Import libraries

In [171…

```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
import seaborn as sns
```

Upload Dataframe

In [74]:

```
df1 = pd.read_csv('Unemployment in India.csv')
df2= pd.read_csv('Unemployment_Rate_upto_11_2020.csv')
```

View some records from each dataframe

In [76]:

```
df1.head(5)
```

Out[76]:

	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 [77]:

```
df2.head(5)
```

Out[77]:

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longi
0	Andhra Pradesh	31-01-2020	M	5.48	16635535	41.02	South	15.
1	Andhra Pradesh	29-02-2020	M	5.83	16545652	40.90	South	15.
2	Andhra Pradesh	31-03-2020	M	5.79	15881197	39.18	South	15.
3	Andhra Pradesh	30-04-2020	M	20.51	11336911	33.10	South	15.
4	Andhra Pradesh	31-05-2020	M	17.43	12988845	36.46	South	15.

Date processing

In [83]:

```
print("first dataframe\n",df1.isnull().sum(),"\n")
print("\n second datafrome \n\n",df2.isnull().sum())
```

```
first dataframe
 Region                                      28
 Date                                       28
 Frequency                                  28
 Estimated Unemployment Rate (%)            28
 Estimated Employed                         28
 Estimated Labour Participation Rate (%)    28
Area                                        28
dtype: int64


 second datafrome 

 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 first dataframe there are some null values so we have to clean for better visualization

In [86]:

```
df1 = df1.dropna()
```

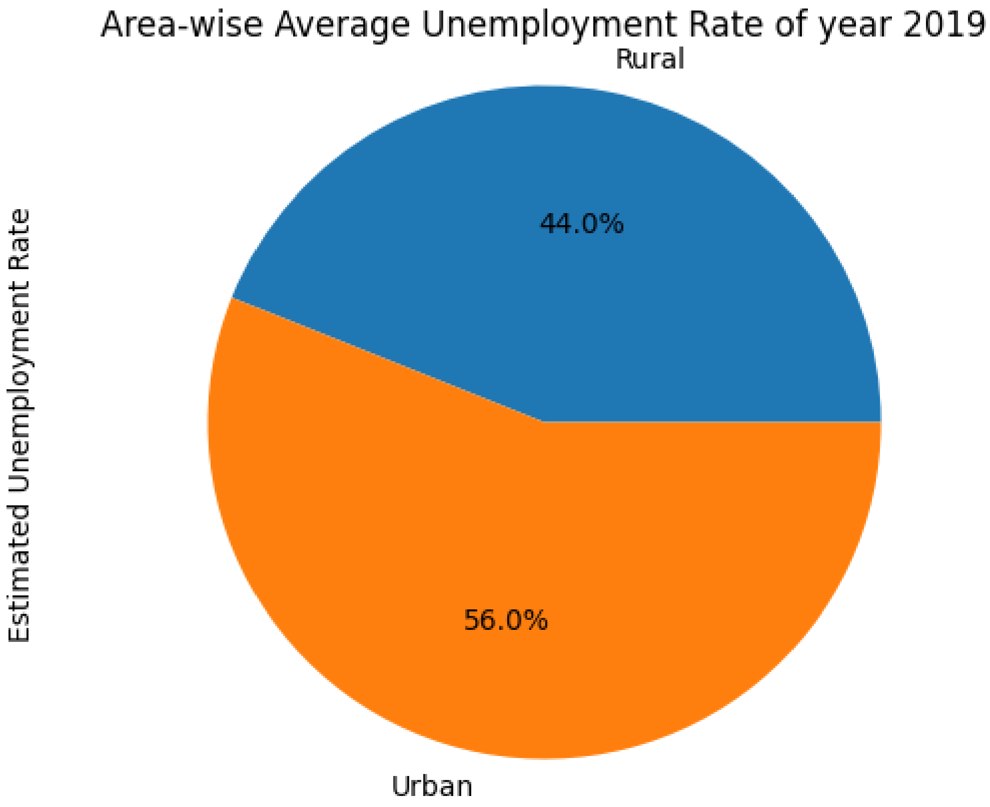
In [87]:
```
print(df1.isnull().sum())
```

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

In [88]:
```
# now we will convert data in columns
df1.columns = ["State","Date","Frequency","Estimated Unemployment Rate","Estimat
df2.columns=["State","Date","Frequency","Estimated Unemployment Rate","Estimatec
```

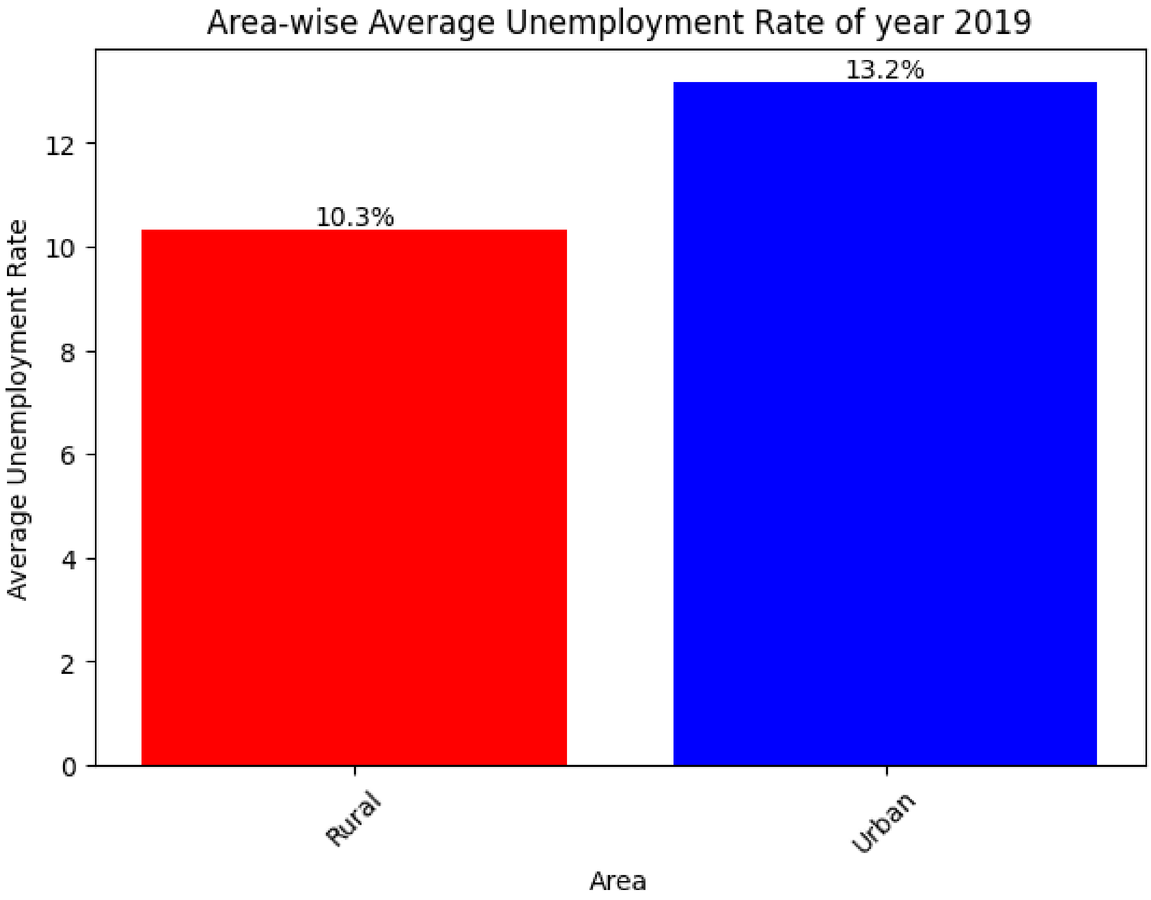
Data Visualization

In [96]:
```
average_of_regions1=df1.groupby("Area")["Estimated Unemployment Rate"].mean()
average_of_regions1.plot(kind='pie',autopct='%1.1f%%')
plt.axis('equal')
plt.title('Area-wise Average Unemployment Rate of year 2019')
plt.show()
```



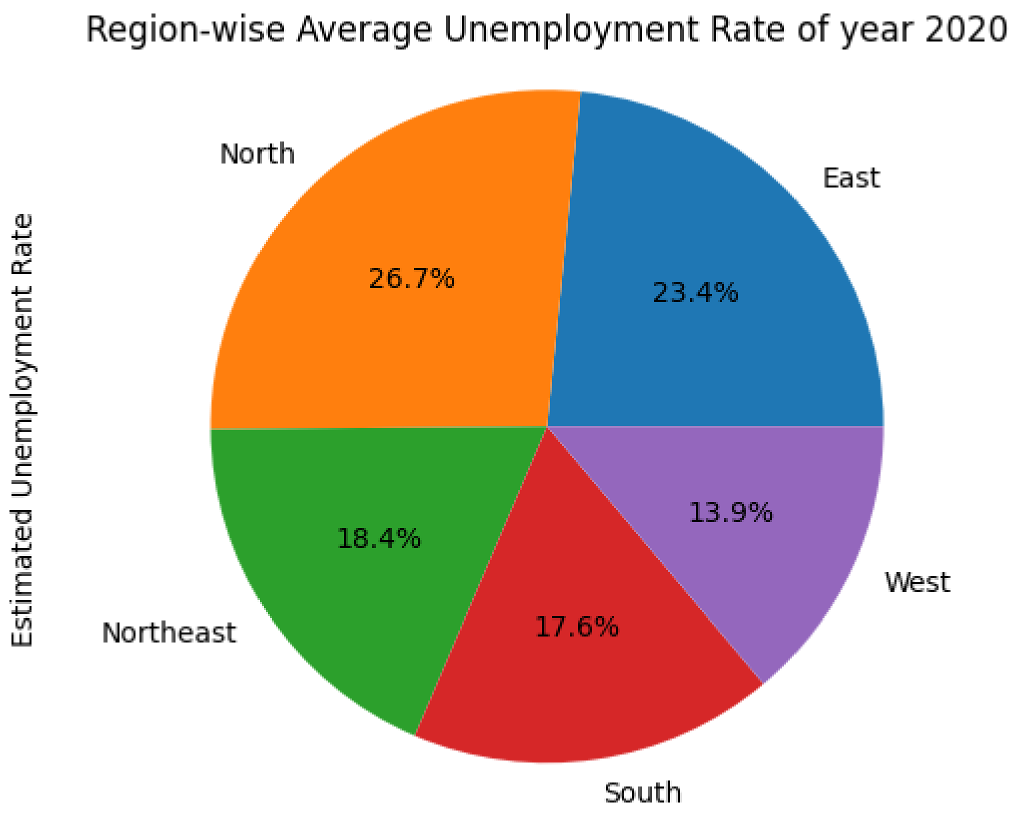
In [111…
```
average_of_regions1 = df1.groupby("Area")["Estimated Unemployment Rate"].mean()

# Create the bar chart

plt.bar(average_of_regions1.index, average_of_regions1.values,width=0.8,color=['

# Add percentage labels on the bars
for index, value in enumerate(average_of_regions1.values):
    plt.text(index, value, f"{value:.1f}%", ha='center', va='bottom')
# Set the title and labels
```

```
plt.title('Area-wise Average Unemployment Rate of year 2019')
plt.xlabel('Area')
plt.ylabel('Average Unemployment Rate')

# Show the plot
plt.tight_layout()  # To adjust spacing and avoid label overlapping (optional)
plt.xticks(rotation=45)  # Rotate x-axis labels for better readability (optional
plt.show()
```



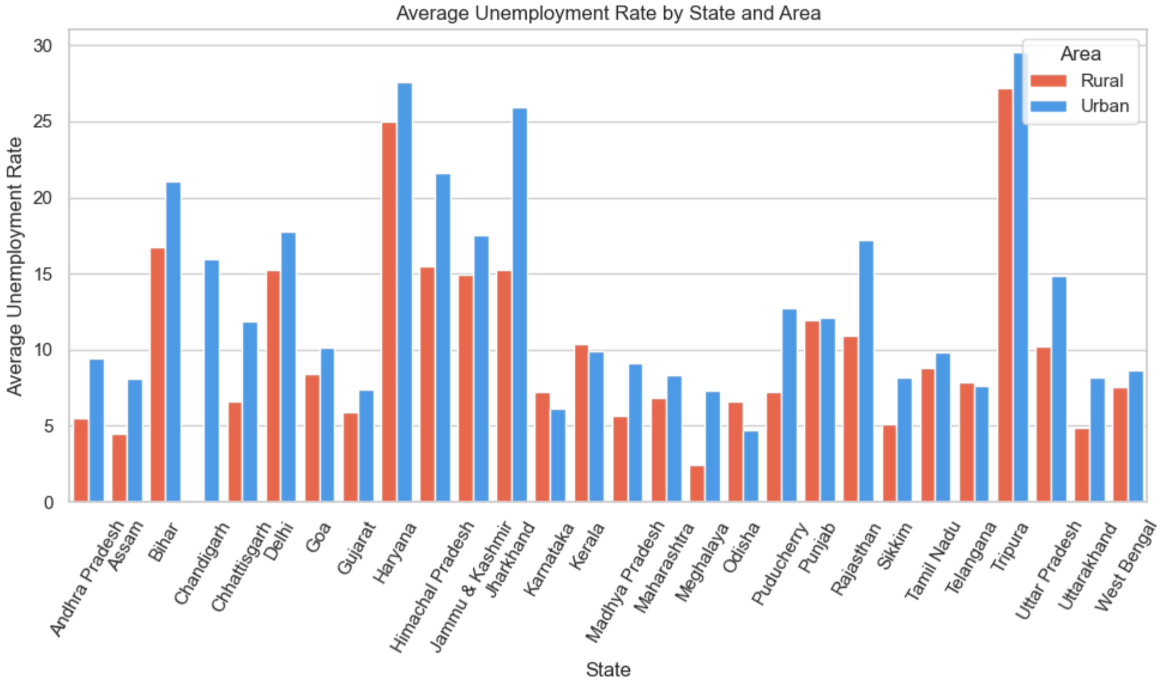
In [112…

```
average_of_regions2=df2.groupby("Region")["Estimated Unemployment Rate"].mean()
average_of_regions2.plot(kind='pie',autopct='%1.1f%%')
plt.axis('equal')
plt.title('Region-wise Average Unemployment Rate of year 2020')
plt.show()
```



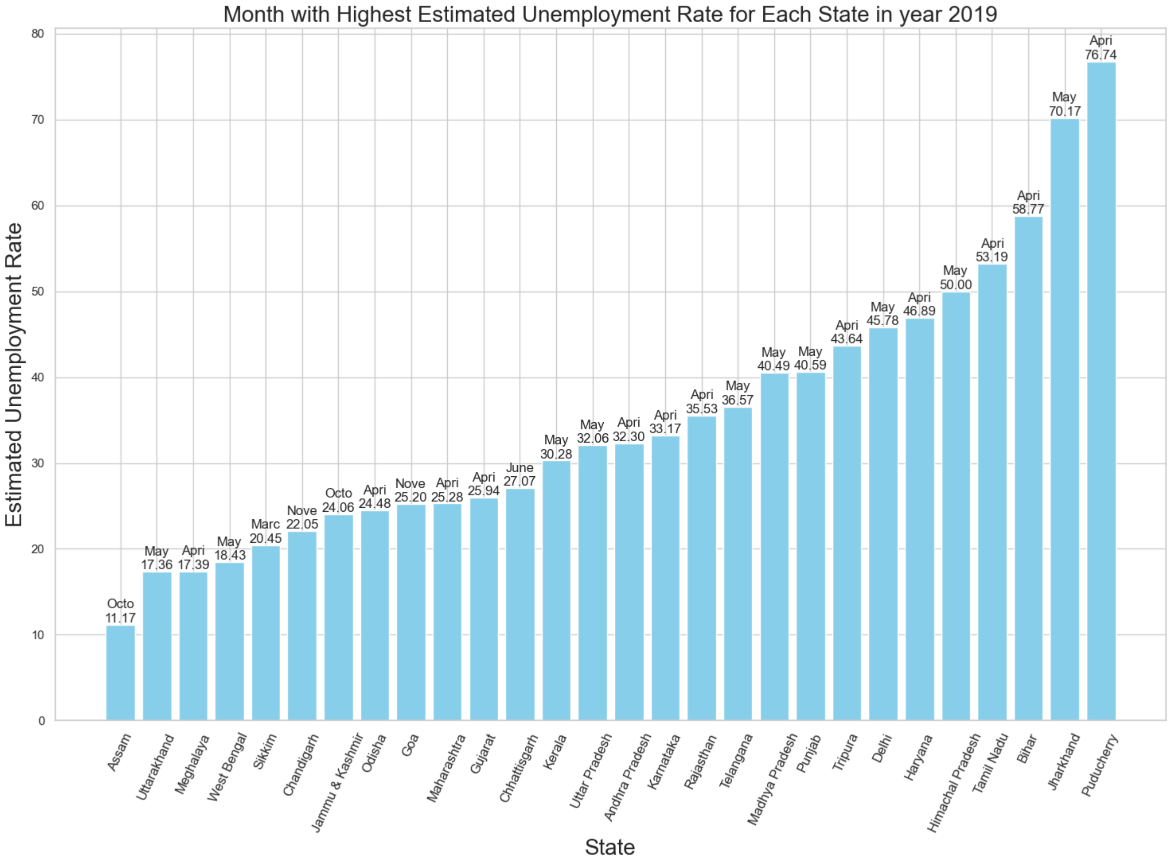
In [123…

