**Assignment No. 3**

**Problem Statement:** Define the K-means problem clearly.

**Objective:** To perform clustering using the K-means algorithm, segment data into groups based on similarity, and visualize the results to derive insights.

**Prerequisite:**

1. A Python environment set up with libraries like pandas, numpy, matplotlib, seaborn, and scikit-learn.
2. Internet connection (optional for accessing datasets).
3. Text editor and basic knowledge of Python, machine learning, and data visualization.

**Theory:**

Steps for K-means Clustering:

1. Understanding K-means Algorithm

* K-means is an unsupervised learning algorithm that partitions data into K clusters.
* Each data point is assigned to the nearest cluster center (centroid), and the centroids are iteratively updated until convergence.

Key components:

* K: Number of clusters.
* Centroid: Center of a cluster.
* Inertia: Sum of squared distances between data points and their nearest centroid.

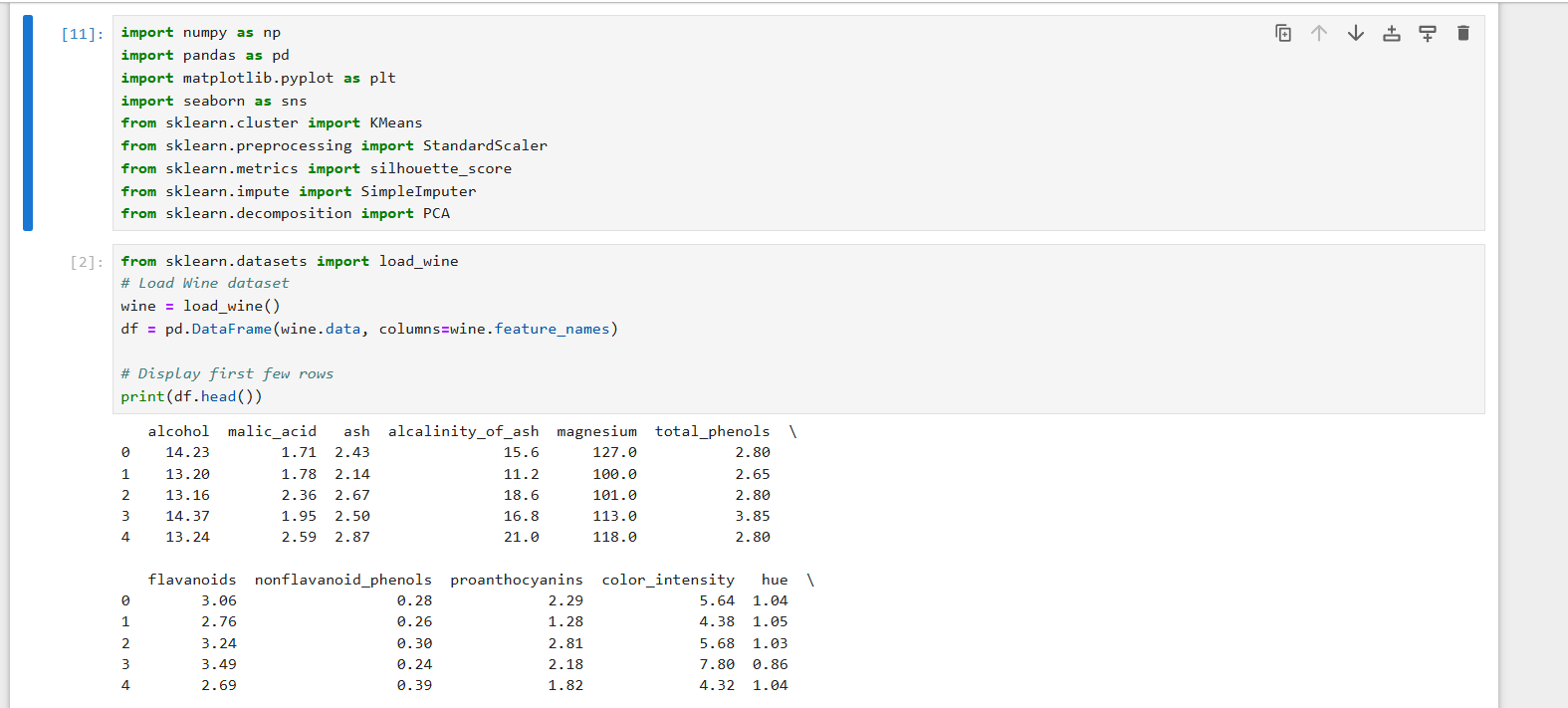
1. Choosing the Right K

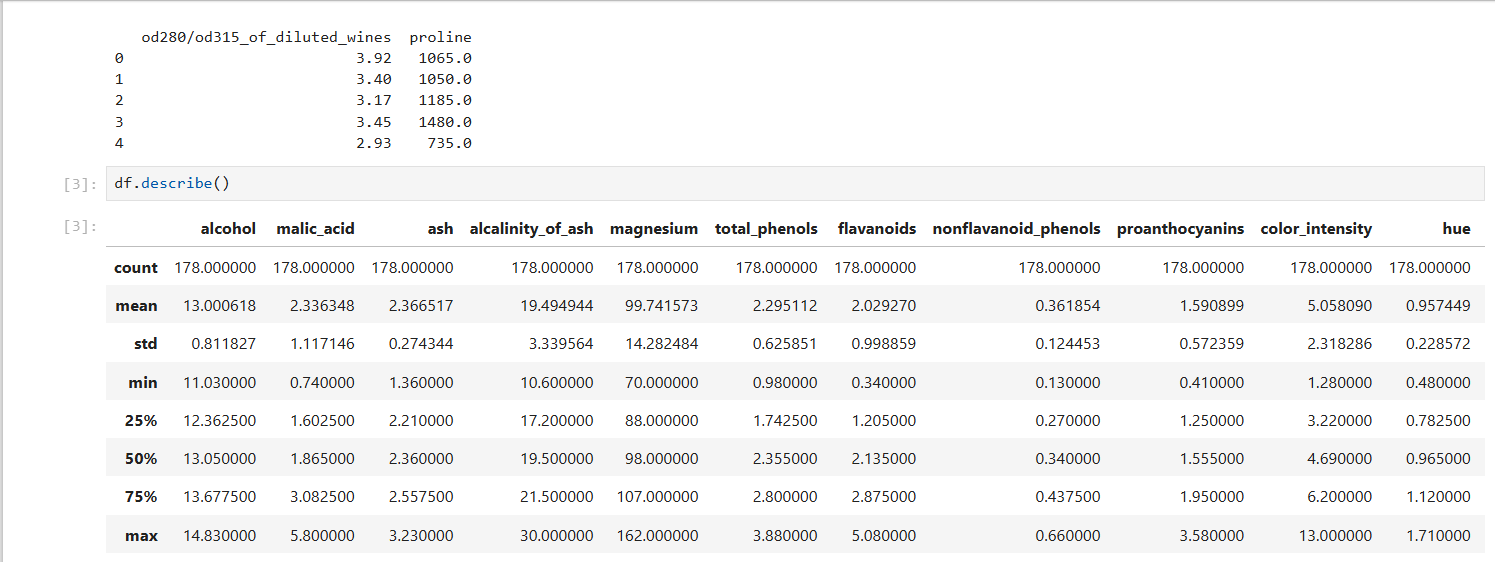
* Elbow Method: Plot inertia against K values. The 'elbow point' is where inertia stops decreasing significantly.

