In [1]:

```python
import numpy as np   #linear algebra
import pandas as pd   #data processing
```

In [3]:

```python
df=pd.read_csv("Unemployment_Rate_upto_11_2020.csv") #read dataset
```

In [4]:

```python
df.head() #returns first 5 entries
```

Out[4]:

	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 [5]:

```python
df.tail() #returns last 5 entries
```

Out[5]:

	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 [6]:

```python
#returns tuple of shape (Rows, columns) of dataframe
df.shape
```

Out[6]:

```
(267, 9)
```

In [7]:

```python
#prints information about the dataframe
df.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 [8]:

```python
#returns numerical description of the data in the dataframe
df.describe()
```

Out[8]:

	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 [9]:

```python
x = df['Region'] #plotting column 'Region' on x-axis
```

In [10]:

```python
x #print x
```

Out[10]:

```
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: Region, Length: 267, dtype: object
```

In [11]:

```python
y=df[' Estimated Unemployment Rate (%)'] #plotting column 'Estimated Unemployment Rate (%)' on y-axi
```

In [12]:

```python
y #print y
```

Out[12]:

```
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:  Estimated Unemployment Rate (%), Length: 267, dtype: float64
```

In [13]:

```python
df2=df.iloc[:,3]
```

In [14]:

```python
df2
```

Out[14]:

```
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:  Estimated Unemployment Rate (%), Length: 267, dtype: float64
```

In [15]:

```python
import plotly.express as px
import matplotlib.pyplot as plt
```

In [16]:

```python
fg = px.bar(df,x='Region',y=' Estimated Unemployment Rate (%)',color='Region',
            title='Unemploymeny Rate (State Wise) by Bar Graph',template='plotly')
fg.update_layout(xaxis={'categoryorder':'total descending'})
fg.show()
```
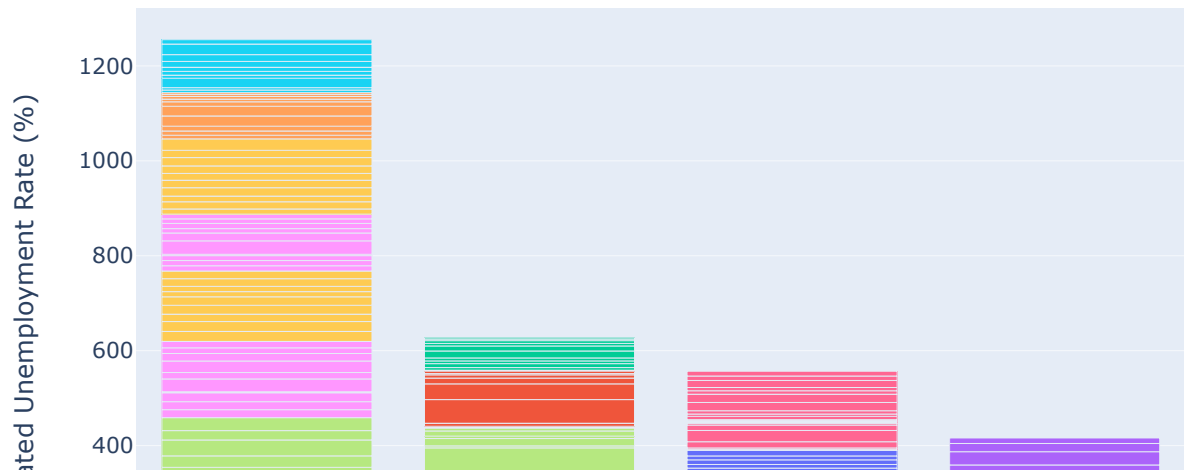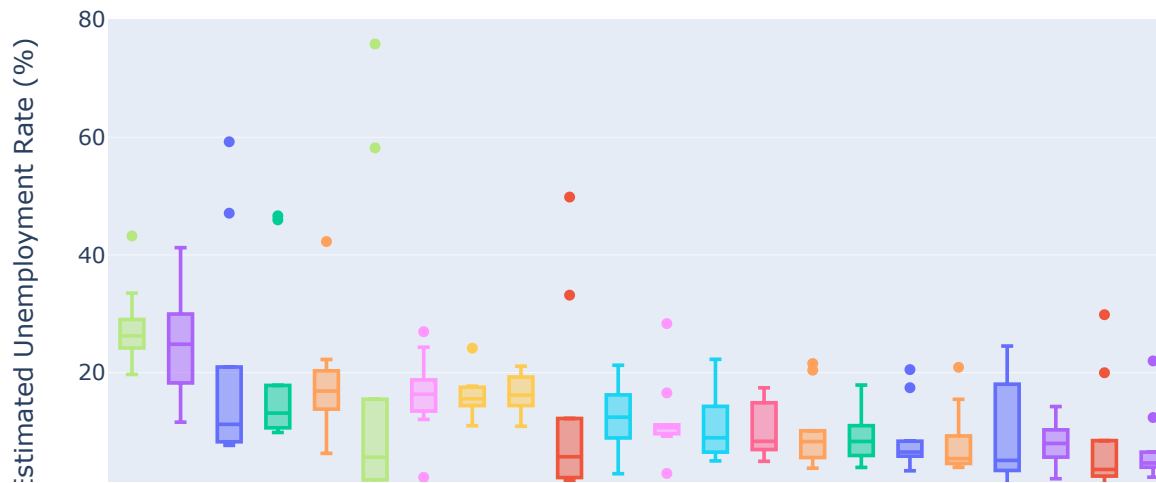
Unemploymeny Rate (State Wise) by Bar Graph
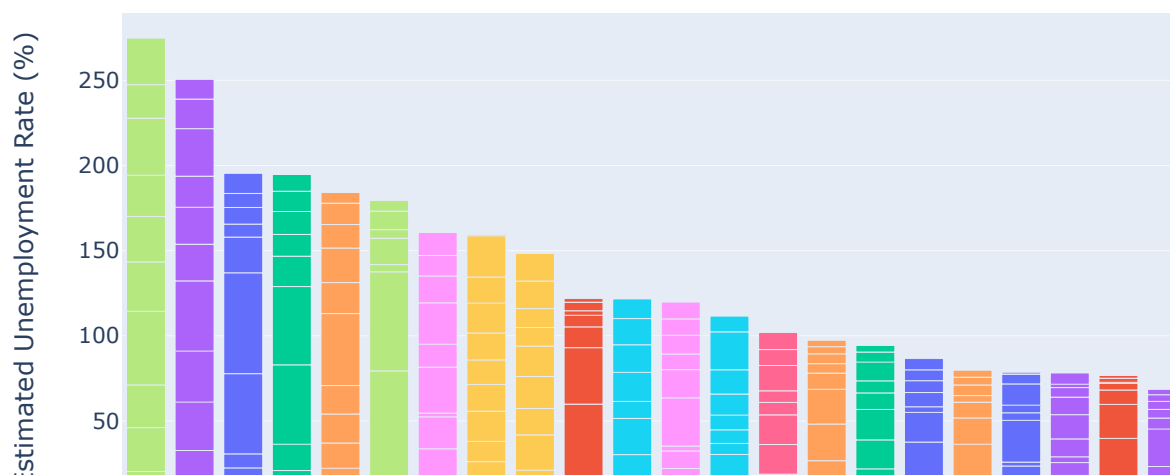
In [17]:

```python
fg = px.bar(df,x='Region.1',y=' Estimated Unemployment Rate (%)',color='Region',
            title='Unemploymeny Rate (Region Wise) by Bar Graph',template='plotly')
fg.update_layout(xaxis={'categoryorder':'total descending'})
fg.show()
```

Unemploymeny Rate (Region Wise) by Bar Graph

In [18]:

