**Clustering Metrics Analysis for the Code**

**1. Optimal Number of Clusters**

* The Davies-Bouldin (DB) index is used to measure clustering quality. A lower DB index indicates better separation and compactness of clusters.
* From the code:
  + The optimal number of clusters, optimal\_k, is determined by the cluster count corresponding to the lowest DB index.
  + Example output:
  + Optimal number of clusters: 3
  + Minimum Davies-Bouldin Index: 0.4 (value may vary based on data).

**2. Additional Clustering Metrics**

Consider calculating other metrics to validate and interpret the clustering:

1. **Silhouette Score**:
   * Measures how similar a point is to its cluster vs. others.
   * Ranges from -1 to 1; higher values indicate better-defined clusters.
   * **Python Code**:
   * from sklearn.metrics import silhouette\_score
   * silhouette\_avg = silhouette\_score(clustering\_data\_scaled, optimal\_labels)
   * print(f"Silhouette Score: {silhouette\_avg}")
2. **Inertia (Within-Cluster Sum of Squares)**:
   * Measures how tightly data points are grouped around centroids.
   * Lower values indicate better clustering but should decrease with more clusters.
   * **Python Code**:
   * inertia = optimal\_kmeans.inertia\_
   * print(f"Inertia: {inertia}")
3. **Calinski-Harabasz Index**:
   * Ratio of the sum of between-cluster dispersion to within-cluster dispersion.
   * Higher values indicate better-defined clusters.
   * **Python Code**:
   * from sklearn.metrics import calinski\_harabasz\_score
   * ch\_score = calinski\_harabasz\_score(clustering\_data\_scaled, optimal\_labels)
   * print(f"Calinski-Harabasz Index: {ch\_score}")

**Output Example**

Assuming 3 clusters (optimal\_k = 3):

Optimal number of clusters: 3

Minimum Davies-Bouldin Index: 0.42

Silhouette Score: 0.63

Inertia: 2500.5

Calinski-Harabasz Index: 1123.45

**3. Recommendations**

* **Validate Results**: Cross-check the cluster quality using silhouette and Calinski-Harabasz scores.
* **Visualize**: Extend PCA visualization with cluster centers and other dimensions to interpret customer segments.
* **Analyze Clusters**: Use descriptive statistics (mean, median) to characterize customer groups for actionable insights.