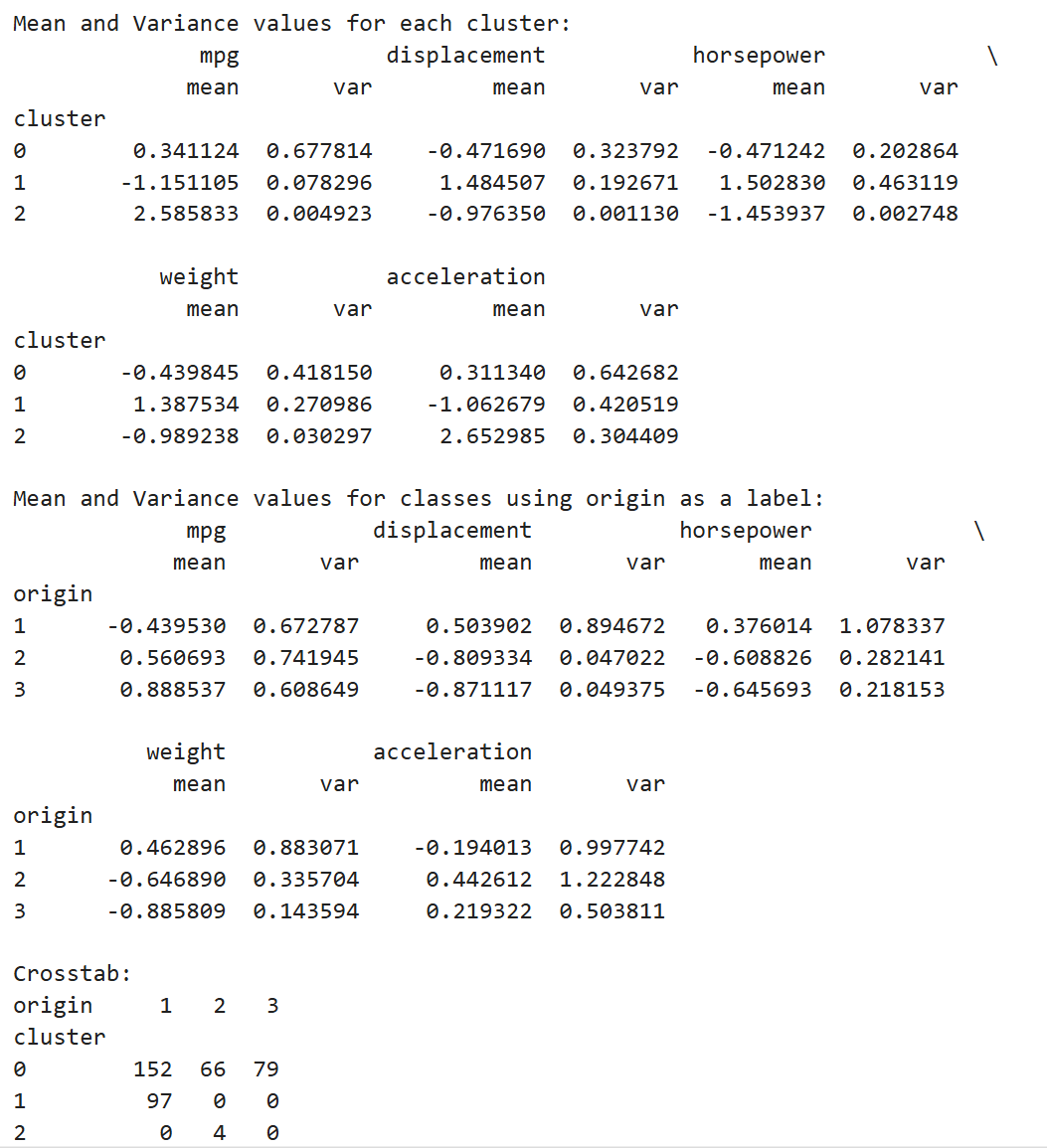
**Problem 1**

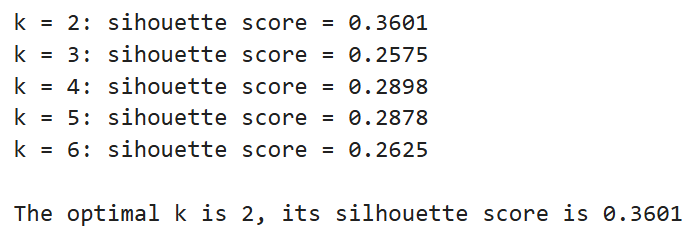
The final output of problem 1 is shown below:

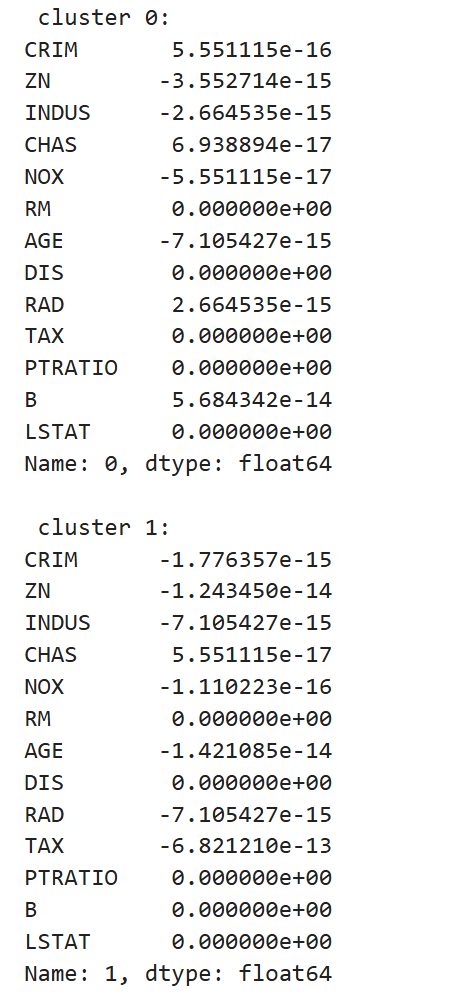


The crosstab indicates that there is no clear relationship between cluster assignment and class label. Specifically, cluster 1 and 2 perfectly corresponds to origin 1 and origin 2 respectively. Despite that, cluster 0 is a mixture of all three origins. Therefore, while a partial relationship exists due to the strong correspondence of clusters 1 and 2 with specific origins, no clear global relationship is evident.

**Problem 2**

The final output of problem 2 is shown below:

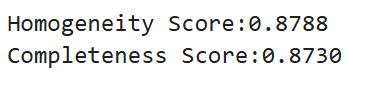




The optimal k is 2. The second image shows the differences between mean values of each cluster and centroid coordinates. These differences are extremely small or exactly zero. These negligible differences may result from floating-point arithmetic errors. Such results confirm that the cluster means and centroid coordinates are effectively equivalent.

**Problem 3**

The final output of problem 3 is shown below:



For Homogeneity Score:

Homogeneity Score reflects the purity of clusters. A high homogeneity score of 0.8788 indicates that the clusters created by K-Means are highly consistent with the true class labels.

For Completeness Score:

Completeness Score reflects the coverage of true classes by clusters. A high completeness score of 0.8730 indicates that K-Means successfully groups most samples of each true class together. However, given that completeness score is a bit lower than homogeneity score, minor instances where samples from the same class may be split across clusters.

Totally, both high homogeneity score and completeness score indicate that the clustering is highly effective.