Unemployment Analysis

In [56]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [57]:
```
a=pd.read_csv("unemployment1.csv")
a.head(5)
```

Out[57]:

	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	79.74
1	Andhra Pradesh	29-02-2020	M	5.83	16545652	40.90	South	15.9129	79.74
2	Andhra Pradesh	31-03-2020	M	5.79	15881197	39.18	South	15.9129	79.74
3	Andhra Pradesh	30-04-2020	M	20.51	11336911	33.10	South	15.9129	79.74
4	Andhra Pradesh	31-05-2020	M	17.43	12988845	36.46	South	15.9129	79.74

In [58]:
```
a.tail(5)
```

Out[58]:

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longitude	latitude
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

In [59]:
```
a.ndim
```

Out[59]:
```
2
```

In [60]:
```
a.shape
```

Out[60]:
```
(267, 9)
```

In [61]:
```
a.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 [62]:
```
a.describe()
```

Out[62]:

	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 [63]:
```
a.isnull()
```

Out[63]:

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longitude	latitude
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...
262	False	False	False	False	False	False	False	False	False
263	False	False	False	False	False	False	False	False	False
264	False	False	False	False	False	False	False	False	False
265	False	False	False	False	False	False	False	False	False
266	False	False	False	False	False	False	False	False	False

267 rows × 9 columns

In [64]:
```
a.isna().sum()
```

Out[64]:
```
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 [65]:
```
a.columns=['State','Data','Frequency','EUR','EE','ELPR','Region','longitude','latitude']
```

In [66]:
```
a
```

Out[66]:

	State	Data	Frequency	EUR	EE	ELPR	Region	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

In [67]:
```
a.dtypes
```

Out[67]:
```
State         object
Data          object
Frequency     object
EUR          float64
EE             int64
ELPR         float64
Region        object
longitude    float64
latitude     float64
dtype: object
```

In [68]:
```
a.nunique()
```

Out[68]:
```
State         27
Data          10
Frequency      1
EUR          252
EE           267
ELPR         248
Region         5
longitude     27
latitude      24
dtype: int64
```

In [69]:
```
a.max()
```

Out[69]:
```
State        West Bengal
Data          31-10-2020
Frequency              M
EUR                75.85
EE              59433759
ELPR               69.69
Region              West
longitude       33.7782
latitude        92.9376
dtype: object
```

In [70]:
```
a.min()
```

Out[70]:
```
State        Andhra Pradesh
Data             29-02-2020
Frequency                 M
EUR                     0.5
EE                   117542
ELPR                  16.77
Region                East
longitude            10.85
latitude             71.19
dtype: object
```

Data Visualization

In [71]:
```
sns.heatmap(a.corr(),annot=True)
```

Out[71]:
```
<AxesSubplot:>
```

In [72]:
```
x=a["State"]
x
```

Out[72]:
```
0      Andhra Pradesh
1      Andhra Pradesh
2      Andhra Pradesh
3      Andhra Pradesh
4      Andhra Pradesh
            ...      
262       West Bengal
263       West Bengal
264       West Bengal
265       West Bengal
266       West Bengal
Name: State, Length: 267, dtype: object
```

In [73]:
```
y=a["EUR"]
y
```

Out[73]:
```
0       5.48
1       5.83
2       5.79
3      20.51
4      17.43
       ...  
262     7.29
263     6.83
264    14.87
265     9.35
266     9.98
Name: EUR, Length: 267, dtype: float64
```

In [ ]:
```

```