2. Based on this graph in particular, a good choice of $K$ would be around 5 to 10, depending on your application. We first used the elbow method to approximate where the “best” $K$ would be, which is around 6. However, for our application, we think that increasing $K$ to 10 makes sense, as there are 196 countries and we want our clusters to be too big. After about 10, the drop off in cost is only marginal compared to how many clusters you end up with.

3. Below is the division of countries created after running k-means with $K = 10$. It seems like the clusters are not very well balanced. Some have a lot of countries and other have only a few. The countries in each of the sections seem like they could definitely be related based on the features given. All of the features seem to be related to freedoms or elements of a modern country, so the pairings below make sense. You can see how the
countries in the clusters might be related based on their cluster centers. Looking at the cluster centers against the average over all countries, cluster 0 for example seems to have lower education on average but higher trustworthy governments. Clusters, averages over all countries and per clusters given below.


1['Dominican Republic', 'Haiti', 'Lesotho', 'Liberia', 'Nauru', 'Zambia']


5['Austria', 'Belgium', 'Canada', 'Denmark', 'Finland', 'France', 'Germany', 'Luxembourg', 'Netherlands', 'New Zealand', 'Norway', 'Sweden', 'Switzerland']


8['Andorra', 'Antigua and Barbuda', 'Cameroon', 'China', 'Iran (Islamic Republic of)', 'Lao People’s Democratic Republic', 'San Marino', 'Sri Lanka']

9['Australia', 'Eritrea', 'Hungary', 'Iceland', 'Ireland', 'Israel', 'Italy', 'Liechtenstein', 'Malta', 'Palau', 'Singapore', 'United Kingdom of Great Britain and Northern Ireland', 'United States of America', 'Uruguay']

overall average: [ 0.04429545 0.04036232 0.04403623 0.07179353 0.09368864 0.11663710 0.07123337 0.05723337 0.07686862 0.04129617 0.05614306]

mu = [array([ 0.04841273, 0.0265020 , 0.0357975 , 0.07223124, 0.09599801, 0.08991681, 0.10703439, 0.03014306, 0.03380248, 0.07688485, 0.07944997, 0.09927951, 0.03557497, 0.0661774, 0.03783992, 0.08464345 0.03522813 0.03197616 0.06964062 0.05723337 0.07686862 0.04129617 0.05614306])]

[0.04429545 0.04036232 0.04403623 0.07179353 0.09368864 0.11663710 0.07123337 0.05723337 0.07686862 0.04129617 0.05614306]
4. The country we looked at is Hungary. This country belongs to the 9\textsuperscript{th} cluster (counting from 0). This also happens to be the cluster that the United States belongs to. These countries are not geographically similar, but they are politically and economically similar. For example, each of the countries in this cluster have economies with on par job opportunities based on the average, and higher than average support for people that can’t work. It is somewhat surprising that Uruguay is also in this cluster, but based on the data Uruguay must be more similar to Hungary and the U.S. then we realized. Using our general knowledge of these places, it seems like this cluster makes a lot of sense because the countries are quite similar in broader categories, even though they are not geographically close.