We initially tried deducing meanings to the groups using the data in the averages sheet of Observations.xlsx. However, we couldn’t really determine the discriminating attributes from this data alone. We hence applied mean normalization and feature scaling using the formula:

Here is the j’th parameter of the i’th observation.

This gave us the data in abcd.xls

We then looked at the averages of the different features in the transformed data (Sheet1, newobv.xlsx) and arrived at a set of discriminating features for each cluster as follows:

Cluster 1- High K, High Fibre, and High sugar.

Cluster 2- low Na, low carbs, and low calories

Cluster3- Average values of nutrients (the largest group)

Cluster 4- High Na, low sugar, High Carbs,