

Meet Barot
Mmb557

Questions and Plot

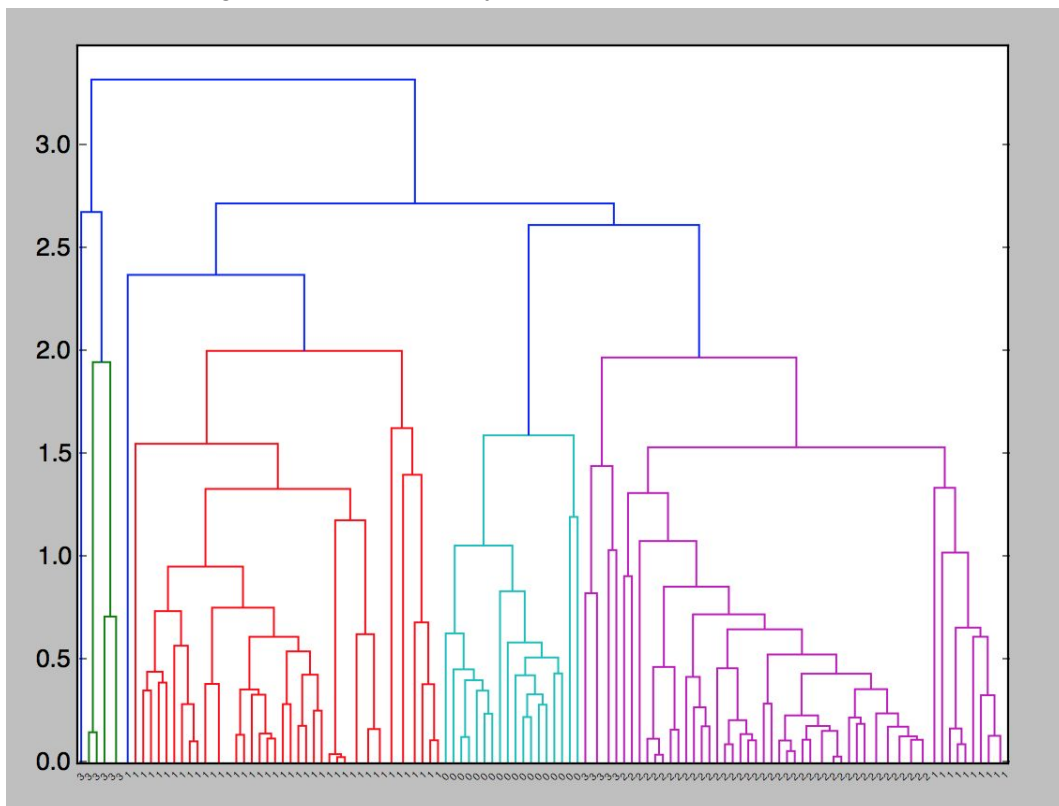
3. Random permutation of truth labels vs. ground truths: 0.5630551034643106
Clustering labels from k-means vs. ground truths: 0.8355445066442508

4. The minimum function value observed was 660.0288115.

Objective function values observed:	Rand Indices:
660.0288115	0.8355445066
779.1936872	0.8000250389
789.3882392	0.8000983671
779.1936872	0.8000250389
19.82915575	0.7998524493

5. Almost all the foods in this random sample clustered into four groups except one. I was not able to get the actual food items showing up in a readable form on the dendrogram.

Dendrogram of 30 Randomly Sampled Foods from Each Group



6. The agglomerative clustering performed very badly (randIndex of 0.2437) because only one cluster was output when $t=3.8$. However, it performed slightly better at exactly $t=3.0$, giving a randIndex of 0.382, with 3 clusters output. Given a $t=.1$, there was a randIndex of 0.754, with many clusters. Given a $t=.01$, there was a randIndex of 0.755, with even more clusters.

