CodeAlpha

DATA SCIENCE INTERNSHIP REPORT

## Project Title: UNEMPLOYMENT ANALYSIS WITH PYTHON

Submitted by

## Aiswarya k

Nehru Institute of Engineering and Technology

Duration: Oct 1 - Oct 31

# ACKNOWLEDGMENT

I would like to express my sincere gratitude to CodeAlpha for offering me the opportunity to work as a Data Science Intern. I am thankful to my mentors and team members for their valuable guidance and constant support throughout the project duration.

# ABSTRACT

This project aims to analyze the unemployment rate across various states in India using Python. Unemployment is a key indicator of economic health and understanding its patterns during the COVID-19 pandemic provides insights into labor market fluctuations. The project involves data cleaning, exploratory data analysis, and visualizations to interpret unemployment trends and their causes.

# INTRODUCTION

Unemployment occurs when individuals who are capable and willing to work cannot find suitable job opportunities. During the COVID-19 pandemic, India experienced a sharp increase in unemployment rates due to lockdowns and economic disruptions. This project analyzes unemployment data across Indian states to identify trends and assess the overall impact of the pandemic using Python for data analysis and visualization.

# DATASET DESCRIPTION

The dataset used in this project contains unemployment statistics for various Indian states. It provides monthly data during the COVID-19 period and includes the following columns:

| Column Name | Description |
| --- | --- |
| Region | Name of the state or union territory in India |
| Date | The observation date (monthly) |
| Frequency | Measurement frequency (Monthly) |
| Estimated Unemployment Rate (%) | Percentage of people unemployed in each state |
| Estimated Employed | Number of employed individuals |
| Estimated Labour Participation Rate (%) | Percentage of people participating in the labor force |
| Area | Classification of the data as Urban or Rural |

# EXPLORATORY DATA ANALYSIS (EDA)

Exploratory Data Analysis (EDA) was performed to understand the structure, patterns, and relationships within the dataset. The following Python libraries were used: pandas, matplotlib, and seaborn.

## Code Snippets:

```python  
import pandas as pd  
import matplotlib.pyplot as plt  
import seaborn as sns  
  
# Load dataset  
df = pd.read\_csv('Unemployment in India.csv')  
  
# Display first few rows  
print(df.head())  
  
# Check for missing values  
print(df.isnull().sum())  
```

The dataset was found to be clean with minimal missing values. Summary statistics and unique regions were identified to understand the distribution of data across states and time.

# DATA VISUALIZATION AND ANALYSIS

Various visualizations were created to analyze the unemployment trends. Line plots and bar charts were used to display how unemployment rates varied across time and regions.

```python  
# Unemployment rate by region  
plt.figure(figsize=(12,6))  
sns.barplot(x='Region', y='Estimated Unemployment Rate (%)', data=df)  
plt.xticks(rotation=90)  
plt.title('Unemployment Rate by State')  
plt.show()  
  
# Unemployment trend over time  
df['Date'] = pd.to\_datetime(df['Date'])  
plt.figure(figsize=(10,5))  
sns.lineplot(x='Date', y='Estimated Unemployment Rate (%)', data=df)  
plt.title('Unemployment Rate Over Time')  
plt.show()  
```

The analysis revealed a significant spike in unemployment during April–May 2020, coinciding with nationwide lockdowns. States like Bihar, Haryana, and Tamil Nadu experienced sharp increases in unemployment rates, while rural areas showed greater recovery in later months compared to urban regions.

# CONCLUSION

The project successfully analyzed the unemployment trends in India during the COVID-19 pandemic. The findings indicate that lockdowns had a substantial impact on employment levels, leading to a spike in unemployment rates during mid-2020. Data visualization provided valuable insights into regional and temporal variations. Future work could include predictive modeling to forecast unemployment trends.

# REFERENCES

1. Centre for Monitoring Indian Economy (CMIE) - https://unemploymentinindia.cmie.com/  
2. Python libraries: pandas, matplotlib, seaborn  
3. CodeAlpha Internship Task Documentation