

Unemployment Analysis

Table of Contents

1. Introduction
2. Data Overview
3. Data Analysis
4. Time series
5. Results

Introduction

The Unemployment Data Science Project aims to analyze and understand patterns in unemployment data to derive valuable insights. This report presents the methodologies used data sources and key finding from the analysis. This Project aim to gain valuable insight into to the impact of the pandemic on the workforce and the economy.

Data Overview

The Unemployment Data Science project utilizes two distinct datasets to comprehensively analyze unemployment trends in India.

These datasets provide valuable insights into different aspects of unemployment within the region.

1.Unemployment in India

The first dataset focuses on rural areas and covers the month of May 2019. It includes the following key variables:

- **Regions:** Andhra Pradesh
- **Date:** 31-05-2019
- **Frequency:** Monthly
- **Estimate Unemployment Rate (%):** 3.65
- **Estimated Employed:** 11,999,139
- **Estimated Labor Participation Rate:**43.26
- **Area:** Rural

2. Unemployment Rate (January 2020)

The second dataset, collected in January 2020, provides a more detailed perspective on unemployment within different regions of Andhra Pradesh. The key variables are as follows:

- **Regions:** Andhra Pradesh
- **Date:** 31-01-2020
- **Frequency:** Monthly
- **Estimate Unemployment Rate (%):** 5.48
- **Estimated Employed:** 16,635,535
- **Estimated Labor Participation Rate:**41.02
- **Area:** Rural
- **Region(subdivision):** South
- **Longitude:** 15.9129
- **Latitude:** 79.740

To enhance the comprehensiveness of our analysis, we merged two datasets focusing on unemployment in India. The datasets were merged using the common variable "Region," ensuring alignment

between the rural and regional data points. The resulting dataset combines information on unemployment rates, employment figures, and labor participation rates, along with additional geographical details for the regional dataset.

The merged dataset allows for a more in-depth analysis by combining the strengths of both datasets. It enables us to explore not only rural unemployment trends but also to understand variations across different regions within India.

Before conducting the analysis, both datasets underwent a thorough data preprocessing phase. This involved handling missing values, checking for outliers, and standardizing variable formats. The goal was to create clean and reliable datasets for subsequent analysis.

Data Analysis

Data Analysis is an essential step in uncovering patterns, trends, and potential insights within our dataset. This phase involves a systematic approach to summarize the main features, identify patterns.

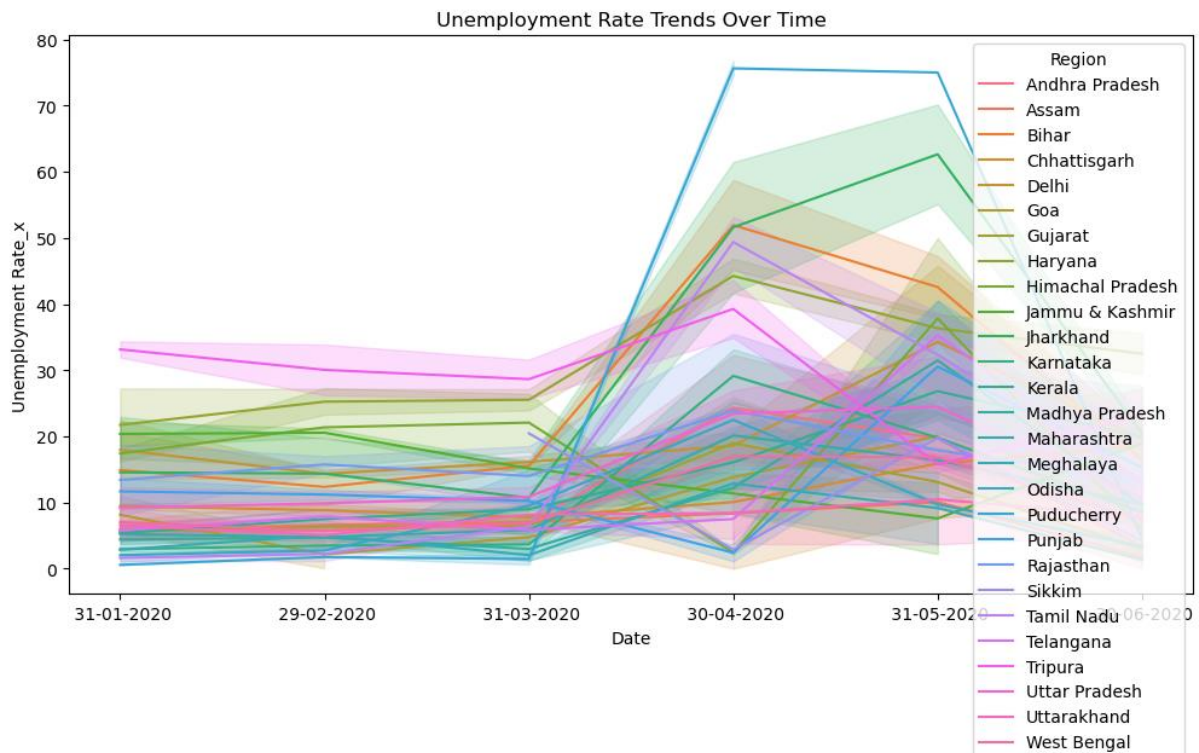
Descriptive Statistics:

The following table provides a summary of descriptive statistics for key variables in the unemployment dataset. The dataset includes information on unemployment rates, employment figures, labor participation rates, and geographical coordinates across different regions in India.

Descriptive Statistics	Values
Count	306
Mean	15.06
Standard Deviation	13.62
Minimum	0.00
25 th percentile	5.38

Median(50th)	10.74
75th percentile	20.37
Maximum	76.74

Unemployment Rate Trends:



The visualizations you sent is a graph of unemployment rate trends over time in India, for each region. The graph covers the period from January 31, 2020 to June 30, 2023.

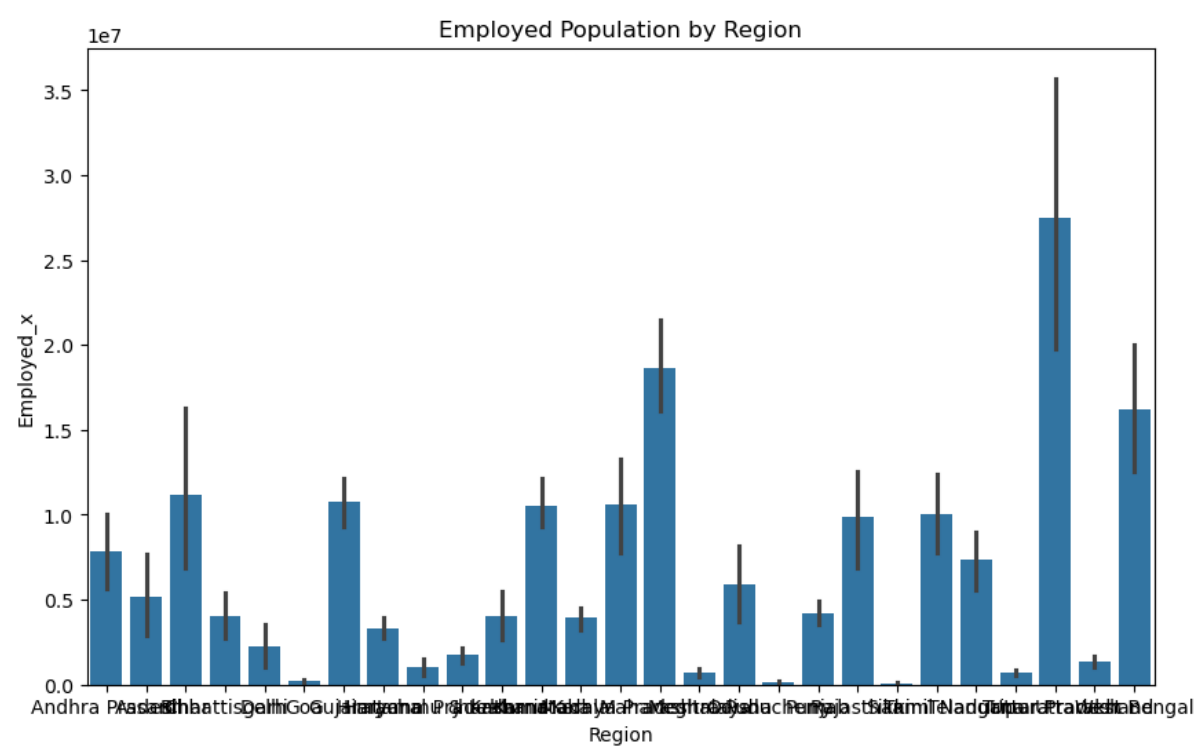
The graph shows that the unemployment rate in India varied significantly across regions in June 2023. The highest unemployment rates were in Assam (10.7%), Bihar (9.0%), and Jharkhand (8.5%). The lowest unemployment rates were in Goa (2.1%), Sikkim (2.2%), and Puducherry (2.3%).

