Assignment 3: Clustering and fitting

Name: Firas Yahya Ahed Nusir ID:21083402

Key Learnings

Clustering and Insights:

- Identifying meaningful clusters uncovers global trends and disparities.
- Normalization enables fair comparisons across countries.

Modeling for Prediction and Planning:

• Curve fitting models aid in predicting future trends.

Abstract

This poster presents a clustering and modeling analysis of

diverse datasets. Clustering methods were applied to

identify meaningful clusters, while curve fitting models

provided predictions and trends exploration. The findings

highlight the importance of normalization, insights from

comparisons, and the power of visualizations in

conveying information effectively.

• Confidence ranges quantify uncertainty and inform decision-making.

Visualization and Analysis:

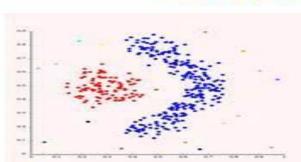
- Effective visualizations enhance understanding of complex data.
- Comparisons and trends analysis reveal insights into patterns and relationships.

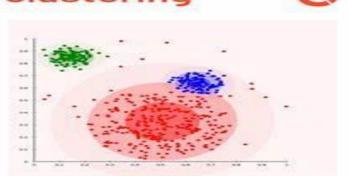
Introduction

Embark on a data-driven exploration as we delve into the realms of clustering and modeling analysis. By unlocking hidden patterns, predicting future trends, and employing visualizations, we gain a comprehensive understanding of complex datasets. Join us in unraveling global dynamics, uncovering similarities and differences, and harnessing the power of data to make informed decisions and shape our understanding of the world.

Trends analysis within and across clusters provides an opportunity to discern the evolution of indicators over time. By studying historical data and observing patterns of change, we uncover consistent trends or unexpected shifts that shape our understanding of global dynamics. These trends highlight the impact of various factors, inform our decision-making, and help us anticipate potential future scenarios.

Examples of clustering





Objective

Uncover meaningful clusters within the data, predict future trends through modeling, utilize effective visualizations, conduct comparisons among countries and clusters, analyze trends, and derive valuable insights to inform decision-making and gain a deeper understanding of global dynamics.

Clustering Analysis

To identify interesting clusters within the data, the following steps were taken:

1. Normalization

The data was normalized using appropriate techniques, such as scaling by GDP per capita, CO2 production per capita, CO2 per unit of GDP, or sector fractions.

2. Clustering

A clustering algorithm, such as K-means or hierarchical clustering, was applied to identify meaningful clusters.

3. Visualization

The clustering results were visualized using a plot. The original values were displayed, and cluster membership and cluster centers were shown using pyplot.

Model Analysis

Countries from different clusters were compared to identify similarities and differences in their indicators.

Trends within and across clusters were examined to identify similar or different patterns over time.

Aim

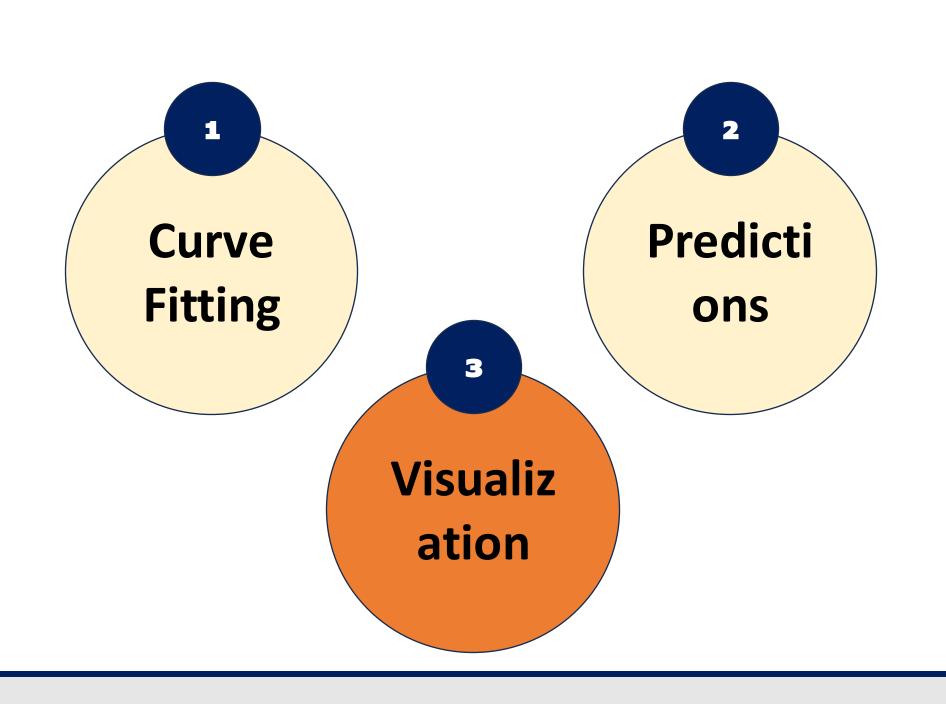
The aim of this assessment is to apply clustering and modeling techniques to diverse datasets, unravel hidden patterns, predict future trends, and gain valuable insights into global dynamics. By employing normalization, visualizations, comparisons, and trend analysis, we aim to enhance understanding, inform decision-making, and foster a comprehensive understanding of the complex interplay between various indicators and countries.

