UNEMPLOYMENT ANALYSIS WITH PYTHON

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

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

```python
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
...	...	...	...	...	...	...	...	...	...
262	West Bengal	30-06-2020	M	7.29	30726310	40.39	East	22.9868	87.855
263	West Bengal	31-07-2020	M	6.83	35372506	46.17	East	22.9868	87.855
264	West Bengal	31-08-2020	M	14.87	33298644	47.48	East	22.9868	87.855
265	West Bengal	30-09-2020	M	9.35	35707239	47.73	East	22.9868	87.855
266	West Bengal	31-10-2020	M	9.98	33962549	45.63	East	22.9868	87.855

267 rows × 9 columns

```python
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
12	Assam	31-03-2020	M	4.77	11542888	46.73	Northeast	26.2006	92.9376
13	Assam	30-04-2020	M	11.06	6830817	29.55	Northeast	26.2006	92.9376
14	Assam	31-05-2020	M	9.55	11367897	48.26	Northeast	26.2006	92.9376
15	Assam	30-06-2020	M	0.60	9095944	35.07	Northeast	26.2006	92.9376
16	Assam	31-07-2020	M	3.77	10286757	40.88	Northeast	26.2006	92.9376
17	Assam	31-08-2020	M	5.53	9781310	39.52	Northeast	26.2006	92.9376
18	Assam	30-09-2020	M	1.19	14107641	54.38	Northeast	26.2006	92.9376
19	Assam	31-10-2020	M	3.02	11949329	46.84	Northeast	26.2006	92.9376

```python
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	8.98	58060531	39.23	North	26.8467	80.9462
239	Uttar Pradesh	31-03-2020	M	10.11	56976338	38.89	North	26.8467	80.9462
240	Uttar Pradesh	30-04-2020	M	21.54	50915056	39.73	North	26.8467	80.9462
241	Uttar Pradesh	31-05-2020	M	20.41	49801902	38.22	North	26.8467	80.9462
242	Uttar Pradesh	30-06-2020	M	9.47	55380649	37.29	North	26.8467	80.9462
243	Uttar Pradesh	31-07-2020	M	5.56	56201654	36.19	North	26.8467	80.9462
244	Uttar Pradesh	31-08-2020	M	5.79	55831744	35.96	North	26.8467	80.9462
245	Uttar Pradesh	30-09-2020	M	4.18	56106836	35.45	North	26.8467	80.9462
246	Uttar Pradesh	31-10-2020	M	3.75	56539521	35.49	North	26.8467	80.9462
247	Uttarakhand	31-01-2020	M	5.49	2711639	32.99	North	30.0668	79.0193
248	Uttarakhand	29-02-2020	M	4.99	3020931	36.48	North	30.0668	79.0193
249	Uttarakhand	31-03-2020	M	19.85	2539302	36.27	North	30.0668	79.0193
250	Uttarakhand	30-04-2020	M	6.48	2720115	33.23	North	30.0668	79.0193
251	Uttarakhand	31-05-2020	M	8.01	2694072	33.38	North	30.0668	79.0193
252	Uttarakhand	30-06-2020	M	8.61	2656071	33.06	North	30.0668	79.0193
253	Uttarakhand	31-07-2020	M	12.38	2938552	38.07	North	30.0668	79.0193
254	Uttarakhand	31-08-2020	M	14.26	2717528	35.90	North	30.0668	79.0193
255	Uttarakhand	30-09-2020	M	22.26	2695230	39.18	North	30.0668	79.0193
256	Uttarakhand	31-10-2020	M	9.23	2739309	34.03	North	30.0668	79.0193
257	West Bengal	31-01-2020	M	6.94	35820789	47.35	East	22.9868	87.8550
258	West Bengal	29-02-2020	M	4.92	36964178	47.74	East	22.9868	87.8550
259	West Bengal	31-03-2020	M	6.92	35903917	47.27	East	22.9868	87.8550
260	West Bengal	30-04-2020	M	17.41	26938836	39.90	East	22.9868	87.8550
261	West Bengal	31-05-2020	M	17.41	28356675	41.92	East	22.9868	87.8550
262	West Bengal	30-06-2020	M	7.29	30726310	40.39	East	22.9868	87.8550
263	West Bengal	31-07-2020	M	6.83	35372506	46.17	East	22.9868	87.8550
264	West Bengal	31-08-2020	M	14.87	33298644	47.48	East	22.9868	87.8550
265	West Bengal	30-09-2020	M	9.35	35707239	47.73	East	22.9868	87.8550
266	West Bengal	31-10-2020	M	9.98	33962549	45.63	East	22.9868	87.8550

Basic information of dataset