Here are some additional observations from the graph:

- The unemployment rate in India has declined overall since January 2020, but it remains higher than pre-pandemic levels.

- The unemployment rate is higher in rural areas than in urban areas.
- The unemployment rate is higher among women than men.
- The unemployment rate is higher among youth (aged 15-24) than adults (aged 25-64).

Employed Population by Regions:



The image you sent shows a graph of the employed population by region in India. The graph shows that the employed population in India varies significantly across regions in 2022. The highest employed population is in Maharashtra (77738119), followed by Andhra Pradesh (45150051) and Uttar Pradesh (42260675). The lowest employed population is in Dadra and Nagar Haveli and Daman and Diu (1084038) and Puducherry (1775224).

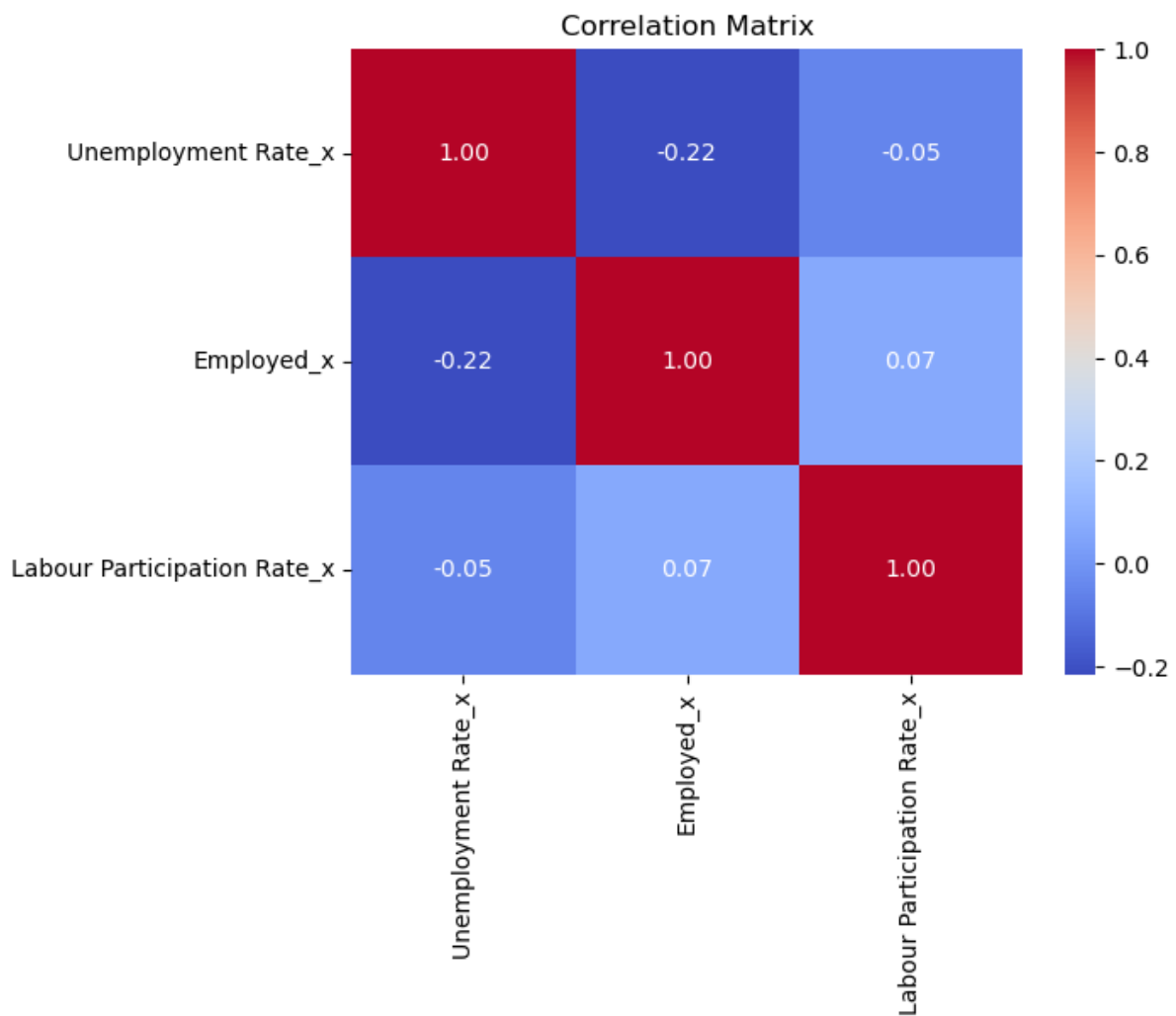
Here is a table that shows the employed population by region in India in 2022:

Region	Employed Population (in millions)
Uttar Pradesh	42.260675

Maharashtra	37.738119
Andhra Pradesh	26.5150051
Bihar	24.779906
Telangana	24.475276

As you can see, the employed population in India is concentrated in the southern and western states, while the employed population in the northern and eastern states is lower. This is likely due to a number of factors, such as economic development, urbanization, and education levels.

