Exploring the Relationship between Forest Area and CO2 Emission: A Cluster Analysis Approach

By

Marasingha Arachchilage Sanduni Sakunthala Jayathilaka

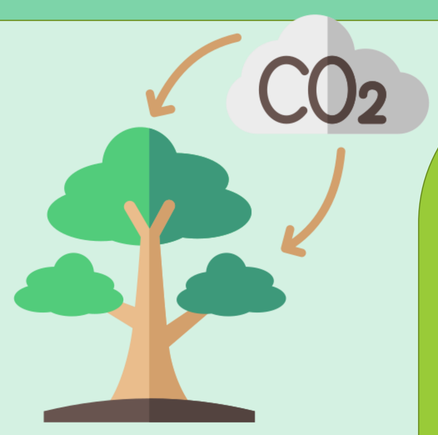
INTRODUCTION

Climate change and its impact on the environment has become an increasingly important issue in recent years. One of the main contributors to climate change is the increase in carbon dioxide (CO2) emissions, which has led to a rise in global temperatures and other environmental challenges. Forest conservation is one of the effective ways to mitigate the impact of CO2 emissions as forests absorb and store carbon dioxide from the atmosphere. In this study, we have performed a clustering analysis on two indicators - CO2 emission and forest area - to identify patterns in environmental impact across countries. Using K-means clustering algorithm, Have grouped countries with similar environmental impact into clusters. We have also selected one country from each cluster to investigate similarities and differences in their environmental impact. Additionally, compared the environmental impact of countries from different clusters and investigated trends in CO2 emissions and forest area across countries.

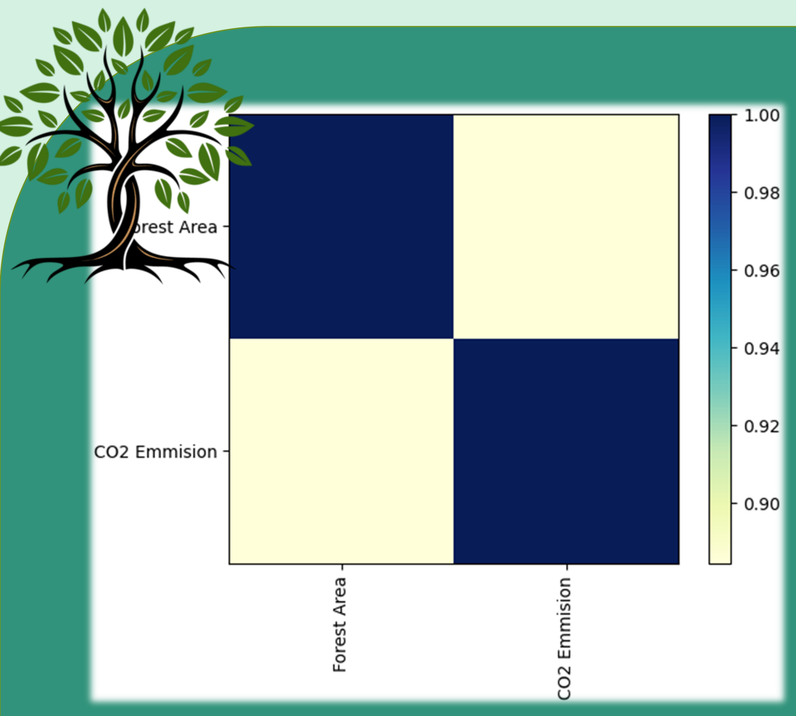
By identifying clusters of countries with similar environmental impact, we can gain insights into global trends and develop strategies to mitigate the impact of climate change. This study provides valuable information on the impact of forest conservation on CO2 emissions and identifies countries that have successfully implemented policies to mitigate their environmental impact.

