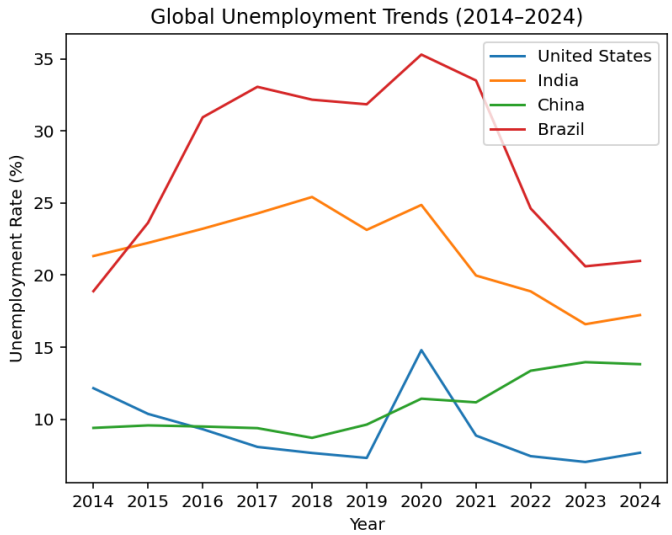
Global Unemployment Trends (2014–2024)

Student Name: Mohammed Abdul Jilani
Student Number: 24168848
Module: Applied Data Science 1 – Statistics and Trends
Instructor: Dr. William Cooper

INTRODUCTION

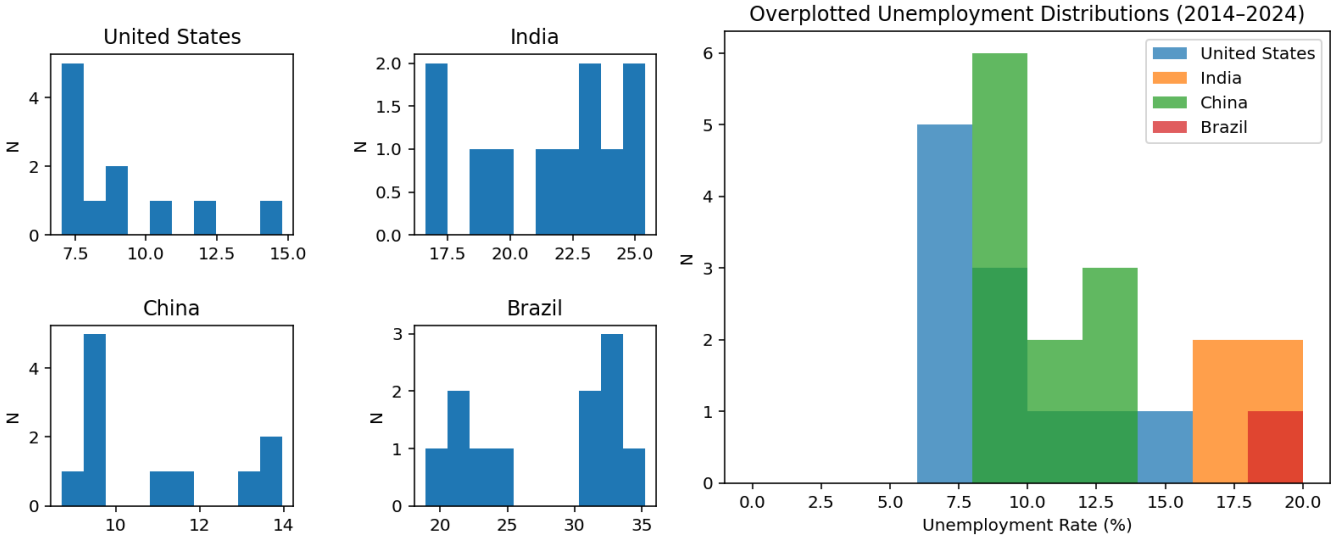
This report analyses data on unemployment rates for four countries: the United States, India, China, and Brazil; from 2014 to 2024. The aim is to explore patterns and relationships using statistical and visual techniques learned during the module. The dataset was obtained from Kaggle and prepared in Python using Jupyter Notebook. Three plots: a histogram, a line plot, and a box plot are used to explain the trends and relationships observed.

Unemployment Rate Over Time (Relational Plot)



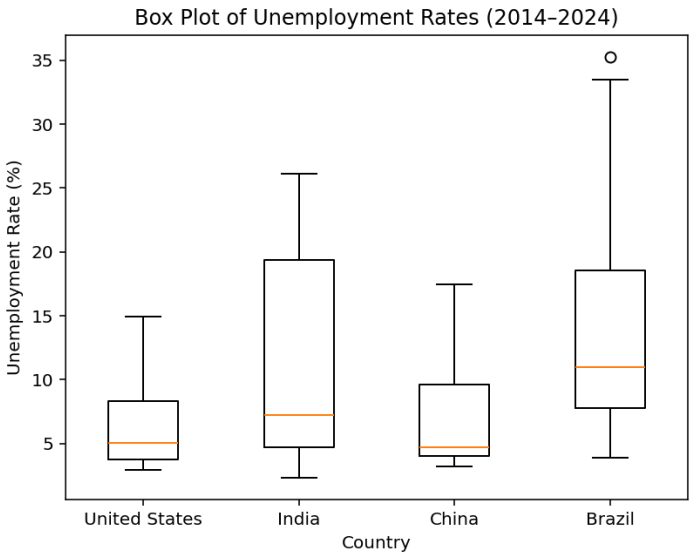
The given above line plot shows the changes in unemployment rates from 2014 to 2024. India and Brazil fluctuated more than the US and China, which remained relatively stable. All four nations display noticeable peaks during the pandemic years (2020–2021), which are followed by recovery. This illustrates how different economies are impacted by external shocks in terms of employment.

Distribution of Unemployment Rates (Categorical Plot)



The distribution of unemployment rates for the countries selected is displayed by the above histograms. Brazil and India have higher frequencies and wider ranges, whereas the United States and China group at lower rates. The data is quite asymmetric and has heavy tails, as indicated by as variance is moderate, skewness is positive, and kurtosis is high, and mean unemployment rate is around 6%. This suggests that more extreme changes in unemployment occur in developing nations.

Comparison of Unemployment Rate Distributions (Statistical Plot)



This box plot displays the changes in unemployment rates for the four nations that were selected between 2014 and 2024.With the whiskers indicating the range and the line inside displaying the median, each box represents the distr ibution of unemployment rates for a single nation.With smaller boxes and fewer outliers, the unemployment rates i n China and the US have been more consistent over time.Brazil and India contain more outlier points and bigger bo xes, indicating that countries have seen larger fluctuations and inconsistent years of high unemployment. This plot helped me compare the overall spread and consistency of unemployment rates across countries in one clear visual.

Conclusion and Reflection

In conclusion, this project helped me understand global unemployment patterns and practise using Python for data analysis. The histogram, line plot, and box plot illustrated that developed nations such as the United States and China exhibit more stable unemployment rates, whereas India and Brazil experience greater fluctuations. By calculating the mean, variance, skewness, and kurtosis helped describe the data shape and spread. Additionally, I also learned to use Git for saving and updating my code, and overall this assignment improved my confidence in data cleaning, creating graphs, and interpreting real-world statistics clearly.

References

Kaggle (2024). Global Unemployment Data. Retrieved from https://www.kaggle.com/datasets/sazidthe1/global-unemployment-data