**Answer 1**

Based on the output of the code, comparing the mean and variance, it can be seen that cluster 1 is similar to origin1, and cluster 2 is very close to origin3 but the mpg is much higher than the mean of origin3. The features of cluster 0 are partially similar to both origin2 and origin3. Furthermore, based on the Contingency Table, Standardized Residuals, and the chi-square p-value, it can be inferred that there is a significant association between cluster assignment and origin labels. This association is particularly strong in clusters 1 and 2. However, due to the mixed nature of cluster 0, it is difficult to confirm this conclusion definitively.

**Answer 2**

Based on the output of the code, it can be seen that k=2 is the optimal number of clusters because its silhouette score of 0.3601 is the highest. The code output provides the mean of all features for each cluster in the optimal clustering. These values are numerically very close to the centroid coordinates, but conceptually, the cluster mean represents the average feature values of actual samples, while the centroid is the theoretical center optimized by K-Means. In some cases, they can be used interchangeably.

**Answer 3**

Based on the output of the code, under the optimal k conditions, the Homogeneity Score is 0.8788 and the Completeness Score is 0.8730. A high homogeneity score indicates that the cluster assignments are pure, meaning that the samples within each cluster primarily belong to the same category. However, a score less than 1 indicates the presence of a small number of samples from other categories. A high completeness score indicates that the vast majority of samples from each category are assigned to the same cluster. However, a score less than 1 indicates that a small number of samples from some categories are assigned to other clusters.