```
# PROBLEM STATEMENT
# The COVID-19 pandemic caused a sharp rise in unemployment, affecting
economies worldwide.
# This project aims to analyze the factors contributing to this
increase and build predictive models to forecast future unemployment
trends.
# By examining labor force data, policies, and industry disruptions,
the goal is to provide insights that can help mitigate future
unemployment spikes.
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import datetime as dt
import warnings
warnings.filterwarnings("ignore")
df = pd.read csv(r"C:\Users\Sarav\OneDrive\Desktop\unemployment
dataset\Unemployment in India.csv")
df
             Region
                            Date Frequency Estimated Unemployment
Rate (%)
     Andhra Pradesh
                      31-05-2019
                                    Monthly
0
3.65
    Andhra Pradesh
                      30-06-2019
1
                                    Monthly
3.05
                      31-07-2019
     Andhra Pradesh
                                    Monthly
3.75
3
     Andhra Pradesh
                      31-08-2019
                                    Monthly
3.32
     Andhra Pradesh
                      30-09-2019
4
                                    Monthly
5.17
. .
. . .
763
                NaN
                             NaN
                                         NaN
NaN
                                         NaN
764
                NaN
                             NaN
NaN
765
                                         NaN
                NaN
                             NaN
NaN
                NaN
                                         NaN
766
                             NaN
NaN
767
                NaN
                             NaN
                                         NaN
NaN
      Estimated Employed Estimated Labour Participation Rate (%)
Area
0
              11999139.0
                                                              43.24
```

11755881.0   42.05					
Rural 2	Rural	11755	001 0		42.05
12086707.0		11/55	001.0		42.03
Rural	2	12086	707.0		43.50
Rural 4	Rural				
12256762.0	3	12285	693.0		43.97
Rural		12256	762 0		44.60
		12250	702.0		44.08
NaN					
NaN	762		NoN		NoN
NaN			waw		Nan
NaN	764		NaN		NaN
NaN	NaN				
NaN	765		NaN		NaN
Man (767 NaN NaN NaN NaN NaN NaN NaN NaN NaN Na	NaN				
NaN			NaN		NaN
Region   Date   Frequency   Estimated   Unemployment			MaN		NaN
Region   Date   Frequency   Estimated   Unemployment			IVAIV		inain
Andhra Pradesh 31-05-2019 Monthly 3.65 L Andhra Pradesh 30-06-2019 Monthly 3.05 L Andhra Pradesh 31-07-2019 Monthly 3.75 B Andhra Pradesh 31-08-2019 Monthly 3.32 A Andhra Pradesh 30-09-2019 Monthly 5.17 Estimated Employed Estimated Labour Participation Rate (%) Area D 11999139.0 43.24 Rural L 11755881.0 42.05 Rural L 12086707.0 43.50 Rural B 12285693.0 43.97	<pre>df.head() Rate (%)</pre>		Date	Frequency	Estimated Unemployment
Andhra Pradesh 30-06-2019 Monthly 3.05 2 Andhra Pradesh 31-07-2019 Monthly 3.75 3 Andhra Pradesh 31-08-2019 Monthly 3.32 4 Andhra Pradesh 30-09-2019 Monthly 5.17  Estimated Employed Estimated Labour Participation Rate (%) Area 9 11999139.0 43.24 Rural 11755881.0 42.05 Rural 2 12086707.0 43.50 Rural 3 12285693.0 43.97	0 Andhra	-	31-05-2019	Monthly	
2 Andhra Pradesh 31-07-2019 Monthly 3.75   3 Andhra Pradesh 31-08-2019 Monthly 3.32   4 Andhra Pradesh 30-09-2019 Monthly 5.17   Estimated Employed Estimated Labour Participation Rate (%) Area   9 11999139.0   43.24   Rural   11755881.0   42.05   Rural   2 12086707.0   43.50   Rural   3 12285693.0   43.97	1 Andhra	Pradesh	30-06-2019	Monthly	
Andhra Pradesh 31-08-2019 Monthly 3.32 Andhra Pradesh 30-09-2019 Monthly 5.17  Estimated Employed Estimated Labour Participation Rate (%) Area D 11999139.0 43.24 Rural L 11755881.0 42.05 Rural D 12086707.0 43.50 Rural B 12285693.0 43.97	2 Andhra	Pradesh	31-07-2019	Monthly	
A Andhra Pradesh 30-09-2019 Monthly 5.17  Estimated Employed Estimated Labour Participation Rate (%) Area  11999139.0 43.24  Rural 11755881.0 42.05  Rural 2 12086707.0 43.50  Rural 3 12285693.0 43.97	3 Andhra	Pradesh	31-08-2019	Monthly	
Area  11999139.0 43.24  Rural  1 11755881.0 42.05  Rural  2 12086707.0 43.50  Rural  3 12285693.0 43.97		Pradesh	30-09-2019	Monthly	
11999139.0 43.24 Rural L 11755881.0 42.05 Rural 2 12086707.0 43.50 Rural 3 12285693.0 43.97		ated Emplo	yed Estima	ted Labour	Participation Rate (%)
Rural L 11755881.0 42.05 Rural 2 12086707.0 43.50 Rural 3 12285693.0 43.97	0	1199913	9.0		43.24
1 11755881.0 42.05  Rural 2 12086707.0 43.50  Rural 3 12285693.0 43.97	Rural	1133313			.3.2.
2 12086707.0 43.50 Rural 3 12285693.0 43.97	1	1175588	1.0		42.05
Rural 3 12285693.0 43.97	Rural		_		
3 12285693.0 43.97	2	1208670	7.0		43.50
Rural	3	1228569	3.0		43.97
	Kural				

4 Rura	l	12256762	. 0				44.68	
df.ta	ail()							
763 764 765 766 767	Region NaN NaN NaN NaN NaN	Date From NaN NaN NaN NaN NaN NaN NaN	equency NaN NaN NaN NaN NaN	Estima	ated Uner	nployment Rate	(%) \ NaN NaN NaN NaN NaN NaN	
Area		ated Empl	oyed Es	timated	l Labour	Participation	Rate (%)	
763 NaN			NaN				NaN	
764 NaN			NaN				NaN	
765			NaN				NaN	
NaN 766			NaN				NaN	
NaN 767			NaN				NaN	
NaN	£.							
df.i		nd DataEr	ame.info	o.f		Pogion	Date	
	uency Andhra		d Unemplo	yment F	Rate (%) Monthly	Region \	Date	
1 3.05	Andhra	Pradesh	30-06-2	019	Monthly			
2 3.75		Pradesh	31-07-2	019	Monthly			
3 3 3.32		Pradesh	31-08-2	019	Monthly			
4 5.17	Andhra	Pradesh	30-09-2	019	Monthly			
763		NaN		NaN	NaN			
NaN 764		NaN		NaN	NaN			
NaN 765		NaN		NaN	NaN			
NaN 766		NaN		NaN	NaN			
NaN 767		NaN		NaN	NaN			

```
NaN
      Estimated Employed Estimated Labour Participation Rate (%)
Area
              11999139.0
                                                                43.24
0
Rural
              11755881.0
                                                                42.05
1
Rural
               12086707.0
                                                                43.50
Rural
              12285693.0
                                                                43.97
Rural
              12256762.0
                                                                44.68
Rural
. . .
                      NaN
763
                                                                   NaN
NaN
764
                      NaN
                                                                   NaN
NaN
                                                                   NaN
765
                      NaN
NaN
766
                      NaN
                                                                   NaN
NaN
                      NaN
                                                                   NaN
767
NaN
[768 \text{ rows } x 7 \text{ columns}] >
df.shape
(768, 7)
df.columns
Index(['Region', ' Date', ' Frequency', ' Estimated Unemployment Rate
(%)',
       ' Estimated Employed', ' Estimated Labour Participation Rate
(%)',
       'Area'],
      dtype='object')
df.describe()
       Estimated Unemployment Rate (%) Estimated Employed \
                             740.000000
                                                7.400000e+02
count
                              11.787946
                                                7.204460e+06
mean
std
                               10.721298
                                                8.087988e+06
min
                               0.000000
                                                4.942000e+04
                               4.657500
25%
                                                1.190404e+06
                               8.350000
                                                4.744178e+06
50%
```

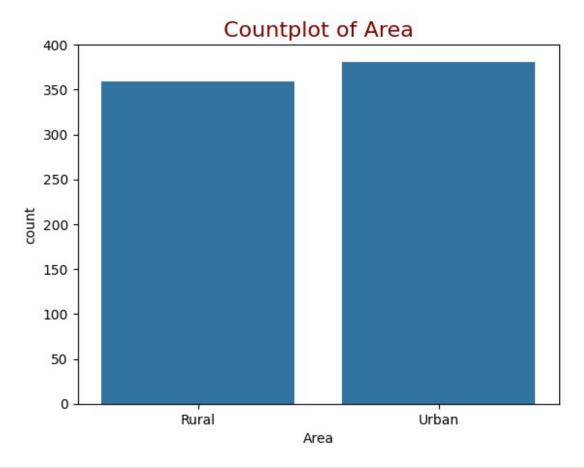
75% max				887500 740000	1.127549e+ 4.577751e+	
count mean std min 25% 50% 75% max		ated Lab	our Particip	ation Rate 740.000 42.630 8.111 13.330 38.062 41.160 45.505	0000 0122 L094 0000 2500 0000	
df.is	null()	# IT CHE	CKS FOR MISS	ING VALUES		
0 1 2 3 4  763 764 765	Region False False False False True True True	Date False False False False True True True	Frequency False False False False True True True	Estimated	Unemployment	Rate (%) \ False False False False False True True True
766 767	True True	True True	True True True			True True True
Area	Estima	ted Empl	oyed Estim	ated Labour	<sup>-</sup> Participatio	n Rate (%)
0 False		F	alse			False
1 False		F	alse			False
2 False		F	alse			False
3 False		F	alse			False
4 False		F	alse			False
763 True			True			True
764 True			True			True
765			True			True
True 766 True			True			True

```
767
                    True
                                                              True
True
[768 rows x 7 columns]
df.isnull().sum() # THERE ARE MISSING VALUES
Region
                                            28
Date
                                            28
Frequency
                                            28
Estimated Unemployment Rate (%)
                                            28
Estimated Employed
                                            28
Estimated Labour Participation Rate (%)
                                            28
Area
                                            28
dtype: int64
df.dropna(inplace=True) # NULL VALUES ARE DELETED
df
             Region
                            Date Frequency Estimated Unemployment
Rate (%)
     Andhra Pradesh
                      31-05-2019
                                    Monthly
0
3.65
     Andhra Pradesh
                      30-06-2019
1
                                    Monthly
3.05
2
     Andhra Pradesh
                      31-07-2019
                                    Monthly
3.75
3
     Andhra Pradesh
                      31-08-2019
                                    Monthly
3.32
4
     Andhra Pradesh
                      30-09-2019
                                    Monthly
5.17
. .