```
sns.set(style="whitegrid")

# Define a custom color palette (you can use any valid matplotlib color names or
custom_palette = ["#FF5733", "#3399FF"]  # Example: Red and Green

# Create the bar chart using seaborn with the custom color palette
plt.figure(figsize=(10, 6))  # Adjust the figure size as per your preference
sns.barplot(x='State', y='Estimated Unemployment Rate', hue='Area', data=avg_une
            palette=custom_palette)

# Set the title and labels
plt.title('Average Unemployment Rate by State and Area')
plt.xlabel('State')
plt.ylabel('Average Unemployment Rate')

# Show the plot
plt.xticks(rotation=60)  # Rotate x-axis labels for better readability (optional
plt.legend(title='Area', loc='upper right')  # Add a legend for the 'Area' categ
plt.tight_layout()  # To adjust spacing and avoid label overlapping (optional)
plt.show()
```



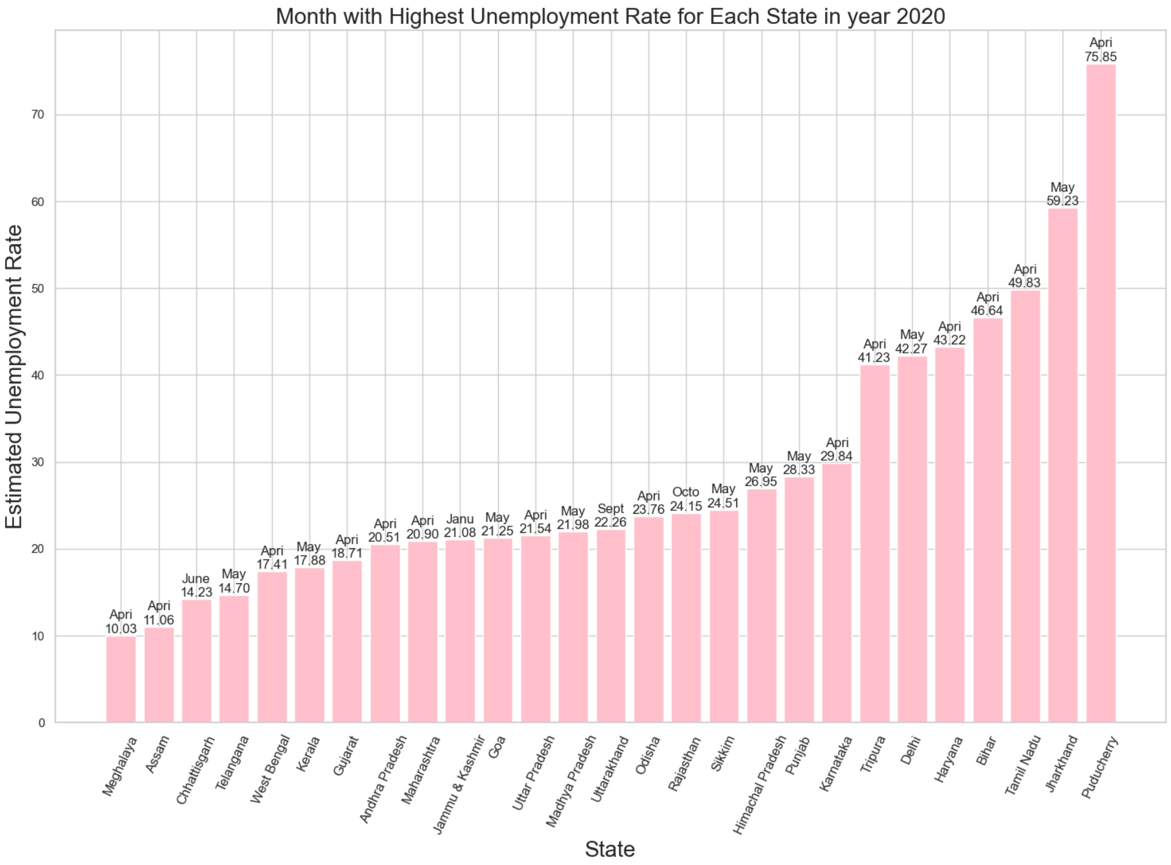
In [153…

```
# Extract the month from the 'Date' column and create a new 'Month' column with
df1['Month'] = df1['Date'].dt.strftime('%B')

# Find the month with the highest estimated unemployment rate for each state
highest_month = df1.groupby('State')['Estimated Unemployment Rate'].idxmax()
highest_month_df = df1.loc[highest_month, ['State', 'Month', 'Estimated Unemploy

# Sort the DataFrame by the highest unemployment rate in ascending order
highest_month_df.sort_values('Estimated Unemployment Rate', inplace=True)

# Create a bar graph to visualize the month with the highest estimated unemploym
plt.figure(figsize=(18, 15))
plt.bar(highest_month_df['State'], highest_month_df['Estimated Unemployment Rate

# Add text labels above each bar showing the first four letters of the month and
for x, y, month in zip(highest_month_df['State'], highest_month_df['Estimated Un
    month_abbreviated = month[:4]  # Get the first four letters of the month
    plt.text(x, y, f'{month_abbreviated}\n{y:.2f}', ha='center', va='bottom', fo

# Set axis labels, title, and rotate x-axis labels for better readability
plt.xlabel('State',fontsize=20)
plt.ylabel('Estimated Unemployment Rate',fontsize=20)
plt.title('Month with Highest Estimated Unemployment Rate for Each State in year
plt.xticks(rotation=65,fontsize=12)

# Adjust the spacing between the subplots to avoid label overlapping
plt.subplots_adjust(bottom=0.3)

plt.show()
```



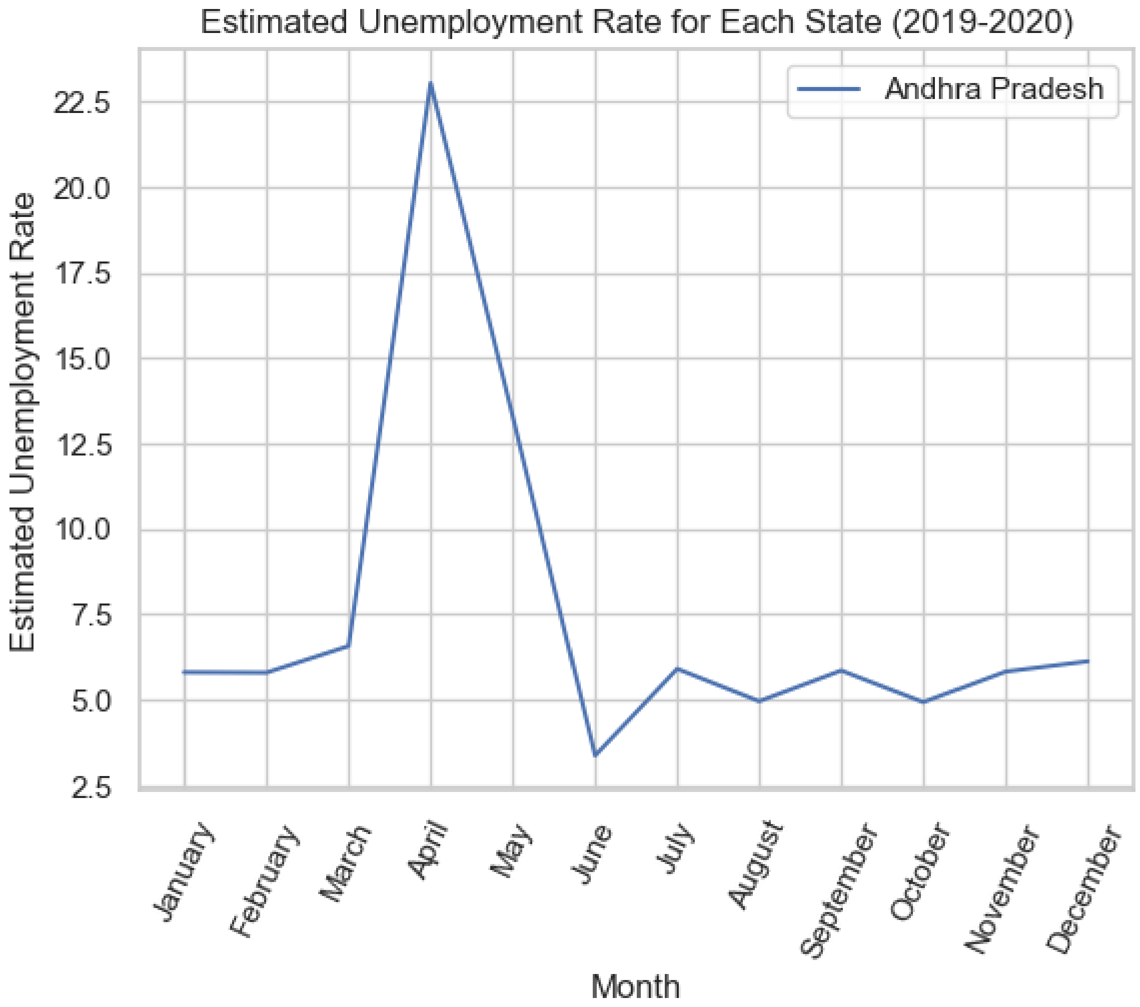
In [169…

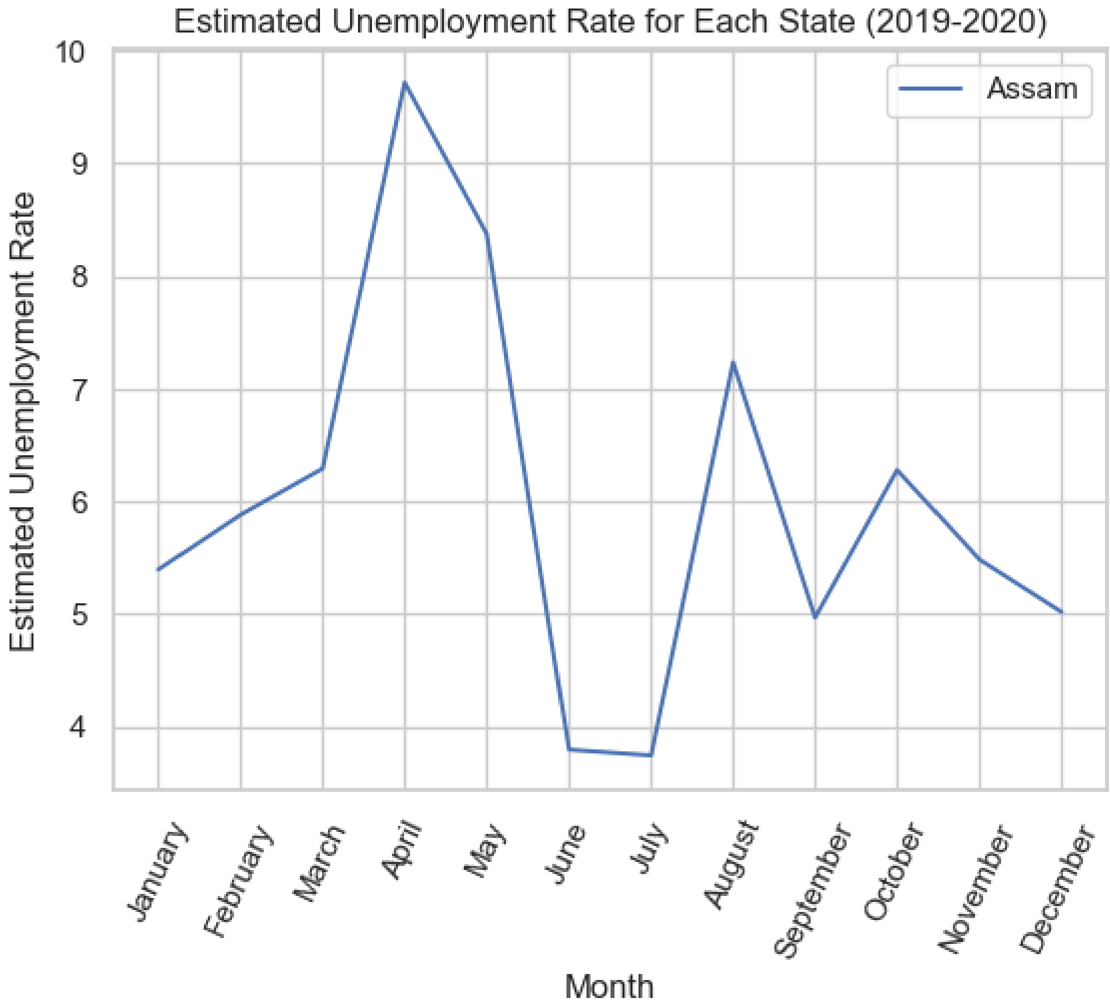
```
# Convert the 'Date' column to datetime format
df2['Date'] = pd.to_datetime(df2['Date'])

# Extract the month from the 'Date' column and create a new 'Month' column with
df2['Month'] = df2['Date'].dt.strftime('%B')

# Find the month with the highest estimated unemployment rate for each state
highest_month2 = df2.groupby('State')['Estimated Unemployment Rate'].idxmax()
highest_month_df2 = df2.loc[highest_month2, ['State', 'Month', 'Estimated Unempl

# Sort the DataFrame by the highest unemployment rate in ascending order
highest_month_df2.sort_values('Estimated Unemployment Rate', inplace=True)

# Create a bar graph to visualize the month with the highest estimated unemploym
plt.figure(figsize=(18, 15))
plt.bar(highest_month_df2['State'], highest_month_df2['Estimated Unemployment Ra

# Add text labels above each bar showing the month and the corresponding unemplo
for x, y, month in zip(highest_month_df2['State'], highest_month_df2['Estimated 
    month_abbreviated = month[:4]  # Get the first four letters of the month
    plt.text(x, y, f'{month_abbreviated}\n{y:.2f}', ha='center', va='bottom', fo

# Set axis labels, title, and rotate x-axis labels for better readability
plt.xlabel('State', fontsize=20)
plt.ylabel('Estimated Unemployment Rate', fontsize=20)
plt.title('Month with Highest Unemployment Rate for Each State in year 2020', fo
plt.xticks(rotation=65, fontsize=12)

