UNEMPLOYMENT ANALYSIS WITH PYTHON

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as pt
import seaborn as sb
```

```
data_unemp = pd.read_csv("/content/Unemployment_Rate_upto_11_2020.csv")
```

```
data_unemp
```

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longitude	latitude
0	Andhra Pradesh	31-01-2020	M	5.48	16635535	41.02	South	15.9129	79.740
1	Andhra Pradesh	29-02-2020	M	5.83	16545652	40.90	South	15.9129	79.740
2	Andhra Pradesh	31-03-2020	M	5.79	15881197	39.18	South	15.9129	79.740
3	Andhra Pradesh	30-04-2020	M	20.51	11336911	33.10	South	15.9129	79.740
4	Andhra Pradesh	31-05-2020	M	17.43	12988845	36.46	South	15.9129	79.740
...	...	...	...	...	...	...	...	...	...

```
data_unemp.head(20)
```

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longitude	latitude
0	Andhra Pradesh	31-01-2020	M	5.48	16635535	41.02	South	15.9129	79.7400
1	Andhra Pradesh	29-02-2020	M	5.83	16545652	40.90	South	15.9129	79.7400
2	Andhra Pradesh	31-03-2020	M	5.79	15881197	39.18	South	15.9129	79.7400
3	Andhra Pradesh	30-04-2020	M	20.51	11336911	33.10	South	15.9129	79.7400
4	Andhra Pradesh	31-05-2020	M	17.43	12988845	36.46	South	15.9129	79.7400
5	Andhra Pradesh	30-06-2020	M	3.31	19805400	47.41	South	15.9129	79.7400
6	Andhra Pradesh	31-07-2020	M	8.34	15431615	38.91	South	15.9129	79.7400
7	Andhra Pradesh	31-08-2020	M	6.96	15251776	37.83	South	15.9129	79.7400
8	Andhra Pradesh	30-09-2020	M	6.40	15220312	37.47	South	15.9129	79.7400
9	Andhra Pradesh	31-10-2020	M	6.59	15157557	37.34	South	15.9129	79.7400
10	Assam	31-01-2020	M	4.66	13051904	52.98	Northeast	26.2006	92.9376
11	Assam	29-02-2020	M	4.41	10088268	40.77	Northeast	26.2006	92.9376
		31-03-							

```
data_unemp.tail(30)
```

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longitude	latitude
237	Uttar Pradesh	31-01-2020	M	7.58	59433759	39.63	North	26.8467	80.9462
238	Uttar Pradesh	29-02-2020	M	6.94	59433759	39.63	North	26.8467	80.9462
239	Uttar Pradesh	03-03-2020	M	10.11	56976338	38.89	North	26.8467	80.9462

Basic information of dataset

```
data_unemp.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 267 entries, 0 to 266
Data columns (total 9 columns):
 #   Column                                   Non-Null Count  Dtype  
---  ------                                   --------------  -----  
 0   Region                                   267 non-null    object 
 1    Date                                    267 non-null    object 
 2    Frequency                               267 non-null    object 
 3    Estimated Unemployment Rate (%)         267 non-null    float64
 4    Estimated Employed                      267 non-null    int64  
 5    Estimated Labour Participation Rate (%) 267 non-null    float64
 6   Region.1                                 267 non-null    object 
 7   longitude                                267 non-null    float64
 8   latitude                                 267 non-null    float64
dtypes: float64(4), int64(1), object(4)
memory usage: 18.9+ KB
```

```
data_unemp.describe()
```

	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	longitude	latitude
count	267.000000	2.670000e+02	267.000000	267.000000	267.000000
mean	12.236929	1.396211e+07	41.681573	22.826048	80.532425
std	10.803283	1.336632e+07	7.845419	6.270731	5.831738
min	0.500000	1.175420e+05	16.770000	10.850500	71.192400
25%	4.845000	2.838930e+06	37.265000	18.112400	76.085600
50%	9.650000	9.732417e+06	40.390000	23.610200	79.019300
75%	16.755000	2.187869e+07	44.055000	27.278400	85.279900
max	75.850000	5.943376e+07	69.690000	33.778200	92.937600

```
data_unemp.size
```

```
2403
```

```
data_unemp.shape
```

```
(267, 9)
```

```
data_unemp.columns
```

```
Index(['Region', ' Date', ' Frequency', ' Estimated Unemployment Rate (%)',
       ' Estimated Employed', ' Estimated Labour Participation Rate (%)',
       'Region.1', 'longitude', 'latitude'],
      dtype='object')
```