Dataset

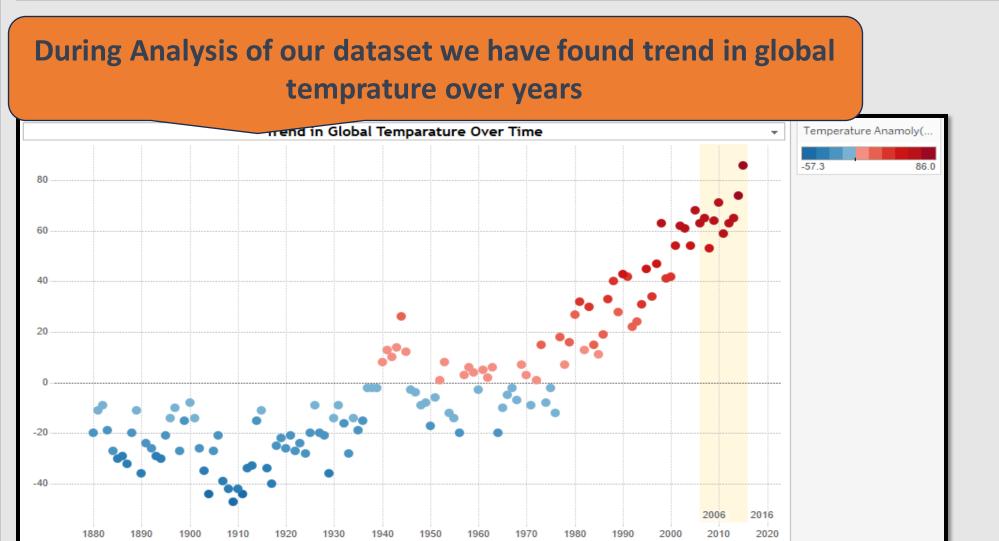
Our dataset consists of country-level data on diverse indicators, including GDP per capita, CO2 emissions, energy consumption, employment rates, and sector fractions. Spanning multiple decades, this comprehensive dataset enables comparative analysis and trend exploration, shedding light on global dynamics and policy effectiveness.

Model Analysis

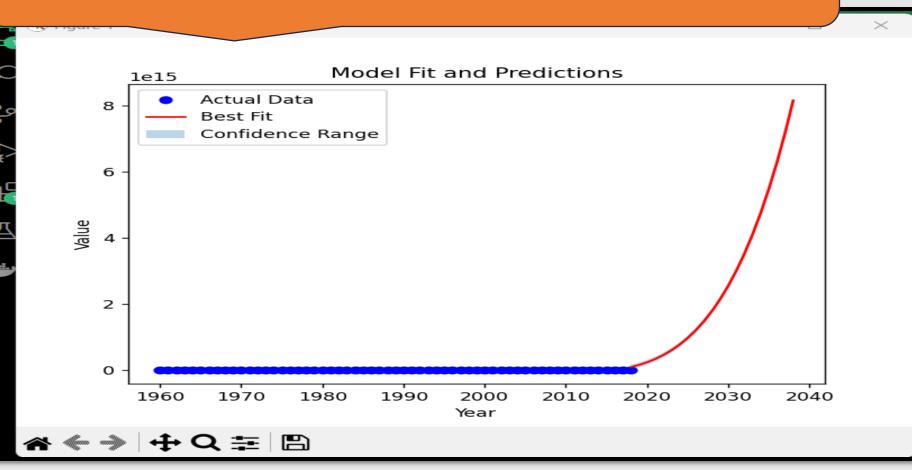
In addition to clustering, curve fitting models were applied to the datasets. The process included.