```python
fg = px.box(df,x='Region',y=' Estimated Unemployment Rate (%)',color='Region',
            title='Unemploymeny Rate (Statewise) by Box Plot',template='plotly')
fg.update_layout(xaxis={'categoryorder':'total descending'})
fg.show()
```
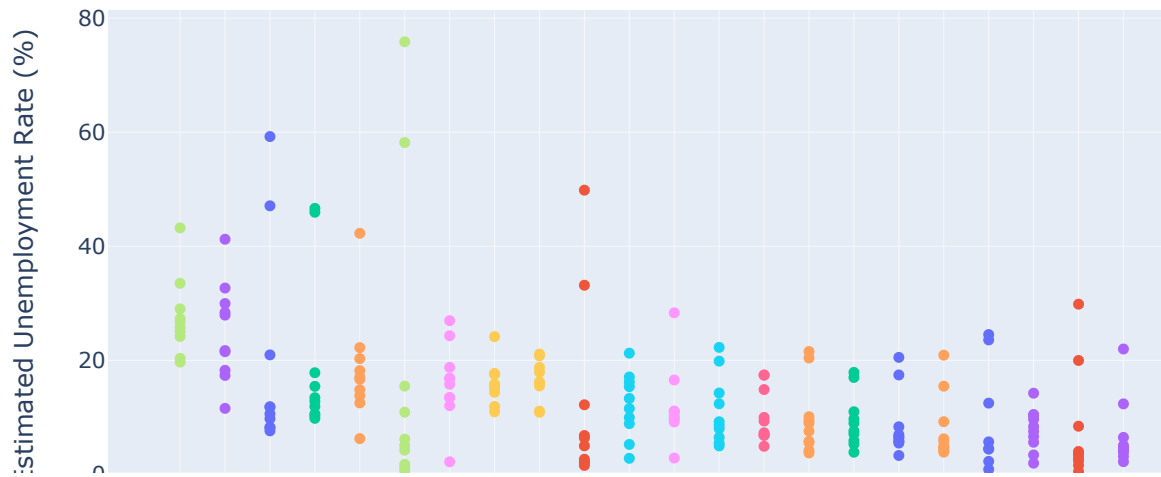
Unemploymeny Rate (Statewise) by Box Plot

In [19]:

```python
fg = px.scatter(df,x='Region',y=' Estimated Unemployment Rate (%)',color='Region',
                title='Unemploymeny Rate (Statewise) by Scatter Plot',template='plotly')
fg.update_layout(xaxis={'categoryorder':'total descending'})
fg.show()
```
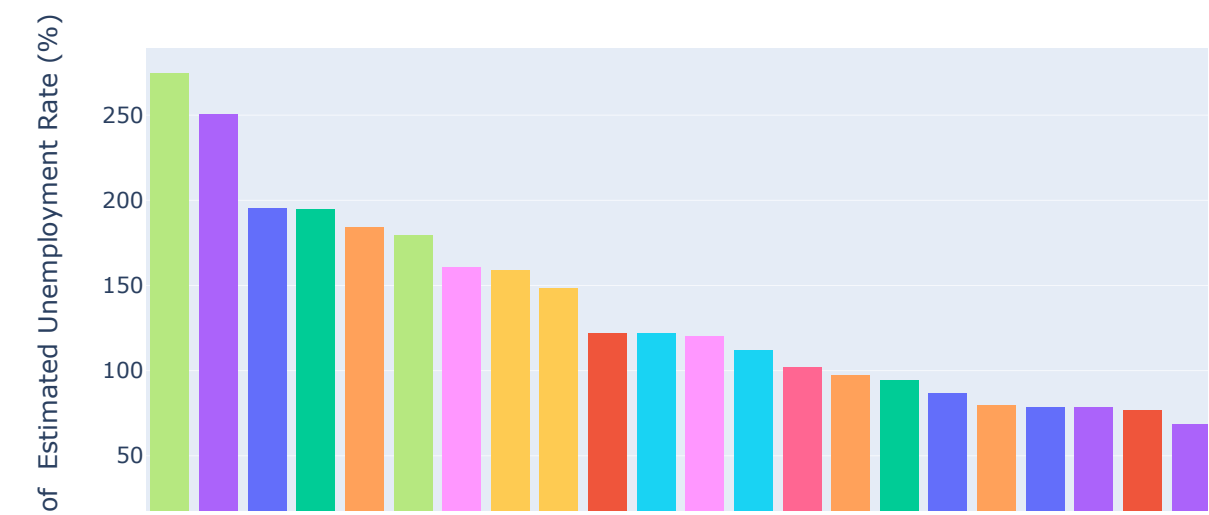
Unemploymeny Rate (Statewise) by Scatter Plot

In [20]:

```
fg = px.histogram(df,x='Region',y=' Estimated Unemployment Rate (%)',color='Region',
                  title='Unemploymeny Rate (Statewise) by Histogram',template='plotly')
fg.update_layout(xaxis={'categoryorder':'total descending'})
fg.show()
```

Unemploymeny Rate (Statewise) by Histogram

In [ ]: