

Project Report - Online Retail II Clustering Analysis

Dataset size : 805,620 rows / 5,881 customers

Clusters : Hierarchical vs K-Medoids, both k=4

Silhouette : Hierarchical 0.94 / K-Medoids 0.28

FP-Growth : 173 item-sets -> 16 rules ($\geq 2\%$ support)

Figure 1: Truncated Dendrogram (Single Linkage)

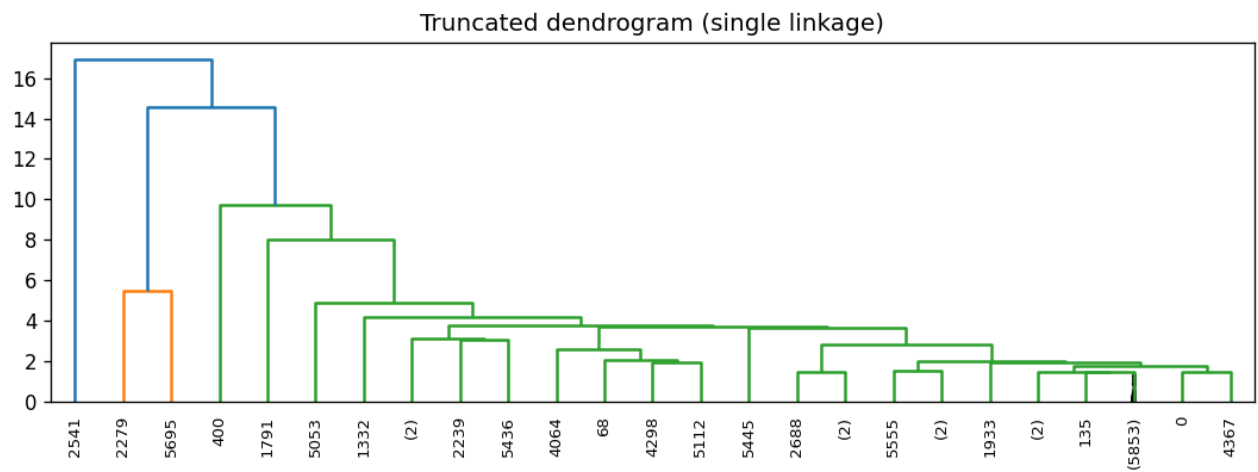


Figure 2: Hierarchical Clusters Visualization (PCA)

Hierarchical Clustering (k=4)