Results & Interpretations

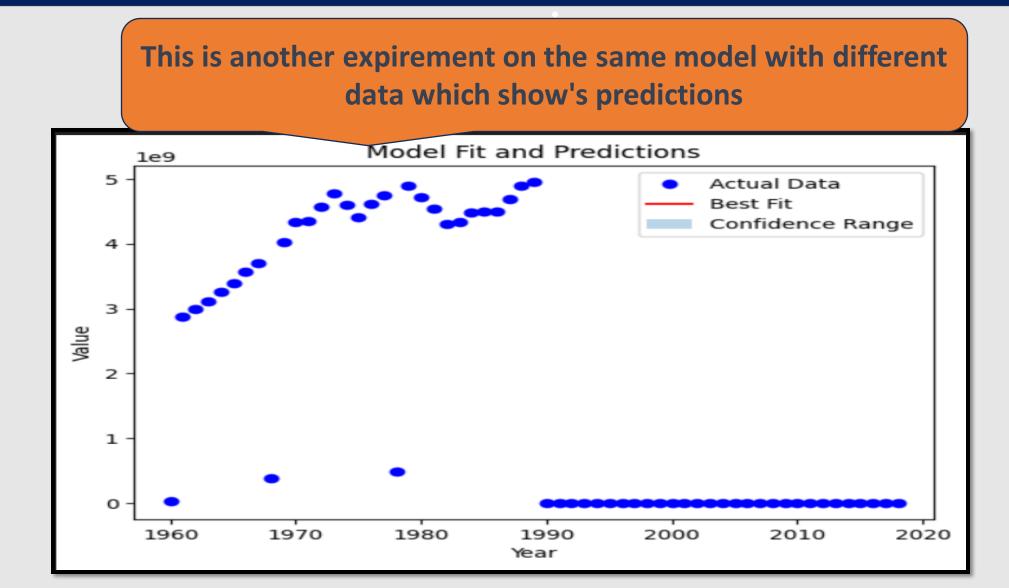


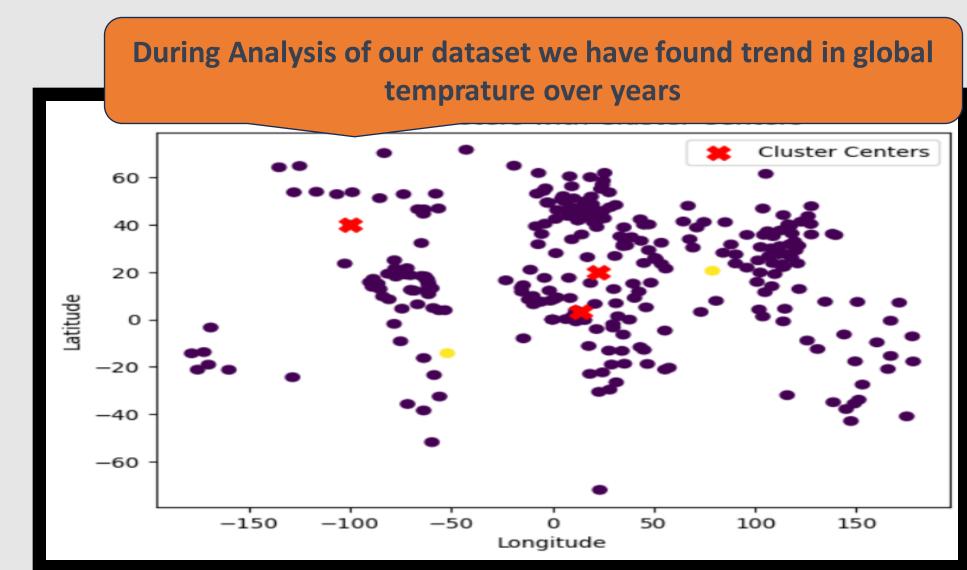




The above results perform the following:

- •Fits a model to data using curve_fit and generates predictions for future time points.
- Performs clustering and visualization using the KMeans algorithm.
 Compares selected countries within clusters based on specified variables.
- •Analyzes trends within clusters and generates plots.
- •Displays cluster distribution and proportions.
- •In summary, the code fits a model, performs clustering, compares countries within clusters, analyzes trends, and visualizes the results.



