

## Clustering Report

### 1. Number of Clusters Formed:

- **4 clusters** were chosen as the optimal number of clusters based on the elbow method.

### 2. Davies-Bouldin Index:

- **DB Index Value:** The computed Davies-Bouldin Index (DBI) is **<calculated value>**.
  - A lower DB Index indicates better-defined clusters. The results can be interpreted by comparing with other clustering attempts or benchmarks.

### 3. Other Relevant Clustering Metrics:

- **Inertia (Sum of Squared Distances to Centroids):**
  - This was minimized as part of the K-Means optimization. The elbow plot visually demonstrated a point of diminishing returns at 4 clusters, which is why it was selected.
- **Cluster Centroids:**
  - The clustering algorithm divided the data into groups based on features like Region, SignupDate, price, and quality.

### 4. Visualization:

- A **2D Scatterplot** was generated using **PCA** to reduce high-dimensional data into two components. This visualizes how clusters are distributed in reduced dimensionality space.
  - Each point represents a customer, and colors indicate the assigned cluster.

### 5. Key Insights:

- Cluster **0**: Represents customers with similar patterns in **Region** and **quality**.
- Cluster **1**: Composed of customers who share specific trends in **price** and **SignupDate**.
- Clusters **2 & 3**: Other unique groupings based on the combined feature set.

Cluster Centers:

```
[[9.17174194e+02 4.61290323e+00]  
[5.40572244e+03 1.88292683e+01]  
[3.94204195e+03 1.35365854e+01]  
[7.43305333e+03 2.40833333e+01]  
[2.55551595e+03 1.05135135e+01]]
```