# Adjust the spacing between the subplots to avoid label overlapping
plt.subplots_adjust(bottom=0.3)

plt.show()
```

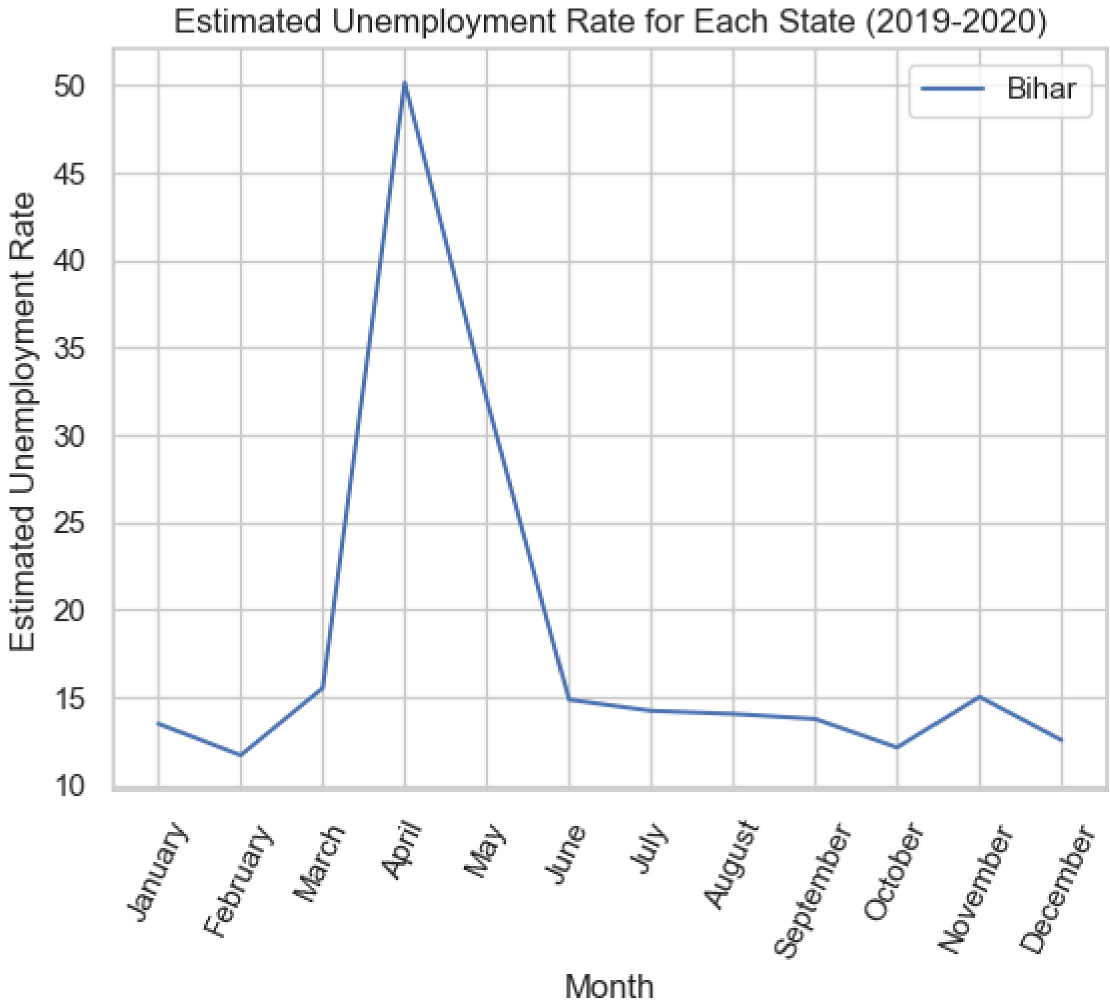


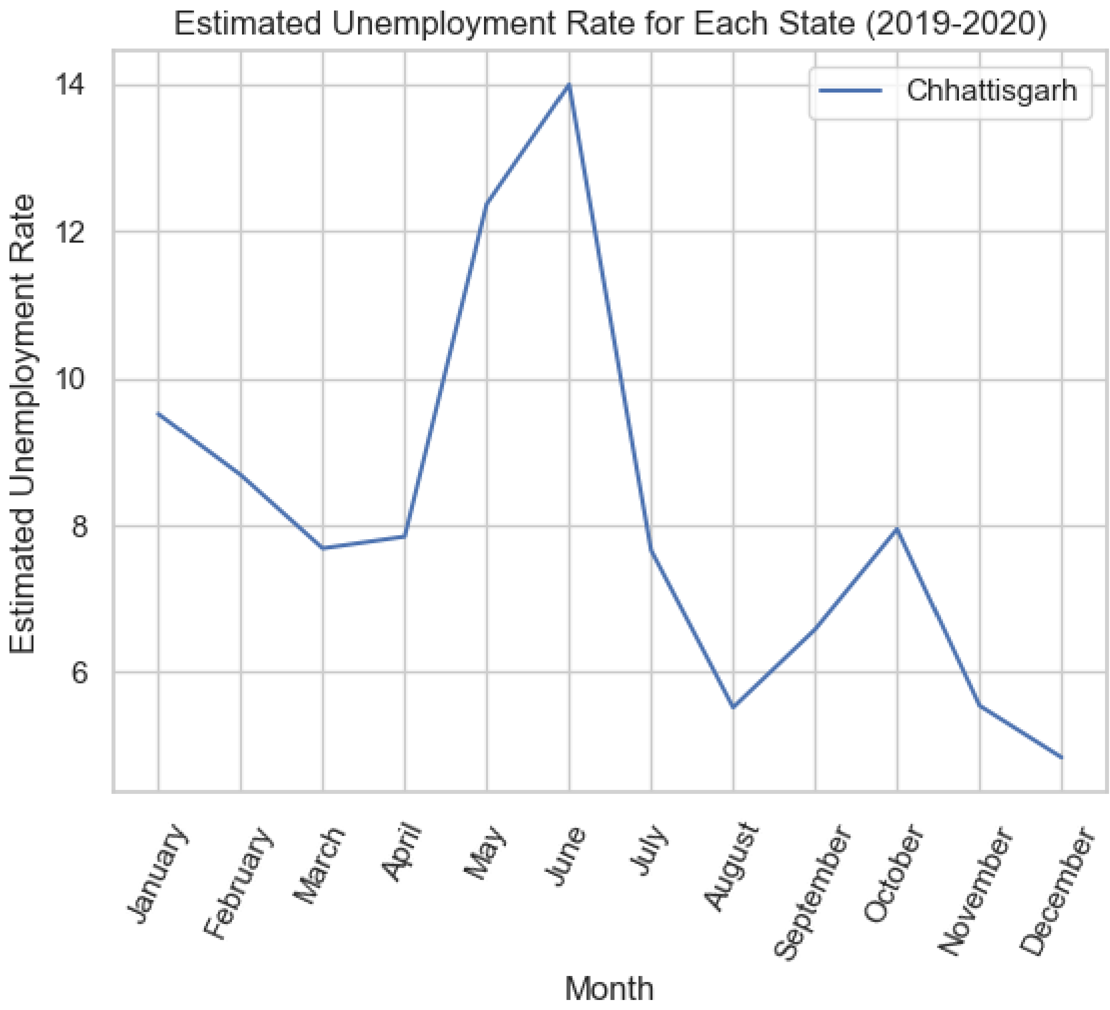
In [170…

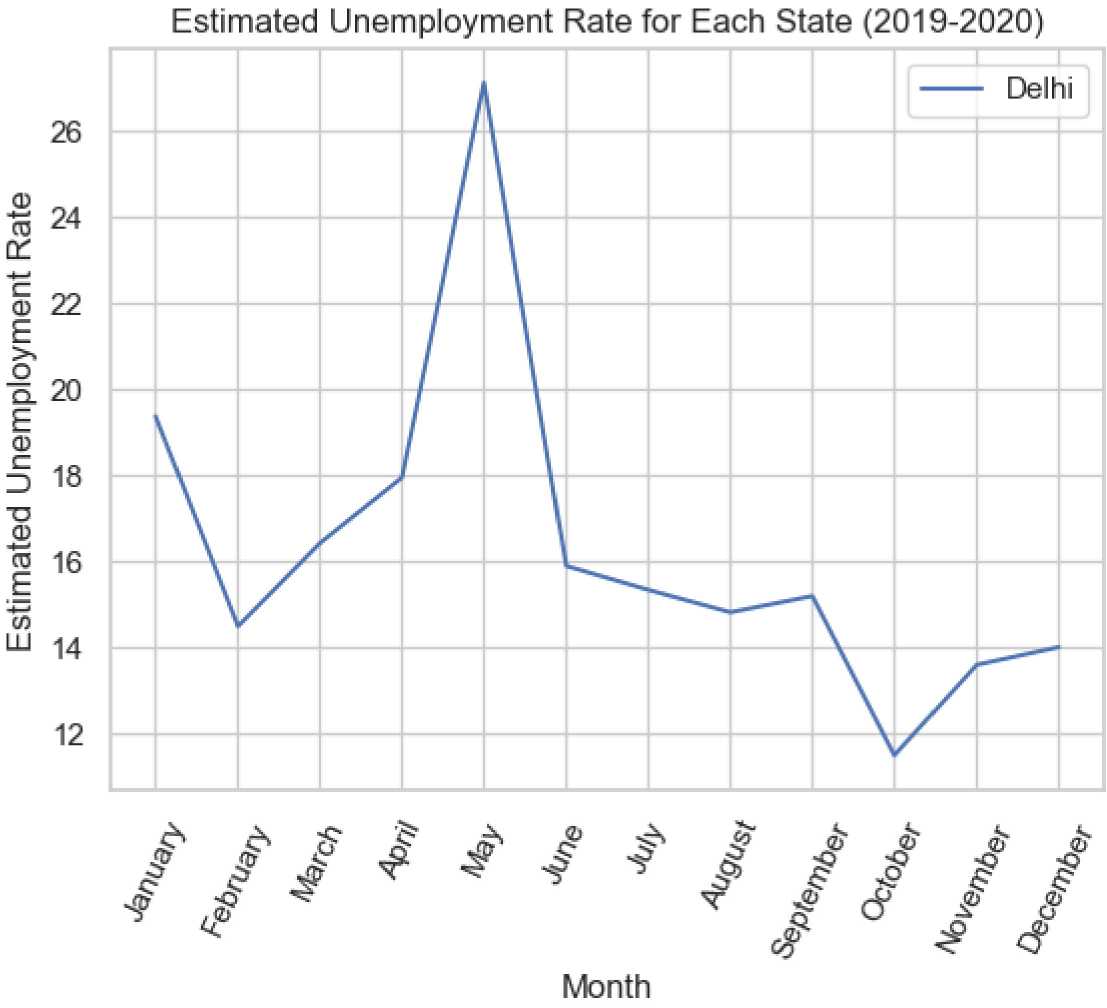
```
# Concatenate the two datasets
combined_data = pd.concat([df1, df2])

# Define the order of the months
month_order = ['January', 'February', 'March', 'April', 'May', 'June', 'July', '

# Convert the 'Month' column to categorical data type with the custom sort order
combined_data['Month'] = pd.Categorical(combined_data['Month'], categories=month

# Calculate the mean Estimated Unemployment Rate for each month and state
monthly_data = combined_data.groupby(['Month', 'State'])['Estimated Unemployment

# Get a list of unique states
states = combined_data['State'].unique()

# Create separate line plot graphs for each state
for state in states:
    state_data = monthly_data[monthly_data['State'] == state]
    plt.plot(state_data['Month'], state_data['Estimated Unemployment Rate'], lab