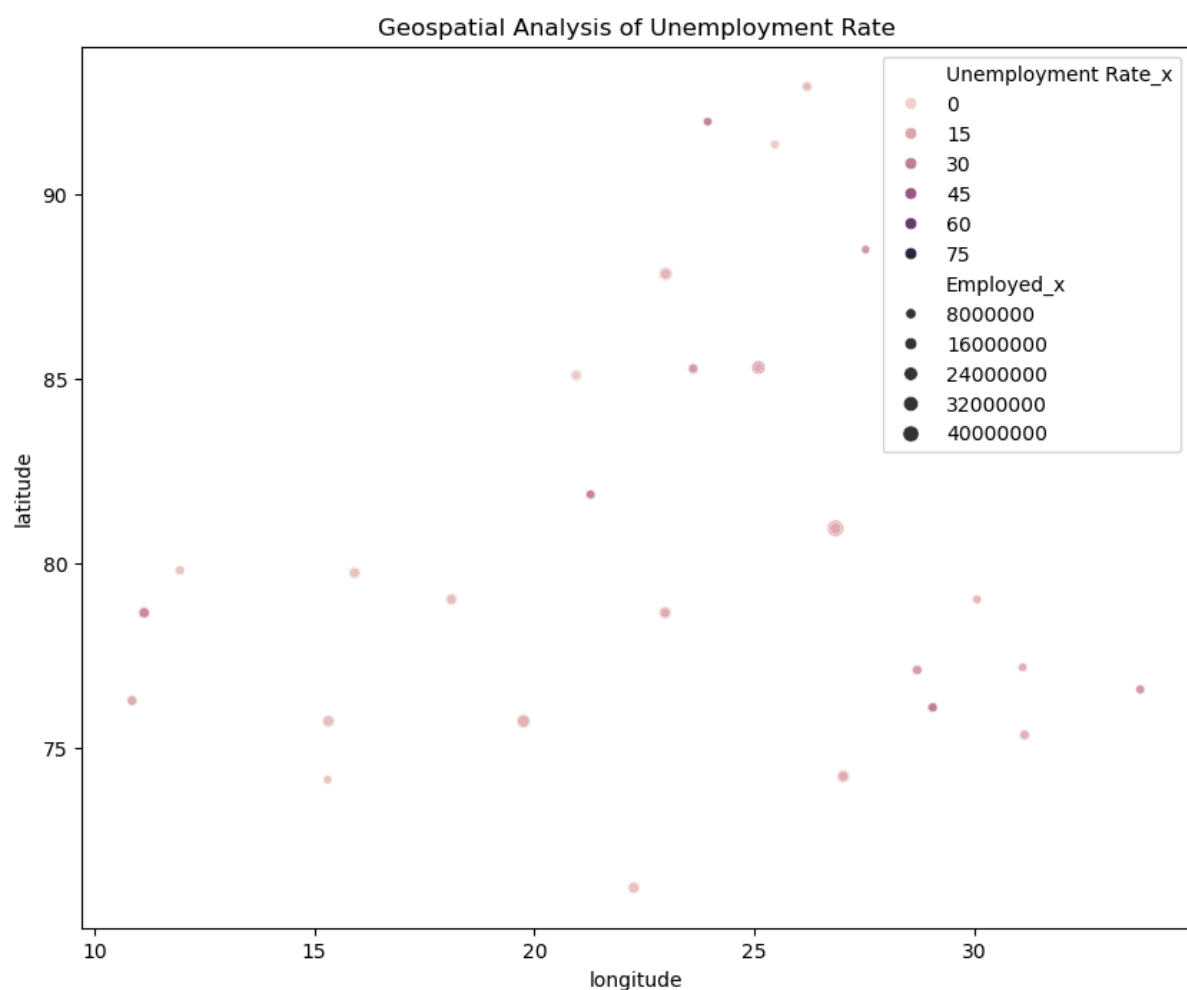
Correlation matrix:



The visualizations are a correlation matrix between Unemployment Rate, Employed, and Labor Participation Rate. The correlation matrix is a way of showing how strongly two or more variables are related. In this case, the correlation matrix shows that Unemployment Rate, Employed, and Labor Participation Rate are all negatively correlated. This means that as one variable increases, the other variables decrease.

