Unemployment Rate Analysis in India During COVID-19

Objective

The purpose of this project was to analyze the unemployment rate in India during the COVID-19 pandemic and understand how nationwide lockdowns affected employment across different regions and time periods. The analysis aimed to identify unemployment patterns, visualize changes, and compare rural versus urban unemployment trends.

Datasets Used

Two real-world datasets were used to ensure consistent results:

Dataset 1 – "Unemployment in India.csv"
• Duration: 2019 to mid-2020
• Entries: 768 rows, 7 columns
• Key fields: Region, Date, Frequency, Estimated Unemployment Rate (%), Estimated Employed, Estimated Labour Participation Rate (%), and Area (Rural/Urban).

Dataset 2 – "Unemployment Rate up to Nov 2020.csv"
• Duration: 2020 to November 2020
• Entries: 267 rows, 9 columns
• Additional attributes: Latitude and Longitude for geographical mapping.

Both datasets revealed a significant increase in unemployment rates during the lockdown period in 2020.

Data Cleaning and Preparation

Before analysis, both datasets were thoroughly cleaned and standardized. Missing values and duplicates were removed, and date formats were standardized for time-based analysis. After cleaning, both datasets were ready for analysis.
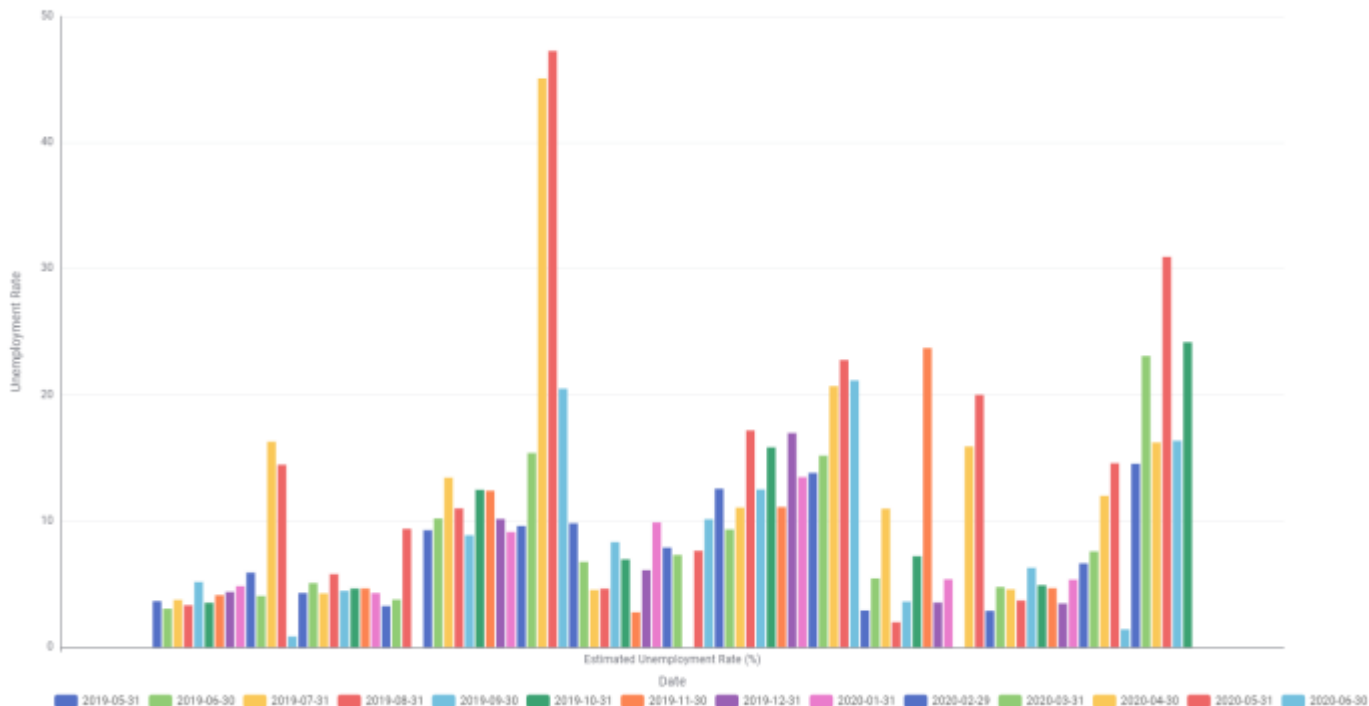
Data Visualization and Analysis

1) Overall Trend (2019–2020): The unemployment rate remained stable between 8–10% before March 2020. During the lockdown (March–May 2020), it spiked above 25% and later declined gradually as restrictions eased.

