

K-Means Clustering of Countries Based on Population Growth and Agricultural Land

Student Name: Sriram

Student ID: 22028171

Report Link : <https://github.com/sriramsonu123/ADS-Assignment-3.git>

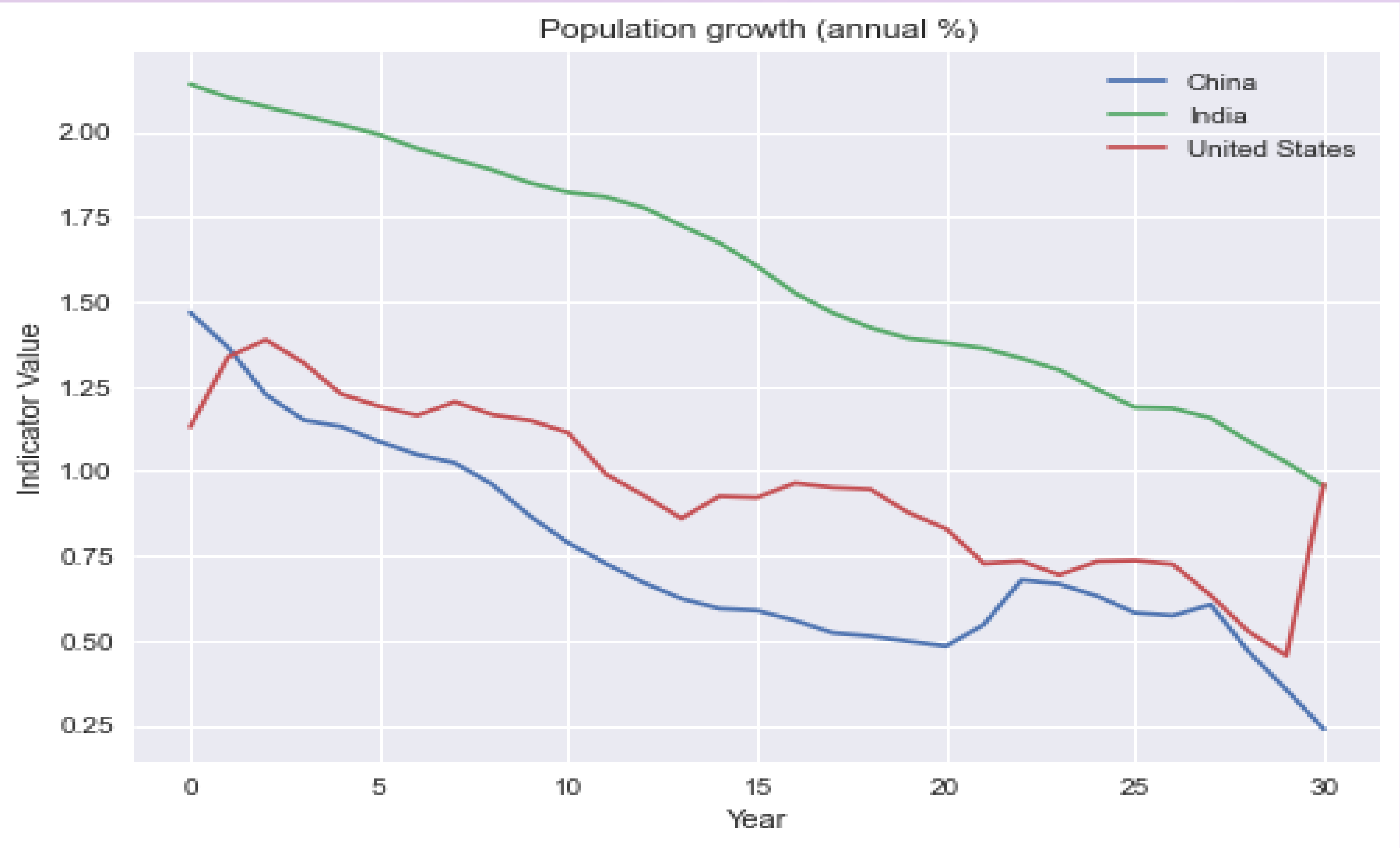
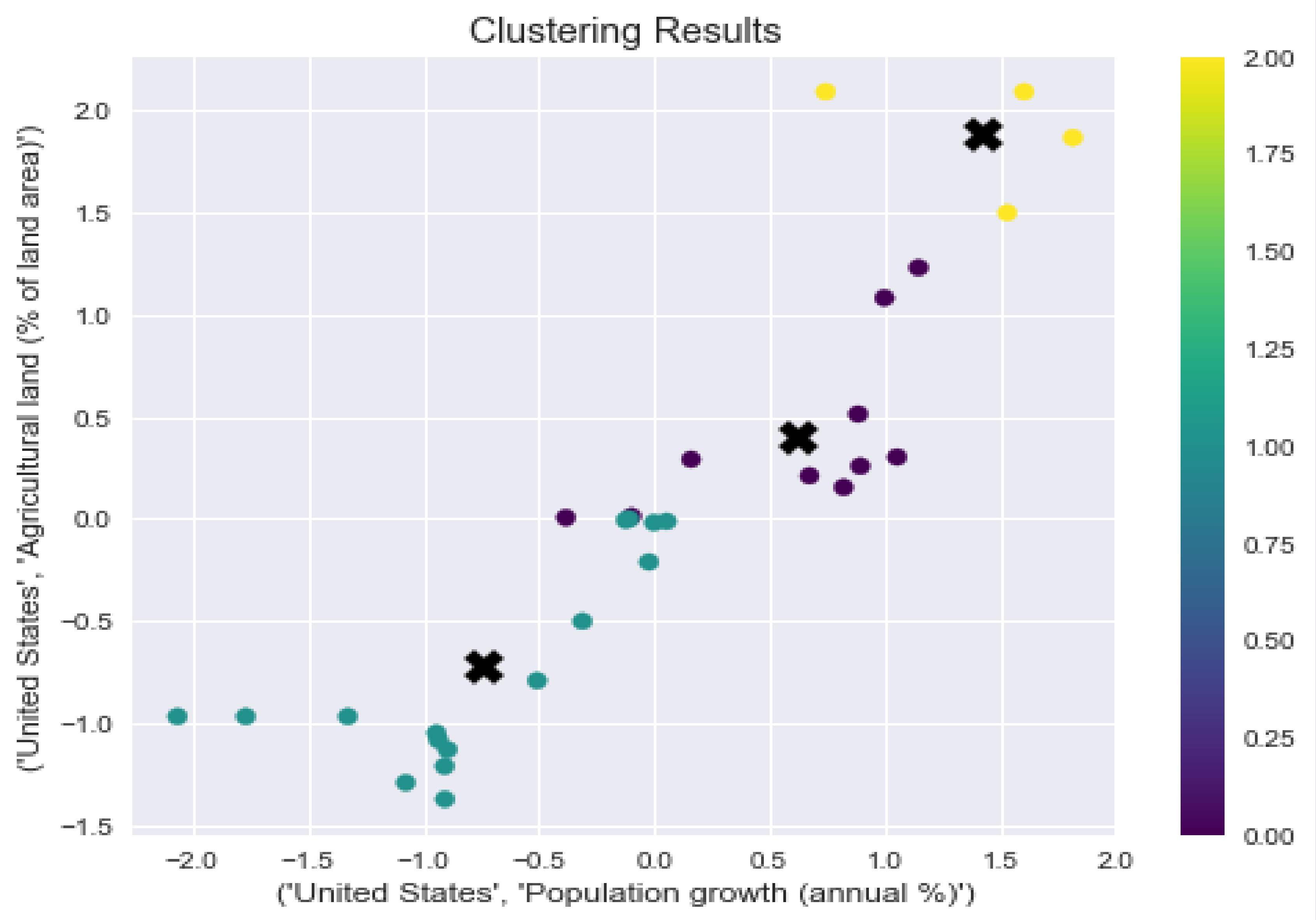
Abstract: This report uses k-means clustering to group four countries - United States, China, India, and Japan - based on two indicators: population growth and agricultural land. The results show that the countries can be clustered into two groups based on their population growth and agricultural land. Cluster 0 consists of the United States and Japan, which have lower population growth rates and lower percentages of agricultural land compared to Cluster 1, which consists of China and India.

Introduction:

K-means clustering is a popular unsupervised machine learning algorithm that is used to group similar data points together. A k-means algorithm is used to cluster countries based on two indicators: population growth and agricultural land. Specifically, clustering of four countries - United States, China, India, and Japan.

Population growth refers to the increase in the number of individuals in a population over time. Agricultural land, on the other hand, refers to the percentage of a country's land area that is used for agriculture. These two indicators provide insights into a country's development and economic activity.

Scikit-learn library in Python is used to perform the k-means clustering.



Conclusion: the k-means clustering of the four countries based on population growth and agricultural land resulted in two clusters: the United States and Japan in Cluster 0, and China and India in Cluster 1. The clustering can be interpreted in various ways, but it suggests a possible relationship between a country's level of development, population growth, and reliance on agriculture.