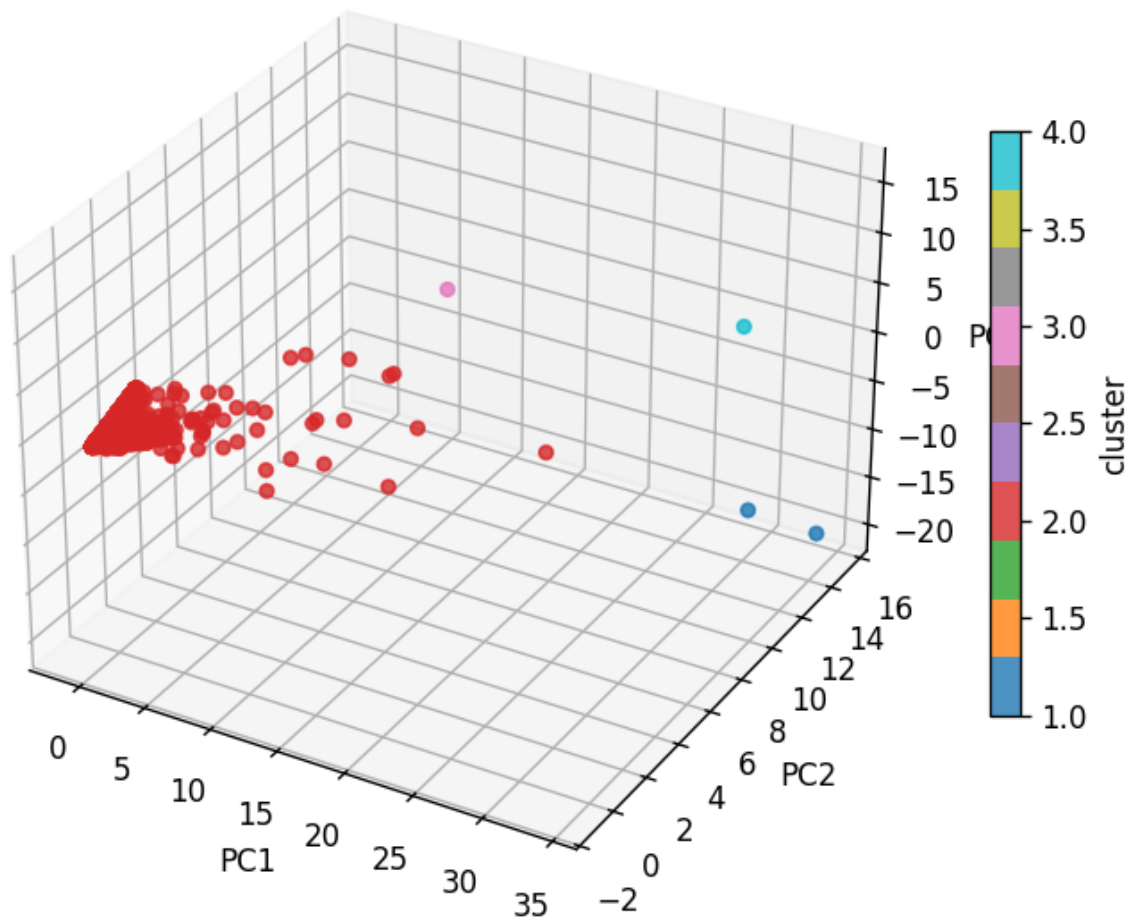


Figure 3: K-Medoids Clusters Visualization (PCA)

K-Medoids Clustering (k=4)

