IMPORTING LIBRARIES

In [1]:
```
import pandas as pd
```

In [2]:
```
import numpy as np
```

In [3]:
```
import matplotlib.pyplot as mlt
```

In [4]:
```
import plotly.express as plt
```

IMPORTING DATASET

In [5]:
```
dataset = pd.read_csv("Unemployment.csv")
```

In [6]:
```
dataset
```

Out[6]:

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longitude	lat
0	Andhra Pradesh	31-01-2020	M	5.48	16635535	41.02	South	15.9129	7
1	Andhra Pradesh	29-02-2020	M	5.83	16545652	40.90	South	15.9129	7
2	Andhra Pradesh	31-03-2020	M	5.79	15881197	39.18	South	15.9129	7
3	Andhra Pradesh	30-04-2020	M	20.51	11336911	33.10	South	15.9129	7
4	Andhra Pradesh	31-05-2020	M	17.43	12988845	36.46	South	15.9129	7
...	...	...	...	...	...	...	...	...	
262	West Bengal	30-06-2020	M	7.29	30726310	40.39	East	22.9868	8
263	West Bengal	31-07-2020	M	6.83	35372506	46.17	East	22.9868	8
264	West Bengal	31-08-2020	M	14.87	33298644	47.48	East	22.9868	8
265	West Bengal	30-09-2020	M	9.35	35707239	47.73	East	22.9868	8
266	West Bengal	31-10-2020	M	9.98	33962549	45.63	East	22.9868	8

267 rows × 9 columns

In [7]:
```
dataset.head(15)
```

Out[7]:

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longitude	lati
0	Andhra Pradesh	31-01-2020	M	5.48	16635535	41.02	South	15.9129	79.
1	Andhra Pradesh	29-02-2020	M	5.83	16545652	40.90	South	15.9129	79.
2	Andhra Pradesh	31-03-2020	M	5.79	15881197	39.18	South	15.9129	79.
3	Andhra Pradesh	30-04-2020	M	20.51	11336911	33.10	South	15.9129	79.
4	Andhra Pradesh	31-05-2020	M	17.43	12988845	36.46	South	15.9129	79.
5	Andhra Pradesh	30-06-2020	M	3.31	19805400	47.41	South	15.9129	79.
6	Andhra Pradesh	31-07-2020	M	8.34	15431615	38.91	South	15.9129	79.
7	Andhra Pradesh	31-08-2020	M	6.96	15251776	37.83	South	15.9129	79.
8	Andhra Pradesh	30-09-2020	M	6.40	15220312	37.47	South	15.9129	79.
9	Andhra Pradesh	31-10-2020	M	6.59	15157557	37.34	South	15.9129	79.
10	Assam	31-01-2020	M	4.66	13051904	52.98	Northeast	26.2006	92.
11	Assam	29-02-2020	M	4.41	10088268	40.77	Northeast	26.2006	92.
12	Assam	31-03-2020	M	4.77	11542888	46.73	Northeast	26.2006	92.
13	Assam	30-04-2020	M	11.06	6830817	29.55	Northeast	26.2006	92.
14	Assam	31-05-2020	M	9.55	11367897	48.26	Northeast	26.2006	92.

In [8]:
```
dataset.tail(15)
```

Out[8]:

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Region.1	longitude
252	Uttarakhand	30-06-2020	M	8.61	2656071	33.06	North	30.0668
253	Uttarakhand	31-07-2020	M	12.38	2938552	38.07	North	30.0668
254	Uttarakhand	31-08-2020	M	14.26	2717528	35.90	North	30.0668
255	Uttarakhand	30-09-2020	M	22.26	2695230	39.18	North	30.0668
256	Uttarakhand	31-10-2020	M	9.23	2739309	34.03	North	30.0668
257	West Bengal	31-01-2020	M	6.94	35820789	47.35	East	22.9868
258	West Bengal	29-02-2020	M	4.92	36964178	47.74	East	22.9868
259	West Bengal	31-03-2020	M	6.92	35903917	47.27	East	22.9868
260	West Bengal	30-04-2020	M	17.41	26938836	39.90	East	22.9868
261	West Bengal	31-05-2020	M	17.41	28356675	41.92	East	22.9868
262	West Bengal	30-06-2020	M	7.29	30726310	40.39	East	22.9868
263	West Bengal	31-07-2020	M	6.83	35372506	46.17	East	22.9868
264	West Bengal	31-08-2020	M	14.87	33298644	47.48	East	22.9868
265	West Bengal	30-09-2020	M	9.35	35707239	47.73	East	22.9868
266	West Bengal	31-10-2020	M	9.98	33962549	45.63	East	22.9868

In [9]:
```
dataset.shape
```

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

In [10]:
```
dataset.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 [11]:
```
X = dataset['Region']
```

In [13]:
```
Y = dataset[' Estimated Unemployment Rate (%)']
```

In [14]:
```
X
```

Out[14]:
```
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 [15]:
```
Y
```

Out[15]:
```
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 [16]:
```
df = dataset.iloc[:,3]
```

In [17]:
```
df
```

Out[17]:
```
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
```

ANALYSIS OF DATA USING GRAPHS

In [24]:
```
dia = plt.bar(dataset, x = 'Region.1', y = ' Estimated Unemployment Rate (%)', col
dia.update_layout(xaxis = {'categoryorder':'total descending'})
```

In [26]:
```
dia = plt.bar(dataset, x = 'Region', y = ' Estimated Unemployment Rate (%)', color
dia.update_layout(xaxis = {'categoryorder':'total descending'})
```

ANALYSIS OF DATA USING BOX PLOT

In [27]:
```
dia = plt.box(dataset, x = 'Region', y = ' Estimated Unemployment Rate (%)', color
dia.update_layout(xaxis = {'categoryorder':'total descending'})
```

ANALYSIS OF DATA USING HISTOGRAM

In [28]:
```
dia = plt.histogram(dataset, x = 'Region', y = ' Estimated Unemployment Rate (%)',
dia.update_layout(xaxis = {'categoryorder':'total descending'})
```

In [ ]: