Unemployment analysis with python

Abstract :

This report presents an analysis of unemployment trends in India, utilizing a dataset covering various states and regions. The study employs Python libraries such as Pandas, Matplotlib, Seaborn, and Plotly for data loading, exploratory data analysis (EDA), and visualization. Key aspects of the analysis include examining overall employment distribution, regional unemployment rates, and a focused investigation into the impact of the COVID-19 lockdown on unemployment rates across different states. The report highlights data preprocessing, visualization techniques, and comparative analysis of unemployment figures before and during the lockdown period, providing insights into the economic impact of the pandemic.

Introduction :

Unemployment is a critical economic indicator that reflects the health of a nation's labor market. In a diverse country like India, understanding unemployment patterns across different states and regions is crucial for policymaking and economic planning. This report delves into an analysis of unemployment data, with a particular focus on visualizing the distribution of estimated employed individuals, regional unemployment rates, and assessing the changes in unemployment due to specific events, such such as the lockdown. The analysis leverages various data science techniques to extract meaningful insights and present them effectively.

Data Loading and Initial Exploration :

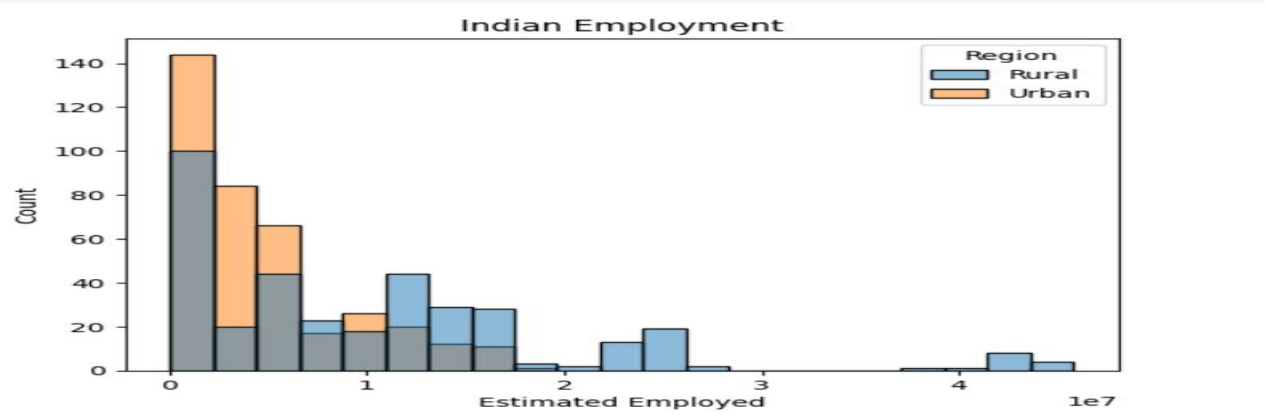
The analysis began by loading the Unemployment in India.csv dataset into a Pandas DataFrame. Initial steps involved understanding the dataset's structure and contents:

- Loading the Dataset: The Unemployment in India.csv file was loaded into a Pandas DataFrame for further processing.
- Dataset Overview
- Correlation Matrix Computation: A correlation matrix was computed using Correlation Heatmap Visualization: The computed correlation matrix was visualized as a heatmap using seaborn.heatmap with a 'coolwarm' colormap. This heatmap provided a quick visual assessment of inter-feature correlations, highlighting strong positive or negative relationships within the data.

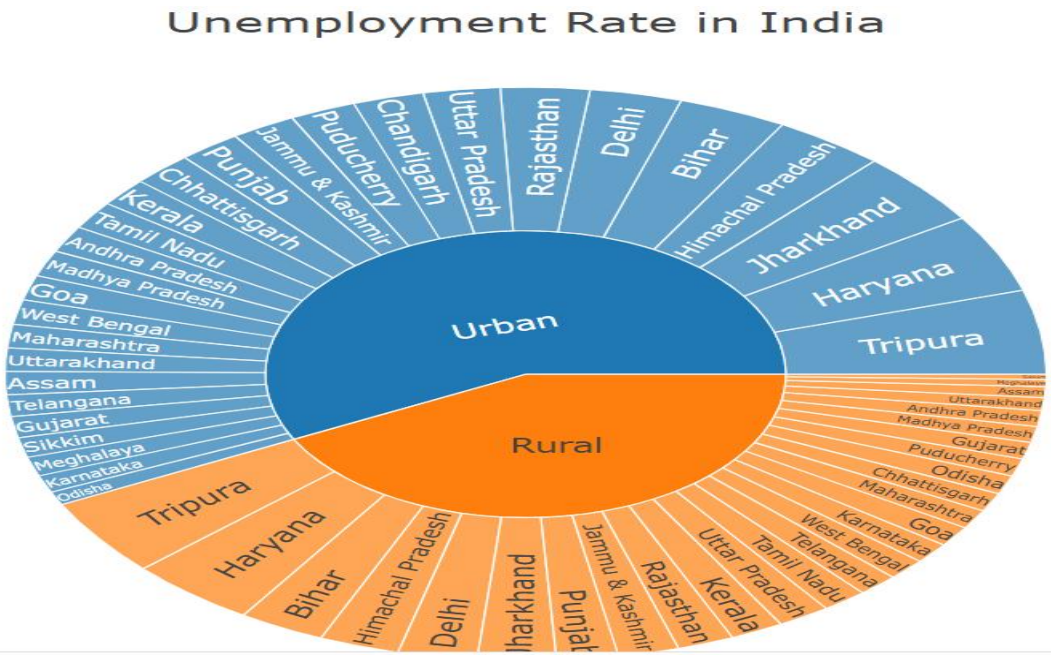
Exploratory Data Analysis (EDA)

Exploratory Data Analysis was performed to uncover patterns, identify anomalies, and establish relationships within the dataset.

- Estimated Employed Distribution by Region: A histogram was generated to visualize the distribution of "Estimated Employed" individuals, segmented by "Region". This plot helps in understanding which regions contribute most to employment and the overall spread of employment figures.



- Unemployment Rate Sunburst Chart: Missing values were handled by dropping rows where all values were null . A plotly.express sunburst chart was then created to interactively visualize the "Estimated Unemployment Rate" across different "Regions" and their respective "States". This hierarchical visualization provides a detailed breakdown of unemployment rates at regional and state levels.



Impact of Lockdown Analysis :

This section focused on analyzing the lockdown's impact on unemployment rates through the following steps:

- **Data Preprocessing:**

	States	Date	Frequency	Estimated Unemployment Rate	Estimated Employed	Estimated Labour Participation Rate	Region	Month_int
0	Andhra Pradesh	2019-05-31	Monthly	3.65	11999139.0	43.24	Rural	5
1	Andhra Pradesh	2019-06-30	Monthly	3.05	11755881.0	42.05	Rural	6
2	Andhra Pradesh	2019-07-31	Monthly	3.75	12086707.0	43.50	Rural	7
3	Andhra Pradesh	2019-08-31	Monthly	3.32	12285693.0	43.97	Rural	8
4	Andhra Pradesh	2019-09-30	Monthly	5.17	12256762.0	44.68	Rural	9

- **Defining Lockdown Periods:** Calculating Mean Unemployment Rates: Mean unemployment rates for each state were calculated for both the 'Before Lockdown' and 'During/After Lockdown' periods to enable direct comparison.

	States	Unemployment Rate before lockdown	Unemployment Rate after lockdown
0	Andhra Pradesh	10.054167	10.756250
1	Assam	6.049091	6.750000
2	Bihar	25.370000	23.690000
3	Chandigarh	14.942500	19.366667
4	Chhattisgarh	11.659167	9.028750

- **Computing Percentage Change:** The 'percentage change in Unemployment' for each state was calculated to quantify the lockdown's effect on unemployment.