**Analysing Logistics Performance Index of World Bank**

**Abstract**

This study utilizes K-Means clustering to analyse the Logistics Performance Index (LPI) data provided by the World Bank. The objective is to group countries based on their logistics performance, uncovering patterns and insights that can guide policy-making and strategic improvements in global logistics. The analysis reveals distinct clusters of countries, highlighting variations in logistics efficiency and identifying potential areas for development.

**Introduction**

Logistics performance is critical for economic growth, trade facilitation, and global competitiveness. The World Bank's Logistics Performance Index (LPI) offers a comprehensive measure of logistics efficiency across countries. Applying K-Means clustering to the LPI data to identify clusters of countries with similar logistics performance, providing a data-driven approach to understanding global logistics trends.

**Methodology**

**Data Collection**

The data for this study was sourced from the World Bank's LPI dataset, which includes six key dimensions: customs, infrastructure, international shipments, logistics quality and competence, tracking and tracing, and timeliness.

**Data Preprocessing**

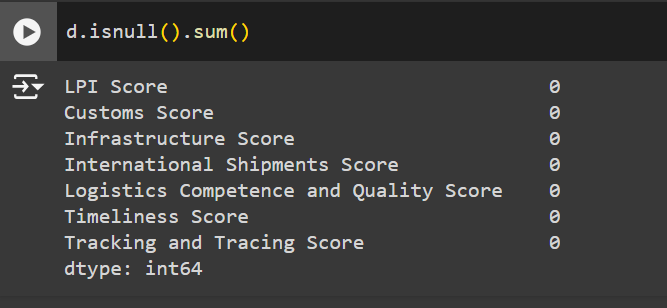
The data was pre-processed to handle missing values and normalize the features for clustering. Each LPI component was scaled to ensure equal weighting in the clustering process.

The features of the data set were found to be initially,

A screenshot of a computer

Description automatically generated

Data Pre-Processing : Removing Null values and outliers was performed.



The columns that are not required and are dependent on one more key in the dataset were removed and correlation map was drawn

A screenshot of a video game

Description automatically generated

**CORRELATION MAP:**

**A screenshot of a graph

Description automatically generated**

**It was found that ALL the features including LPI itself have a strong correlation with it.**

**FEATURES INCLUDED for the model ARE:**

A black screen with green and orange text

Description automatically generated

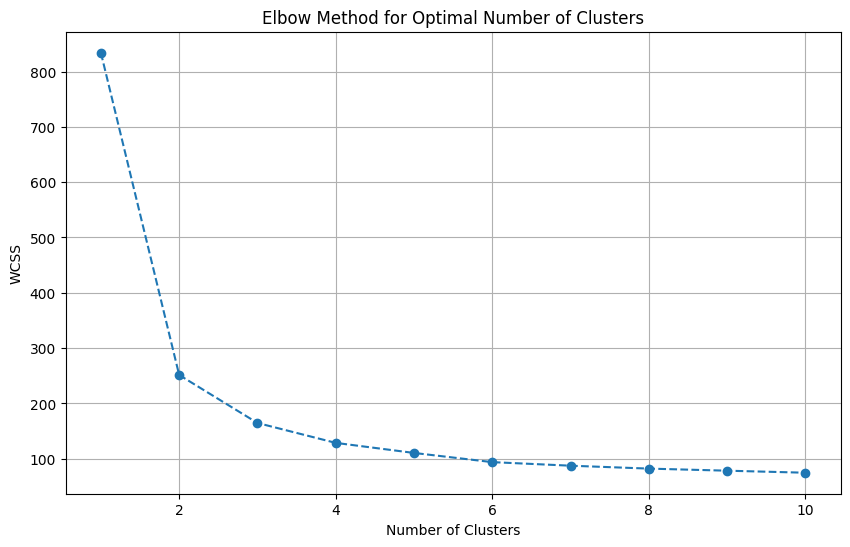
**K-Means Clustering**

K-Means clustering, a partition-based method, was chosen for its simplicity and effectiveness in handling large datasets. The algorithm partitions the dataset into k clusters by minimizing the within-cluster sum of squares.

1. **Choosing the Number of Clusters**: The optimal number of clusters (k) was determined using the Elbow Method, which involves plotting the explained variance against the number of clusters and selecting the point where the variance reduction diminishes.
2. **Implementation**: The clustering algorithm was implemented using Python’s Scikit-learn library.

**Results**

**Optimal Number of Clusters:** The Elbow Method suggested that four clusters (k=4) provided the best balance between variance explained and model simplicity.