```
data_unemp.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
```

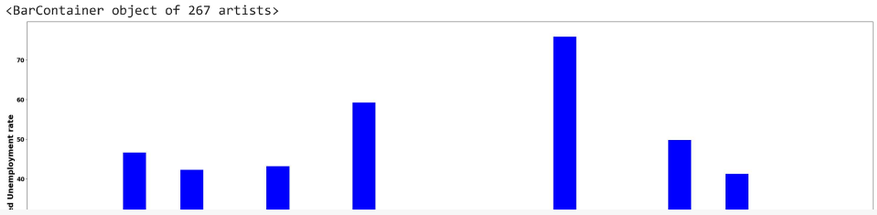
```
#for checking duplicacy
data_unemp.duplicated().sum()
```

```
0
```

Data Visualization

```
data_unemp=pd.DataFrame(data_unemp)
y=data_unemp[' Estimated Unemployment Rate (%)']
x=data_unemp['Region']
pr= pt.figure(figsize=(40, 15))
pt.xlabel("States",fontweight='bold',fontsize=20)
pt.ylabel("Estimated Unemployment rate",fontweight='bold',fontsize=20)
pt.xticks(fontweight='bold',rotation='vertical',fontsize=20)
pt.yticks(fontweight='bold',fontsize=15)
pt.bar(x,y, color='b',align='center')
```

```
<BarContainer object of 267 artists>
```



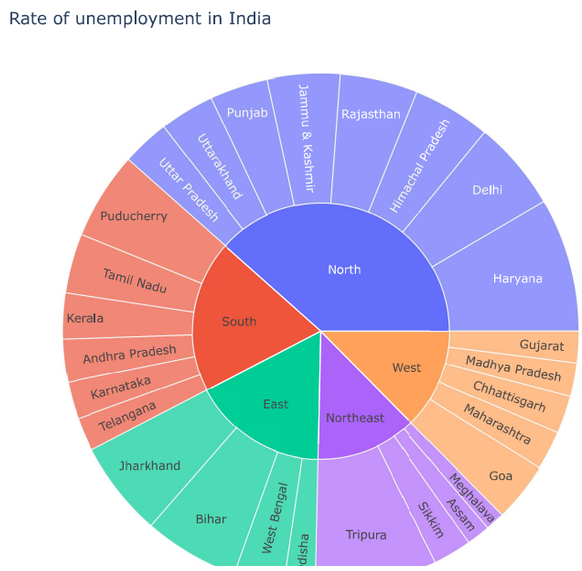

```
# State wise rate of unemplyement

u_emp= data_unemp[['Region',' Estimated Unemployment Rate (%)']].groupby('Region').sum().sort_values(by=' Estimated Unemployment Rate (%)', ascending =False)
u_emp
```

Region	Estimated Unemployment Rate (%)
Haryana	274.77
Tripura	250.55
Jharkhand	195.39
Bihar	194.71
Delhi	184.14
Puducherry	179.42
Himachal Pradesh	160.65
Rajasthan	158.68
Jammu & Kashmir	148.30
Tamil Nadu	121.87
Goa	121.67
Punjab	119.81
Uttarakhand	111.56
West Bengal	101.92
Uttar Pradesh	97.37
Kerala	94.34
Andhra Pradesh	86.64
Maharashtra	79.79
Sikkim	78.34
Chhattisgarh	78.19
Karnataka	76.68
Madhya Pradesh	68.54
Telangana	68.33
Odisha	64.62
Gujarat	63.76
Assam	48.56
Meghalaya	38.66

Display unemployement rate

```
import plotly.express as pl
```

```
unemp_data= data_unemp[["Region", "Region.1", ' Estimated Unemployment Rate (%)']]
figure= pl.sunburst(unemp_data, path=["Region.1", "Region"],values=' Estimated Unemployment Rate (%)',width=700, height=700, color_continuous_scale="spectral",title="Rate of u

figure.show(renderer='colab')
figure.show(renderer='notebook')
```



Scatterplot

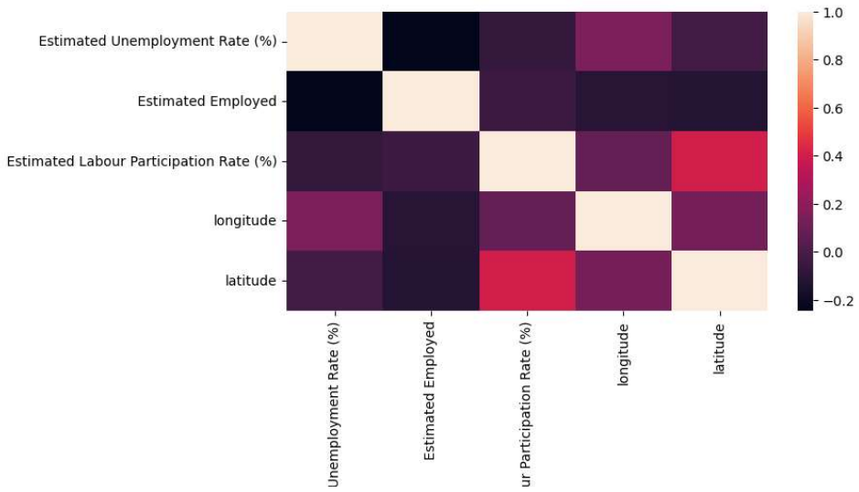
```
import plotly.express as px
```