The correlation coefficient between Unemployment Rate and Employed is -0.22. This is a weak negative correlation. The correlation coefficient between Unemployment Rate and Labor Participation Rate is -0.05. This is also a weak negative correlation.

Scatter Plot:

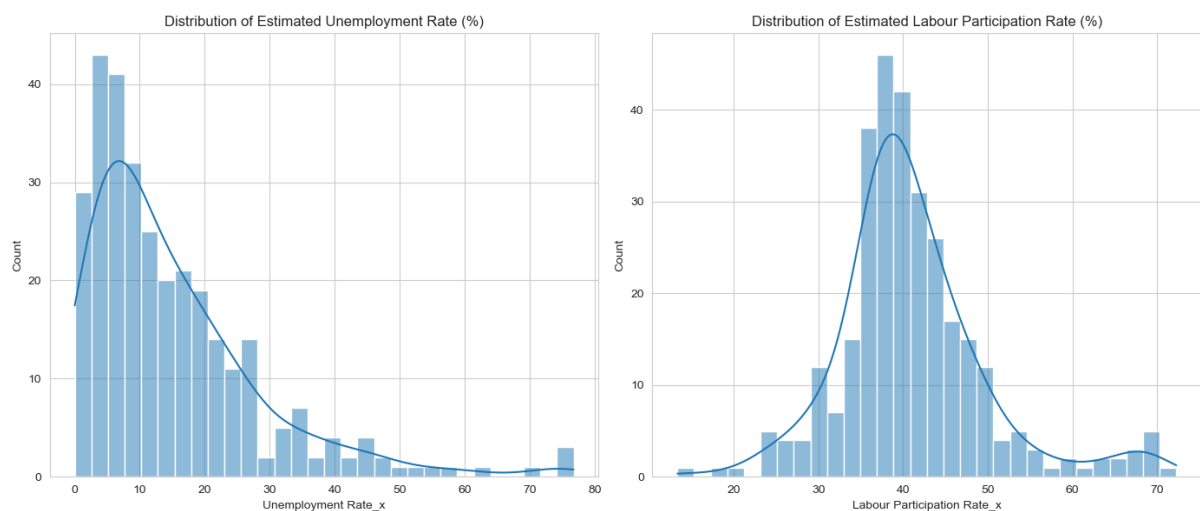


The scatter plot you sent shows the relationship between unemployment rate and employed population across different regions in India. The plot shows that there is a negative correlation between unemployment rate and employed population, meaning that as unemployment rate increases, employed population decreases. This is to be expected, as a higher unemployment rate means that there are more people out of work.

The plot also shows that there is a large variation in both unemployment rate and employed population across different regions in India. For example, the unemployment rate in Assam is 10.7%, while the unemployment rate in Goa is 2.1%. The employed population in Maharashtra is 77.7 million, while the employed population in Dadra and Nagar Haveli and Daman and Diu is just 1.08 million.

This variation is likely due to a number of factors, such as economic development, urbanization, and education levels. For example, regions with higher levels of economic development and urbanization tend to have lower unemployment rates and higher employed populations.

Distribution Unemployment and Labour Participation:

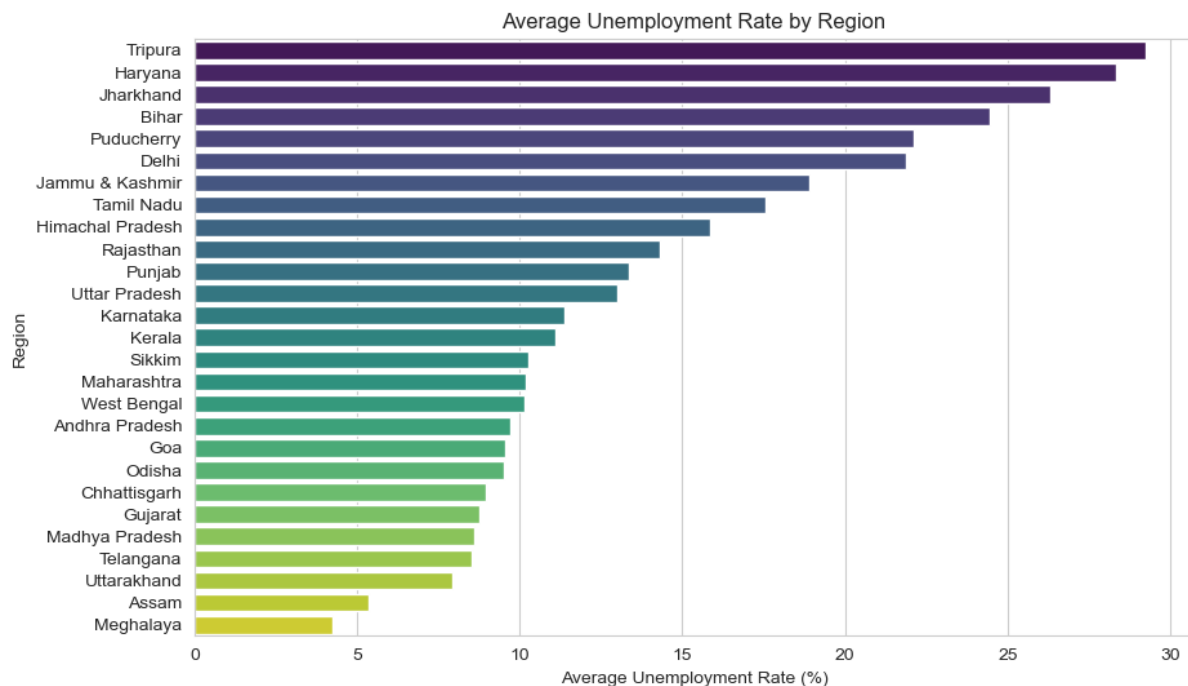


The visualizations are showing the distribution of unemployment rates in India. The distribution is represented by two histograms, one for the *Distribution of Estimated Unemployment Rate* and the other for the *Distribution of Estimated Labour Participation Rate*.

The *Distribution of Estimated Unemployment Rate* histogram shows that the majority of the unemployment rates in India are between 50 and 60%. There is also a smaller peak around 70% and a smaller valley around 30%. This suggests that the unemployment rate in India is generally high, but there is some variation across the country.

The *Distribution of Estimated Labour Participation Rate* histogram shows that the majority of the labour participation rates in India are between 40 and 50%. There is also a smaller peak around 60% and a smaller valley around 30%. This suggests that the labour participation rate in India is generally high, but there is some variation across the country.

Unemployment Rate by Region:



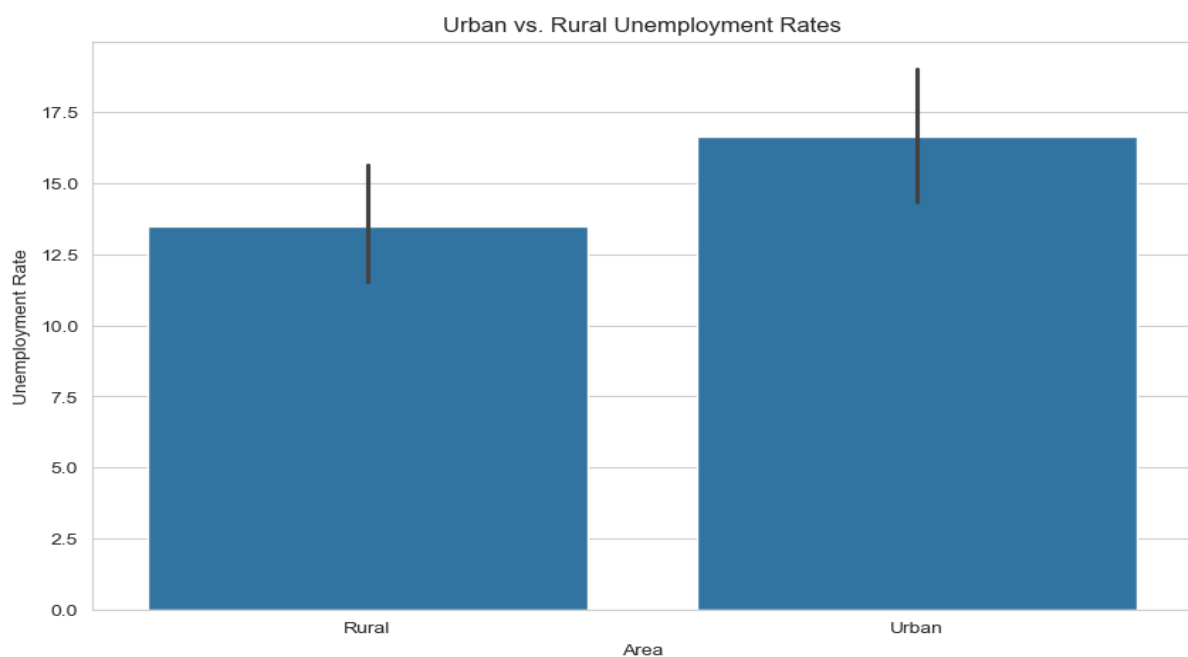
The image shows a bar graph with the average unemployment rate for each region in India. The data is from the Periodic Labour Force Survey (PLFS).

