.46    South
..    ...    ...    ...
262    30726310    40.39    East
263    35372506    46.17    East
264    33298644    47.48    East
265    35707239    47.73    East
266    33962549    45.63    East

    longitude    latitude
0    15.9129    79.740
1    15.9129    79.740
2    15.9129    79.740
3    15.9129    79.740
4    15.9129    79.740
..    ...    ...
262    22.9868    87.855
263    22.9868    87.855
264    22.9868    87.855
265    22.9868    87.855
266    22.9868    87.855
```

[267 rows x 9 columns]>

In [4]: `print(data.describe)`

```
<bound method NDFrame.describe of
ed Unemployment Rate (%) \
0    Andhra Pradesh    31-01-2020    M    5.48
1    Andhra Pradesh    29-02-2020    M    5.83
2    Andhra Pradesh    31-03-2020    M    5.79
3    Andhra Pradesh    30-04-2020    M    20.51
4    Andhra Pradesh    31-05-2020    M    17.43
..    ...    ...    ...    ...
262    West Bengal    30-06-2020    M    7.29
263    West Bengal    31-07-2020    M    6.83
264    West Bengal    31-08-2020    M    14.87
265    West Bengal    30-09-2020    M    9.35
266    West Bengal    31-10-2020    M    9.98

    Estimated Employed    Estimated Labour Participation Rate (%) Region.1 \
0    16635535    41.02    South
1    16545652    40.90    South
2    15881197    39.18    South
3    11336911    33.10    South
4    12988845    36.46    South
..    ...    ...    ...
262    30726310    40.39    East
263    35372506    46.17    East
264    33298644    47.48    East
265    35707239    47.73    East
266    33962549    45.63    East

    longitude    latitude
0    15.9129    79.740
1    15.9129    79.740
2    15.9129    79.740
3    15.9129    79.740
4    15.9129    79.740
..    ...    ...
262    22.9868    87.855
263    22.9868    87.855
264    22.9868    87.855
265    22.9868    87.855
266    22.9868    87.855
```

[267 rows x 9 columns]>

In [5]: `# see if the dataset contains missing null values or not`  
`print(data.isnull().sum())`

```
Region    0
Date    0
Frequency    0
Estimated Unemployment Rate (%)    0
Estimated Employed    0
Estimated Labour Participation Rate (%)    0
Region.1    0
longitude    0
latitude    0
dtype: int64
```

```
In [6]: # rename all the columns
data.columns = ["States", "Date", "Frequency", "Estimated Unemployment Rate", "Estimated
               "Estimated Labour Participation Rate", "Region", "Longitude", "Latitude"]
```