2) Extended Trend (Up to November 2020): Using the second dataset, unemployment rose sharply during lockdown and recovered toward 8–10% by late 2020.

3) Bar Chart Visualization: A bar chart compared unemployment rates across states, revealing that industrialized states like Maharashtra and Delhi were most affected, while rural states like Bihar saw smaller changes.

Unemployment Rate Trend in India During COVID-19

Unemployment Rate Trend in India During COVID-19 (Up to Nov 2020)
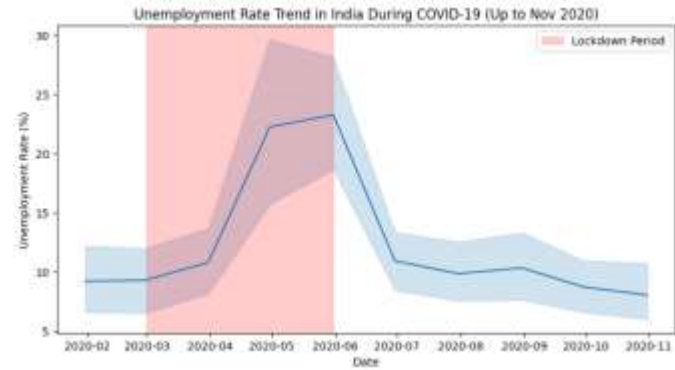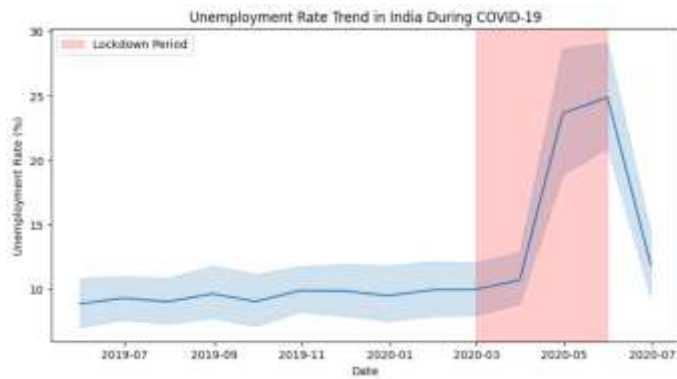
Bar Chart

KNIME Analytics Platform

Alongside Python, the project was also implemented using KNIME, a visual data analytics platform.

Steps in KNIME Workflow:
• CSV Reader Node: Imported datasets
• Missing Value Node: Handled missing data
• GroupBy Node: Aggregated unemployment by region and date
• Line Plot and Bar Chart Nodes: Created visualizations
• Joiner Node: Combined datasets for extended analysis

Benefits:
• No coding required
• Interactive and reproducible workflows
• Easy visualization generation

Insights and Conclusions

- The unemployment rate in India rose dramatically during the COVID-19 lockdown (March–May 2020).
- Urban areas faced more severe job losses than rural ones.
- A gradual recovery began after mid-2020.
- KNIME visually confirmed the same patterns observed in Python.
- Both datasets showed consistent results, confirming data accuracy.

Tools and Technologies Used

Tool / Platform	Purpose
Python (Pandas, Matplotlib, Seaborn)	Data Cleaning, Analysis, and Visualization
KNIME Analytics Platform	Visual Workflow for Data Analysis and Charting
Google Colab	Running Python Notebooks
CSV Datasets	Real-world unemployment data for India

Final Remarks

This project demonstrates how data science techniques can uncover real-world economic impacts during crises. By combining both Python programming and KNIME visual workflows, the project achieved a comprehensive understanding of how the COVID-19 pandemic affected employment in India. The integration of line plots, bar charts, and visual workflows helped reveal key insights about unemployment trends and recovery patterns.