# ANALYSIS AND FORECASTING OF AGRICULTURE, FORESTRY, AND FISHING IMPACT ON CO2 EMISSIONS

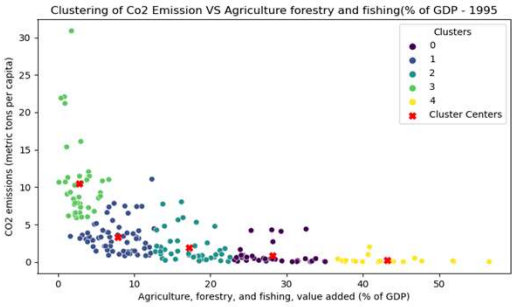

## Abstract

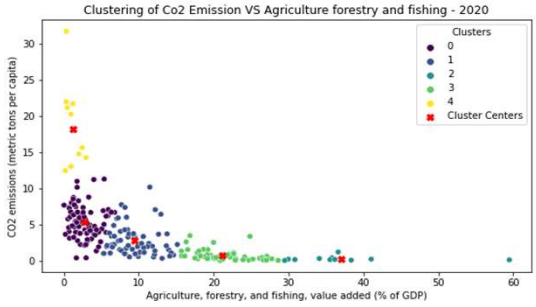
This Python script conducts a comprehensive analysis on the relationship between agriculture, forestry, and fishing value added as a percentage of GDP and CO2 emissions per capita. The code involves reading and filtering data, performing KMeans clustering on different datasets, and generating forecast plots using polynomial regression for selected countries. The visualizations include combined elbow plots for clustering, cluster plots for specific years, and forecasts for CO2 emissions up to 2025. The script provides insights into the clustering patterns and future trends related to the specified indicators.


## Introduction

This research investigates the complex relationship between a nation's per capita CO2 emissions and its economic dependence on forestry, fisheries, and agriculture. By employing resilient data analysis methods such as clustering and predictive modelling, we reveal past patterns and offer perspectives on possible future directions. The selected indicators function as surrogates for comprehending how economic activities affect the environment. By focusing on selected countries, this analysis aims to identify patterns, drivers, and actionable insights, emphasizing the pivotal role of data-driven strategies in fostering a more environmentally conscious and resilient world.

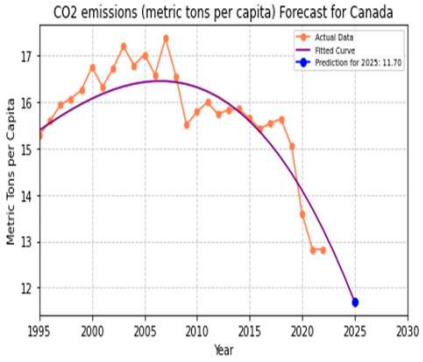
## Cluster Analysis of Environmental Indicators for 1995 and 2020



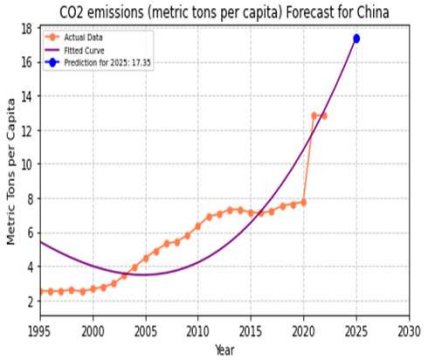




Using k-means clustering to examine the association between CO2 emissions (metric tonnes per capita) and "Agriculture, forestry, and fishing, value added (% of GDP)" over the years 1995 to 2020 finds interesting trends. By using this clustering technique, nations are grouped into clusters that have comparable characteristics for various metrics. The clustered data is then visually represented by scatter plots, which highlight the typical coordinates indicated by cluster center's. Each group's environmental dynamics can be better understood by utilizing the distinct differences within clusters, which makes focused analysis and decision-making possible. Proactive steps towards a sustainable future are based on this comprehensive understanding of nations' environmental footprints

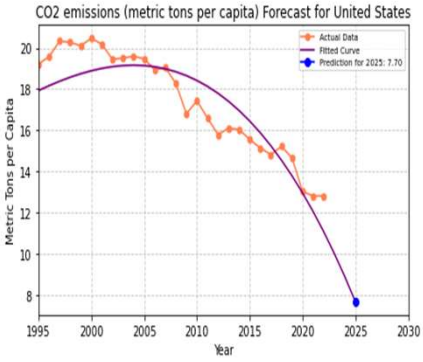
## Predictive C02 Emission forecast for year 2025




Polynomial regression analysis of Canada's projected CO2 emissions shows a consistent decrease with prediction value of 11.70. Historical data is displayed by the coral-colored markers, which show Canada's previous emissions. The underlying patterns are displayed by the fitted purple curve. The projected CO2 emissions for 2025 are indicated by the blue marker.




China's predicted CO2 emissions show a steady rising trend. This increasing tendency is projected to continue through 2025 with prediction value of 17.35. Markers with a coral color indicate historical data. The fitted purple curve illustrates the trends in China's historical emissions. The projected carbon dioxide emissions for 2025 are shown by the blue marker. The necessity of sustainable behaviors is shown by this projection.




Variations in CO2 emissions are seen in the historical trend for the United States .The forecast for 2025 indicates that emissions may decrease with prediction value of 7.70 . The historical CO2 emissions per person for the US are shown by the markers with coral colors. The polynomial regression's collected trends are depicted by the purple fitted curve. The projected CO2 emissions for the US in 2025 are shown by the blue marker.

## Conclusion

In summary, exploration of the relationship between 'Agriculture, forestry, and fishing, value added (% of GDP)' and 'CO2 emissions (metric tons per capita)' for 1995 and 2020 have given a trend. Clustering reveals distinct patterns in the joint distribution of these indicators, suggesting potential correlations. Additionally, polynomial regression forecasts future CO2 emissions based on historical data and highlights how changes in the agriculture sector may impact emissions., when the agriculture, forestry and fishing increases the Co2 emission also increased is predicted for year 2020 for countries Canada, China and United Staes. The forecasting aspect offers a glimpse into the predicted CO2 emissions for selected countries based on historical data and polynomial regression. The clustering and forecasting analyses offer valuable insights into the intricate interplay between economic activities in agriculture and environmental outcomes.

**Github link:**
https://github.com/Aswinisiva/ADSI-3-Clustering-and-Fitting

**Dataset link:**
https://data.worldbank.org/topic/climate-change


**ASWINI SIVAKUMAR**
**22027143**