```
df = pd.read_csv('/content/Unemployment_Rate_upto_11_2020.csv', encoding='UTF-8')

fig = px.scatter(df, x="Region", y=' Estimated Unemployment Rate (%)', color=' Estimated Labour Participation Rate (%)',
                 title="Scatterplot")


fig.show(renderer='colab')
fig.show(renderer='notebook')
```

Heatmap

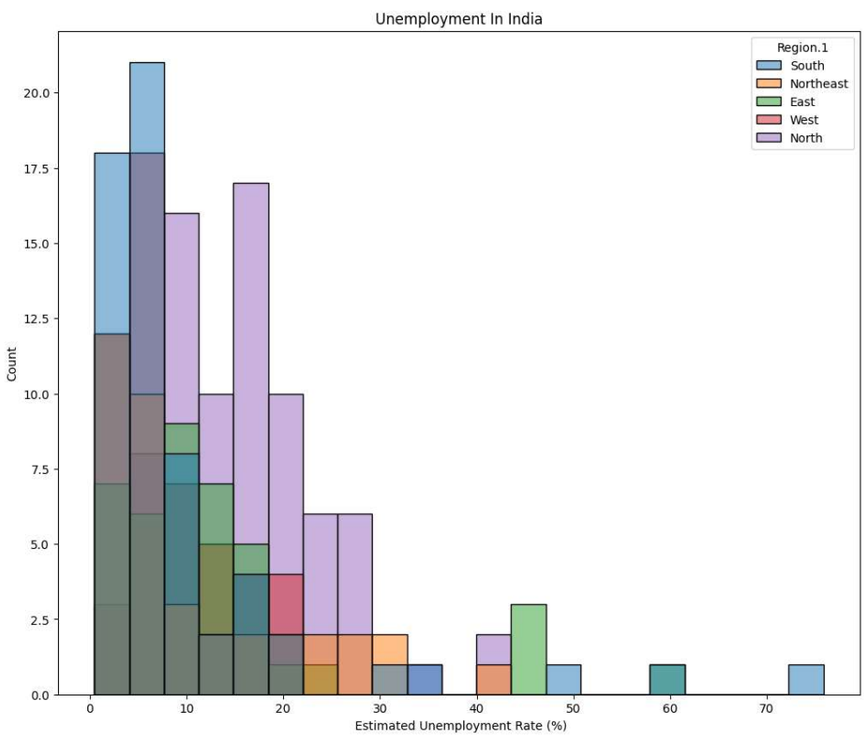
```
pt.figure(figsize=(8,4))
sb.heatmap(data_unemp.corr())
pt.show()
```

```
<ipython-input-21-1f1227181cc7>:2: FutureWarning:

The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False.
```



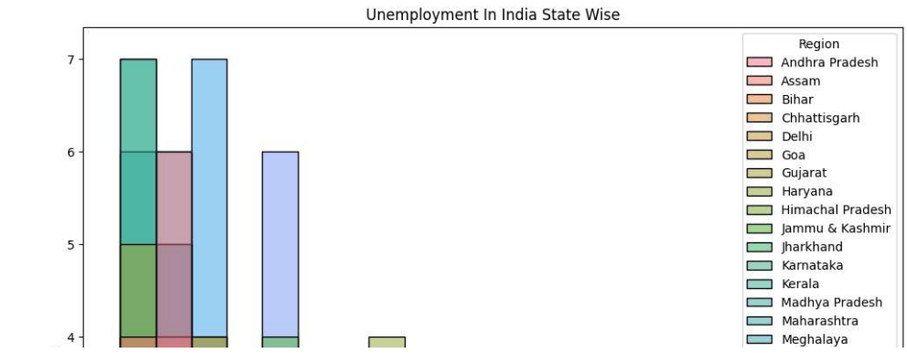
Histogram Plot

```
pt.figure(figsize=(12,10))
pt.title('Unemployment In India')
sb.histplot(x=' Estimated Unemployment Rate (%)', hue="Region.1", data=data_unemp)
pt.show()
```



State wise unemployemnt rate

```
pt.figure(figsize=(12,10))
pt.title('Unemployment In India State Wise')
sb.histplot(x=' Estimated Unemployment Rate (%)', hue="Region", data=data_unemp)
pt.show()
```



**So Here we concluded the states which has the biggest amount of unemployment in the India **

The project by: Abhishek Wagh