```
In [ ]: 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
```

```
In [ ]: data_unemp.describe()
```

Out[ ]:

	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

```
In [ ]: data_unemp.size
```

Out[ ]: 2403

```
In [ ]: data_unemp.shape
```

Out[ ]: (267, 9)

```
In [ ]: data_unemp.columns
```

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

```
In [ ]: data_unemp.isnull().sum()
```

```
Out[ ]: 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 [ ]: #for checking duplicacy
        data_unemp.duplicated().sum()
```

Out[ ]: 0

Data Visualization

```
In [ ]: 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')
```

Out[ ]: <BarContainer object of 267 artists>

```
In [ ]:    # 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
```

Out[ ]:

	Estimated Unemployment Rate (%)
Region	
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

```
In [ ]:    import plotly.express as pl
```

```
In [70]:   !pip install kaleido
```

```
Requirement already satisfied: kaleido in /usr/local/lib/python3.10/dist-packages (0.2.1)
```

```
In [71]:   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 unemployment in India")

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

Rate of unemployment in India

Scatterplot

```
In [ ]: import plotly.express as px
```

```
In [72]: 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')
```
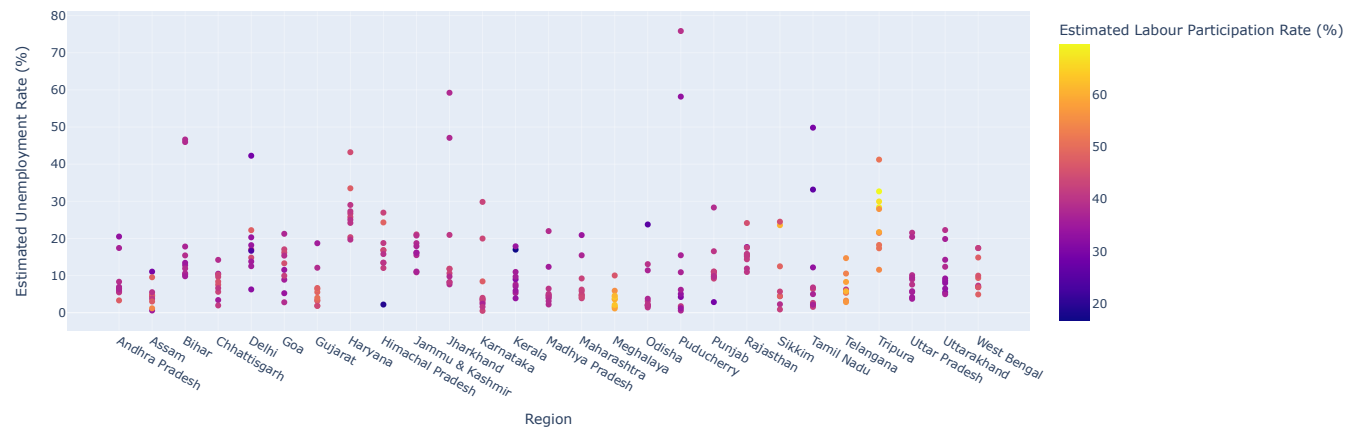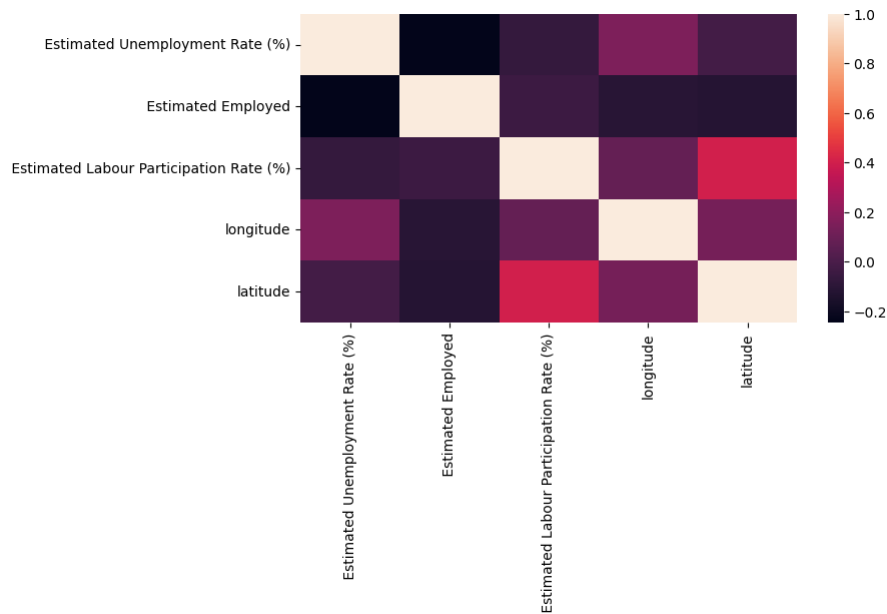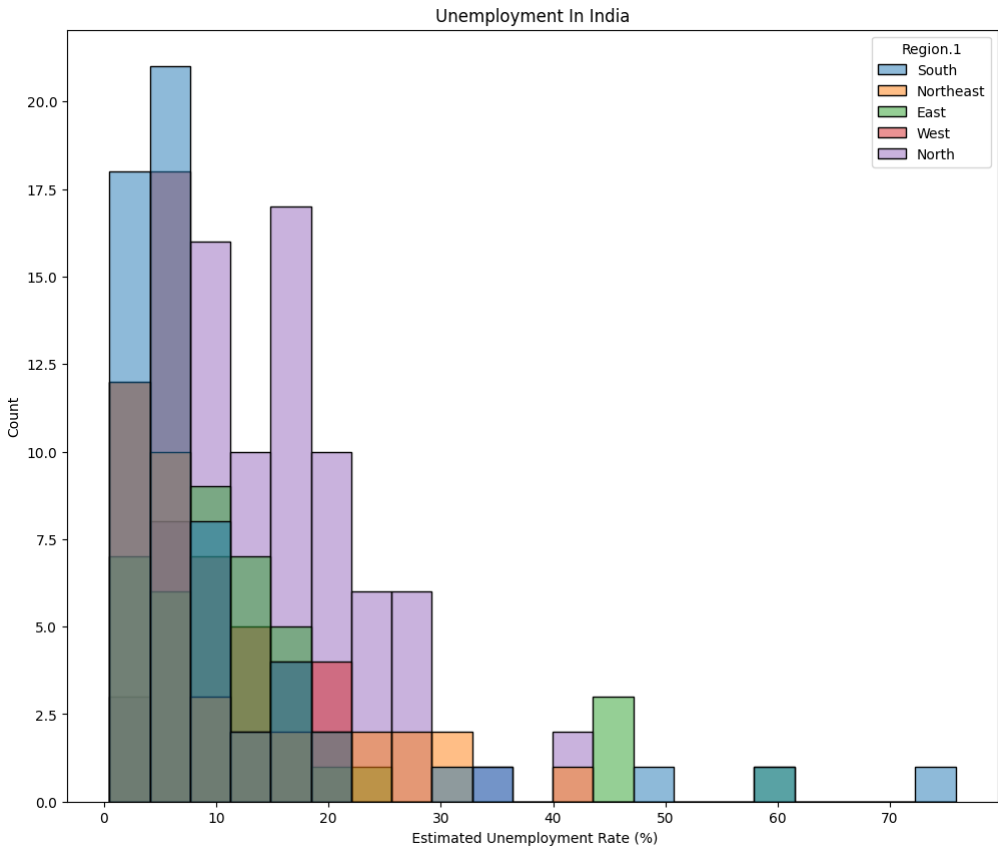
Scatterplot

