

```
In [4]: import pandas as pd
data=pd.read_csv("Unemployment in India.csv")
data.head()
```

Out[4]:

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Area
0	Andhra Pradesh	31-05-2019	Monthly	3.65	11999139.0	43.24	Rural
1	Andhra Pradesh	30-06-2019	Monthly	3.05	11755881.0	42.05	Rural
2	Andhra Pradesh	31-07-2019	Monthly	3.75	12086707.0	43.50	Rural
3	Andhra Pradesh	31-08-2019	Monthly	3.32	12285693.0	43.97	Rural
4	Andhra Pradesh	30-09-2019	Monthly	5.17	12256762.0	44.68	Rural

```
In [6]: data.columns
```

```
Out[6]: Index(['Region', ' Date', ' Frequency', ' Estimated Unemployment Rate (%)',
              ' Estimated Employed', ' Estimated Labour Participation Rate (%)',
              'Area'],
              dtype='object')
```

```
In [7]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 7 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Region                                740 non-null    object
1   Date                                  740 non-null    object
2   Frequency                             740 non-null    object
3   Estimated Unemployment Rate (%)       740 non-null    float64
4   Estimated Employed                    740 non-null    float64
5   Estimated Labour Participation Rate (%) 740 non-null    float64
6   Area                                  740 non-null    object
dtypes: float64(3), object(4)
memory usage: 42.1+ KB
```

```
In [10]: data.dtypes
```

```
Out[10]: Region          object
        Date            object
        Frequency        object
        Estimated Unemployment Rate (%) float64
        Estimated Employed float64
        Estimated Labour Participation Rate (%) float64
        Area             object
        dtype: object
```

```
In [17]: data.isna().sum()
```

```
Out[17]: Region          28
        Date            28
        Frequency        28
        Estimated Unemployment Rate (%) 28
        Estimated Employed 28
        Estimated Labour Participation Rate (%) 28
        Area             28
        dtype: int64
```

```
In [19]: print("Drop rows with all NaN values")
        data = data.dropna(how='all')
        data.isna().sum()
```

Drop rows with all NaN values

```
Out[19]: Region          0
        Date            0
        Frequency        0
        Estimated Unemployment Rate (%) 0
        Estimated Employed 0
        Estimated Labour Participation Rate (%) 0
        Area             0
        dtype: int64
```

```
In [26]: data = data.dropna()
```

