Create a Jupyter notebook file that demonstrates clustering on the wine dataset using Hierarchical Clustering (Agglomerative Clustering) and KMeans. You may use the iris\_clustering.ipynb as a guide/template.

1. Select <u>only two features</u> from the list below. Create a dataframe with those two features from the wine dataset, and use that dataframe for the rest of the exercise

Total Phenols
Flavanoids
Proanthocyanins
Wine Dilution

## Dendrogram

- a. Plot the dendrogram based on the two features you chose. What is a possible y value that will give you <u>4 clusters</u>? Show your answer by plotting a horizontal line on the dendrogram.
- b. Perform agglomerative clustering on your dataset for 4 clusters and plot the results.

## 3. KMeans

- a. Silhouette Score
  - i. Get the best k for your dataset based on the silhouette scores from k=2 to k=16. Print your answer and its silhouette score.
  - ii. Perform KMeans from the best k you induced, and plot the results.

## b. Elbow Method

- i. Compute and plot the inertias for k=2 to k=16.
- ii. Which k is the possible inflection point ("elbow") in the graph? Show your answer by plotting a vertical line for the k-value.
- iii. Perform KMeans from the k value you chose, and plot the results.
- iv. Is the k from the elbow method consistent with the k from the silhouette score? (You can put your answer in the last cell of your Jupyter notebook)