    # Set the x-axis label as 'Month'
    plt.xlabel('Month')

    # Set the y-axis label as 'Estimated Unemployment Rate'
    plt.ylabel('Estimated Unemployment Rate')

    # Set the title of the graph
    plt.title('Estimated Unemployment Rate for Each State (2019-2020)')

    # Add a legend to differentiate the states
    plt.legend()

    # Rotate the x-axis labels for better visibility
    plt.xticks(rotation=65)

    # Display the line plot
    plt.show()

# Print additional analysis and observations based on the plotted data
print("Based on the data shown above, we can observe that the Unemployment rate
```

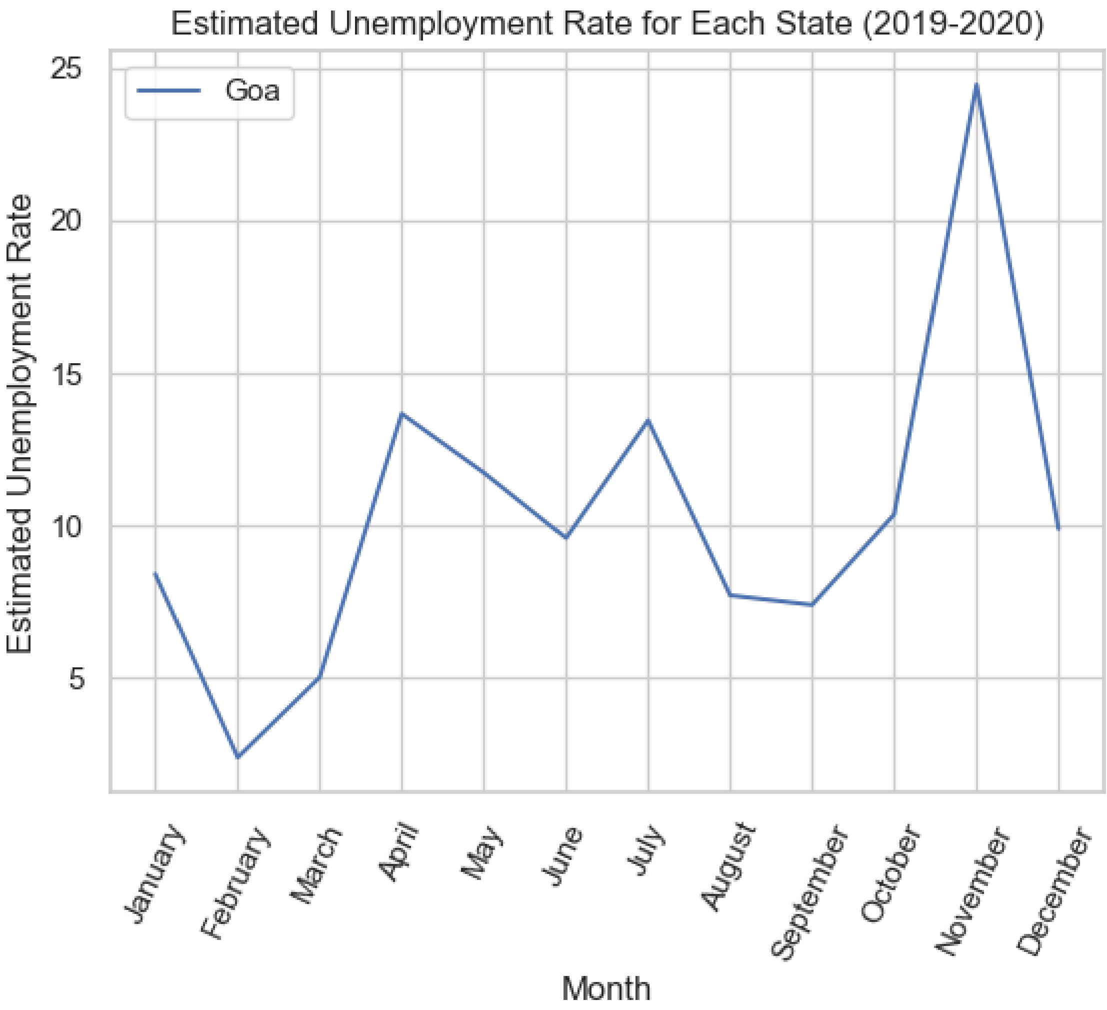


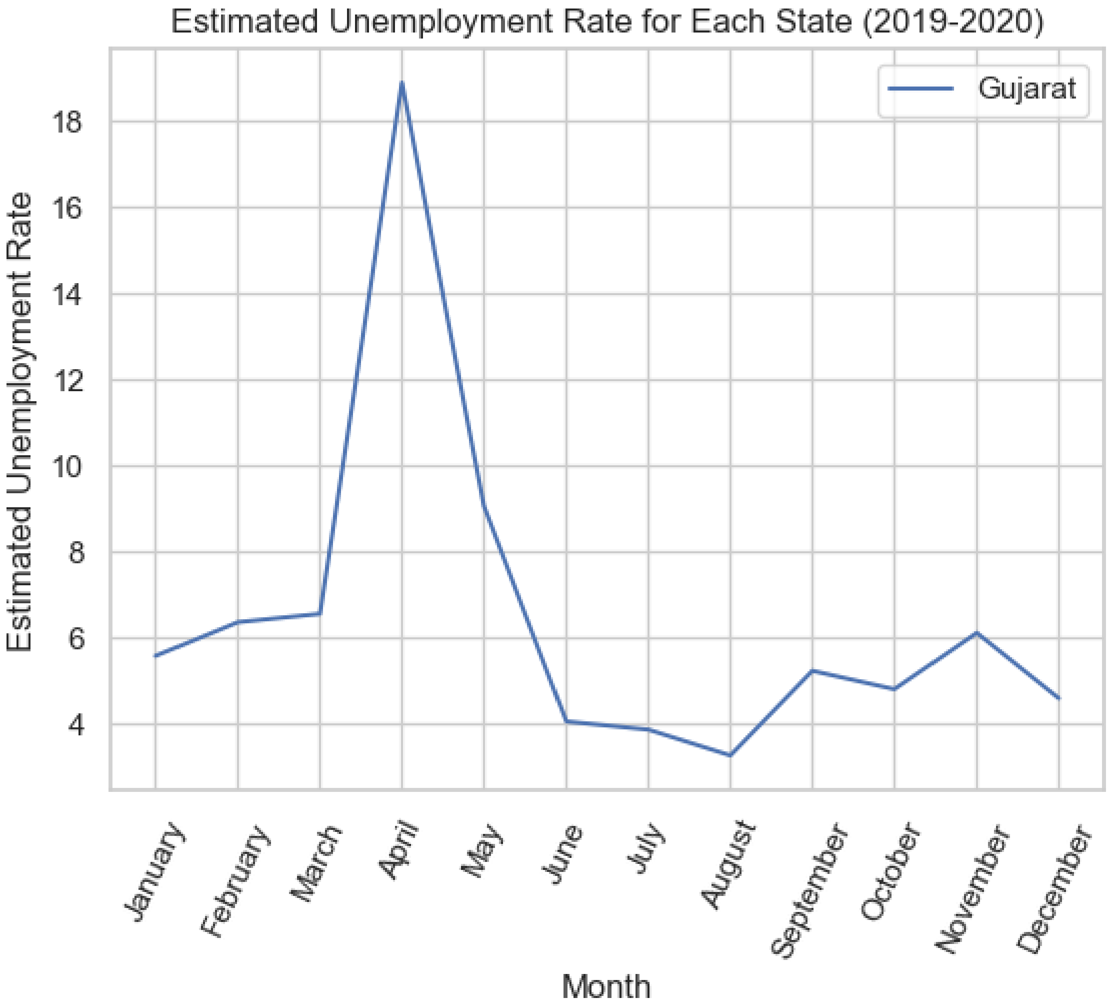

Estimated Unemployment Rate for Each State (2019-2020)
Assam
Estimated Unemployment Rate
10
9
8
7
6
5
4
January
February
March
April
May
June
July
August
September
October
November
December
Month

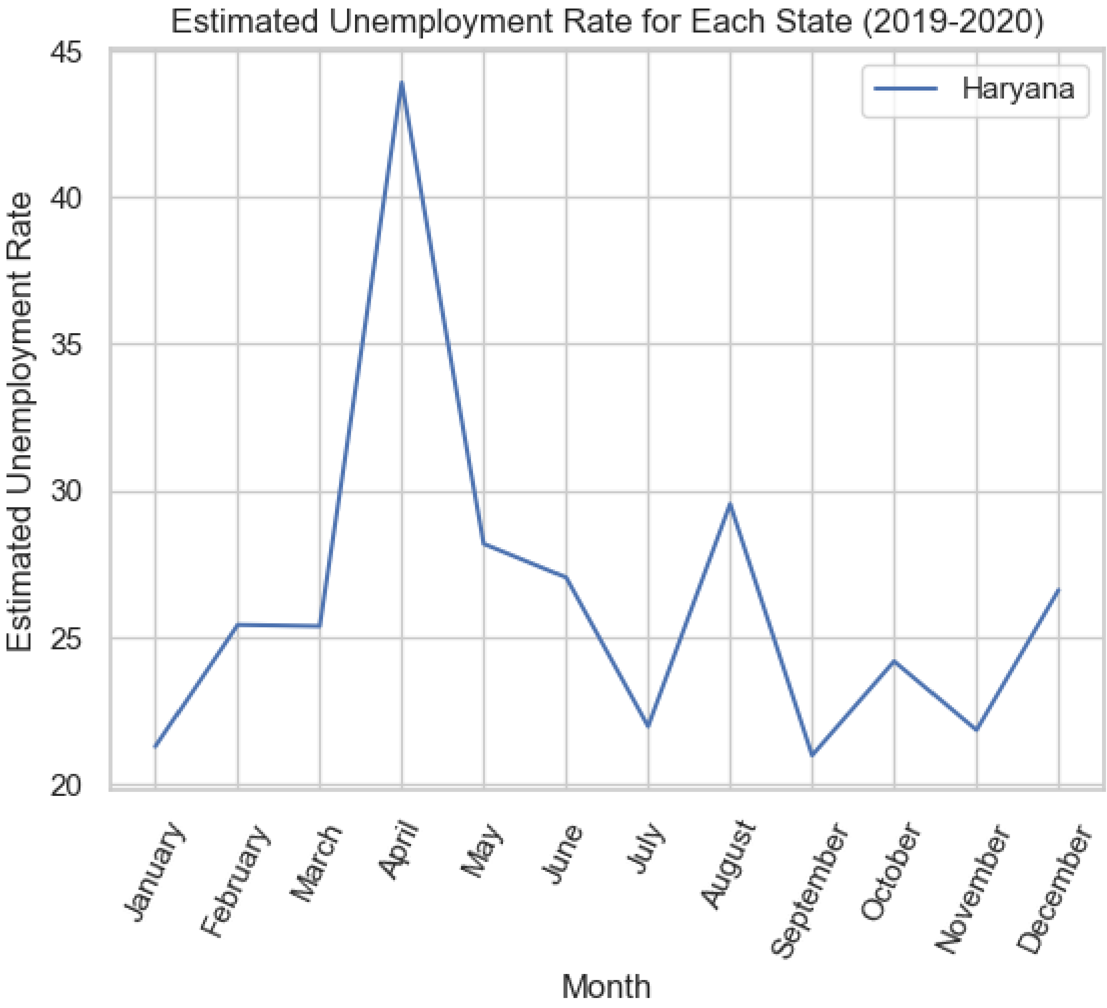

Estimated Unemployment Rate for Each State (2019-2020)
Bihar
Estimated Unemployment Rate
50
45
40
35
30
25
20
15
10
January
February
March
April
May
June
July
August
September
October
November
December
Month

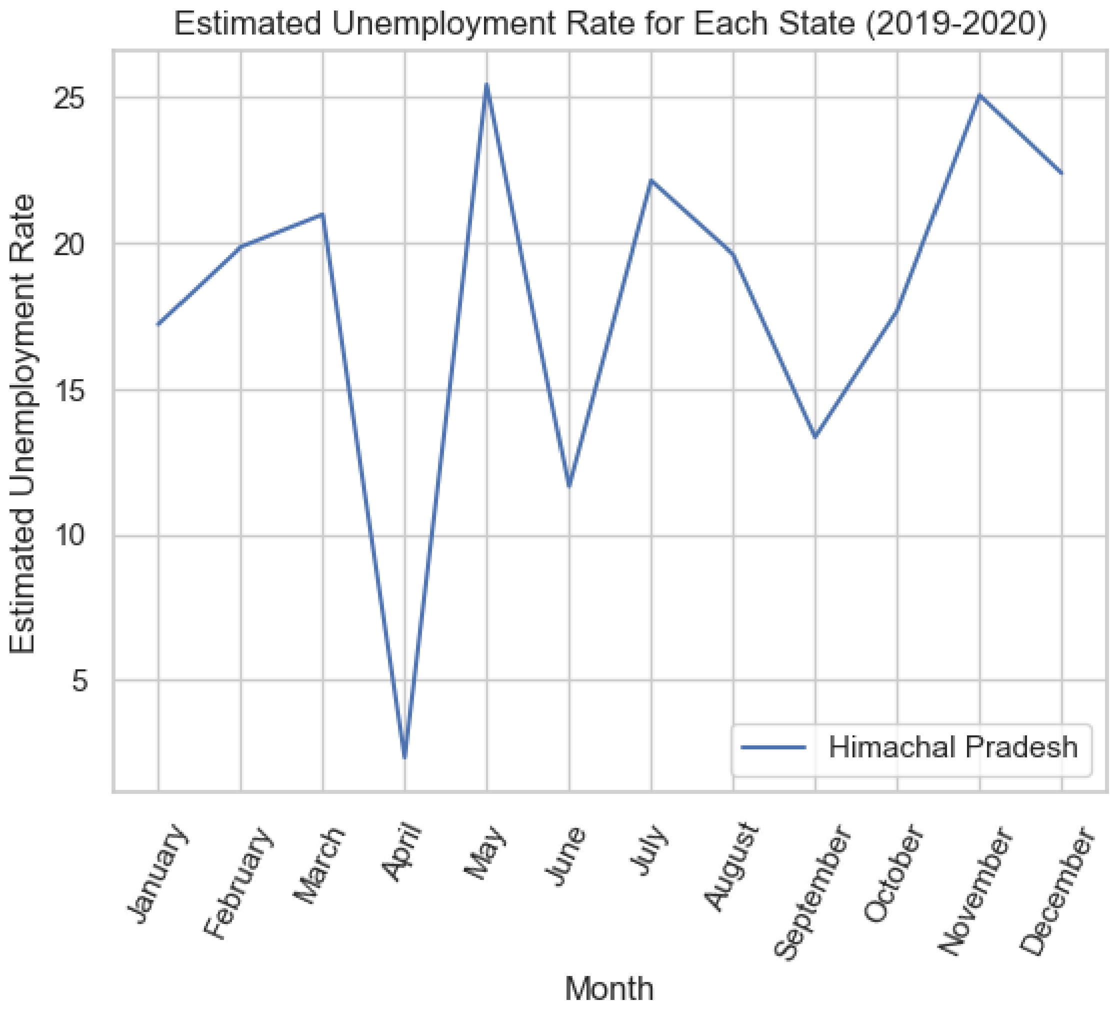

Estimated Unemployment Rate for Each State (2019-2020)
Chhattisgarh
Estimated Unemployment Rate
14
12
10
8
6
January
February
March
April
May
June
July
August
September
October
November
December
Month

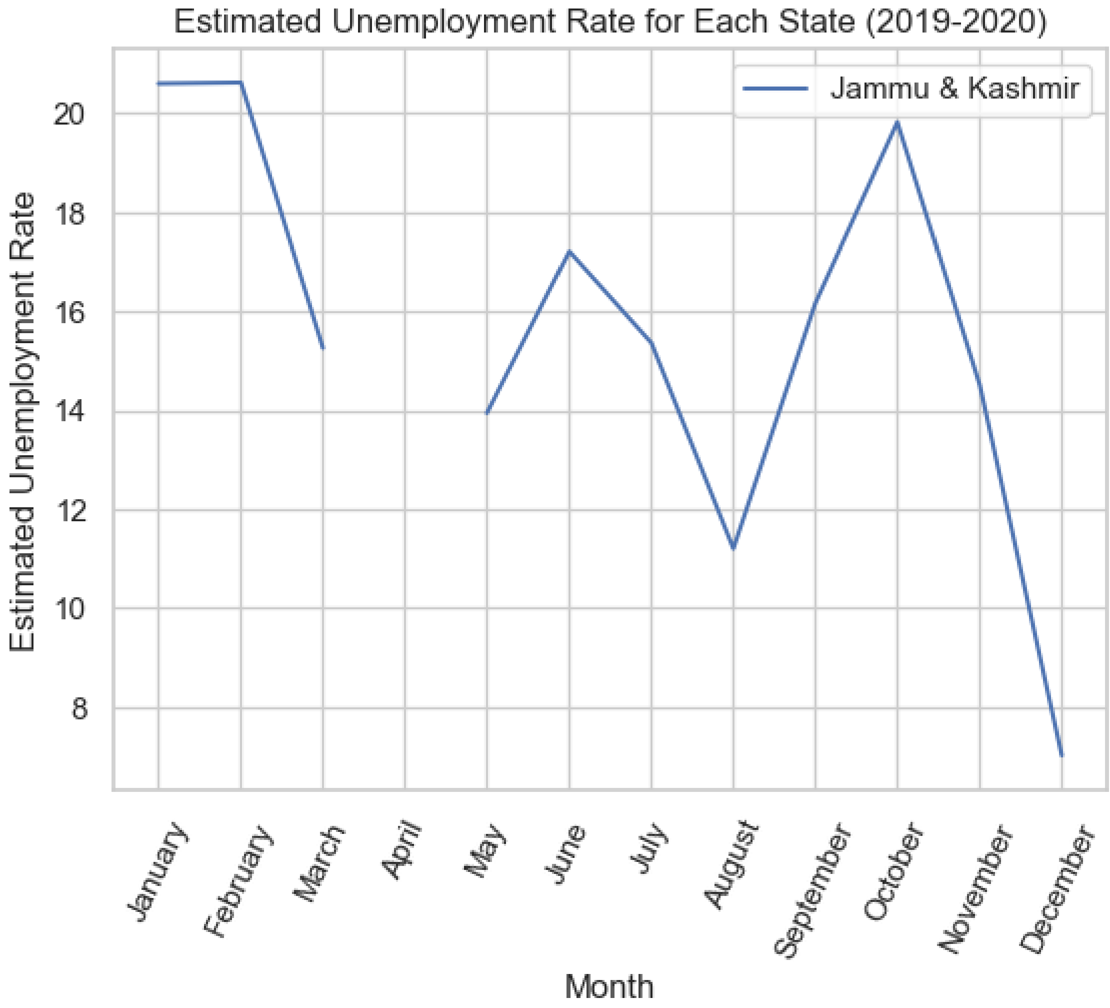