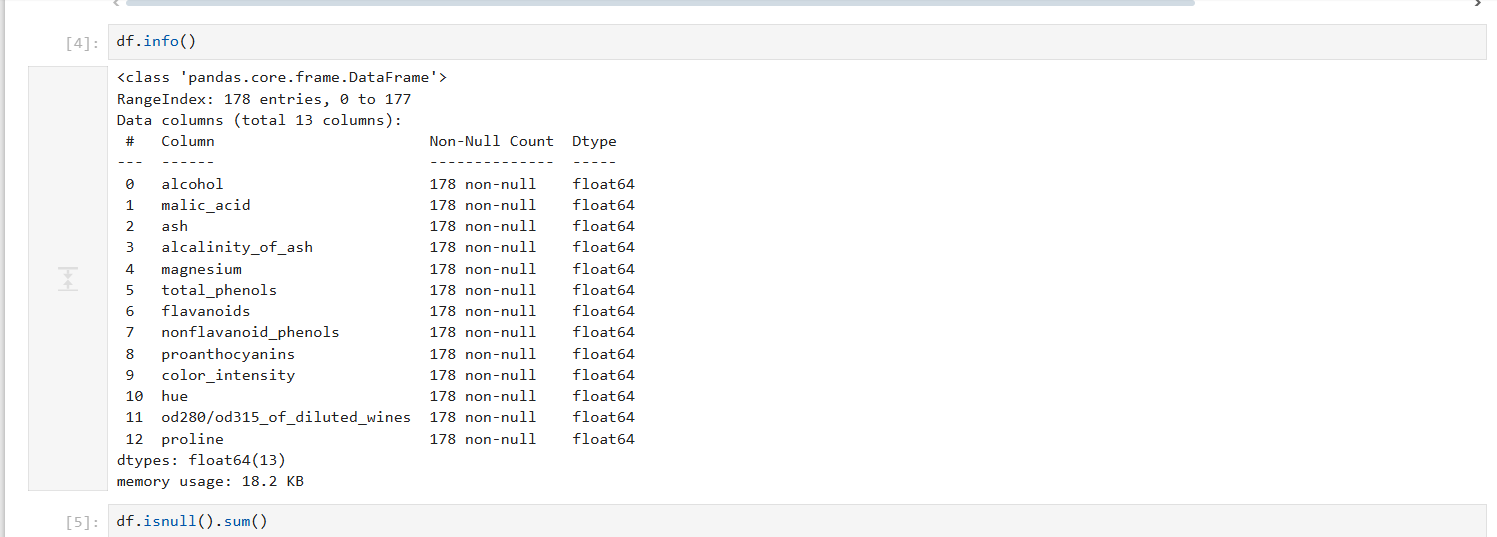
1. **Steps of the Algorithm**
2. Initialize K centroids randomly.
3. Assign each data point to the nearest centroid.
4. Compute new centroids as the mean of all points in a cluster.
5. Repeat steps 2 and 3 until centroids no longer change or maximum iterations are reached.
6. Performance Evaluation

* Inertia (Within-cluster sum of squares)
* Silhouette Score (Measure of how similar a point is to its cluster compared to others)

