In [28]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px

data = pd.read_csv("https://raw.githubusercontent.com/amankharwal/Website-data
print(data.head())
```

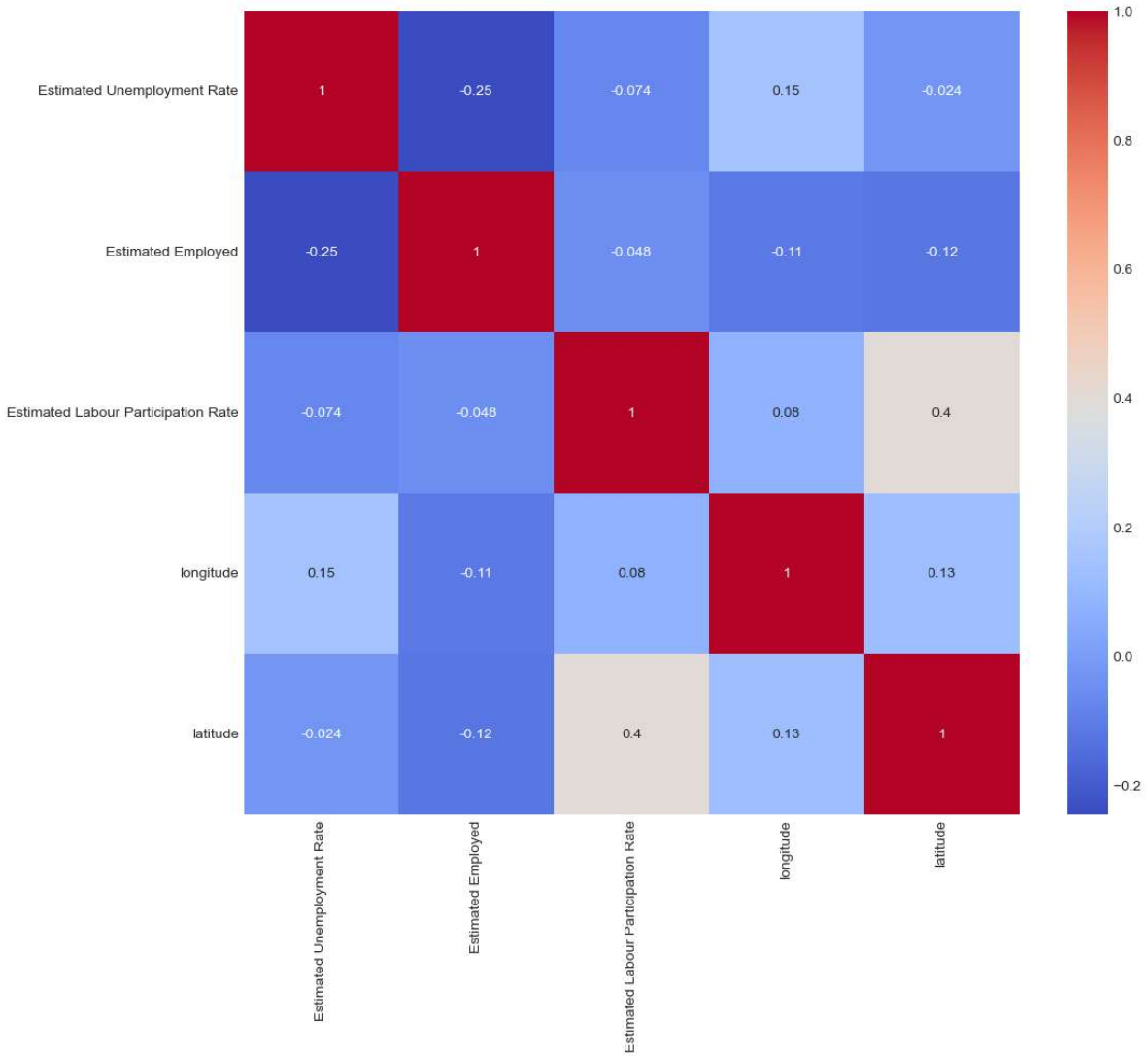
```
           Region         Date  Frequency   Estimated Unemployment Rate (%)
\
0  Andhra Pradesh   31-01-2020          M                              5.48
1  Andhra Pradesh   29-02-2020          M                              5.83
2  Andhra Pradesh   31-03-2020          M                              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 [29]:
```
print(data.isnull().sum())
```

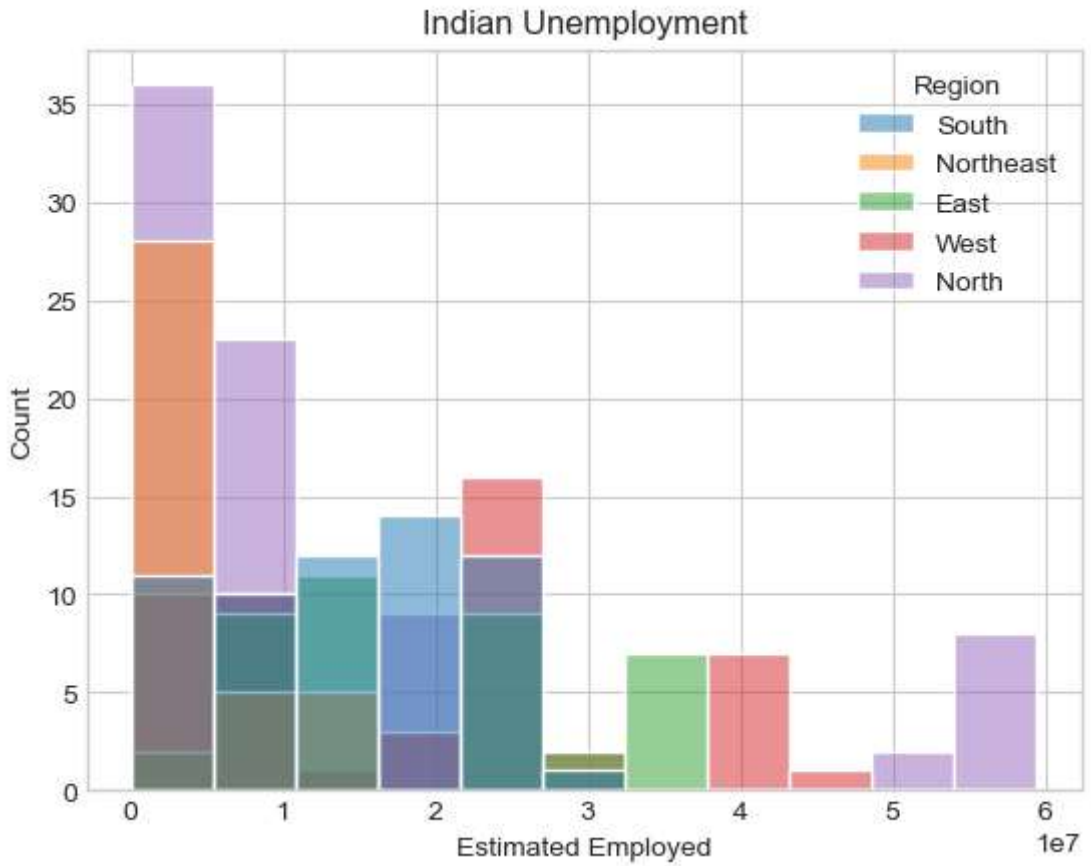
```
Region                                    0
Date                                      0
 Frequency                                0
 Estimated Unemployment Rate (%)          0