```
In [7]: print(data)
```

	States	Date	Frequency	Estimated Unemployment Rate	\
0	Andhra Pradesh	31-01-2020	M	5.48	
1	Andhra Pradesh	29-02-2020	M	5.83	
2	Andhra Pradesh	31-03-2020	M	5.79	
3	Andhra Pradesh	30-04-2020	M	20.51	
4	Andhra Pradesh	31-05-2020	M	17.43	
..	...	...	...	...	
262	West Bengal	30-06-2020	M	7.29	
263	West Bengal	31-07-2020	M	6.83	
264	West Bengal	31-08-2020	M	14.87	
265	West Bengal	30-09-2020	M	9.35	
266	West Bengal	31-10-2020	M	9.98	

	Estimated employed	Estimated Labour Participation Rate	Region	\
0	16635535	41.02	South	
1	16545652	40.90	South	
2	15881197	39.18	South	
3	11336911	33.10	South	
4	12988845	36.46	South	
..	...	...	...	
262	30726310	40.39	East	
263	35372506	46.17	East	
264	33298644	47.48	East	
265	35707239	47.73	East	
266	33962549	45.63	East	

	Longitude	Latitude
0	15.9129	79.740
1	15.9129	79.740
2	15.9129	79.740
3	15.9129	79.740
4	15.9129	79.740
..	...	...
262	22.9868	87.855
263	22.9868	87.855
264	22.9868	87.855
265	22.9868	87.855
266	22.9868	87.855

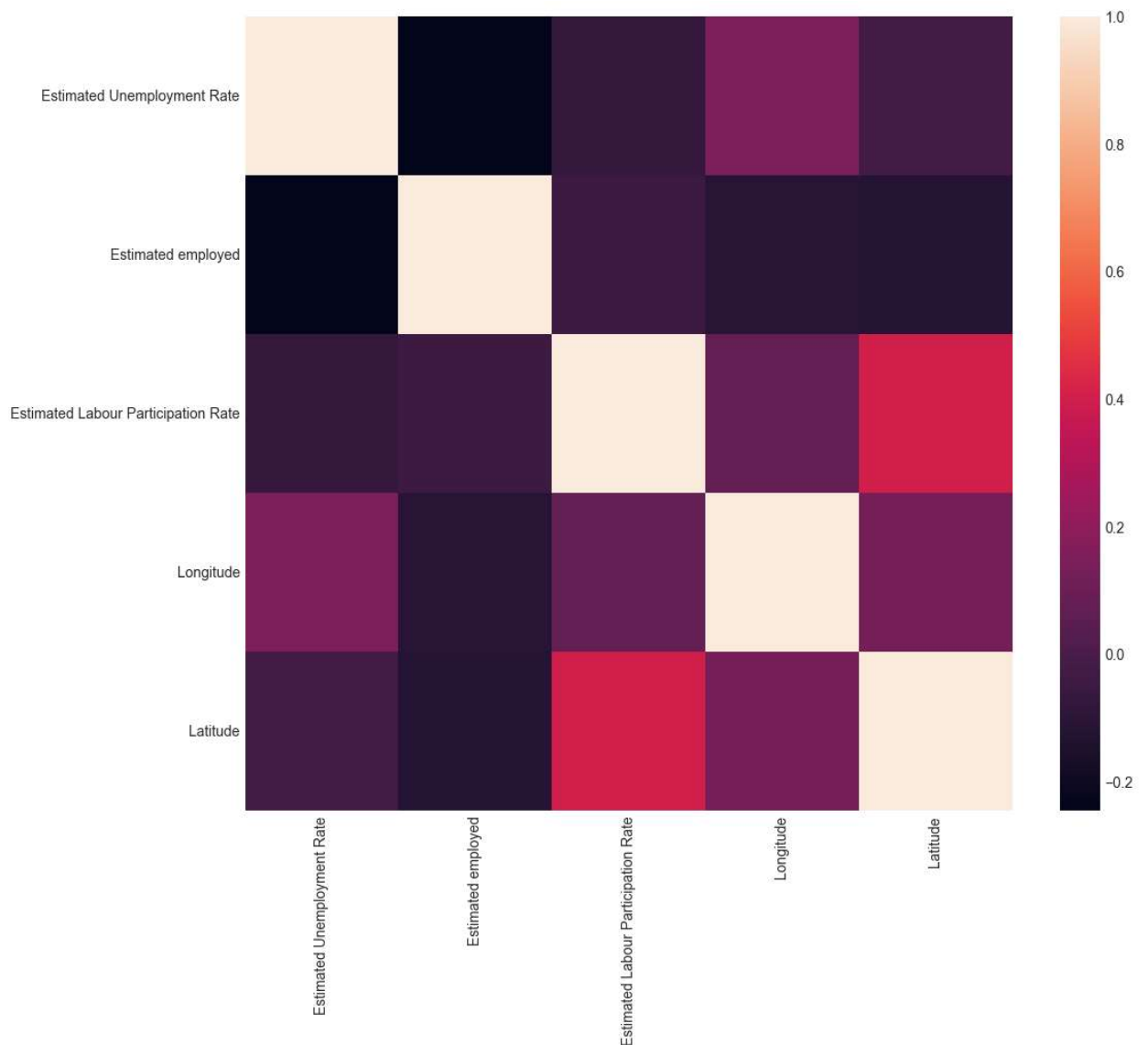
[267 rows x 9 columns]

```
In [8]: #Look at the correlation b/w the features of this dataset.
# Select only numeric columns from the dataset
numeric_data = data.select_dtypes(include=np.number)

