MATH 437/537 ASSIGNMENT X Due April 15, 2020

The dataset `www.mines.edu/~wnavidi/math437537/cereal.csv` contains data on 43 brands of cereal from three manufacturers: General Mills (G), Kellogg's (K), and Quaker (Q). For each brand, the number of calories, and amounts of protein, fat, sodium, fiber, carbohydrates, sugar, and potassium per serving are given.

1. Cluster the brands into 2, 3, and 4 clusters using complete linkage hierarchical clustering with Euclidean distance. For each number of clusters, report the number of items assigned to each cluster.

2. Cluster the brands into 2, 3, and 4 clusters using k-means clustering. For each number of clusters, report the number of items assigned to each cluster. Specify `nstart = 10`. This performs the clustering with 10 initial choices of centroids, and reports the one with the smallest within-groups sum of squares.

3. Construct a scree plot for the k-means clusters, using two through six clusters. Does the plot make a clear suggestion as to the number of clusters to use? If so, how many does it suggest?