

Country Economic Factors for Geographic Clustering Case Study

Dataset: Country-data.csv

Context

The study of socio-economic factors is foundational to understanding and shaping the future of societies and hence of extreme interest to various government and non-government institutions. While GDP is one of the important measures used in one of the popular economic vernacular, it is not the only measure of the growth and the state of an economy. This case study aims to deep dive into one such dataset that contains various socio-economic attributes for countries around the world.

Objective

To identify if there exist various clusters of countries that are more similar to each other than others, in terms of certain socio-economic factors using KMeans, Hierarchical Clustering and DBSCAN algorithms.

Data Dictionary

The data has the following attributes:

- country: Name of the country
- child_mort: Death of children under 5 years of age per 1000 live births
- exports - Exports in % of the GDP per capita
- health - The total spend on health given as % of GDP
- imports - The value of imports given as % of GDP per capita
- income - The net income per person
- inflation - Inflation rate %
- life_expec - Average life expectancy in years
- total_fer - The fertility rate - Average children per woman in the country
- gdpp - GDP per capita