# Plot correlation heatmap
plt.style.use("seaborn-whitegrid")
plt.figure(figsize=(12, 10))
sns.heatmap(numeric_data.corr())
plt.show()
```

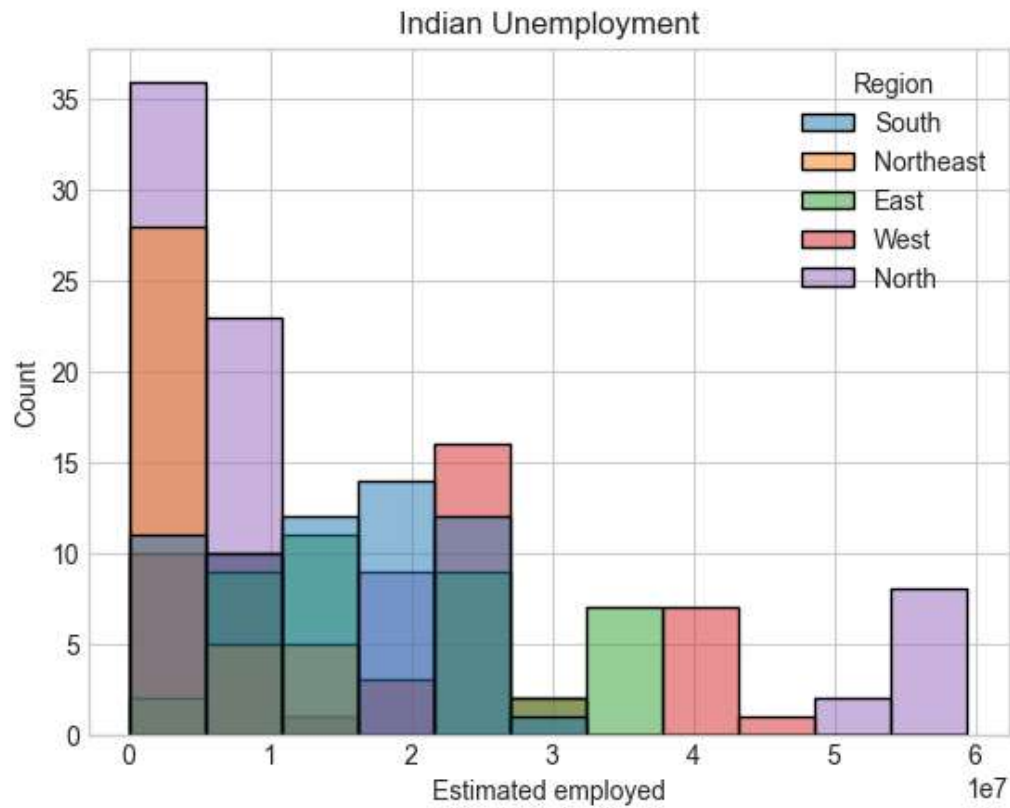
C:\Users\lochita gupta\AppData\Local\Temp\ipykernel\_25220\4284435352.py:6: MatplotlibDeprecationWarning: The seaborn styles shipped by Matplotlib are deprecated since 3.6, as they no longer correspond to the styles shipped by seaborn. However, they will remain available as 'seaborn-v0\_8-<style>'. Alternatively, directly use the seaborn API instead.