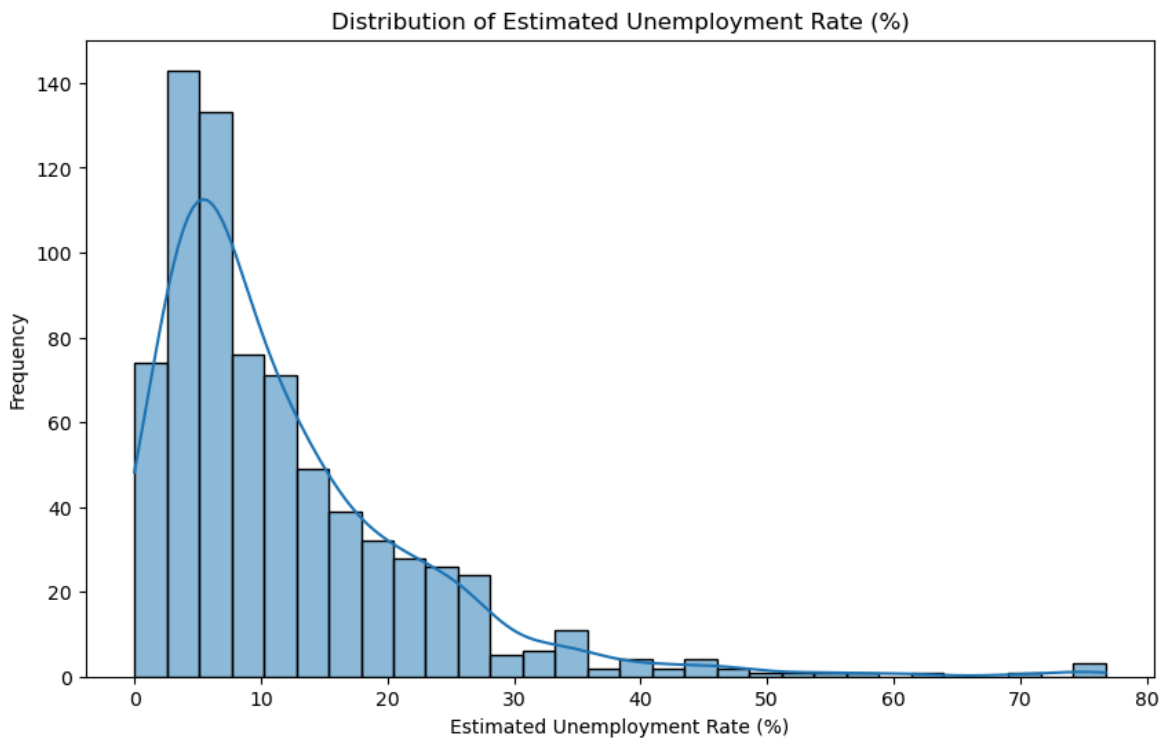
```
In [36]: data[' Date'] = pd.to_datetime(data[' Date'])
        data[' Date']
```

```
Out[36]: 0      2019-05-31
        1      2019-06-30
        2      2019-07-31
        3      2019-08-31
        4      2019-09-30
        ...
        749    2020-02-29
        750    2020-03-31
        751    2020-04-30
        752    2020-05-31
        753    2020-06-30
        Name: Date, Length: 740, dtype: datetime64[ns]
```

```
In [38]: import matplotlib.pyplot as plt
        import seaborn as sns
        # Example 1: Distribution of Estimated Unemployment Rate (%)
        plt.figure(figsize=(10, 6))
        sns.histplot(data[' Estimated Unemployment Rate (%)'], bins=30, kde=True)
        plt.title('Distribution of Estimated Unemployment Rate (%)')
        plt.xlabel('Estimated Unemployment Rate (%)')
```

```
plt.ylabel('Frequency')
plt.show()
```

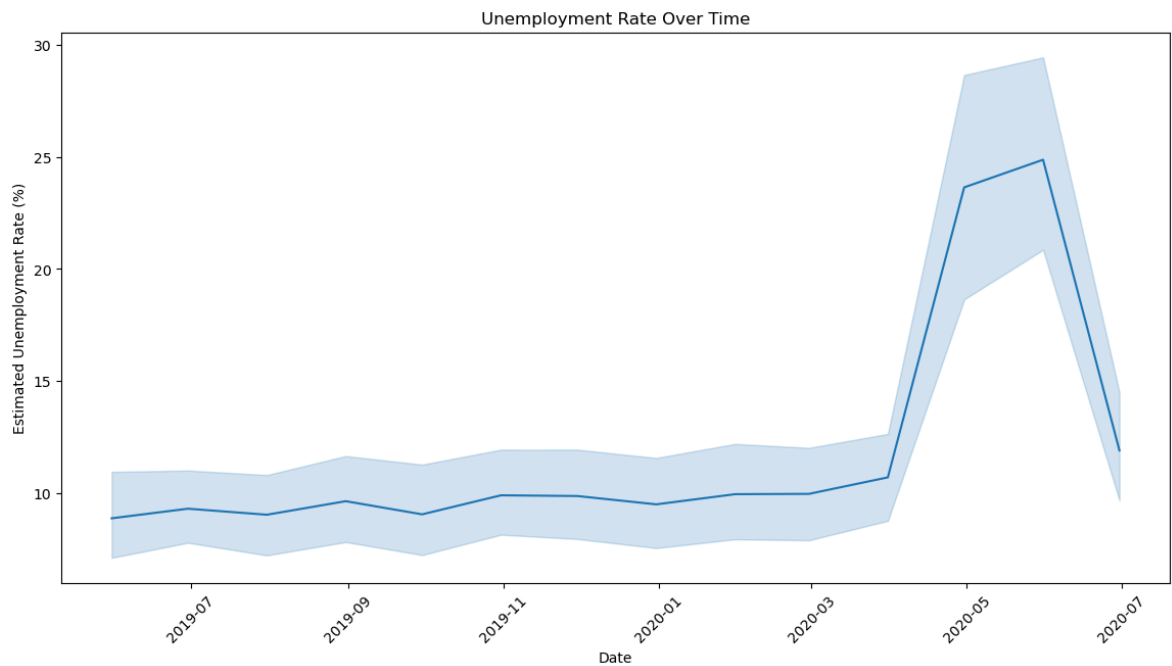
C:\Users\Admin\anaconda3\Lib\site-packages\seaborn_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
with pd.option_context('mode.use_inf_as_na', True):



```
In [64]: print("Unemployment Rate Over Time")
plt.figure(figsize=(14, 7))
sns.lineplot(x=' Date', y=' Estimated Unemployment Rate (%)', data=data)
plt.title('Unemployment Rate Over Time')
plt.xlabel('Date')
plt.ylabel('Estimated Unemployment Rate (%)')
plt.xticks(rotation=45)
plt.show()
```