Estimated Unemployment Rate for Each State (2019-2020)
Delhi
Estimated Unemployment Rate
26
24
22
20
18
16
14
12
January
February
March
April
May
June
July
August
September
October
November
December
Month

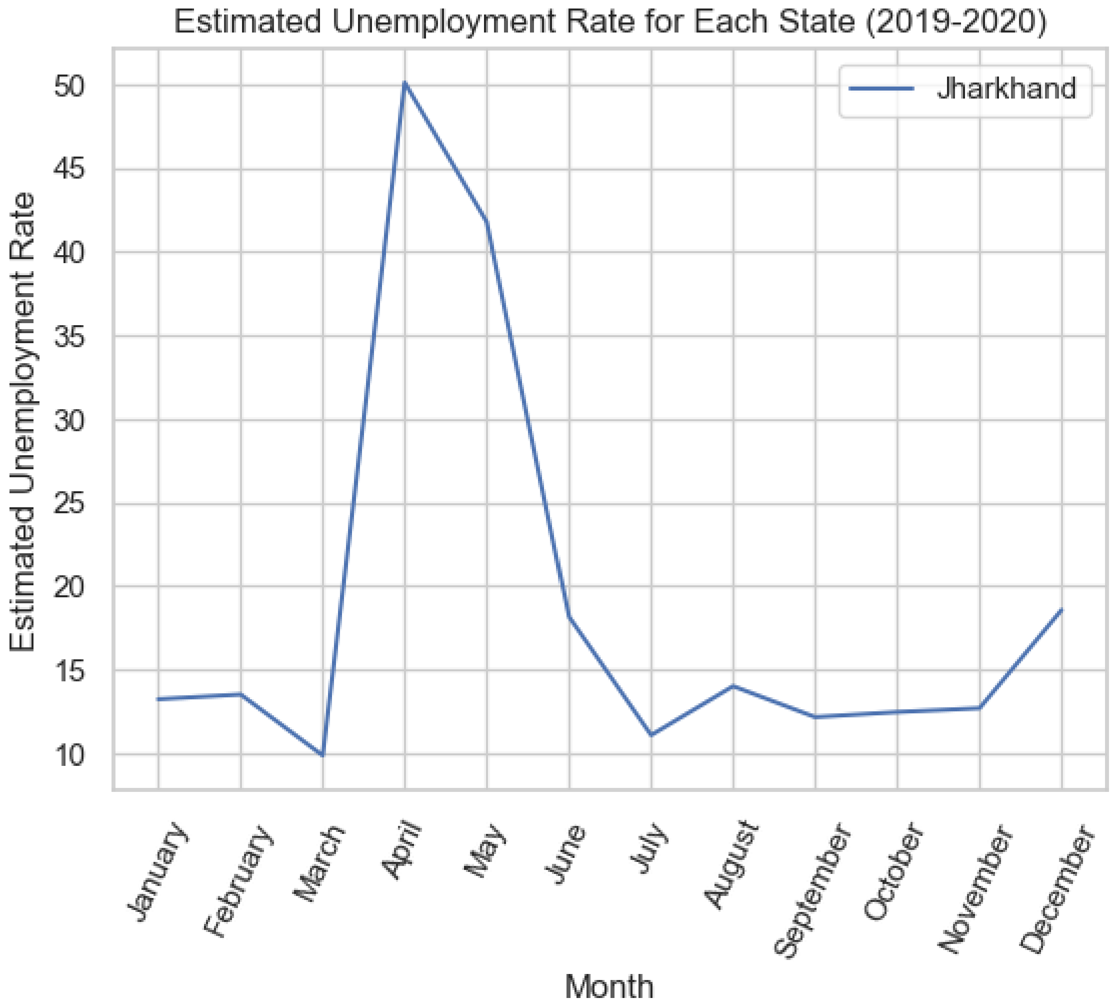

Estimated Unemployment Rate for Each State (2019-2020)
Goa
Estimated Unemployment Rate
25
20
15
10
5
January
February
March
April
May
June
July
August
September
October
November
December
Month

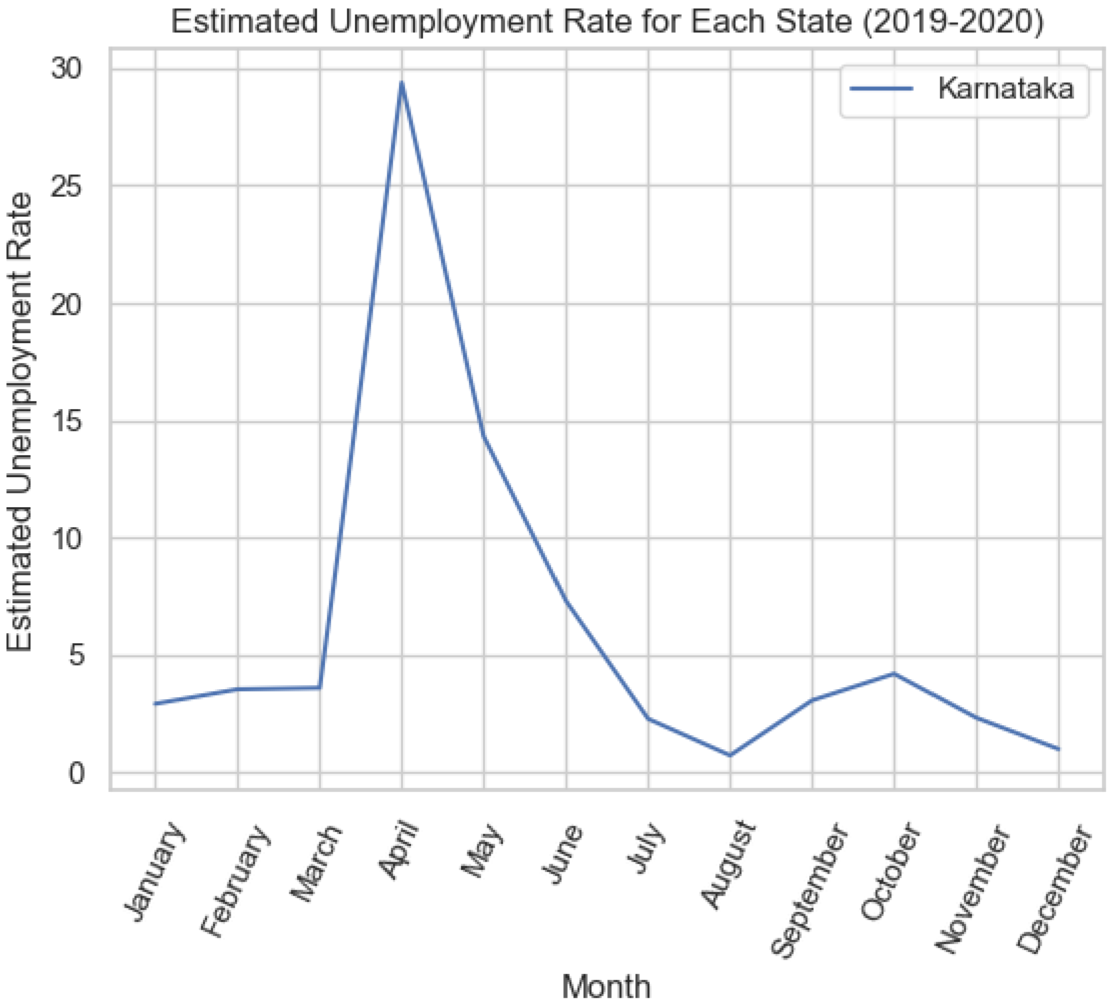

Estimated Unemployment Rate for Each State (2019-2020)
Gujarat
Estimated Unemployment Rate
18
16
14
12
10
8
6
4
January
February
March
April
May
June
July
August
September
October
November
December
Month

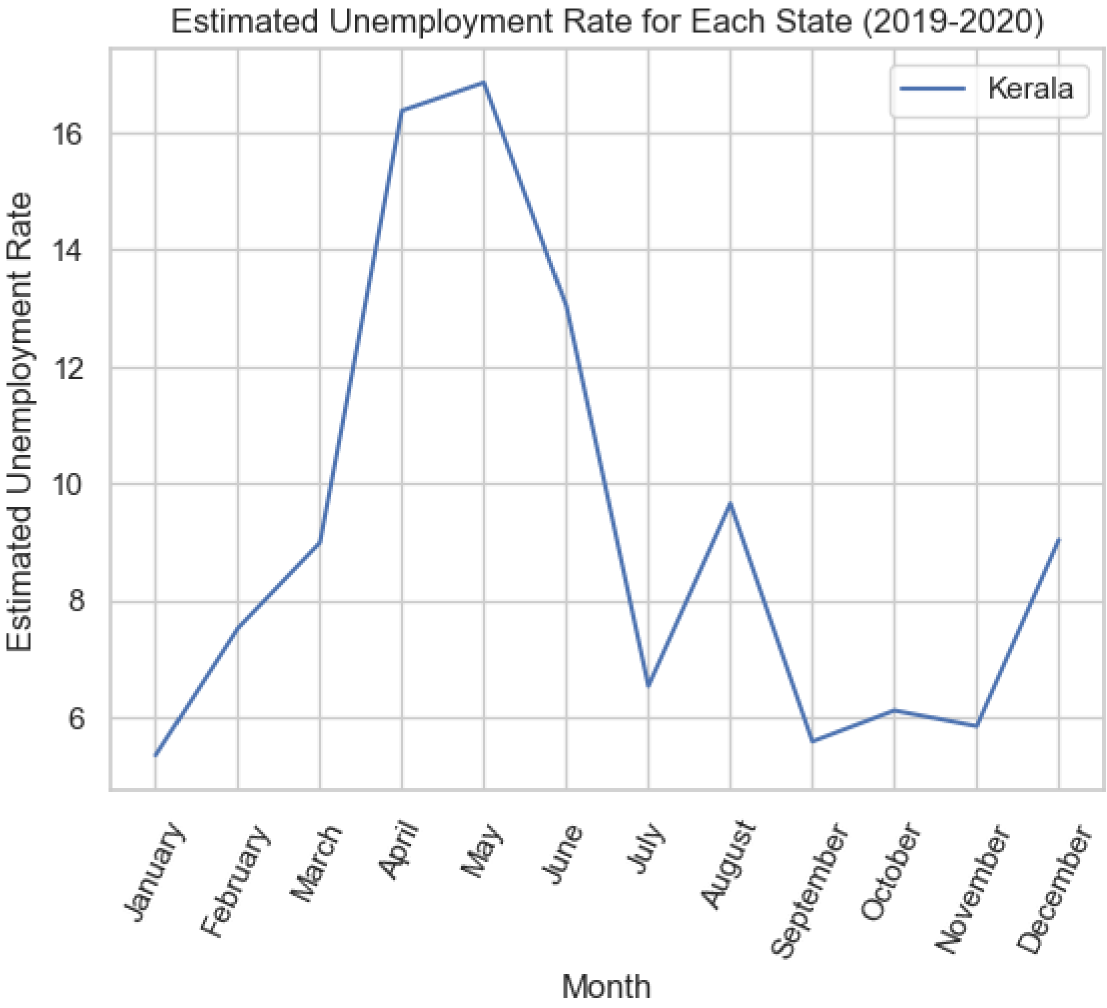

Estimated Unemployment Rate for Each State (2019-2020)
Haryana
Estimated Unemployment Rate
45
40
35
30
25
20
January
February
March
April
May
June
July
August
September
October
November
December
Month

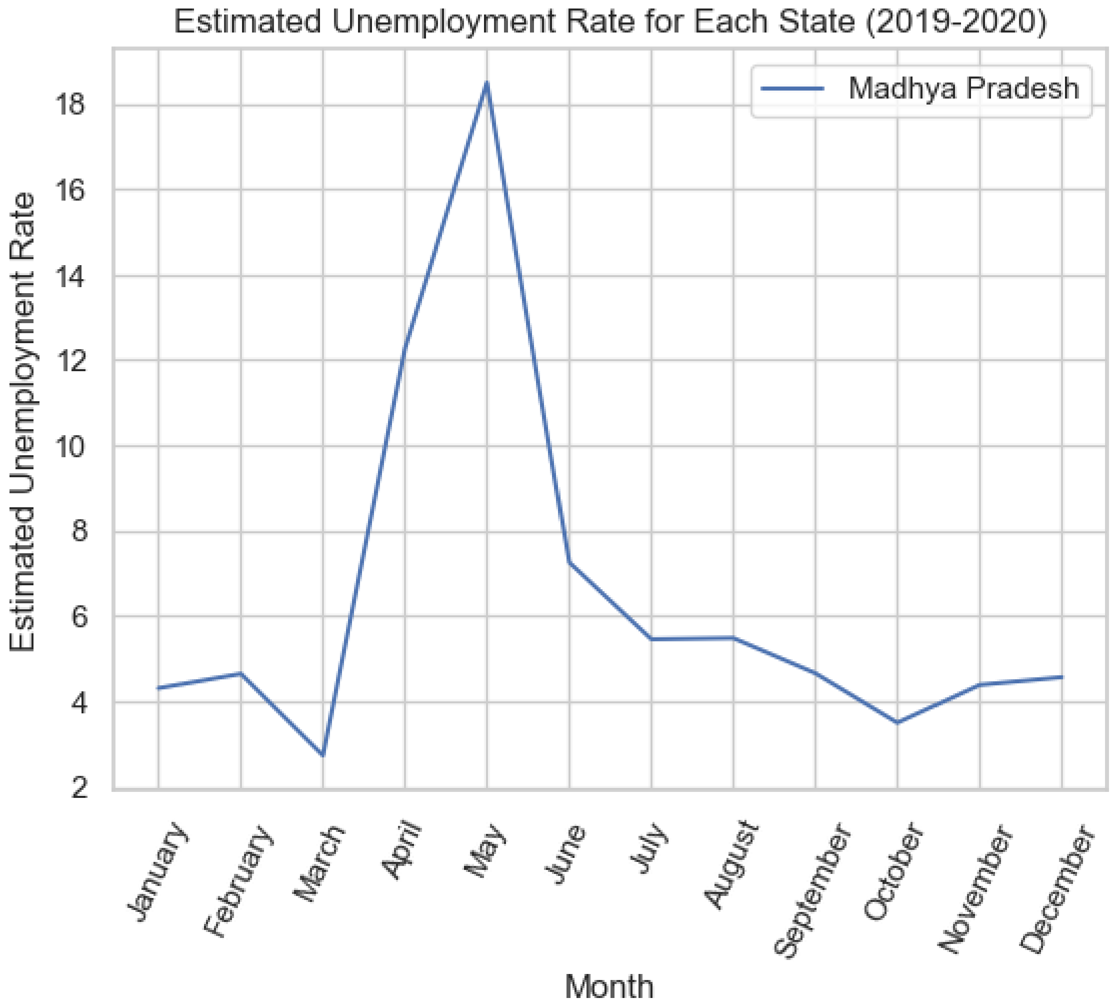

Estimated Unemployment Rate for Each State (2019-2020)
Estimated Unemployment Rate
25
20
15
10
5
Himachal Pradesh
January
February
March
April
May
June
July
August
September
October
November
December
Month

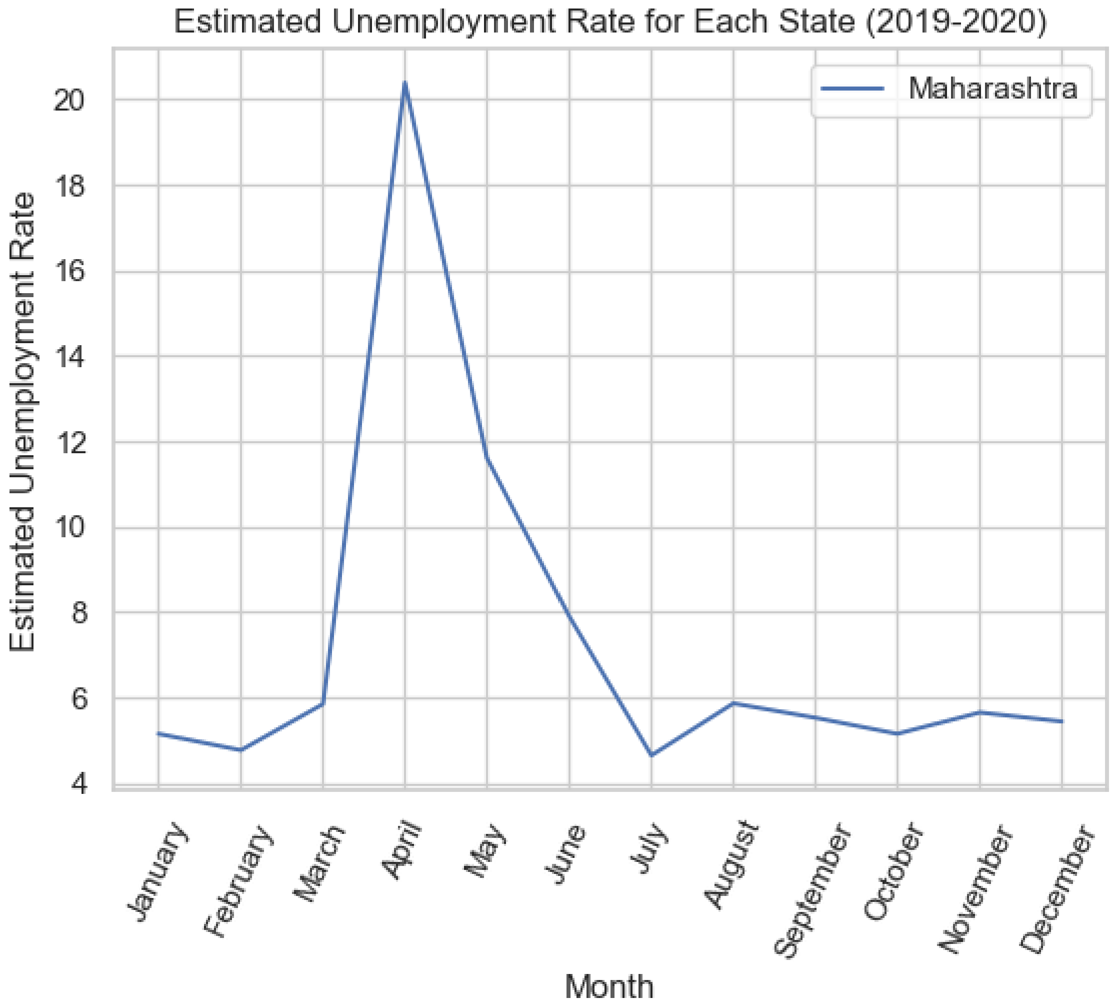

Estimated Unemployment Rate for Each State (2019-2020)
Jammu & Kashmir
Estimated Unemployment Rate
20
18
16
14
12
10
8
January
February
March
April
May
June
July
August
September
October
November
