**k-means on eating habits dataset**

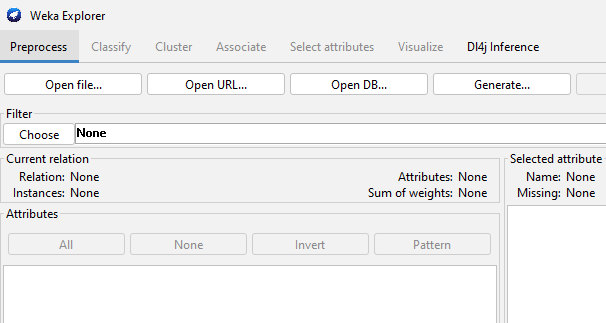
In this tutorial, we will apply k-means on eating habits dataset to classify eating habits based on country. To do this, I have following dataset.



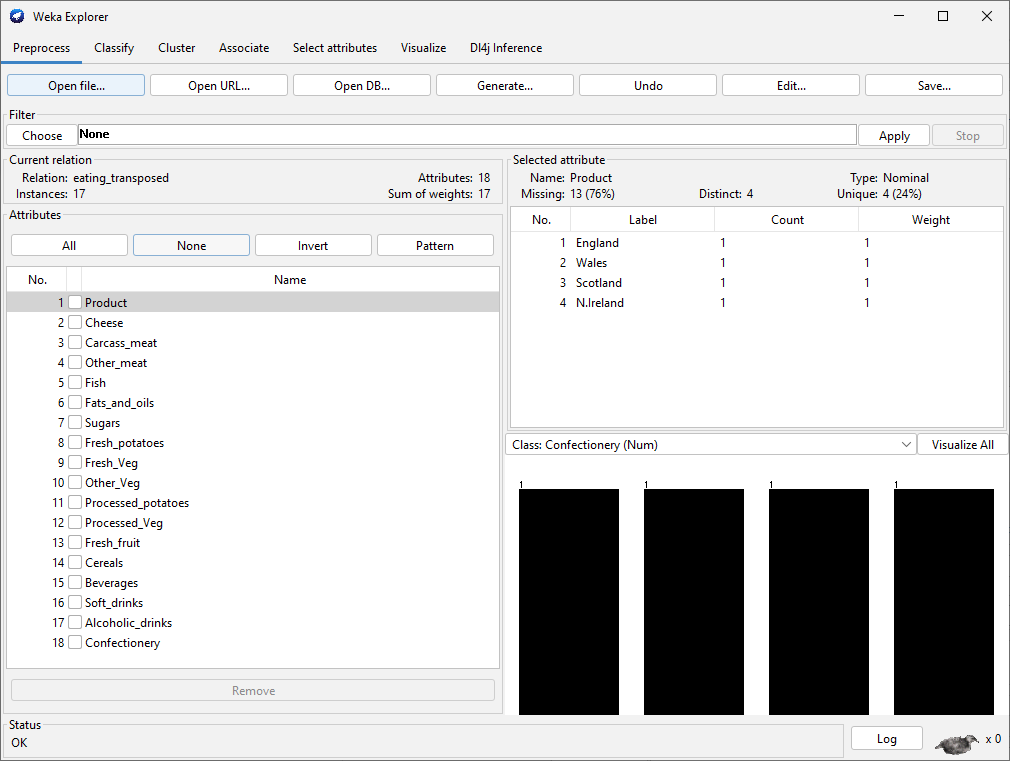
It consists of four instances and each instance shows different country. To apply k-means on our dataset, go to the WEKA explorer tab.



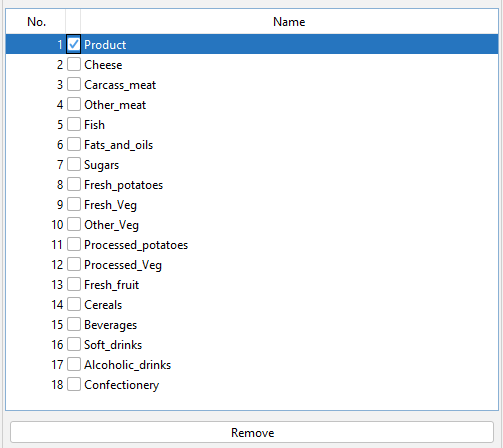
Now, load the dataset using the preprocess tab of WEKA explorer.



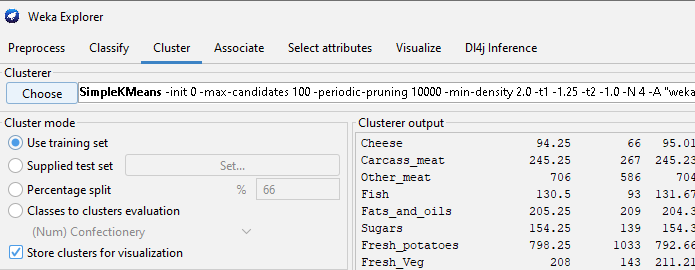
After opening he dataset, it looks like as in the following figure.