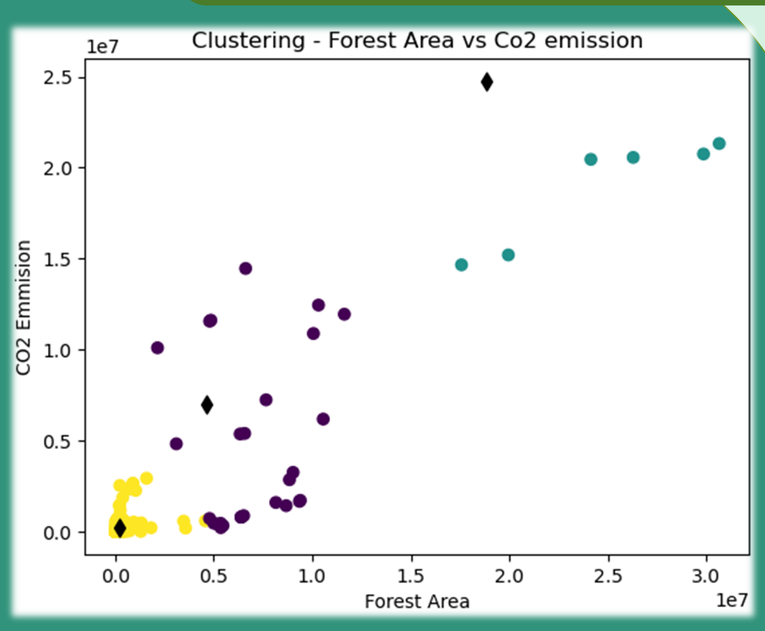
ABSTRACT

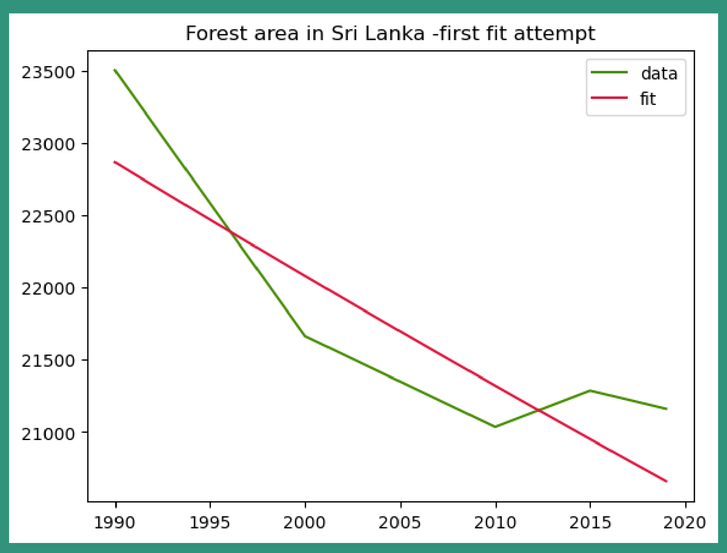
This study performed a clustering analysis on CO2 emissions and forest area to identify patterns in environmental impact across countries. K-means clustering algorithm was used to group countries with similar environmental impact into clusters. Findings reveal that countries with similar environmental impact can be clustered together based on their CO2 emission and forest area. The study highlights successful policies and areas for improvement in environmental practices and provides insights into global trends in environmental impact.

