Freedom, Democracy, and Well-Being: A Comparative Analysis of Global Progress Indexes Using K-Means Clustering

Maxim Farkhat

IPHS300 AI for the Humanities (Spring 2022) Prof Elkins and Chun, Kenyon College

IPHS 300: AI for the Humanities
Spring 2022
Professors Elkins & Chun

Introduction

There are numerous indexes that currently measure a country's success through different perspectives and ideologies. One of the most accepted ways of measuring global success and progress is through a country's Gross Domestic Product (GDP) and economic growth, which is in accordance with the current neoliberal economic doctrine that is dominating the world.[1]

Many scholars, most notably the Nobel prize-winning economist Amartya Sen, believe that when looking at development, GDP is an important means for expanding freedom. However, Sen believes that GDP is not enough. Development, growth, and success should be looked through the lens of expanding human freedom, as well as economic. Freedom is also an important pillar of democracy, which is another critical factor in determining success and growth.[2]

Some scholars argue that we should look even further than freedom. Nick Meynen, an environmentalist journalist from Belgium, believes that we should use different forms of progress indicators that take into account environmental sustainability and human well-being. He believes that the growth mindset with GDP is unsustainable and the world should move towards a "degrowth" economy that puts less emphasis on economic growth and more emphasis on freedom, sustainability, and well-being.[3]

Thesis

The purpose of this project is to do a comparative analysis of different indexes that measure a country's success.

This will be achieved by using Good Old-Fashioned Artificial Intelligence (GOFAI) in the form of clustering to find patterns amongst all the data. The project will look at indexes that measure economic freedom, human freedom, democracy, and alternative measures like the ones described by Meynen.

My **hypothesis** is that the clustering will show a stark difference between the indexes that measure economic freedom and those that look at alternative measures like sustainability and well-being.

Acknowledgements

I'm very grateful to Professor Chun and Elkins for helping with this project and teaching me about AI in the humanities. I'm also grateful to CSAD, particularly Professor Powers, for teaching me and helping me with the indexes for this project.

Materials

This project analyzes **169** countries from the years **2013-2019** using **16** different indexes that were broken down into the following categories:

1. **Economic Freedom** (6 indexes)
2. **Human Freedom** (3 indexes)
3. **Democracy** (4 indexes)
4. **Well-being Alternatives** (3 indexes)

The **Economic Freedom** category includes indexes that measure economic, business, monetary, investment, and financial freedoms from the Fraser Institute[4] and Heritage Foundation.[5]

The **Human Freedom** category includes indexes that measure general freedom from Freedom House,[6] as well as personal and human freedom from the Cato Institute.[7]

The **Democracy** category includes indexes that measure political rights and civil liberties from Freedom House and two general democracy indexes from the Economist[8] and V-Dem Institute.[9]

The **Well-being Alternatives** category includes indexes that measure income inequality based on the Gini coefficient[10] and human development from the United Nations Development Programme, which is inspired by Amartya Sen's work.[11] The category also includes a Happy Planet Index that measures sustainable well-being based on a country's well-being, life expectancy, and ecological footprint.[12]

The **year range** and **number of countries** are based on when and where the indexes intersected. The database doesn't include data from contested territories like Transnistria. The only exception is Hong Kong and Taiwan because out of all the contested territories, they're the most "independent" in the sense that they differ from China in their governance, economies, and culture.

Methodology

Prior to clustering, the project required compiling all the indexes together. To make sure that there was minimal amount of missing data, the year range and number of countries were reduced. Following that reduction, I had to impute the missing data by correlating the indexes in each category. For example, for Democracy, I found the average and median Economist Democracy Index score for each score in the V-Dem Democracy Index (1-5) and used the averages to fill in the missing data. If a country was missing a lot of data in a category, I would use data from a matching country based on their similarities in governance, economy, culture, and geography.

Following the imputation, I ran the data through a k-means clustering algorithm using the PyCaret library. PyCaret is an open-source, low-code machine learning library in Python that was made with the purpose of analyzing citizen data science, which makes it perfect for this project.

Results

Fig. 1: K-means clustering of all the data

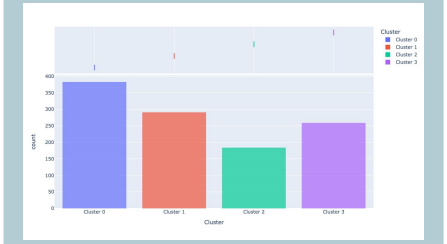

Fig. 2: Distribution plot of each cluster

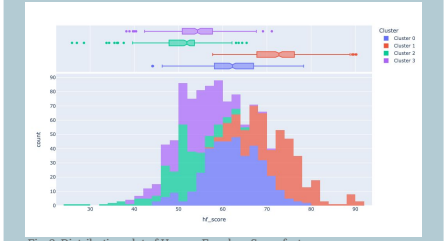

Fig. 3: Distribution plot of Human Freedom Score feature

Discussion

One of the hyperparameters for the unsupervised clustering was the number of clusters. After experiments with multiple amounts, four clusters was the ideal amount because it produced the least overlap and had clear limits as **Figure 1** shows. **Figure 2** shows the distribution of each cluster and the distribution is very even. A perfect distribution rarely happens. **Figure 3** shows the distribution of one of the features, Human Freedom Score. The plot shows a near perfect Gaussian distribution, which displays the range of data that is available in that particular feature. A majority of the features in this dataset display a Gaussian distribution.

After looking at the different clusters, I was able to categorize them into the following categories going from **High Well-being (1) to Low Well-Being (4)**:

- **1** = Cluster 1
- **2** = Cluster 0
- **3** = Cluster 2
- **4** = Cluster 3

The categorization was based purely on my personal analysis based on my expertise as an International Studies major. I looked at the countries in each category and I looked at whether the countries had high or low well-being based on my knowledge of the country's governance, culture, and geography. This is not a perfect categorization and can be up for debate.

What this clustering displays is that the alternative indexes many scholars, like Nick Meynen and Amartya Sen, are calling for us to use are able to differentiate between different countries. The clustering was able to categorically organize the data from a high to low well-being using a variety of indexes. What is still required is to compare the performance of the model on economic indexes versus alternative ones. Either way, this shows that globally we are ready to move beyond only using economic measures of success and growth and incorporate alternative indexes as well.

Bibliography

[1] Nick Meynen, *Frontlines: Stories of Global Environmental Justice*, 255.
[2] Amartya Sen, *Development as Freedom*, 3.
[3] Nick Meynen, 211.
[4] Fraser Institute, *Economic Freedom of the World Reports*, 2014-2020.
[5] The Heritage Foundation, *Index of Economic Freedom Reports*, 2014-2020.
[6] Freedom House, *Freedom in the World Reports*, 2014-2020.
[7] Cato Institute, *The Human Freedom Index Reports*, 2014-2020.
[8] The Economist Intelligence Unit, *Democracy Index Reports*, 2014-2020.
[9] Institute of Democracy and Electoral Assistance, *The Global State of Democracy Reports*, 2014-2020.
[10] Gapminder, *Gini Coefficient Index*, 2014-2020.
[11] United Nations Development Programme, The Human Development Index, 2014-2020.
[12] Happy Planet Index, The Happy Planet Index Reports, 2014-2020.