 Figure 1: Optimal number of clusters using Elbow Method

**Cluster Analysis**

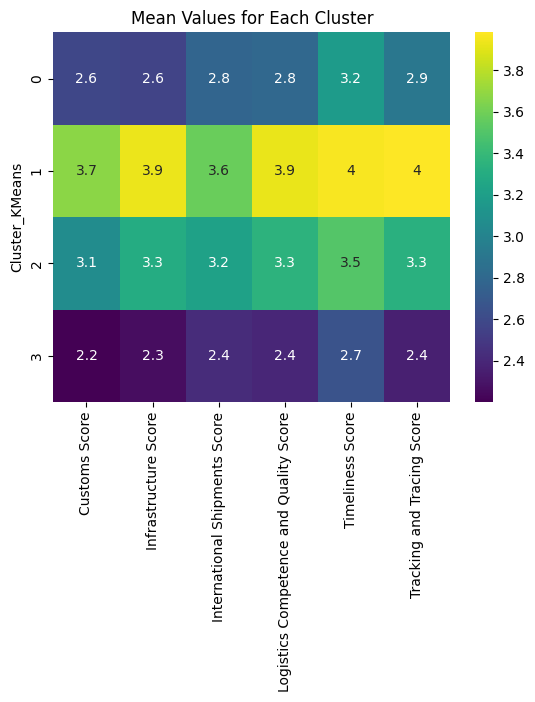
To analyse the characteristics of each cluster, the mean values of the features of each cluster is taken. Each row in the heat map represents a cluster, and each column represents a feature. The values in the table represent the average value of that feature for all countries in that cluster.

Cluster 0 has moderate scores for all features, with Timeliness Score being slightly higher than others.

Cluster 1 has the highest scores for all features, indicating that this cluster represents the best performance across all metrics.

Cluster 2 has above-average scores across all features but does not outperform Cluster 1. It indicates good performance but not the highest.

Cluster 3 has the lowest scores across all features, indicating poor performance in all areas.

Among the 139 countries, India comes under cluster 2. This clustering analysis helps identify the strengths and weaknesses of different groups based on their performance in logistics. Figure 2: Heat Map of mean values for each cluster

Countries in Cluster 0: Honduras, Chile, Indonesia, Peru, Uruguay, Antigua and Barbuda, Benin, Colombia, Costa Rica, Mexico, Namibia, Argentina, Montenegro, Rwanda, Bosnia and Herzegovina, Serbia, Sri Lanka, Bahamas, The, Belarus, Djibouti, El Salvador, Georgia, Kazakhstan, Papua New Guinea, Paraguay, Ukraine, Congo, Rep., Dominican Republic, Mali, Solomon Islands.

Countries in Cluster 1: Poland, Singapore, Malaysia, Iceland, Estonia, United Kingdom, South Africa, Norway, Italy, Greece, China, Australia, United States, Korea, Rep., Taiwan, China, Ireland, Spain, France, United Arab Emirates, Sweden, Hong Kong SAR, China, Canada, Belgium, Austria, New Zealand, Netherlands, Germany, Denmark, Finland, Japan, Israel, Switzerland, Luxembourg

Countries in Cluster 2: Bahrain, Latvia, Qatar, Thailand, India, Lithuania, Portugal, Saudi Arabia, TCrkiye, Czech Republic, Malta, Oman, Philippines, Panama, Slovenia, North Macedonia, Egypt, Arab Rep., Slovak Republic, Romania, Kuwait, Hungary, Botswana, Cyprus, Bulgaria, Brazil, Vietnam, Croatia.

Countries in Cluster 3: Guyana, Iran, Islamic Rep., Iraq, Lao PDR, Liberia, Sudan, Burkina Faso, Fiji, Gambia, The, Gabon, Haiti, Madagascar, Mauritania, Syrian Arab Republic, Venezuela, RB, Cuba, Yemen, Rep., Angola, Cameroon, Cambodia, Somalia, Kyrgyz Republic, Bolivia, Central African Republic, Trinidad and Tobago, Afghanistan, Bangladesh, Guatemala, Guinea-Bissau, Nigeria, Russian Federation, Uzbekistan, Albania, Algeria, Armenia, Zimbabwe, Bhutan, Ghana, Grenada, Guinea, Jamaica, Mauritius, Moldova, Mongolia, Nicaragua, Tajikistan, Togo, Congo, Dem. Rep., Libya.

**Conclusion**

Countries in Cluster 1, which includes Singapore, outperform India (in Cluster 2) across various economic indicators, political stability, geographical factors, and industry-specific domains. Singapore's high GDP per capita, advanced infrastructure, and strategic location as a global trade hub contribute to its superior performance in logistics and quality scores. Politically, Singapore's stable and efficient governance contrasts with India's occasional political instability and policy unpredictability. Additionally, Singapore's small, strategically located territory enables streamlined logistics and advanced port facilities, unlike India's diverse and extensive geography that presents infrastructure development challenges. Industry-wise, Singapore excels in finance, technology, biotechnology, and manufacturing, supported by robust logistics competence and quality. India, though strong in IT services, pharmaceuticals, textiles, and agriculture, lags in logistics and infrastructure quality. To bridge this gap, India must focus on simplifying customs procedures, investing in infrastructure, enhancing port efficiency, and adopting best practices from leading logistics hubs like Singapore.