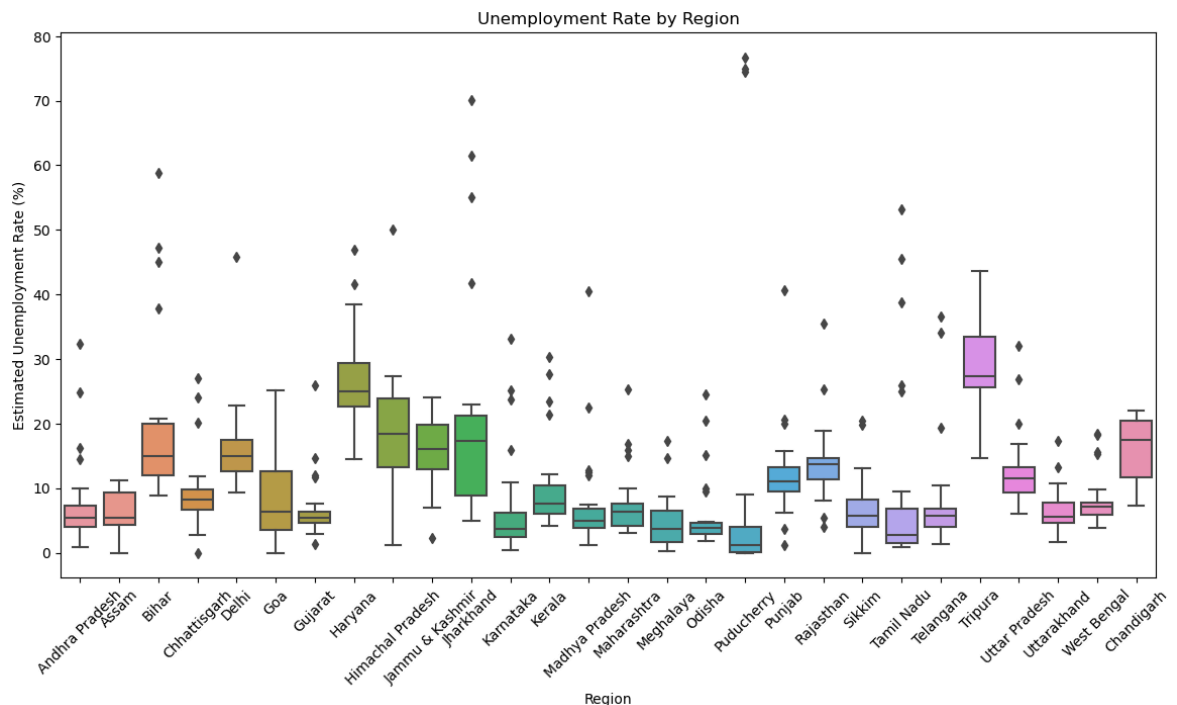
Unemployment Rate Over Time

C:\Users\Admin\anaconda3\Lib\site-packages\seaborn_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
with pd.option_context('mode.use_inf_as_na', True):
C:\Users\Admin\anaconda3\Lib\site-packages\seaborn_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
with pd.option_context('mode.use_inf_as_na', True):



```
In [63]: print("Unemployment Rate by Region (assuming 'Region' is a valid column name)")
plt.figure(figsize=(14, 7))
sns.boxplot(x='Region', y=' Estimated Unemployment Rate (%)', data=data)
plt.title('Unemployment Rate by Region')
plt.xlabel('Region')
plt.ylabel('Estimated Unemployment Rate (%)')
plt.xticks(rotation=45)
plt.show()
```

