Applied Empirical Economics
Cluster and Principal Component Analysis

Gizem Kilicgedik
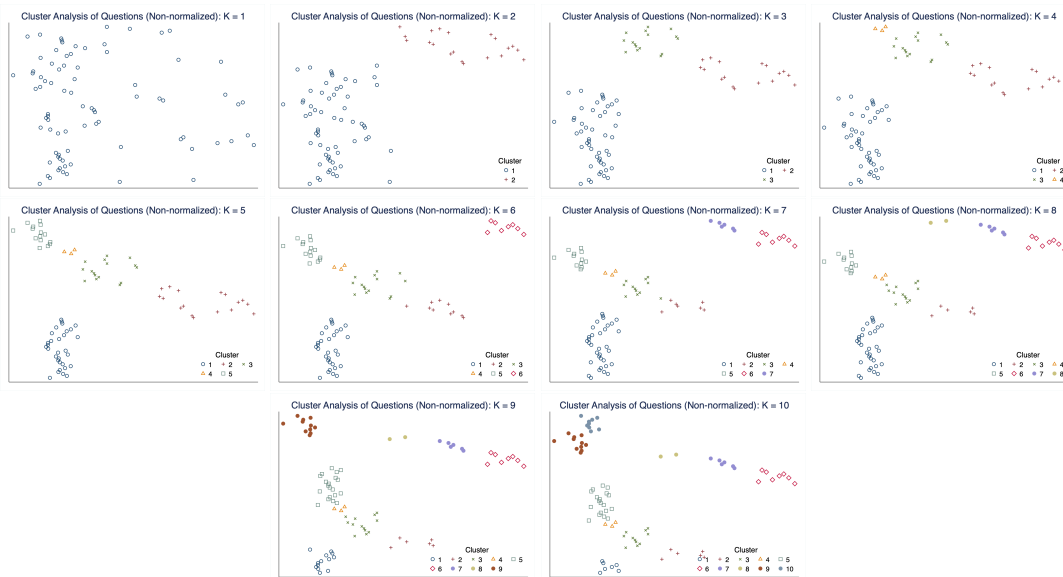
September 20, 2023

1. Cluster Analysis

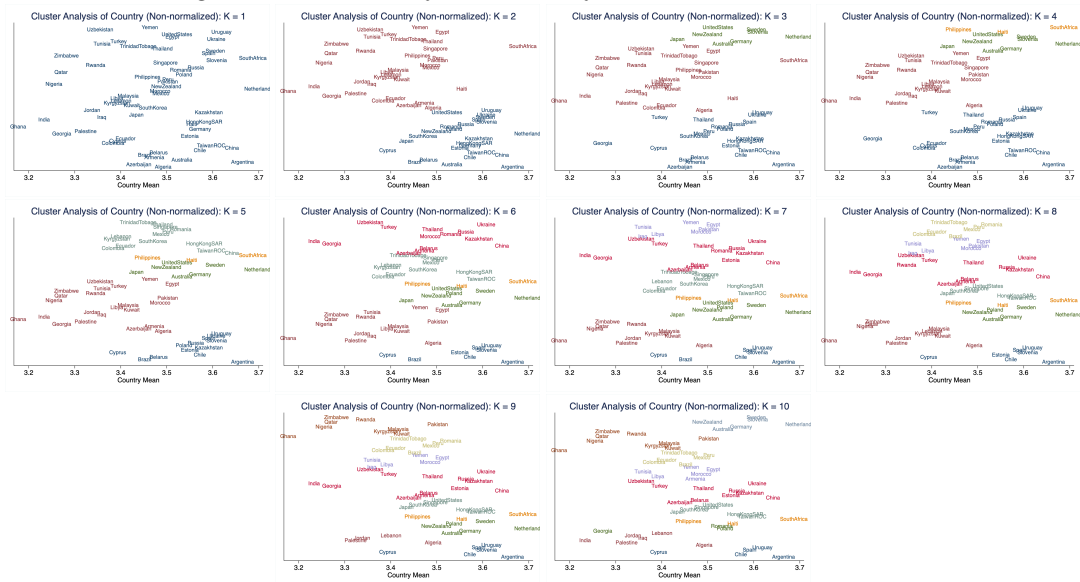### 1.1. Non-normalized data

1.1.1. Question clustering

Figure 1: Cluster analysis of questions with non-normalized data

Figure 1 displays the results of the cluster analysis for varying numbers of clusters. The horizontal axis represents the average values of the questions, while the vertical axis represents a random assignment of numbers based on the groups. The initial figure in the top left depicts the data before any clustering has occurred. As we progress starting from K=4, we observe a clear and effective clustering of the data. Based on the results, it can be seen that the cluster analysis successfully grouped similar questions together.

1.1.2. Country clustering

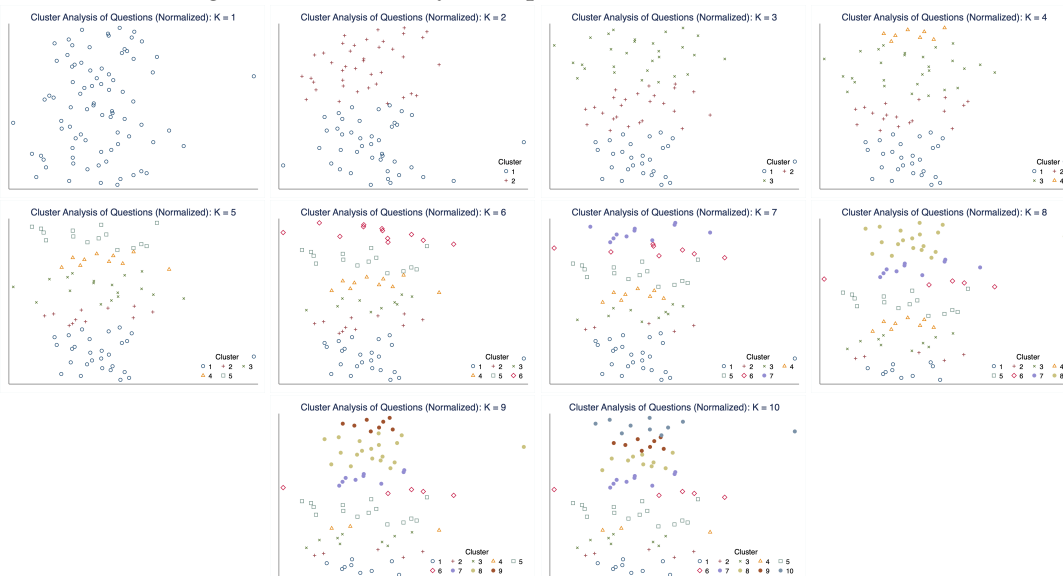Figure 2: Cluster analysis of country with non-normalized data

Just like the process of grouping similar questions, Figure 2 shows how countries are grouped together into clusters.

1.2. Normalized data
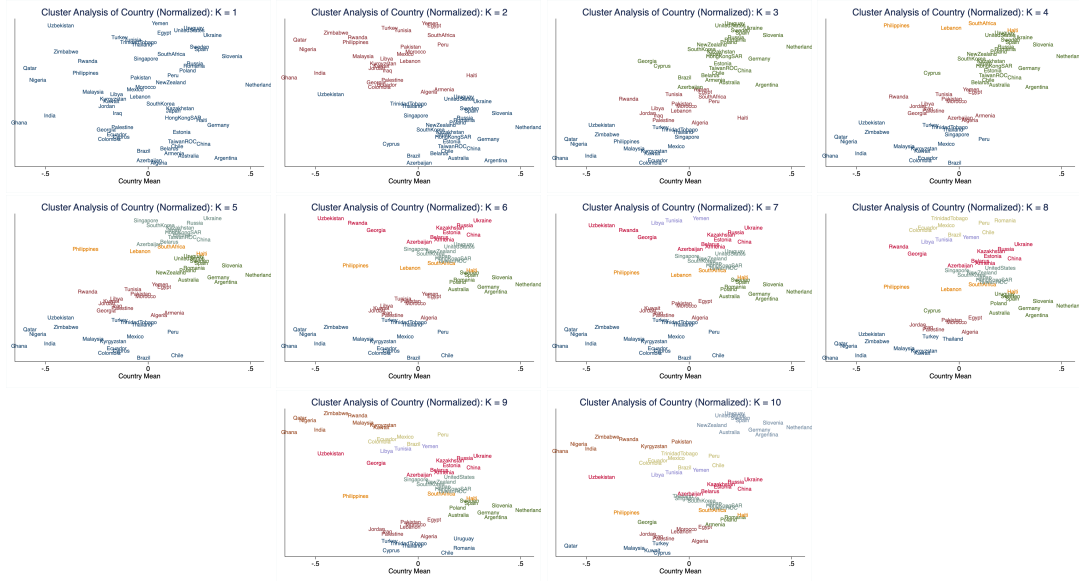
1.2.1. Question clustering

Figure 3: Cluster analysis of questions with normalized data

By organizing the questions into clusters based on the normalized data, we can observe that the resulting groups have a consistent spread of average values. This similarity arises from the method used to normalize the data, which involved centering it around a mean of zero and a standard deviation of one. The cluster analysis successfully identified patterns and effectively clustered the data in a manner comparable to that of non-normalized data. Consequently, the resulting groups exhibit similarities to those derived from non-normalized data.

1.2.2. Country clustering

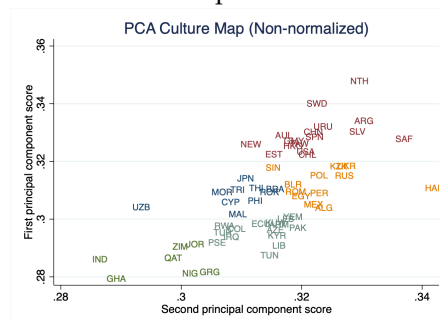Figure 4: Cluster analysis of country with normalized data

We observe a similar pattern of clustering with the non-normalized data. This, in essence, makes sense due to the normalization that was done at the question level. Hence, unlike the normalized question clustering, we see that each cluster is grouped in a rather concentrated means.

2. Principal Component Analysis

2.1. Non-normalized data

Figure 5: PCA culture map with non-normalized data

2.2. Normalized data

Figure 6: PCA culture map with normalized data