Annabel Keppel-Palmer | Github : bit.ly/akeppelpalmer | Notebook : bit.ly/AKPCapstone
Pace University, Seidenberg School of Computer Science and Information Systems | Advisor: Christelle Scharff

Measuring the attitudes of governments towards their female population :

A cluster analysis on discriminatory policies

PACE UNIVERSITY

Abstract

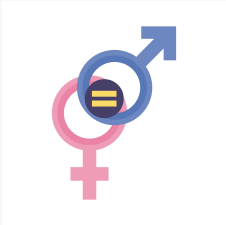
Gender discrimination is one of the biggest problems that we still face in the 21st century. In this research, we used clustering machine learning techniques to analyze discriminatory laws and regulations across 179 different countries. Using the K-Means clustering algorithm, we were able to group countries into one of three clusters to show which countries share similar levels of gender equality and their attitudes towards the equal treatment of their citizens.

Research Questions

1) By assessing discriminatory regulations, can we state which governments are actively working to lessen the gender gap?
2) Which governments/countries have the best policies in place to protect and empower their female citizens?

Related Work

Existing research has shown that the gender gap is decreasing, however at a very slow rate [2]. Many academic papers solely focus on the gender pay gap, rather than discrimination, which this project focuses on. From the research studied, methodology included econometric, regression and clustering algorithms [1][3][6]. Combinations of these models have been the most effective at helping to group and rank together countries with similar outlooks on the wellbeing of their female citizens.



Dataset

The dataset used in this study was extracted from *The Gender, Institutions and Development Database (GID-DB)*. This dataset contains 180 rows (countries) and 35 columns (containing the SIGI score for laws, practices and attitudes of each country). The values in the dataset contain values using the Social Institutions and Gender Index (SIGI) to assign scores for each category. In general these values fall between 0 and 1, where the higher to score, the more discrimination there is in the country.

Methodology

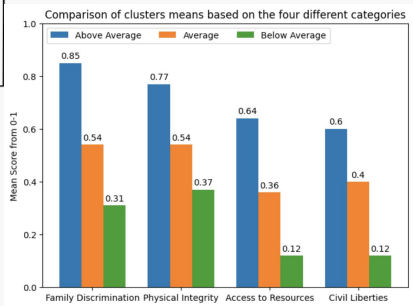
K-means clustering is one of the most applied unsupervised learning models to use when you do not have a specific outcome to predict. In this study, K-Means was used to find patterns and observe similarities between data points, which allowed us to group together (cluster) similar countries.

Below are the results of the least discriminatory countries in each category:

- **Discrimination in the Family** : Switzerland (Mean of 0.20)
- **Restricted Physical Integrity** : Austria, Croatia and Sweden (Mean of 0.3125)
- **Restricted Access to Productive and Financial Resources** : France, Malta, Luxembourg, Portugal, Sweden, Netherlands, Finland, Norway, Denmark, Peru, Belgium, Austria (Mean of 0.20)
- **Restricted Civil Liberties** : 20 Countries (Mean of 0.00)

Results

From the clustering model formed, we have been able to cluster the data into 3 groups for each category. Below is a graph summarizing the mean scores of each cluster:



Conclusion and Future Work

Overall, there is still a lot to be improved when it comes to gender equality. It's surprising to see that even some of the most powerful and economically advanced countries (United States, United Kingdom, Germany) were not amongst the top performing selection in their clusters.

To improve this study, I would like to take the 'practice' column and compare this to the law. This would give us a greater representation of how the countries are actually implementing their laws in action, and we can assess to see if this drastically changes the results.

References

1. H. Alatrista-Salas, B. Esposito, M. Nunez-del-Prado, M. Valdivieso (2017)
2. K. Gupta, R. Rani, A. Sharma, P. Bansal, A. Dev, R. Gandhi (2023)
3. Strittmatter A, Wunsch C (2021)
4. Michael Danquah, Abdul Malik Iddrisu, Ernest Owusu Boakye, Solomon Owusu (2021)
5. Jaanika Meriküll, Merike Kukk,Tairi Rõõm (2020)
6. Albanesi S, Şahin A (2018)
7. Hugo Alatrista-Salas, Pilar Hidalgo-Leon, Miguel Nunez-del-Prado (2018)
8. Andrey Shastri (2014)