

## Report for Project 2 - Clustering Countries

In many ways, the findings from this project are relatively unsurprising. For the most part (and with the knowledge I have about geography and culture), many similar countries ended up in the same clusters.

For example, one cluster contains the United States, the UK, and many other countries which I would consider to be first world, developed countries with a capitalist economic system and democratic governments. This cluster is characterized by a low PDI and a high IDV, which makes sense according to the characteristics of these countries.

Another cluster is made up of countries such as China and Bangladesh, countries with highly concentrated political power and less economic freedom. Unsurprisingly, this cluster is characterized by the highest value of PDI, and relatively average values for all other variables.

Another cluster comprises mostly Latin American countries, and what is notable about this cluster is that the mean value of UAI is higher than the other clusters. Given the often turbulent pasts relating to regime changes and economic crises in these countries, a high intolerance for uncertainty and change makes sense for many of these countries.

Besides these 3 clusters, the next 2 clusters were somewhat surprising to me. The first of these surprising clusters comprises a seemingly random mix of countries (Colombia, El Salvador, Ghana, Mexico, Nigeria, Trinidad and Tobago, and Venezuela), with the notable stat being a very high level of IVR, something I would not have predicted for these countries. The last cluster is also a seemingly random mix of countries, with no real notable high or low values of any statistic. This makes me believe this cluster would be considered an “other” category, for countries that did not really fit into any cluster with our level of K.

Overall, I was impressed with how relatively accurate and sensible the clusters were, especially given the relatively small amount of data used in this project. I would be curious to see an extension of this project with more data points and perhaps with less subjective measurements. For example, clustering countries according to the already existing values, but also adding economic measures such as GDP per capita would be fascinating.