December
Month

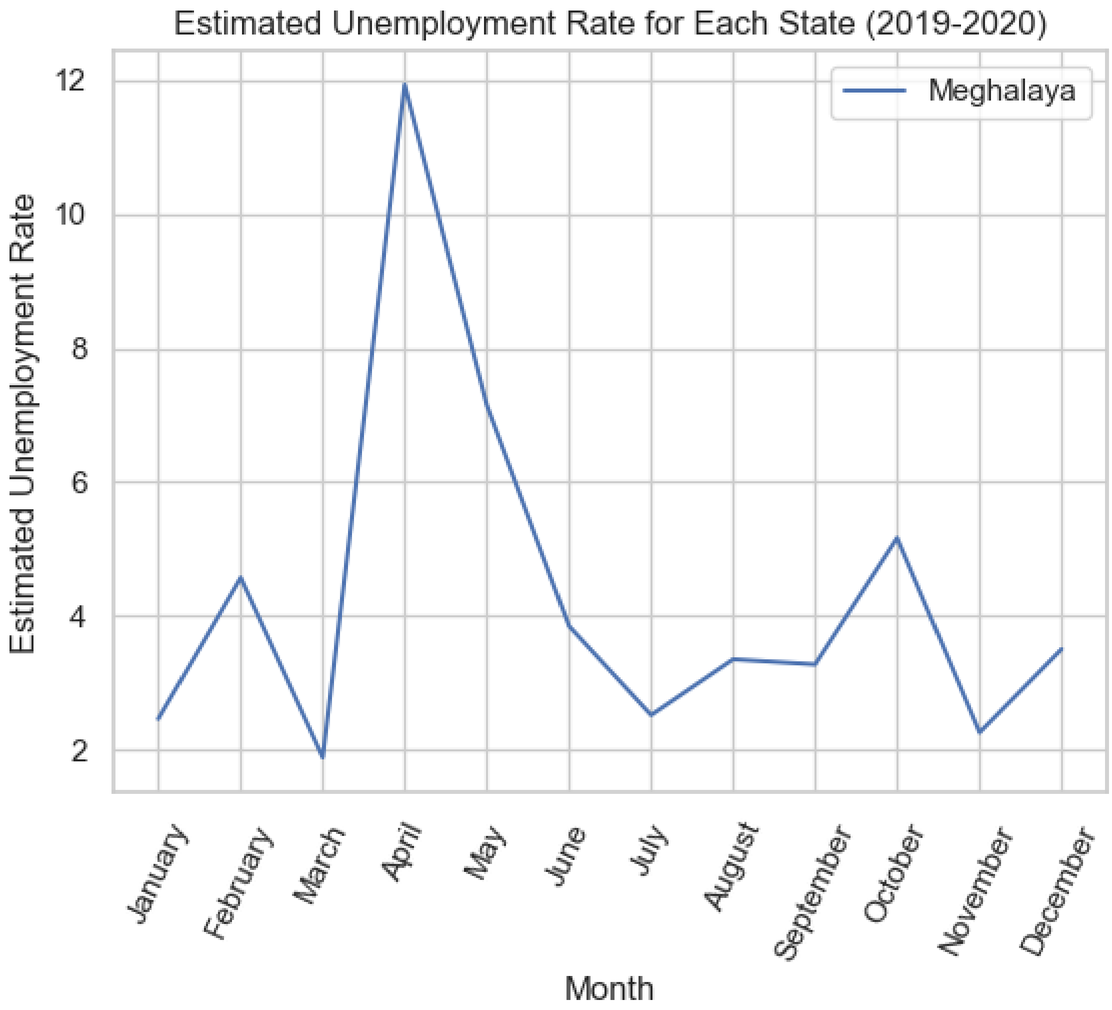

Estimated Unemployment Rate for Each State (2019-2020)
Jharkhand
Estimated Unemployment Rate
50
45
40
35
30
25
20
15
10
January
February
March
April
May
June
July
August
September
October
November
December
Month

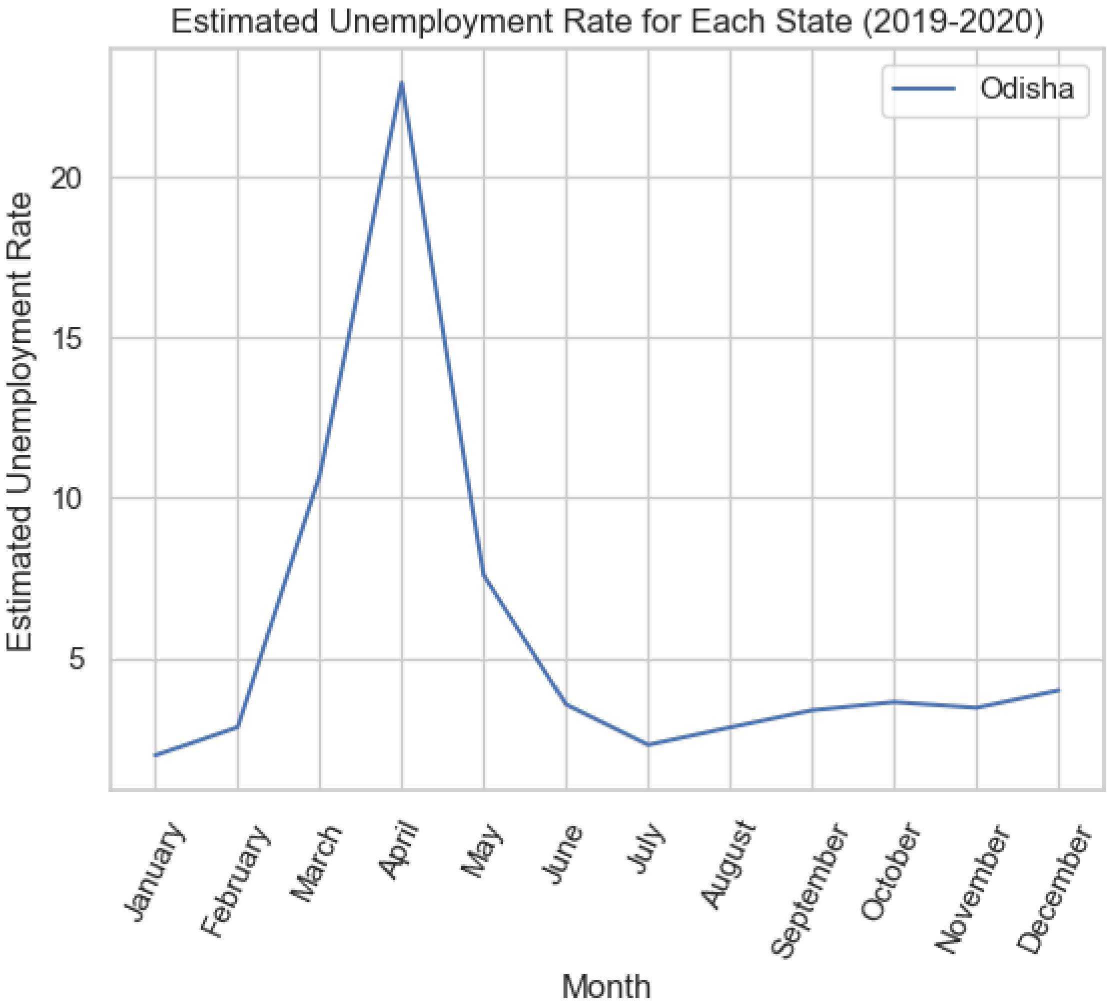

Estimated Unemployment Rate for Each State (2019-2020)
Karnataka
Estimated Unemployment Rate
30
25
20
15
10
5
0
January
February
March
April
May
June
July
August
September
October
November
December
Month

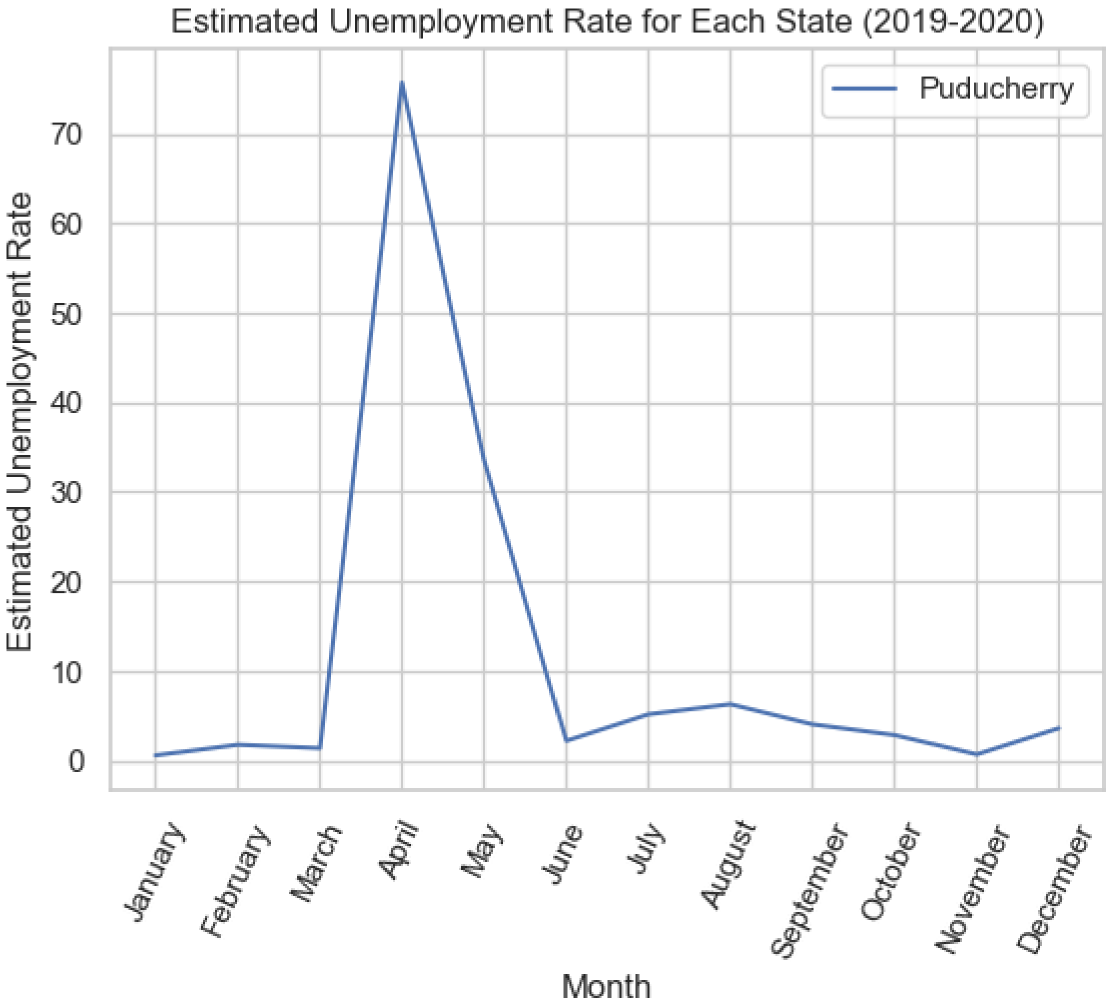

Estimated Unemployment Rate for Each State (2019-2020)
Kerala
Estimated Unemployment Rate
16
14
12
10
8
6
January
February
March
April
May
June
July
August
September
October
November
December
Month

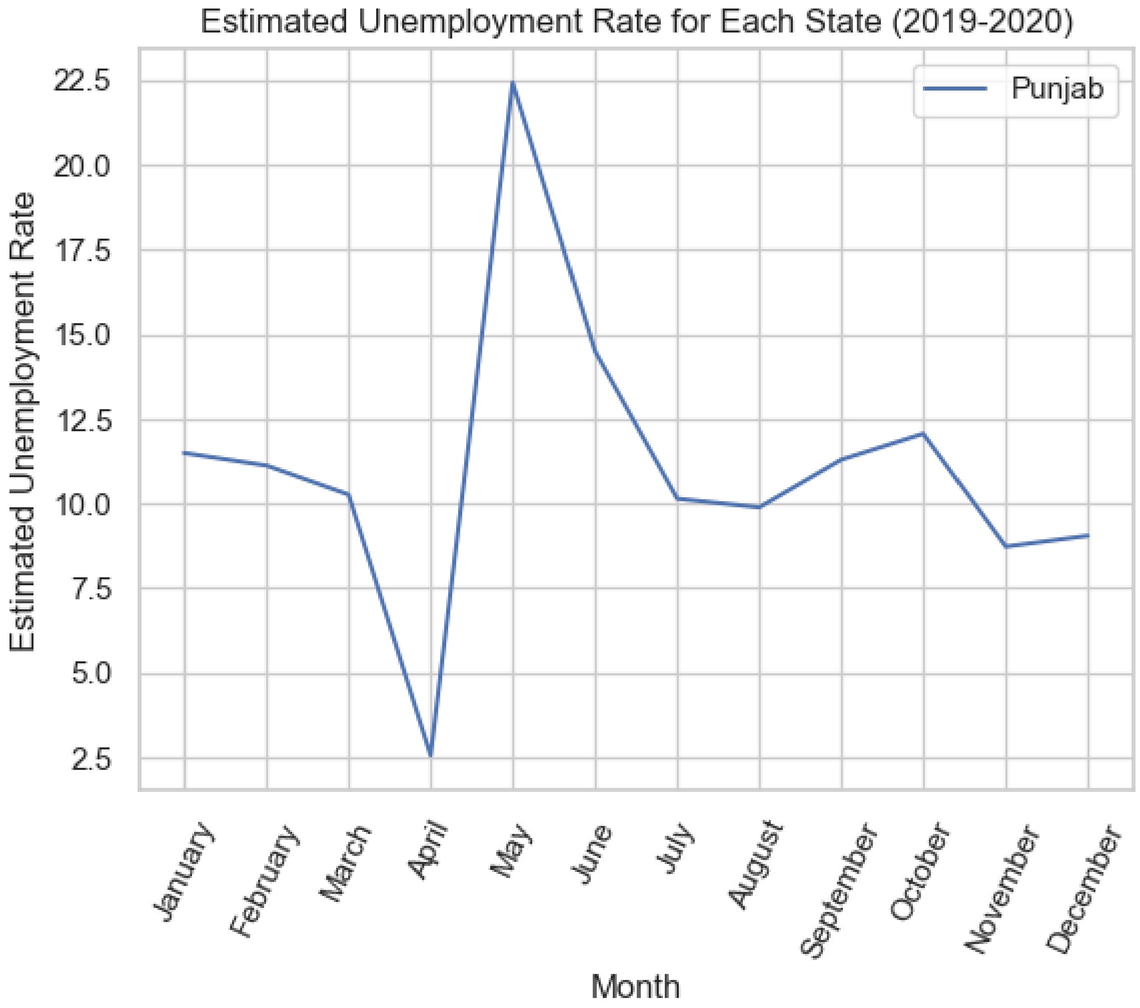

Estimated Unemployment Rate for Each State (2019-2020)
Madhya Pradesh
Estimated Unemployment Rate
18
16
14
12
10
8
6
4
2
January
February
March
April
May
June
July
August
September
October
November
December
Month

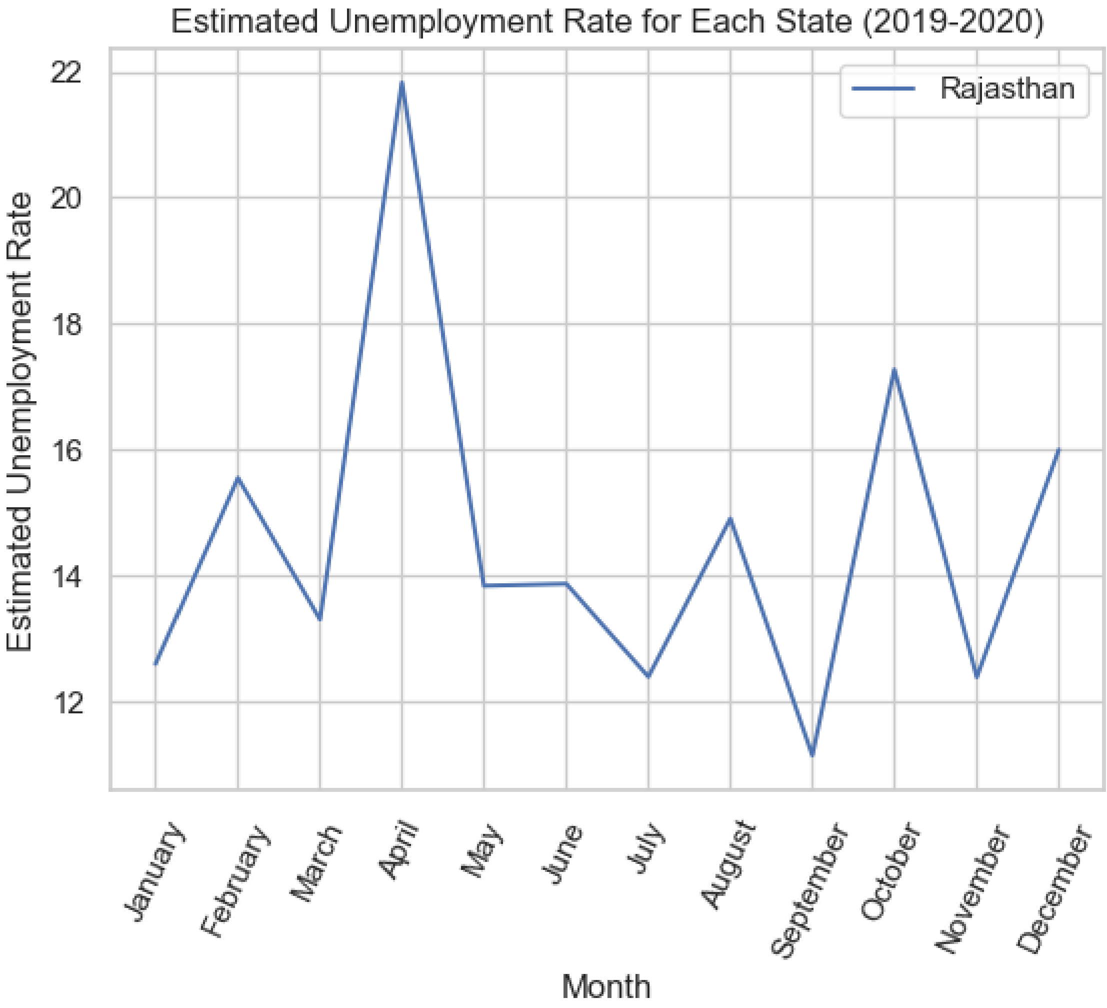

Estimated Unemployment Rate for Each State (2019-2020)
Maharashtra
Estimated Unemployment Rate
20
18
16
14
12
10
8
6
4
January
February
March
April
May
June
July
August
September
October
November
December
Month

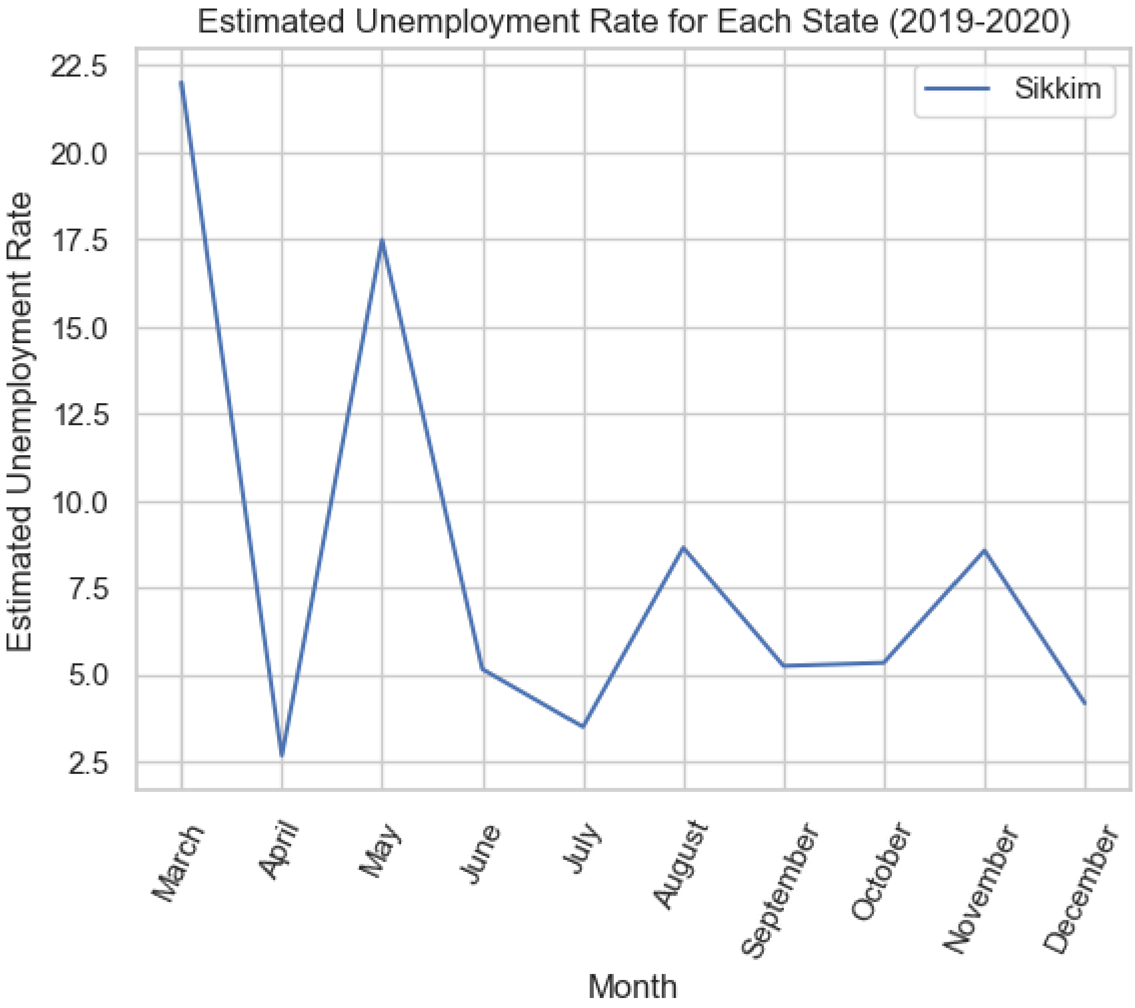

Estimated Unemployment Rate for Each State (2019-2020)
Meghalaya
Estimated Unemployment Rate
12
10
8
6
4
2
January