749
        West Bengal
                      29-02-2020
                                    Monthly
7.55
750
        West Bengal
                      31-03-2020
                                    Monthly
6.67
        West Bengal
                      30-04-2020
751
                                    Monthly
15.63
        West Bengal
                      31-05-2020
752
                                    Monthly
15.22
753
        West Bengal
                      30-06-2020
                                    Monthly
9.86
      Estimated Employed Estimated Labour Participation Rate (%)
Area
              11999139.0
                                                             43.24
Rural
              11755881.0
                                                             42.05
Rural
```

2	12086707	7.0			•	43.50
Rural 3	12285693	3.0			,	43.97
Rural						
4	12256762	2.0			•	44.68
Rural						
• •						
749	10871168	3.0				44.09
Urban						
750	10806105	5.0			•	43.34
Urban		_				
751	9299466	5.0			•	41.20
Urban 752	9240903	2 0				40.67
Urban	924090.	5.0			•	40.07
753	9088933	1.0				37.57
Urban						
[740 rows x 7	7 columns	1				
·						
(740, 7)						
<pre>df = df.renam Frequency':'F 'Est_Unemp_Ra ' Esti Participation df</pre>	Frequency ate', imated Emp	', ' Estin oloyed':'E	nated Unemplo Est_Emp_Rate'	yment Rate , ' Estimat	(%) :	ur.
				_	dex(drop	
	Region	Dat	ce Frequency	_	·	
				_	·	
0 Andhra F 11999139.0 1 Andhra F	Pradesh		19 Monthly	_	_Rate	
0 Andhra F 11999139.0 1 Andhra F 11755881.0 2 Andhra F	Pradesh Pradesh	31-05-201	Monthly Monthly	_	Rate	
0 Andhra F 11999139.0 1 Andhra F 11755881.0 2 Andhra F 12086707.0 3 Andhra F	Pradesh Pradesh Pradesh	31-05-201 30-06-201	Monthly Monthly Monthly Monthly	_	Rate 3.65 3.05	
0 Andhra F 11999139.0 1 Andhra F 11755881.0 2 Andhra F 12086707.0 3 Andhra F 12285693.0 4 Andhra F	Pradesh Pradesh Pradesh Pradesh	31-05-201 30-06-201 31-07-201	Monthly Monthly Monthly Monthly Monthly Monthly	_	Rate 3.65 3.05 3.75	
0 Andhra F 11999139.0 1 Andhra F 11755881.0 2 Andhra F 12086707.0 3 Andhra F 12285693.0	Pradesh Pradesh Pradesh Pradesh	31-05-201 30-06-201 31-07-201 31-08-201	Monthly Monthly Monthly Monthly Monthly Monthly	_	Rate 3.65 3.05 3.75 3.32	
0 Andhra F 11999139.0 1 Andhra F 11755881.0 2 Andhra F 12086707.0 3 Andhra F 12285693.0 4 Andhra F 12256762.0 	Pradesh Pradesh Pradesh Pradesh	31-05-201 30-06-201 31-07-201 31-08-201	Monthly Monthly Monthly Monthly Monthly Monthly Monthly Monthly Monthly	_	Rate 3.65 3.05 3.75 3.32	
0 Andhra F 11999139.0 1 Andhra F 11755881.0 2 Andhra F 12086707.0 3 Andhra F 12285693.0 4 Andhra F 12256762.0  735 West 10871168.0	Pradesh Pradesh Pradesh Pradesh Pradesh Pradesh Pradesh	31-05-201 30-06-201 31-07-201 31-08-201 30-09-201	Monthly	_	Rate 3.65 3.05 3.75 3.32 5.17	

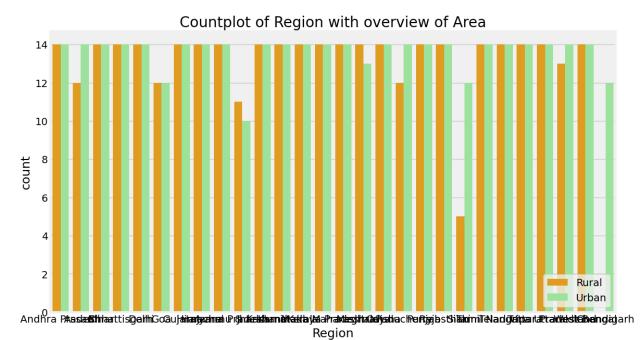
```
737
        West Bengal
                      30-04-2020
                                   Monthly