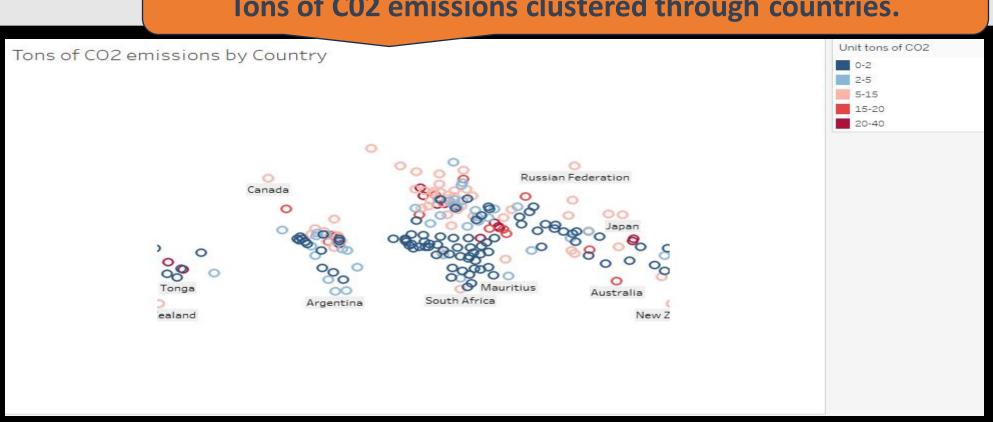


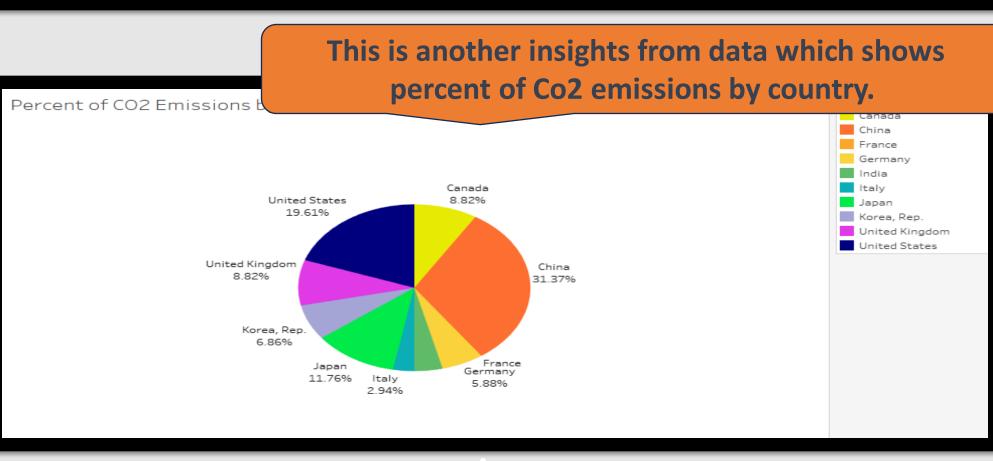
References

- World Bank. (2021). World Development Indicators.
- https://www.youtube.com/watch?v=_WnholbfcoM
 Mathlotlib Development Team (2021) Mathlotlib: \
 - Matplotlib Development Team. (2021). Matplotlib: Visualization with Python.

 Pandas Development Team. (2021). pandas: Powerful Data
- Analysis Toolkit.
 Scikit-learn Developers. (2021). Scikit-learn: Machine Learning in
- Python.
 Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. ACM Computing Surveys (CSUR), 31(3), 264-323.

Here again we have gain some insights from data and found Tons of CO2 emissions clustered through countries.





Conclusion

- Cluster analysis revealed distinct groups of countries based on climate
- change indicators, suggesting varying levels of vulnerability and risk.
 The fitting models demonstrated a clear exponential growth pattern for CO2 emissions over time.
- Comparison of countries within clusters highlighted both similarities and differences in their responses to climate change, emphasizing the need for tailored approaches to address specific challenges.
- The poster effectively communicated the findings, combining informative visualizations and concise text to present a compelling narrative on climate change and its implications for different regions.
- Overall, the analysis and visualization of the data provided valuable insights into the complex relationship between economic factors, carbon emissions, and climate change, underscoring the importance of sustainable development strategies.