February
March
April
May
June
July
August
September
October
November
December
Month

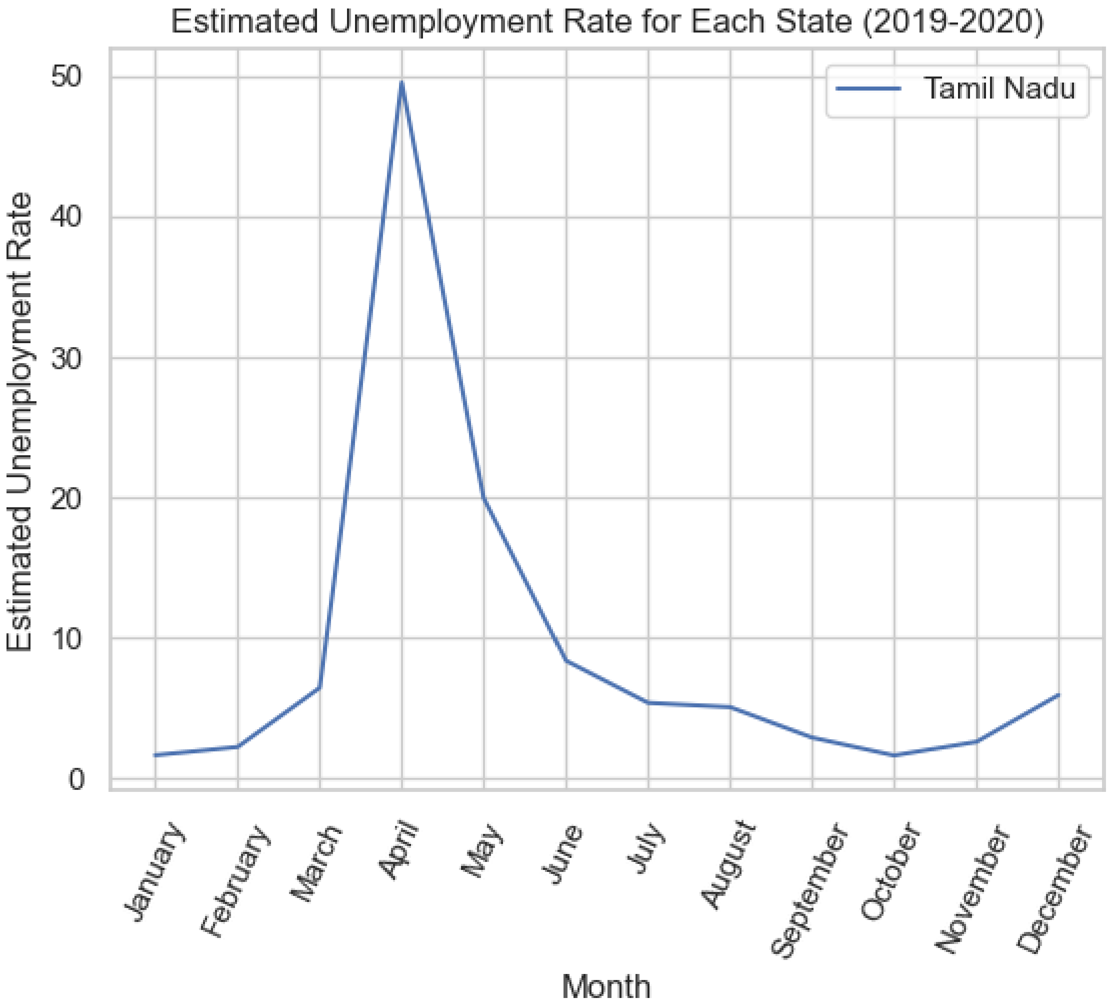

Estimated Unemployment Rate for Each State (2019-2020)
Odisha
Estimated Unemployment Rate
20
15
10
5
January
February
March
April
May
June
July
August
September
October
November
December
Month

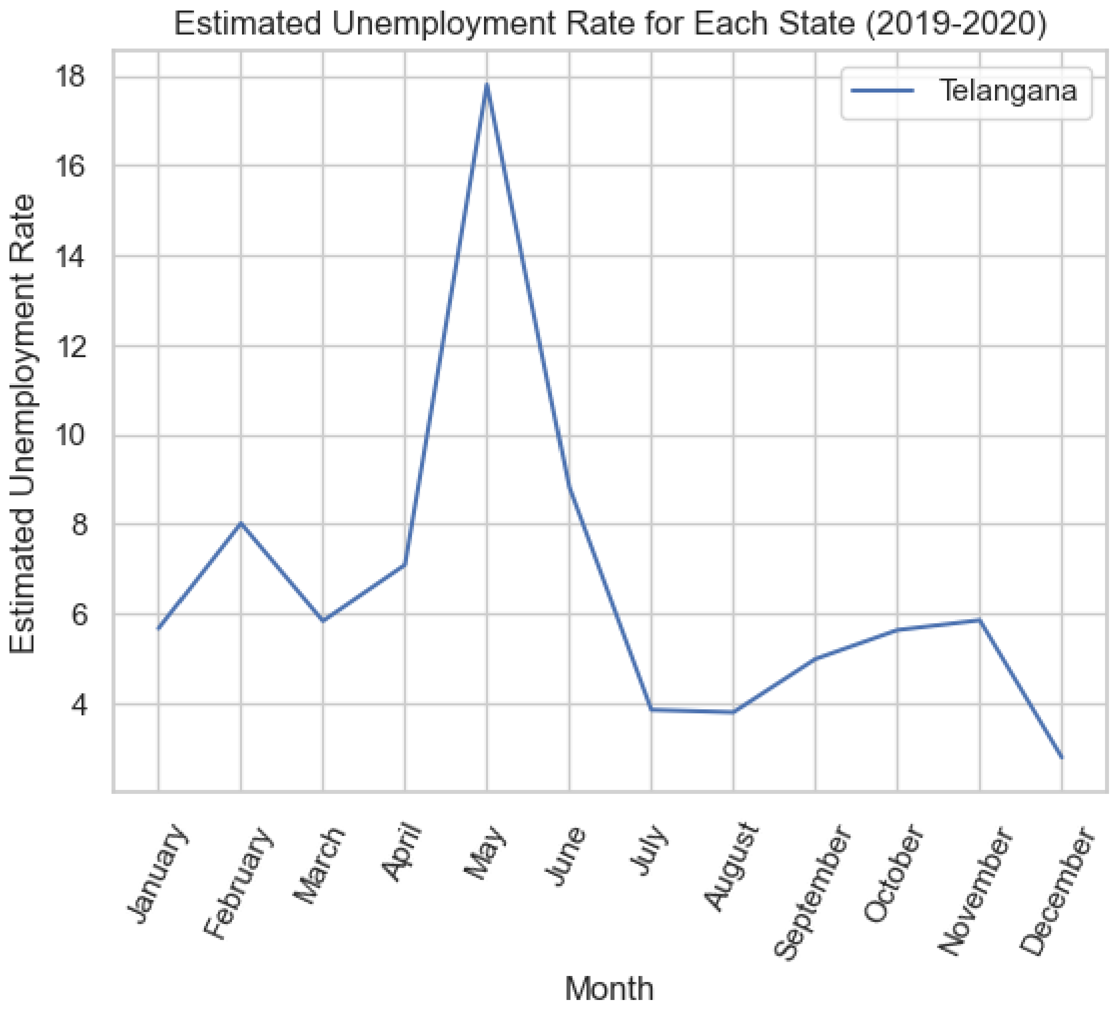

Estimated Unemployment Rate for Each State (2019-2020)
Puducherry
Estimated Unemployment Rate
0
10
20
30
40
50
60
70
January
February
March
April
May
June
July
August
September
October
November
December
Month

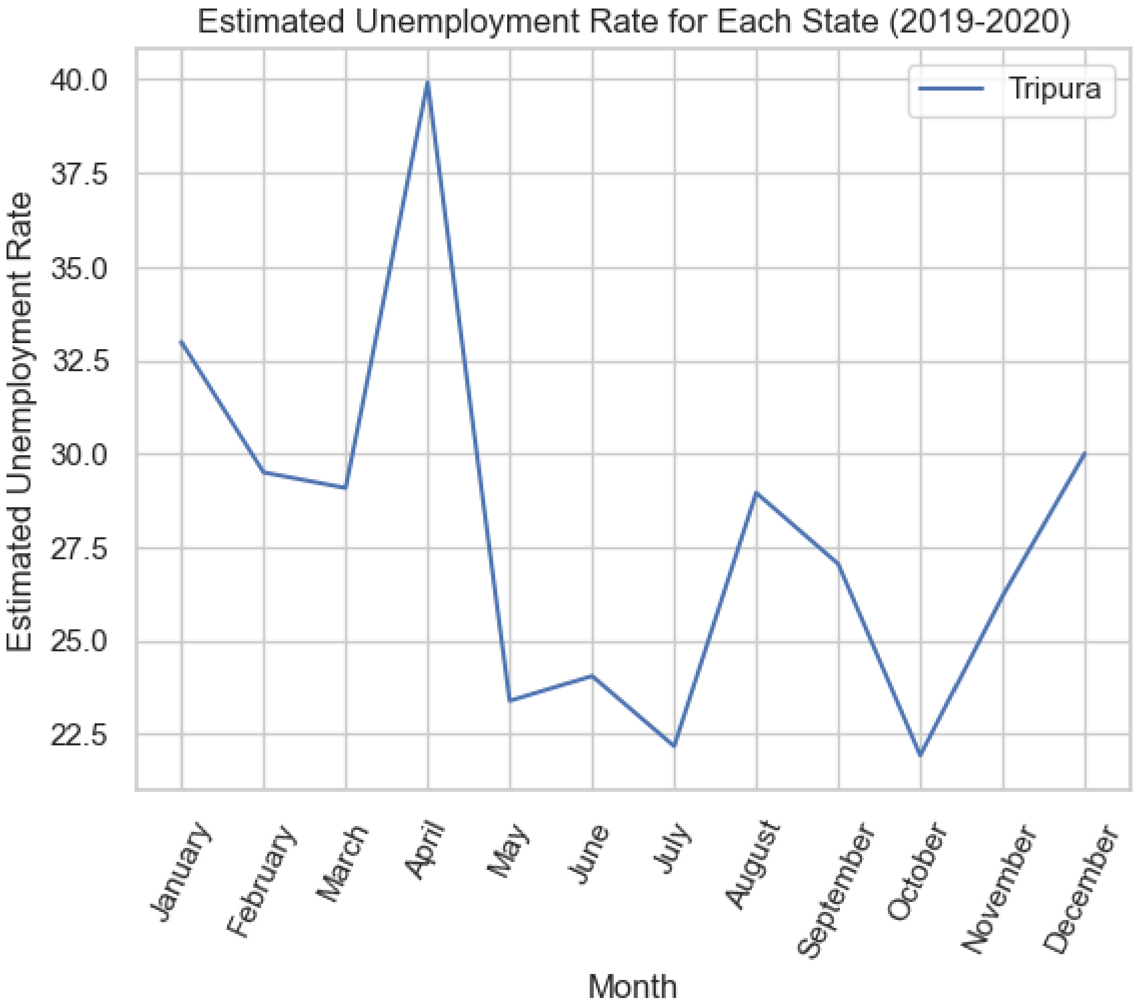

Estimated Unemployment Rate for Each State (2019-2020)
Estimated Unemployment Rate
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
Punjab
January
February
March
April
May
June
July
August
September
October
November
December
Month

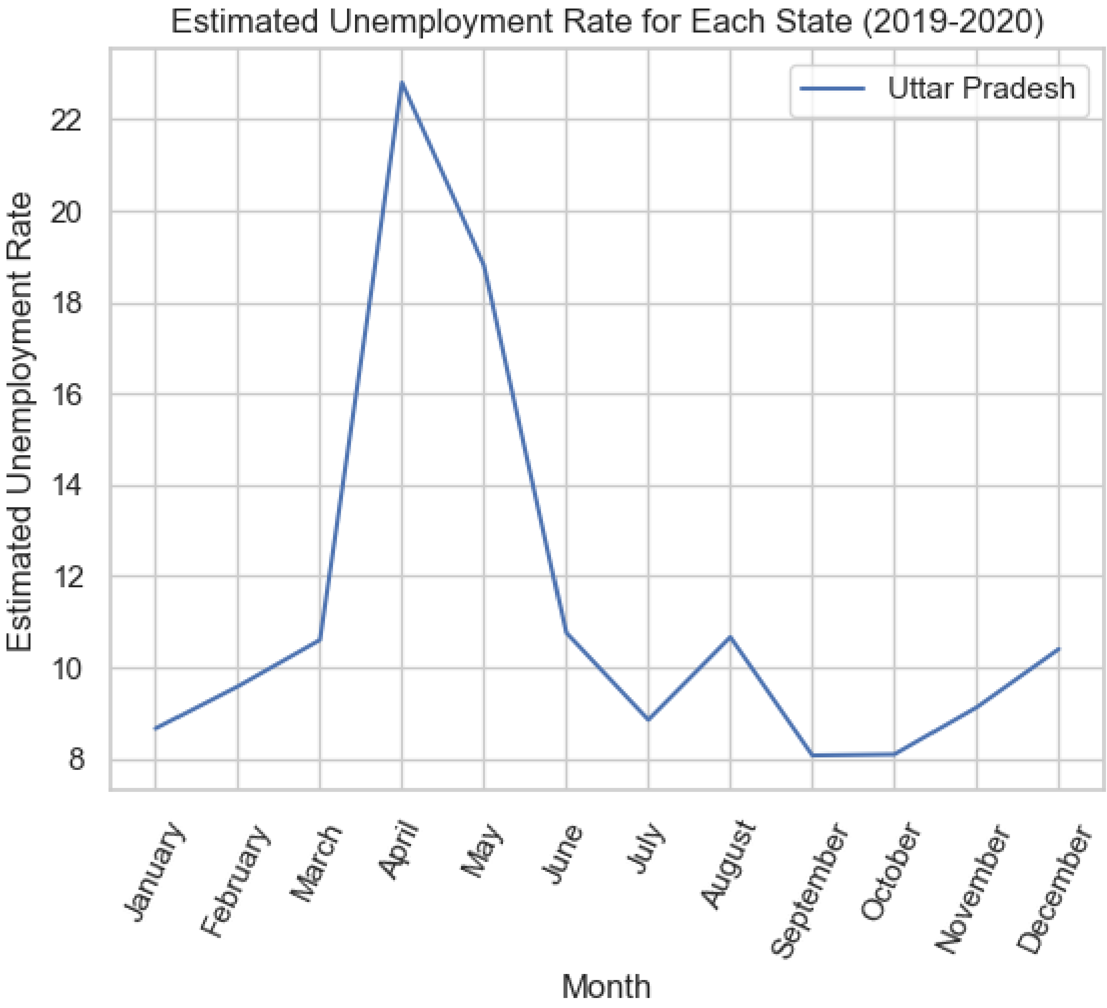

Estimated Unemployment Rate for Each State (2019-2020)
Rajasthan
Estimated Unemployment Rate
22
20
18
16
14
12
January
February
March
April
May
June
July
August
September
October
November
December
Month

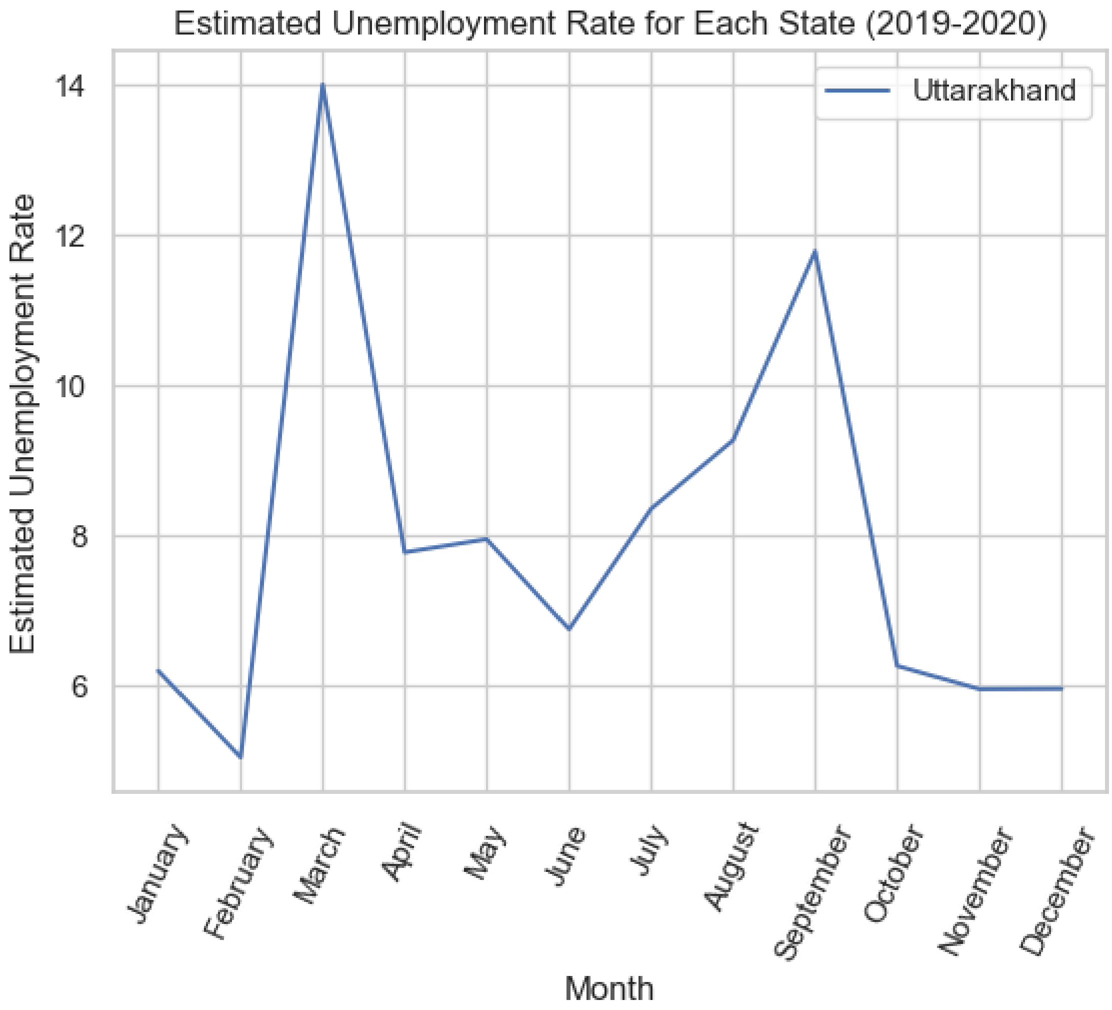

Estimated Unemployment Rate for Each State (2019-2020)
Sikkim
Estimated Unemployment Rate
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
March
April
May
June
July
August
September
October
November
December
Month

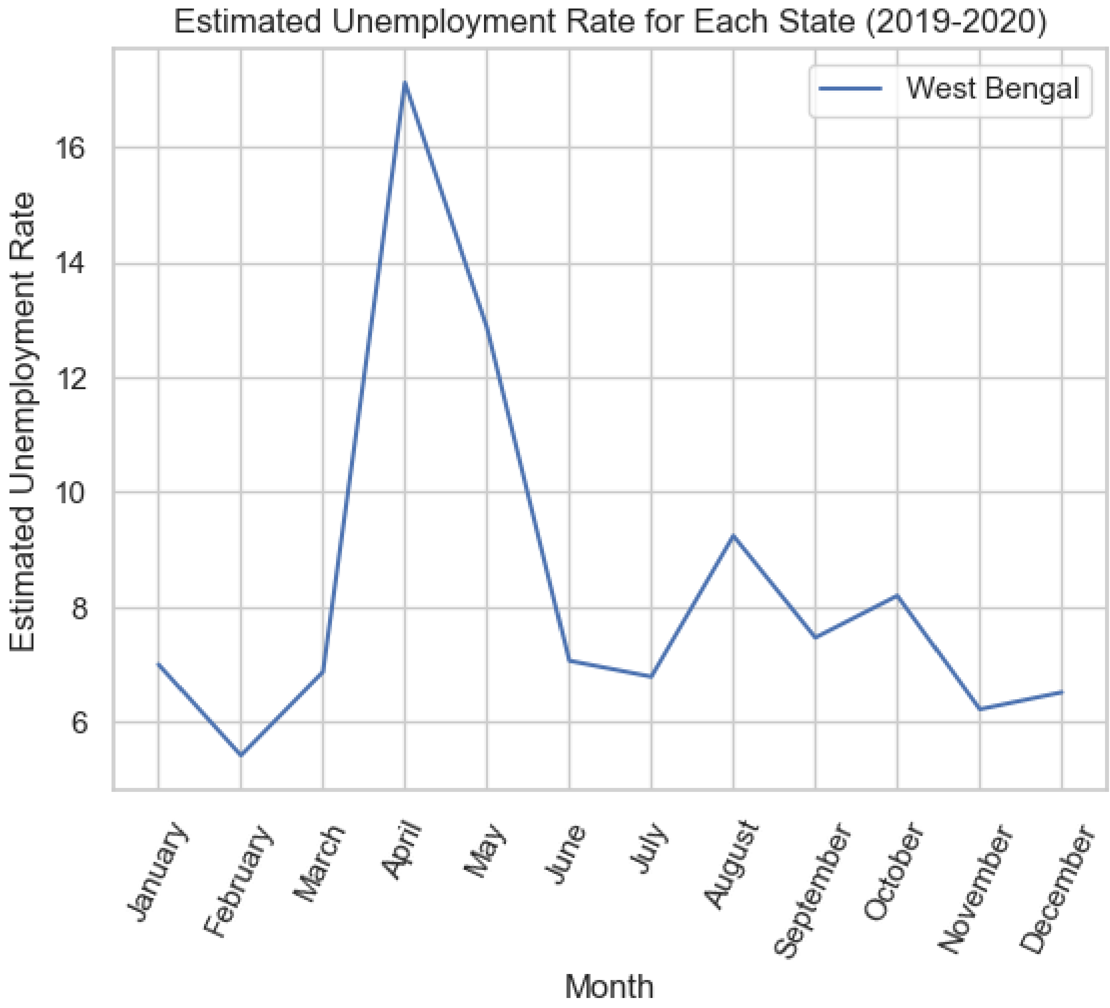

Estimated Unemployment Rate for Each State (2019-2020)