                                                      15.63
9299466.0
738
        West Bengal
                      31-05-2020
                                   Monthly
                                                      15.22
9240903.0
739
        West Bengal 30-06-2020
                                   Monthly
                                                       9.86
9088931.0
     Est Labour Rate
                       Area
0
               43.24
                      Rural
1
               42.05
                      Rural
2
               43.50
                      Rural
3
               43.97
                      Rural
4
               44.68
                      Rural
735
               44.09
                      Urban
736
               43.34
                      Urban
               41.20
                     Urban
737
738
               40.67 Urban
739
               37.57 Urban
[740 rows x 7 columns]
# CHANGE THE DTYPE OF DATE COLUMN
df['Date'] = pd.to datetime(df['Date'])
df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 740 entries, 0 to 739
Data columns (total 7 columns):
#
     Column
                      Non-Null Count
                                       Dtype
- - -
     -----
 0
     Region
                      740 non-null
                                       object
 1
                      740 non-null
     Date
                                       datetime64[ns]
 2
     Frequency
                      740 non-null
                                       object
 3
     Est Unemp Rate
                      740 non-null
                                       float64
4
                      740 non-null
                                       float64
     Est Emp Rate
 5
     Est_Labour_Rate
                      740 non-null
                                       float64
 6
     Area
                      740 non-null
                                       object
dtypes: datetime64[ns](1), float64(3), object(3)
memory usage: 40.6+ KB
# CHECK FOR NO DUPLICATED ROWS IN DATASET
df.duplicated().sum()
np.int64(0)
df.Frequency.value_counts()
Frequency
Monthly
            381
```

```
Monthly
           359
Name: count, dtype: int64
df['Year'] = df['Date'].dt.year
df['Month'] = df['Date'].dt.month
df.head()
                       Date Frequency Est Unemp Rate
          Region
Est_Emp_Rate \
0 Andhra Pradesh 2019-05-31
                              Monthly
                                                 3.65
                                                         11999139.0
1 Andhra Pradesh 2019-06-30
                              Monthly
                                                 3.05
                                                         11755881.0
2 Andhra Pradesh 2019-07-31
                                                 3.75
                                                         12086707.0