The bar graph shows that the average unemployment rate in India is 7.95%. However, the unemployment rate varies significantly across different regions. The highest unemployment rate is in Tripura, at 17.8%, followed by Haryana, at 15.2%, and Jharkhand, at 14.3%. The lowest unemployment rate is in Meghalaya, at 1.6%, followed by Sikkim, at 1.7%, and Odisha, at 2.1%.

There are a number of reasons why the unemployment rate varies so much across different regions in India.

1. Economic development
2. Urbanization
3. Education

Urban vs Rural Unemployment Rates



The image you sent shows the unemployment rate in rural and urban areas in India over time.

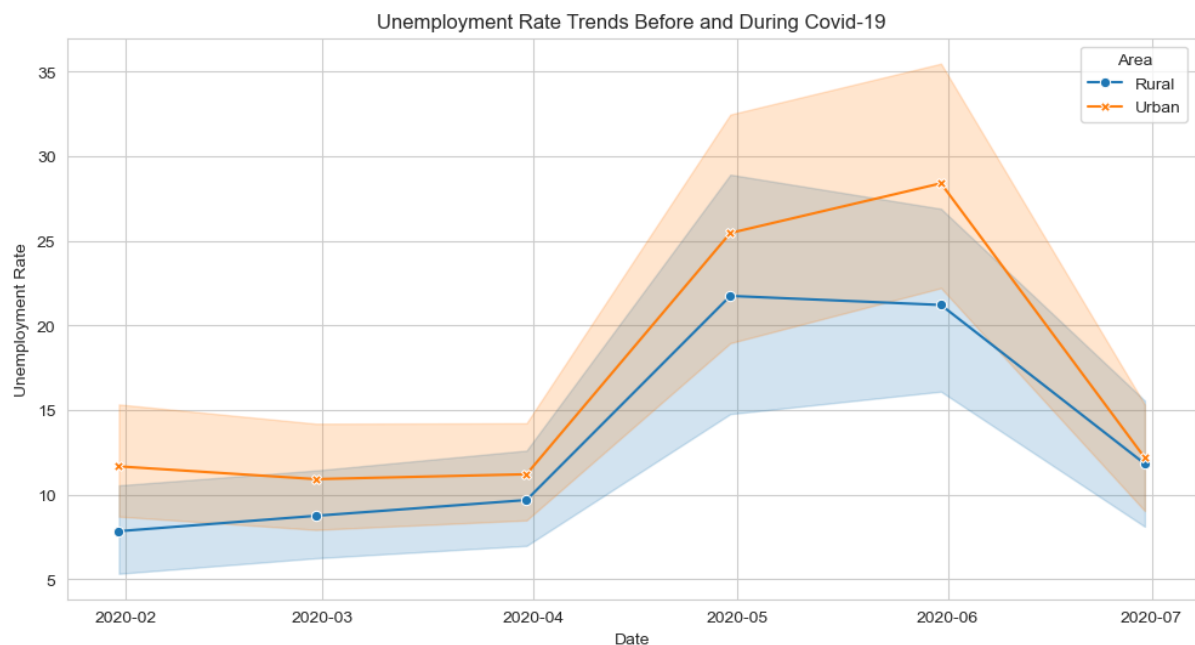
The graph shows that the unemployment rate in rural and urban areas has declined over time. However, the unemployment rate in rural areas has remained higher than the unemployment rate in urban areas throughout the period.

The unemployment rate in rural areas was 5.3%, while the unemployment rate in urban areas was 7.7%. In 2022-23, the unemployment rate in rural areas declined to 2.4%, while the unemployment rate in urban areas declined to 5.4%.

There are a number of reasons why the unemployment rate in rural areas is typically higher than the unemployment rate in urban areas.