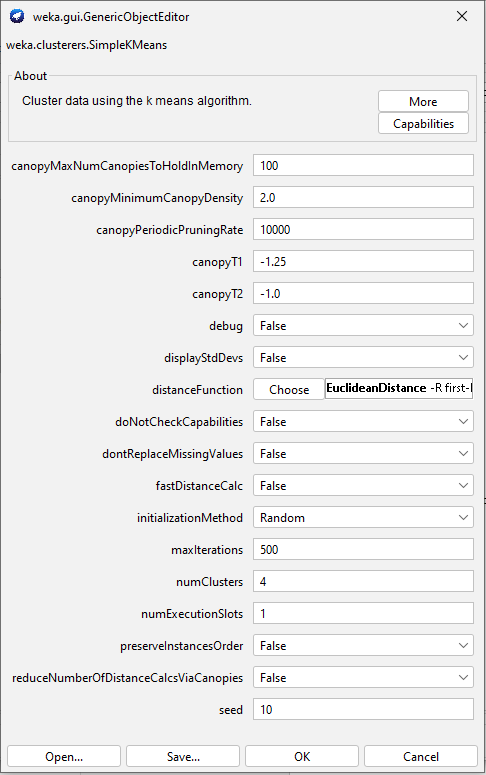


Simply remove the Product attribute by selecting this attribute and pressing remove because it is of nominal type and it contains only the names of countries.

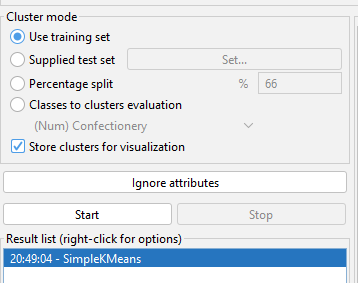


Now, go to the cluster tab and select kmeans algorithm. By default, it makes only two clusters. We will update it to 4 because our dataset consists of 4 countries. You can do this by simply clicking on its name and updating settings as

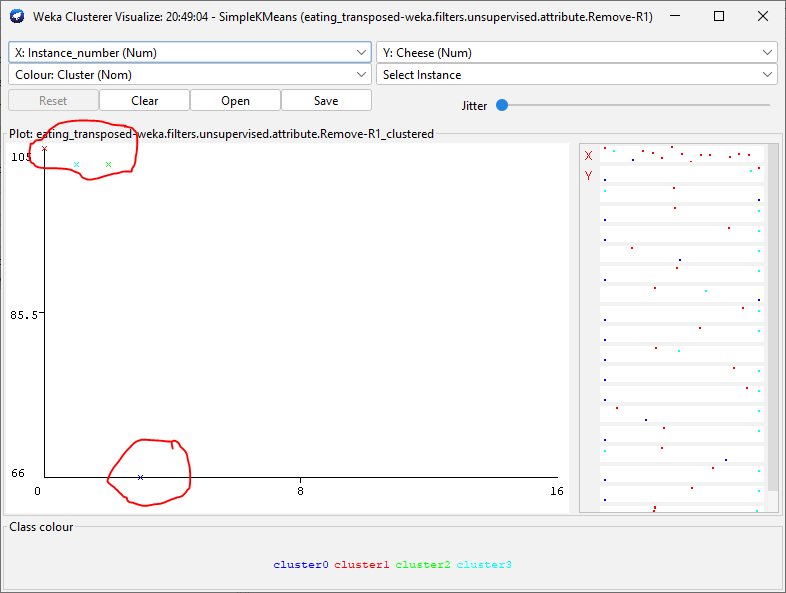




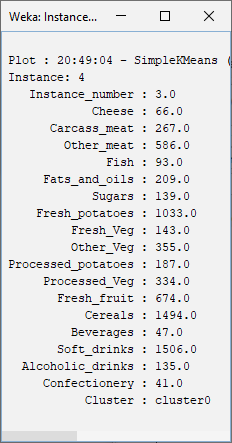
Click on start to apply clustering algorithm on dataset. Then, from result list right click on it and click “visualize cluster assignment” to visualize clusters.



The visualization of clusters are as follows:



You can see that three instances makes clusters very close to each other and one instances is far away that is showing it is from the remaining three instances. Now, click on the different one to see which attributes it contains.



So, the fourth instance is very different that is N. Ireland. Thus, we have been finally proved that N. Ireland eating habits are very different from the remaining three countries.