                              Monthly
3 Andhra Pradesh 2019-08-31
                              Monthly
                                                 3.32
                                                         12285693.0
4 Andhra Pradesh 2019-09-30
                              Monthly
                                                 5.17
                                                         12256762.0
   Est_Labour_Rate
                   Area Year
                                Month
0
            43.24 Rural 2019
                                    5
1
            42.05 Rural
                          2019
                                    6
2
            43.50 Rural
                                    7
                          2019
3
                                    8
            43.97 Rural 2019
4
                                    9
            44.68 Rural 2019
df.Area.value counts()
Area
Urban
        381
        359
Rural
Name: count, dtype: int64
x = sns.countplot(data=df,x='Area')
plt.title("Countplot of Area",fontdict={'fontsize': 16, 'fontweight' :
12, 'color' : 'Maroon'})
plt.show()
```

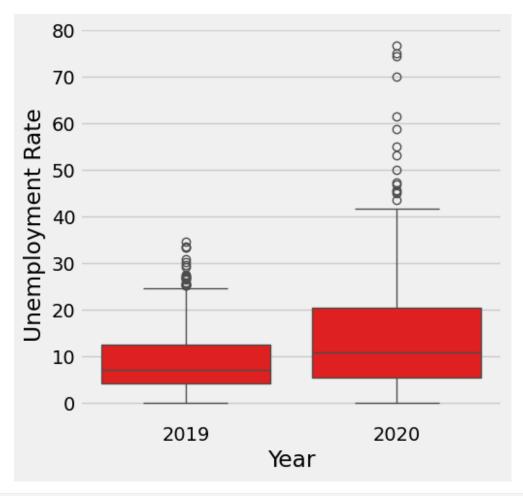


df.Region.value_co	unts()	
Region		
Andhra Pradesh	28	
Kerala	28	
West Bengal	28	
Uttar Pradesh	28	
Tripura	28	
Telangana	28	
Tamil Nadu	28	
Rajasthan Punjab	28 28	
Odisha	28	
Madhya Pradesh	28	
Maharashtra	28	
Karnataka	28	
Jharkhand	28	
Himachal Pradesh	28	
Haryana	28	
Gujarat	28	
Delhi	28	
Chhattisgarh	28	
Bihar	28	

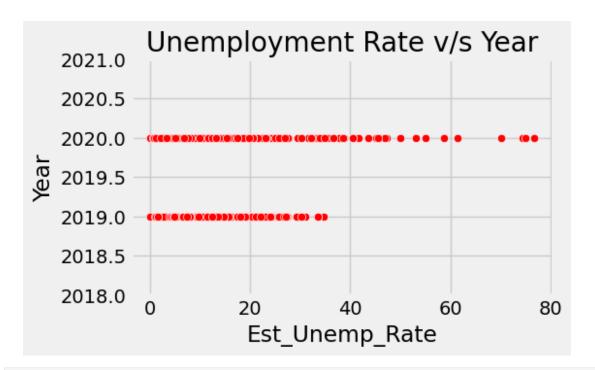
```
Meghalaya
                    27
Uttarakhand
                    27
Assam
                    26
Puducherry
                    26
Goa
                    24
Jammu & Kashmir
                    21
Sikkim
                    17
Chandigarh
                    12
Name: count, dtype: int64
plt.figure(figsize=(14,7))
plt.style.use('fivethirtyeight')
ax = sns.countplot( x=df['Region'], hue = df['Area'], palette =
['orange','lightgreen'] )
plt.legend(loc = 'lower right')
plt.title('Countplot of Region with overview of Area')
plt.show()
```



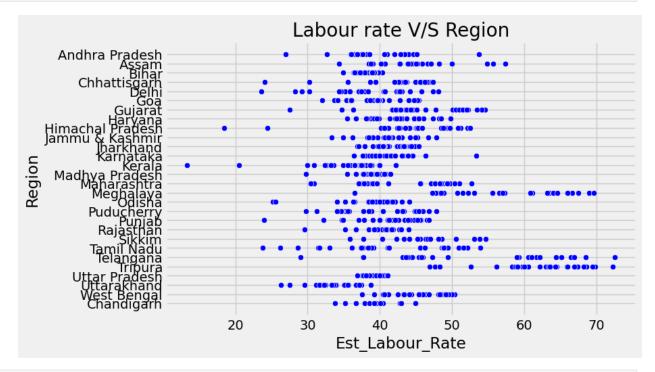
```
# BOXPLOT
plt.figure(figsize=(5,5))
sns.boxplot(y=df['Est_Unemp_Rate'],x=df['Year'],color='red')
plt.ylabel('Unemployment Rate')
plt.show()
```



```
plt.figure(figsize=(5,3))
sns.scatterplot(x=df['Est_Unemp_Rate'],y=df['Year'],color='red')
plt.title('Unemployment Rate v/s Year')
plt.ylim(2018,2021)
plt.show()
# HERE 2020 YEAR HAS A RAPID INCREASE OF UNEMPLOYMENT RATE AFTER COVID
```



```
plt.figure(figsize=(8,5))
sns.scatterplot(x=df['Est_Labour_Rate'], y=df['Region'],color='blue')
plt.title('Labour rate V/S Region')
plt.show() # here meghalaya,tripura and telengana has rapid high of
labour rate
```



```
plt.figure(figsize=(8,5))
sns.scatterplot(x=df['Est_Unemp_Rate'],y=df['Region'],palette=('blue')
```

```
plt.title('Unemployment Rate v/s Region')
plt.show()
```