7. Running K-means for 5, 10, 25, 50, and 75 clusters:

Run for $k=5$

Random item from largest cluster: MACARONI,WHOLE-WHEAT,CKD

Random item from largest cluster: MACARONI,VEG,CKD,ENR

Random item from largest cluster: HOMINY,CANNED,WHITE

Random item from largest cluster: HOMINY,CANNED,WHITE

Random item from largest cluster: SPAGHETTI,PROTEIN-FORTIFIED,CKD,ENR (N X

6.25)

Random item from largest cluster: QUINOA,CKD

Random item from largest cluster: MACARONI,WHOLE-WHEAT,CKD

Random item from largest cluster: QUINOA,CKD

Random item from largest cluster: MACARONI,PROTEIN-FORTIFIED,CKD,ENR,(N X

5.70)

Random item from largest cluster: RICE,BROWN,LONG-GRAIN,CKD

Size of largest cluster: 57

Run for $k=10$

Random item from largest cluster: SPAGHETTI,WHOLE-WHEAT,CKD

Random item from largest cluster: RICE,WHITE,SHORT-GRAIN,CKD,UNENR

Random item from largest cluster: RICE,WHITE,STMD,CHINESE RESTAURANT

Random item from largest cluster: MACARONI,WHOLE-WHEAT,CKD

Random item from largest cluster: RICE,WHITE,SHORT-GRAIN,CKD

Random item from largest cluster: BULGUR,COOKED

Random item from largest cluster: HOMINY,CANNED,WHITE

Random item from largest cluster: NOODLES,EGG,SPINACH,CKD,ENR

Random item from largest cluster: OAT BRAN,COOKED

Random item from largest cluster: SPAGHETTI,WHOLE-WHEAT,CKD

Size of largest cluster: 57

Run for $k=25$

Random item from largest cluster: WILD RICE,COOKED

Random item from largest cluster: HOMINY,CANNED,WHITE

Random item from largest cluster: RICE,WHITE,SHORT-GRAIN,CKD,UNENR

Random item from largest cluster: RICE,BROWN,LONG-GRAIN,CKD

Random item from largest cluster: RICE,WHITE,LONG-GRAIN,PARBLD,UNENR,CKD

Random item from largest cluster: RICE,BROWN,LONG-GRAIN,CKD

Random item from largest cluster: RICE,WHITE,GLUTINOUS,CKD

Random item from largest cluster: PASTA,CORN,COOKED

Random item from largest cluster: AMARANTH GRAIN,CKD

Random item from largest cluster: BULGUR,COOKED

Size of largest cluster: 30

Run for k= 50

Random item from largest cluster: RICE,WHITE,MEDIUM-GRAIN,CKD

Random item from largest cluster: BARLEY,PEARLED,COOKED

Random item from largest cluster: RICE,WHITE,SHORT-GRAIN,CKD

Random item from largest cluster: RICE,WHITE,SHORT-GRAIN,CKD,UNENR

Random item from largest cluster: RICE,WHITE,MEDIUM-GRAIN,CKD

Random item from largest cluster: RICE,BROWN,LONG-GRAIN,CKD

Random item from largest cluster: RICE,BROWN,LONG-GRAIN,CKD

Random item from largest cluster:

RICE,WHITE,LONG-GRAIN,REG,CKD,UNENR,WO/SALT

Random item from largest cluster: RICE,WHITE,SHORT-GRAIN,CKD,UNENR

Random item from largest cluster: RICE,WHITE,MEDIUM-GRAIN,CKD,UNENR

Size of largest cluster: 16

Run for k= 75

Random item from largest cluster: RICE,BROWN,MEDIUM-GRAIN,CKD

Random item from largest cluster: RICE,BROWN,MEDIUM-GRAIN,CKD

Random item from largest cluster: BARLEY,PEARLED,COOKED

Random item from largest cluster: RICE,WHITE,LONG-GRAIN,REG,CKD,ENR

Random item from largest cluster: RICE,WHITE,SHORT-GRAIN,CKD,UNENR

Random item from largest cluster: RICE,WHITE,MEDIUM-GRAIN,CKD

Random item from largest cluster: RICE,WHITE,LONG-GRAIN,REG,CKD,ENR,W/SALT

Random item from largest cluster: RICE NOODLES,CKD

Random item from largest cluster: RICE,WHITE,MEDIUM-GRAIN,CKD,UNENR

Random item from largest cluster: BARLEY,PEARLED,COOKED

Size of largest cluster: 14

From these runs, we can see the size of largest cluster decreasing with a larger number of clusters that the function is forced to make, which makes sense. The foods in the runs with fewer clusters (k=5, k=10) are fairly different from one another. However, when there are a larger number of clusters, fewer foods are included in each one, resulting in many foods being the same with the slight variation of preparation or fortification of nutrients. This was expected.