

Exploring Economic Prosperity: A Clustering Analysis of Global GDP per Capita

Name: Umair
Student ID: 22098012

Introduction

Gross Domestic Product (GDP) per capita is a critical measure, widely used as an indicator of a country's economic health and its citizens' living standards. By dividing a nation's GDP by its population, it reflects the average economic output per person, offering insights into prosperity and productivity. This study embarks on a journey to explore the economic prosperity across nations using clustering analysis to categorize countries into distinct economic tiers based on their GDP per capita. Understanding these variations is crucial for policymakers, economists, and international organizations aiming to gauge economic well-being, identify growth opportunities, and address disparities.

The Dataset

The "GDP per Capita" dataset is a collection of annual economic data representing the average income per person across various countries worldwide. It is derived from the World Development Indicators, providing a comprehensive and reliable source of economic metrics. The dataset spans several decades, offering a longitudinal view of economic changes and trends. It includes a wide range of countries, from developed economies to emerging markets and developing nations, enabling a global analysis of economic prosperity.

Methodology

Our methodology is grounded in statistical analysis and data visualization techniques:

Clustering Analysis: We employ the KMeans algorithm, a well-established method in machine learning, to group countries with similar GDP per capita trajectories. This approach helps us identify clusters of nations with comparable economic characteristics.

Trend Analysis: We conduct a trend analysis to explore the economic trajectory of each cluster over time, identifying patterns of growth, stability, or decline.

References

World Bank. (2023) "World Development Indicators: GDP per capita (constant 2010 US\$)," [Online]. Available: <https://data.worldbank.org/indicator/NY.GDP.PCAP.KD>. [Accessed: 26-Dec-2023].

J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations," in Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, California, 1967, vol. 1, pp. 281–297.

J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Morgan Kaufmann, 2011.

The Project

Objective: The project aims to reveal the underlying patterns in global economic prosperity by clustering countries based on GDP per capita and analyzing their economic trends. By visualizing these patterns, we seek to offer a clear and concise representation of complex economic data.

Findings Overview:

Clustering Insights: We have identified several distinct clusters, each representing a different tier of economic prosperity.

Trend Similarities and Differences: Within clusters, some countries exhibit similar growth trajectories, while others show unique patterns, highlighting the diversity of economic paths even among similar income levels.

Economic Status Distribution: The box plots reveal the spread and central tendency of GDP per capita within each cluster, showing the variability in economic status among countries sharing similar overall characteristics.

Visuals:

The visuals include:

Cluster Distribution Histogram: Provides a quick understanding of how many countries fall into each economic cluster.

Trend Lines Plot: Offers a comparative view of GDP per capita trends for countries representing each cluster.

Box Plot: Summarizes the economic status of countries within each cluster, revealing the range of prosperity levels.

Preliminary Result and Analysis