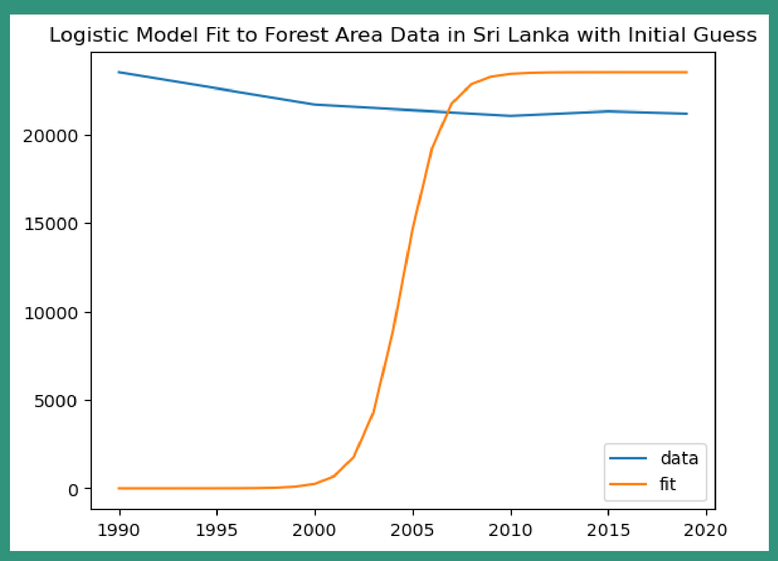


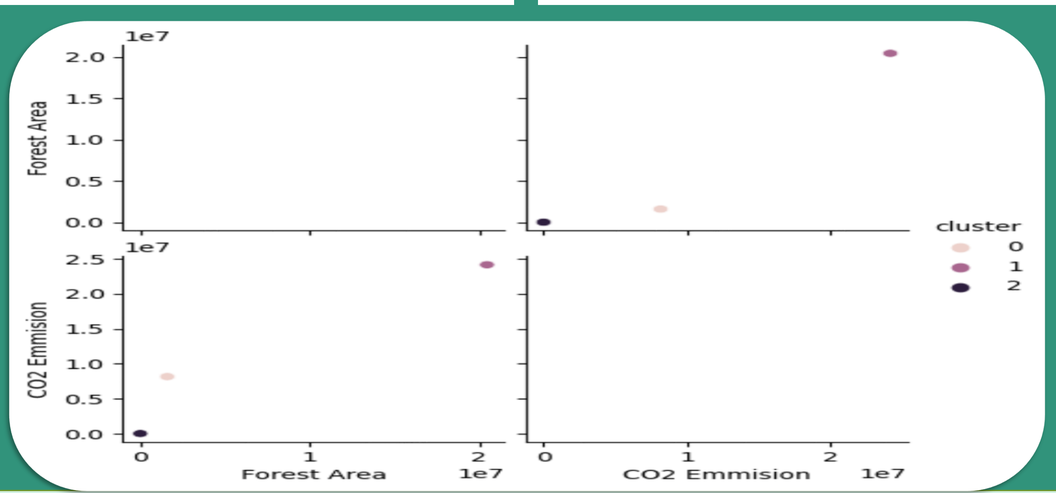
CHARTS OF RESULTS











ANALYSIS FLOW

Select two indicators from https://data.worldbank.org/topic/climate-change

↓

Create data frames from Indicators :Co2 Emission, Forest Area

↓

Merge the data frames based on the "Country Name" column

↓

Compute the correlation matrix and plot it as a heatmap.

↓

Perform K-means clustering with a range of different cluster numbers and calculate silhouette scores to determine the optimal number of clusters

↓

Select 20217 Plot the data and the for selected two indicators

↓

Use the fitted model to make predictions for the forest area in Sri Lanka in 2030, 2040, and 2050

↓

Plot the logistic function using the fitted parameters

↓

Find similarities and Differences, compare countries from different clusters

CONCLUTION

Based on the clustering analysis of countries using forest area and CO2 emission as indicators, we have identified three distinct clusters of countries with different levels of forest area and CO2 emissions. Our analysis shows that countries within the same cluster tend to have similar characteristics in terms of forest area and CO2 emissions. Moreover, our comparison of countries from different clusters has revealed significant differences in forest area and CO2 emissions. Overall, these findings emphasize the need for effective policies and initiatives to address the issue of deforestation and reduce CO2 emissions, which can have significant environmental and socio-economic impacts globally.

Github Linkfor code- https://github.com/SanduniMarasingha/Assignment-3-Clustering_and-_fitting_poster