Tamil Nadu
Estimated Unemployment Rate
0
10
20
30
40
50
January
February
March
April
May
June
July
August
September
October
November
December
Month

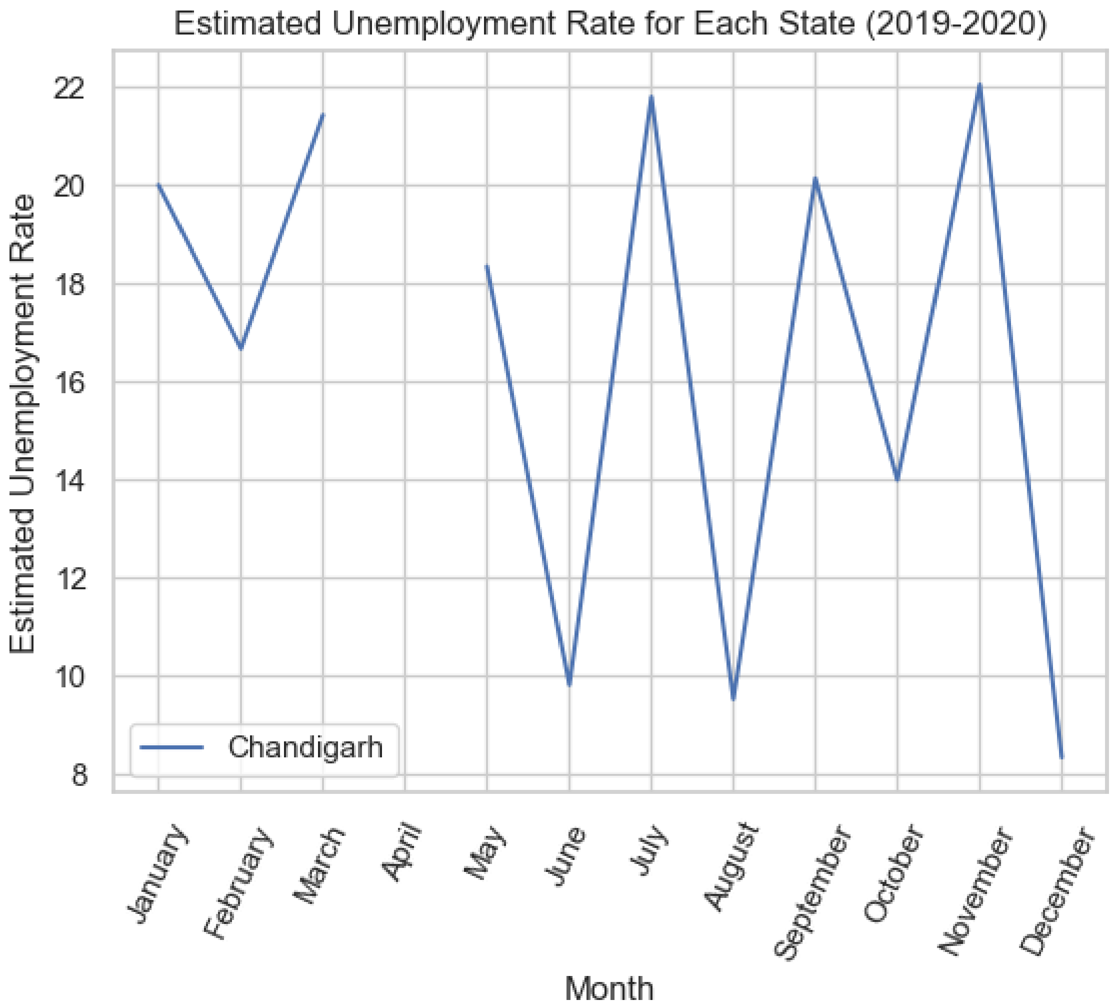

Estimated Unemployment Rate for Each State (2019-2020)
Telangana
Estimated Unemployment Rate
18
16
14
12
10
8
6
4
January
February
March
April
May
June
July
August
September
October
November
December
Month


Estimated Unemployment Rate for Each State (2019-2020)
Estimated Unemployment Rate
40.0
37.5
35.0
32.5
30.0
27.5
25.0
22.5
Tripura
January
February
March
April
May
June
July
August
September
October
November
December
Month


Estimated Unemployment Rate for Each State (2019-2020)
Uttar Pradesh
Estimated Unemployment Rate
22
20
18
16
14
12
10
8
January
February
March
April
May
June
July
August
September
October
November
December
Month


Estimated Unemployment Rate for Each State (2019-2020)
Uttarakhand
Estimated Unemployment Rate
14
12
10
8
6
January
February
March
April
May
June
July
August
September
October
November
December
Month


Estimated Unemployment Rate for Each State (2019-2020)
West Bengal
Estimated Unemployment Rate
16
14
12
10
8
6
January
February
March
April
May
June
July
August
September
October
November
December
Month



```
Based on the data shown above, we can observe that the Unemployment rate tends to
be higher in the months of April and May, and lower in June and July.
```