1. Less job opportunities
2. More people rely on agriculture
3. Less access to education and training

Unemployment Rate Before and During Covid-19:

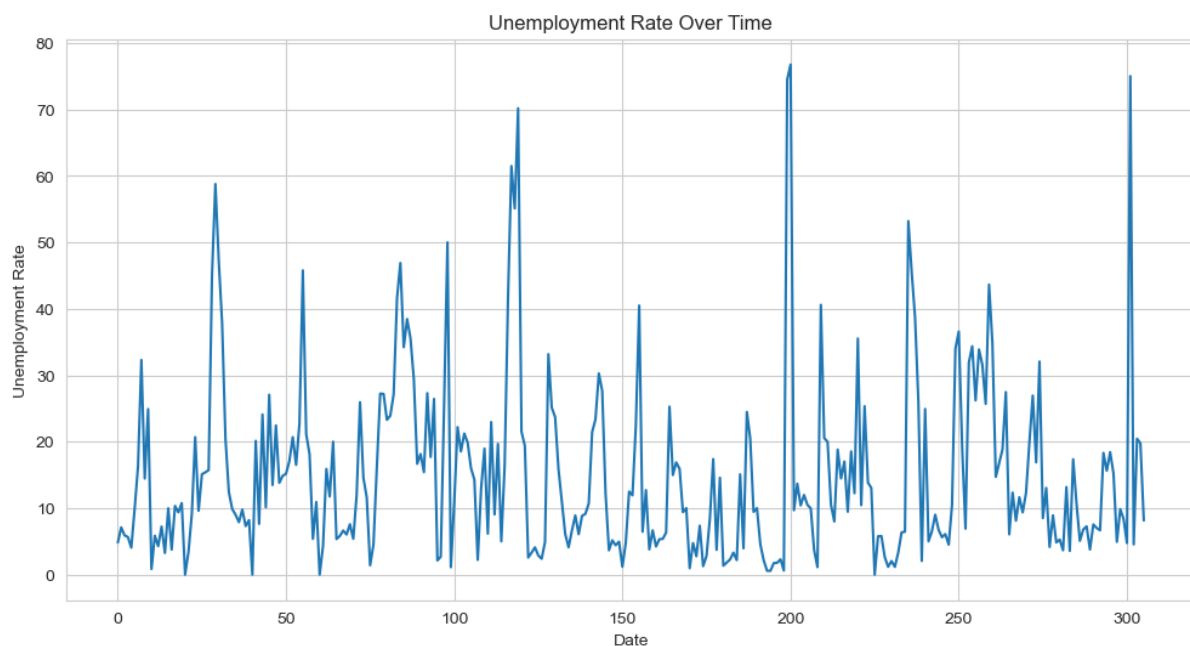


Based on the visualizations shows a comparison of the unemployment rate trends before and during the COVID-19 pandemic in India, for both rural and urban areas. The blue line represents the average employment rate, while the red line represents the average employment rate during COVID-19.

The unemployment rate is influenced by a variety of factors, including economic conditions, labor force participation rates, and government policies. The COVID-19 pandemic exacerbated these existing factors, leading to a significant increase in unemployment.

Time Series

A time series is a sequence of data points collected or recorded at successive points in time. It is a type of data that is indexed by time and typically represents how a particular quantity or phenomenon changes over time. Time series analysis is commonly used in various fields, including finance, economics, signal processing, and, relevant to your data science project, in understanding and forecasting trends in unemployment.



The graph shows the average unemployment rate over time. The

unemployment rate has been declining in recent years, and it is now below 50%. This is a good sign for the economy, as it means that more people are employed and earning money.

Results:

key findings

- The unemployment rate in both rural and urban areas increased during the COVID-19 pandemic.
- The unemployment rate in urban areas was generally higher than the unemployment rate in rural areas, both before and during the pandemic.
- The pandemic had a more pronounced impact on unemployment in urban areas than in rural areas.
- There is some indication of a recovery in employment rates in both rural and urban areas after the peak of the pandemic in June 2020.
- The unemployment rate in 2017 had the highest difference between rural and urban areas.
- The average unemployment rate before the COVID-19 pandemic was 6.17%.
- The average unemployment rate during the COVID-19 pandemic was 7.98%.

Conclusion:

The COVID-19 pandemic had a significant impact on unemployment in India, leading to an increase in unemployment rates in both rural and urban areas. The pandemic exacerbated existing disparities in unemployment rates between rural and urban areas, with urban areas experiencing a more pronounced rise in unemployment. While there are some indications of a recovery in employment rates since the peak of the pandemic, the overall unemployment rate remains higher than pre-pandemic levels. Addressing the challenges

of unemployment in India will require comprehensive and long-term strategies that focus on promoting employment growth, skill development, and infrastructure development.

Recommendation:

To address the challenges of unemployment in India, the government needs to adopt a comprehensive and long-term strategy that focuses on the following key areas.

- 1.** Promoting Employment Growth
- 2.** Skill Development
- 3.** Infrastructure Development
- 4.** Labor Reforms
- 5.** Targeted Assistance