Heatmap

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

Histogram Plot

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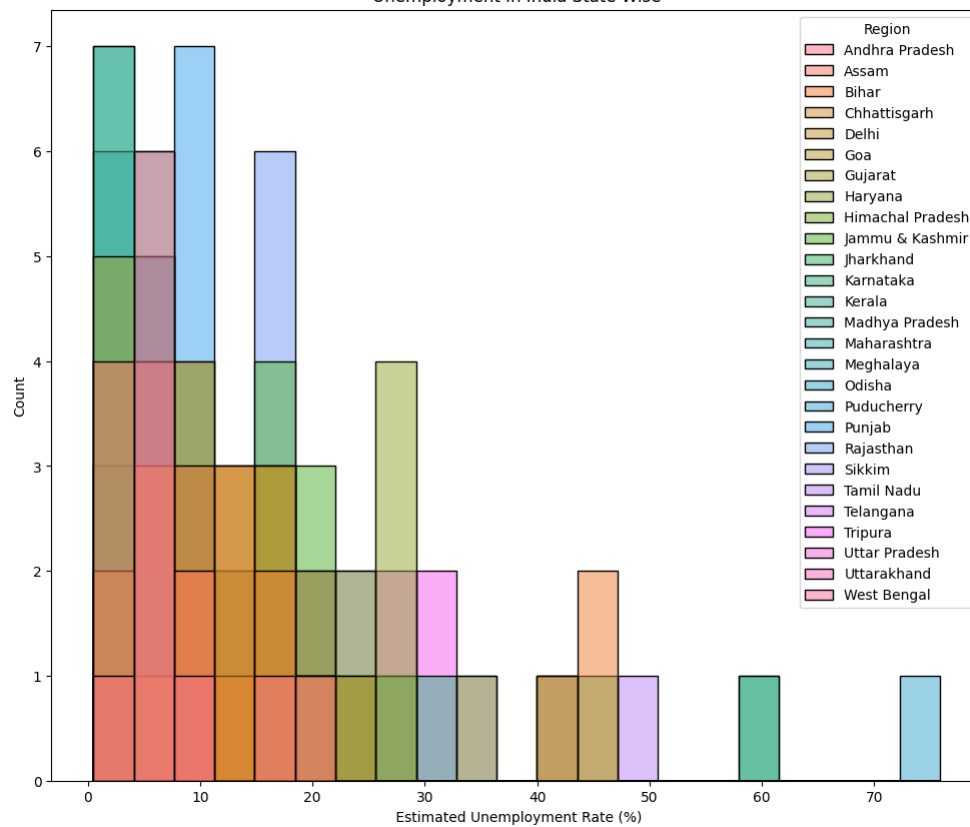
```
In [ ]: pt.figure(figsize=(12,10))
        pt.title('Unemployment In India State Wise')
        sb.histplot(x=' Estimated Unemployment Rate (%)', hue="Region", data=data_unemp)
        pt.show()
```

Unemployment In India State Wise

Region
- Andhra Pradesh
- Assam
- Bihar
- Chhattisgarh
- Delhi
- Goa
- Gujarat
- Haryana
- Himachal Pradesh
- Jammu & Kashmir
- Jharkhand
- Karnataka
- Kerala
- Madhya Pradesh
- Maharashtra
- Meghalaya
- Odisha
- Puducherry
- Punjab
- Rajasthan
- Sikkim
- Tamil Nadu
- Telangana
- Tripura
- Uttar Pradesh
- Uttarakhand
- West Bengal

Export to html

!sudo apt-get install texlive-xetex texlive-fonts-recommended texlive-plain-generic

!jupyter nbconvert --to html /content/UNEMPLOYMENT_ANALYSIS_WITH_PYTHON.ipynb

```
In [65]: !jupyter nbconvert --to html /content/UNEMPLOYMENT_ANALYSIS_WITH_PYTHON.ipynb

[NbConvertApp] Converting notebook /content/UNEMPLOYMENT_ANALYSIS_WITH_PYTHON.ipynb to html
/usr/local/lib/python3.10/dist-packages/nbconvert/filters/widgetsdatatypefilter.py:71: UserWarning: Your element with mimetype(s) dict_keys(['application/vnd.plotly.v1+json']) is not able to be represente
d.
  warn(
[NbConvertApp] Writing 1057890 bytes to /content/UNEMPLOYMENT_ANALYSIS_WITH_PYTHON.html
```