 Estimated Employed                       0
 Estimated Labour Participation Rate (%)  0
Region.1                        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

In [30]:
```
data.columns= ["States","Date","Frequency",
               "Estimated Unemployment Rate",
               "Estimated Employed",
               "Estimated Labour Participation Rate",
               "Region","longitude","latitude"]
```

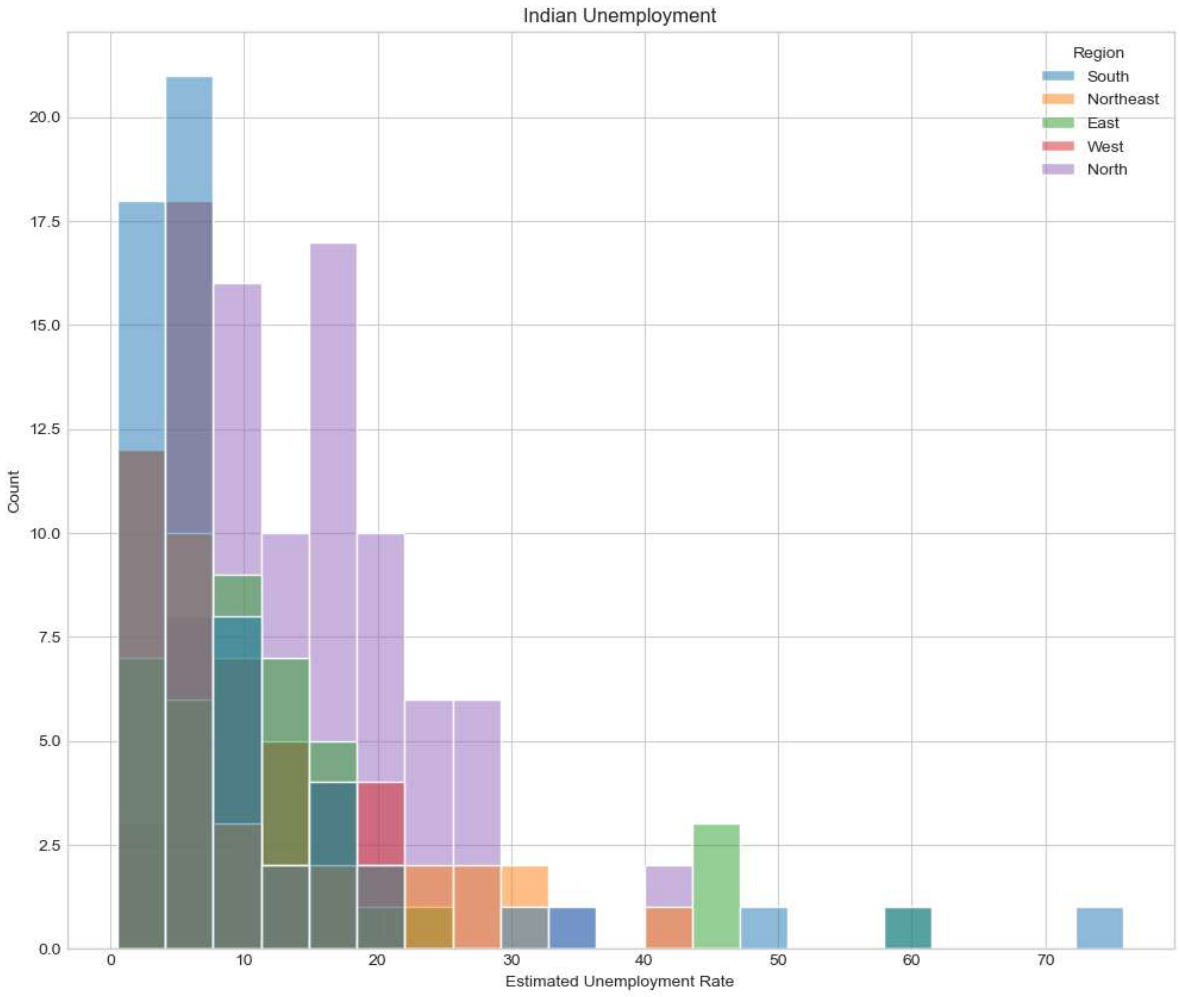
In [31]:
```
numeric_data = data.select_dtypes(include='number')
correlation_matrix = numeric_data.corr()

plt.figure(figsize=(12, 10))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
plt.show()
```



In [32]:
```
data.columns= ["States","Date","Frequency",
               "Estimated Unemployment Rate","Estimated Employed",
               "Estimated Labour Participation Rate","Region",
               "longitude","latitude"]
plt.title("Indian Unemployment")
sns.histplot(x="Estimated Employed", hue="Region", data=data)
plt.show()
```



In [33]:
```
plt.figure(figsize=(12, 10))
plt.title("Indian Unemployment")
sns.histplot(x="Estimated Unemployment Rate", hue="Region", data=data)
plt.show()
```



In [39]:
```
unemploment = data[["States", "Region", "Estimated Unemployment Rate"]]
figure = px.sunburst(unemploment, path=["Region", "States"],
                     values="Estimated Unemployment Rate",
                     width=700, height=700, color_continuous_scale="Reds",
                     title="Unemployment Rate in India")
figure.update_layout(
    sunburstcolorway=["#FF5733", "#FFBD33", "#FF5733", "#FF5733", "#FFBD33", "
    uniformtext=dict(minsize=12, mode='hide'),  # Adjust text size for better
)
figure.show()
```