The Unseen Atlas: Discovering Policy Clusters Globally

Frances LeMond Glasser | Frances.Glasser.Data@gmail.com | LinkedIn: Frances Glasser

College of Humanities and Social Sciences | DSA 412 Exploring Machine Learning

Project Aims

Create a policy tool that generates clusters or pairings of nations based on global development indicators without direct use of geographical, political, or ethnic data. **Github Repository** →

Project Requirements

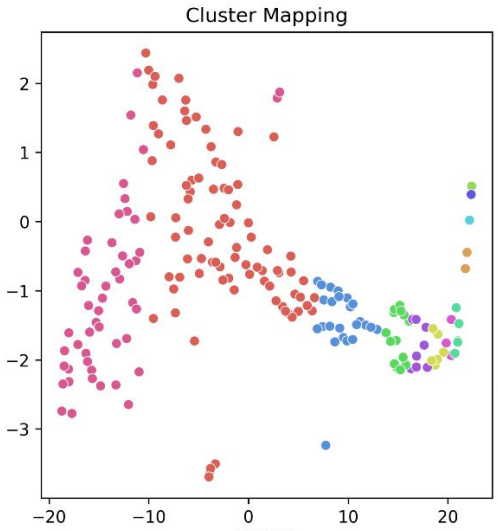
Create a machine learning model using techniques from the course and write an analysis using a dataset of your choice. Report any insights from the model.

Project Method

Data: development indicators from 2016, World Bank.org

Cleaning: manage missing data from features, standardize select features, and reduce multicollinearity.

Model: K-Means Analysis



ADAPT: Data Discoveries

Success! The K-Means model was a highly adaptable model that paired unconventional, authentic case studies:

1. Singapore & Switzerland
2. Iceland & Seychelles
3. Brazil, Indonesia, Saudi Arabia, and Viet Nam.

ADAPT: Data Perspectives

World Bank data was not accurate. The data is self-reported by the nation and there is a lack of data strategy. This project brought my attention the complex nature of trying to record accurate statistics while protecting participant's dignity and global data governance.