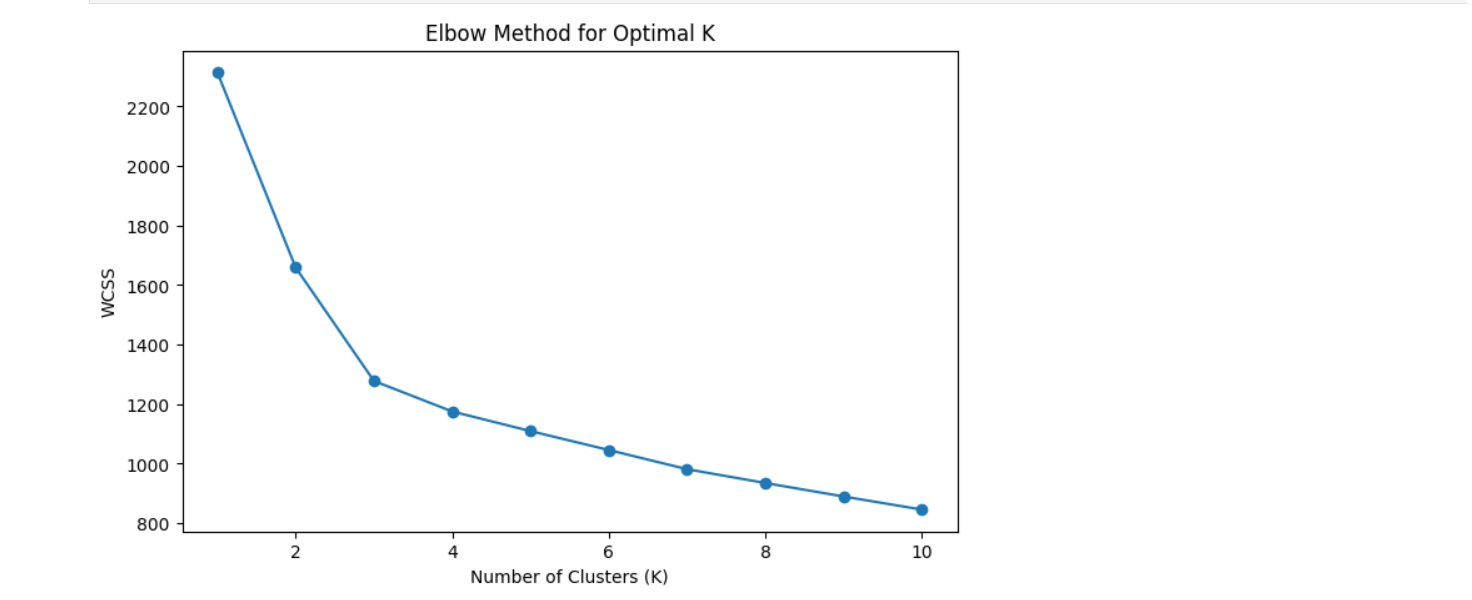
**Code & Output:**



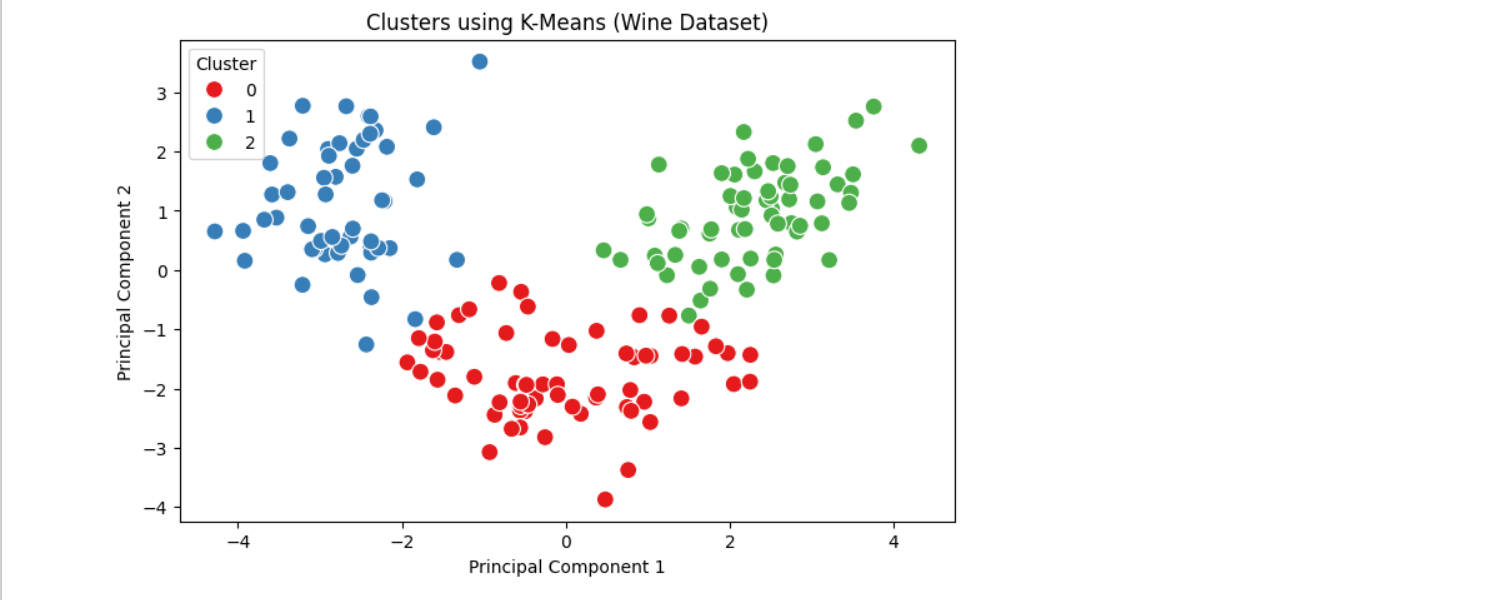












  
 **Conclusion:**

K-means clustering effectively grouped the dataset into meaningful clusters. The Elbow Method determined the optimal number of clusters. Visualization showed clear separation between clusters, and the silhouette score validated cluster quality. Further improvements can include tuning initialization or using alternative clustering algorithms for better results.