Unemployment Rate by Region (assuming 'Region' is a valid column name)



```
In [61]: print(" Labour Participation Rate Over Time (assuming ' Estimated Labour Partici")
plt.figure(figsize=(14, 7))
sns.lineplot(x=' Date', y=' Estimated Labour Participation Rate (%)', data=data,
plt.title('Labour Participation Rate Over Time')
plt.xlabel('Date')
plt.ylabel('Estimated Labour Participation Rate (%)')
plt.xticks(rotation=45)
```

```
plt.legend(title='Region', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.show()
```

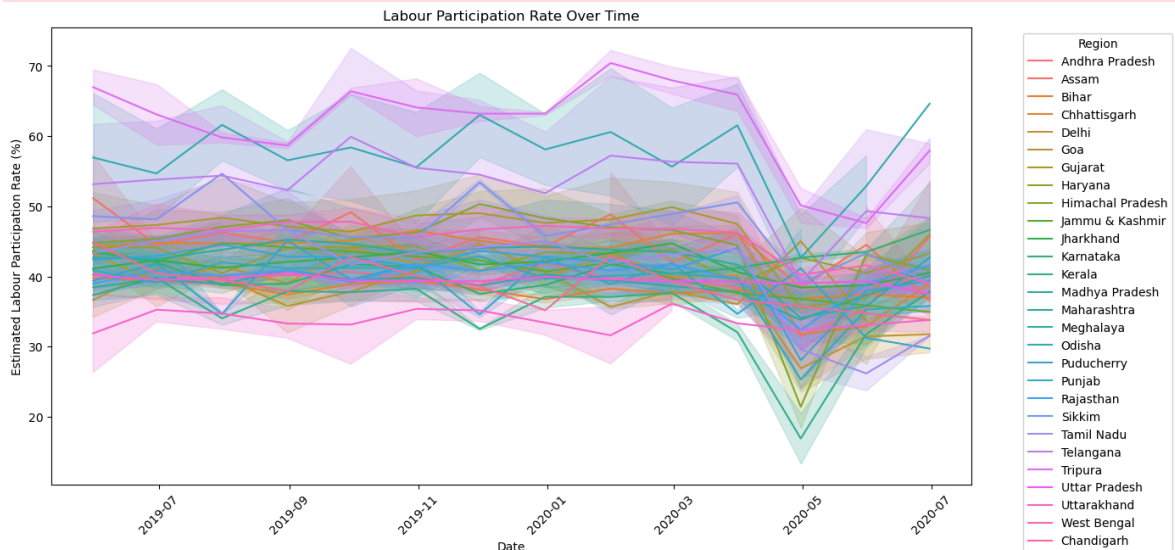
Labour Participation Rate Over Time (assuming 'Estimated Labour Participation Rate (%)' is a valid column name)

C:\Users\Admin\anaconda3\Lib\site-packages\seaborn_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.

```
with pd.option_context('mode.use_inf_as_na', True):
```

C:\Users\Admin\anaconda3\Lib\site-packages\seaborn_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.

```
with pd.option_context('mode.use_inf_as_na', True):
```



```
In [62]: print(" Relationship Between Unemployment Rate and Labour Participation Rate")
plt.figure(figsize=(10, 6))
sns.scatterplot(x=' Estimated Labour Participation Rate (%)', y=' Estimated Unem
plt.title('Unemployment Rate vs Labour Participation Rate')
plt.xlabel('Estimated Labour Participation Rate (%)')
plt.ylabel('Estimated Unemployment Rate (%)')
plt.show()
```

Relationship Between Unemployment Rate and Labour Participation Rate