```
plt.style.use("seaborn-whitegrid")
```



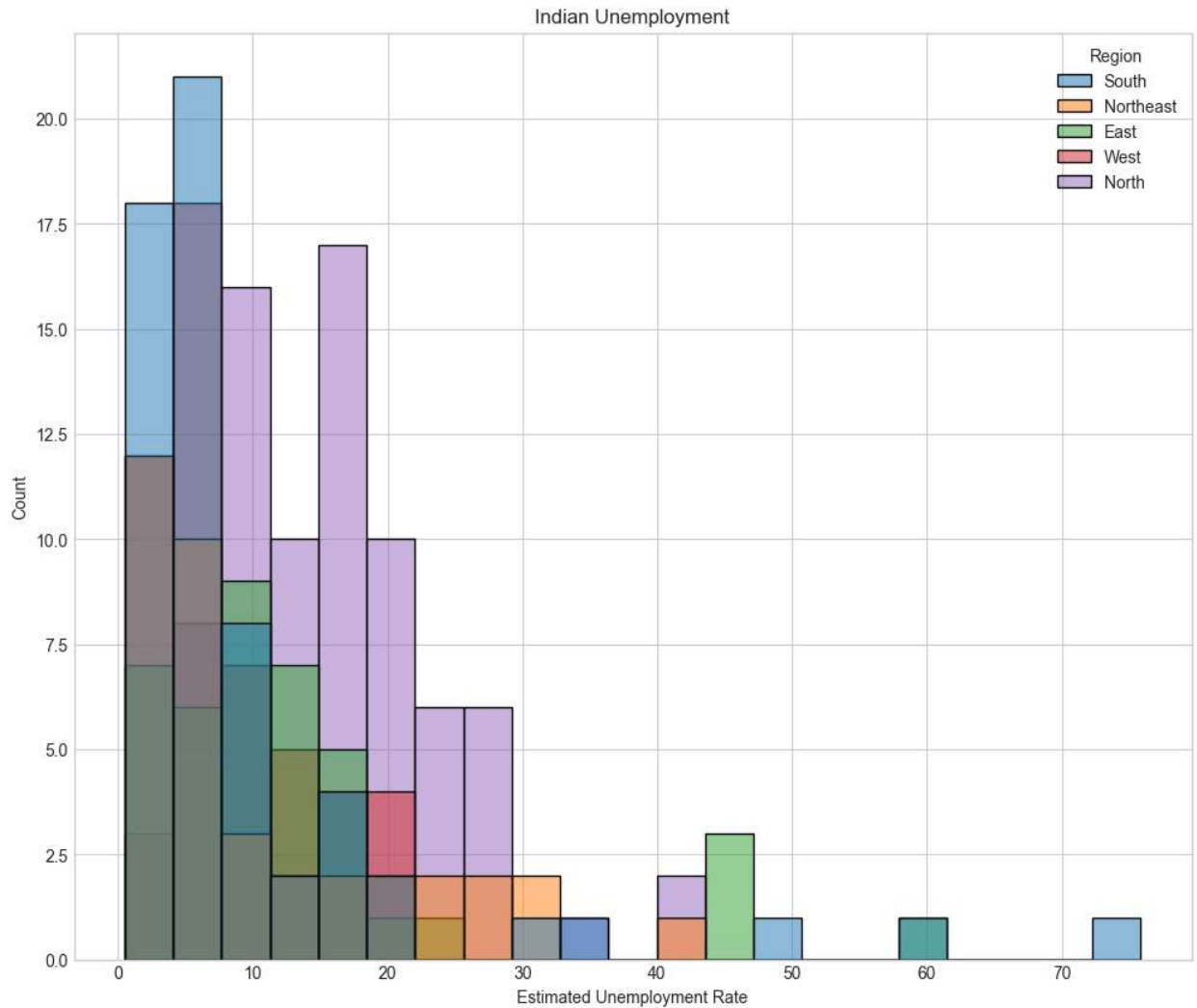
```
In [9]: #Look at the estimated number of employees according to different regions of india.
```

```
In [10]: data.columns = ["States", "Date", "Frequency", "Estimated Unemployment Rate", "Estimated  
          "Estimated Labour Participation Rate", "Region", "Longitude", "Latitude"  
plt.title("Indian Unemployment")  
sns.histplot(x="Estimated employed", hue="Region", data=data)  
plt.show()
```



```
In [11]: # see the unemployment rate according to different regions of india.
```

```
In [12]: plt.figure(figsize=(12,10))  
plt.title("Indian Unemployment")  
sns.histplot(x="Estimated Unemployment Rate", hue="Region", data=data)  
plt.show()
```



```
In [13]: #create a dashboard to analyze the unemployment rate of each Indian state by region  
#use sunburst plot.
```

```
In [14]: Unemployment = data[["States", "Region", "Estimated Unemployment Rate"]]
figure = px.sunburst(Unemployment, path=["Region", "States"],
                    values="Estimated Unemployment Rate",
                    width=700, height=700, color_continuous_scale="RdYlGn",
                    title="Unemployment Rate in India")
figure.show()
```

## Unemployment Rate in India



## Summary :-

In conclusion, the project successfully analyzed the unemployment rate during the COVID-19 period in India. The visualizations and dashboard provided valuable insights into regional variations in unemployment and the estimated number of employed individuals. This analysis can aid policymakers, researchers, and organizations in understanding the impact of the pandemic on employment in different regions of India.