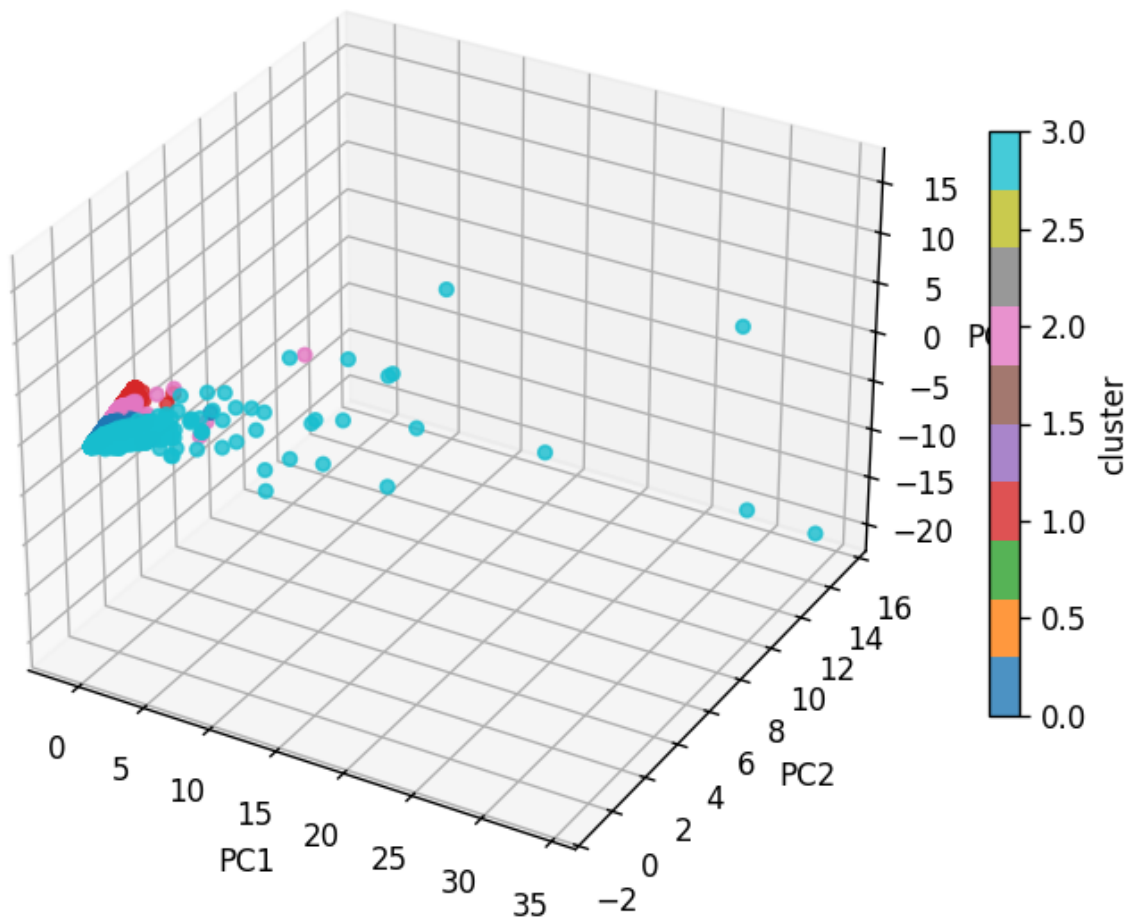


Figure 4: Silhouette Scores Comparison

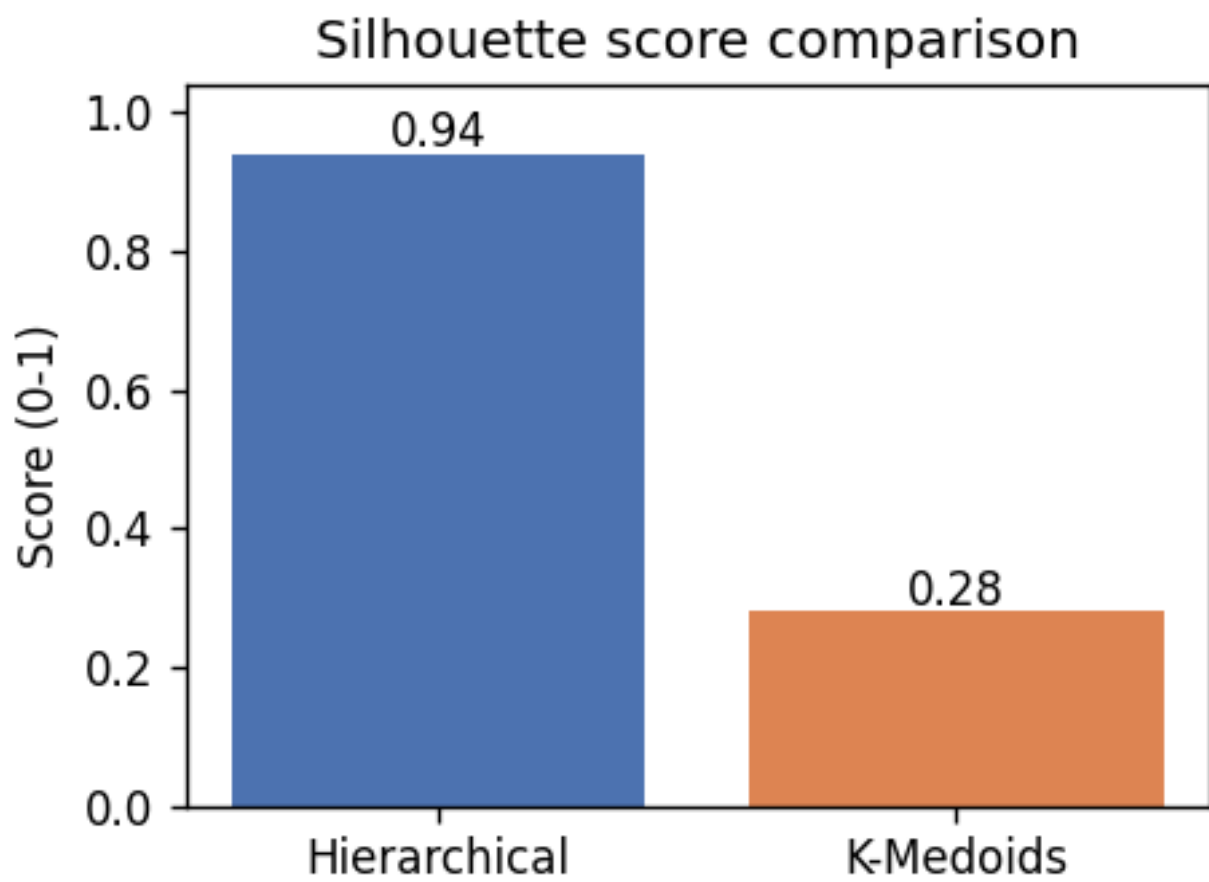


Figure 5: Elbow Curve (KMeans Baseline)

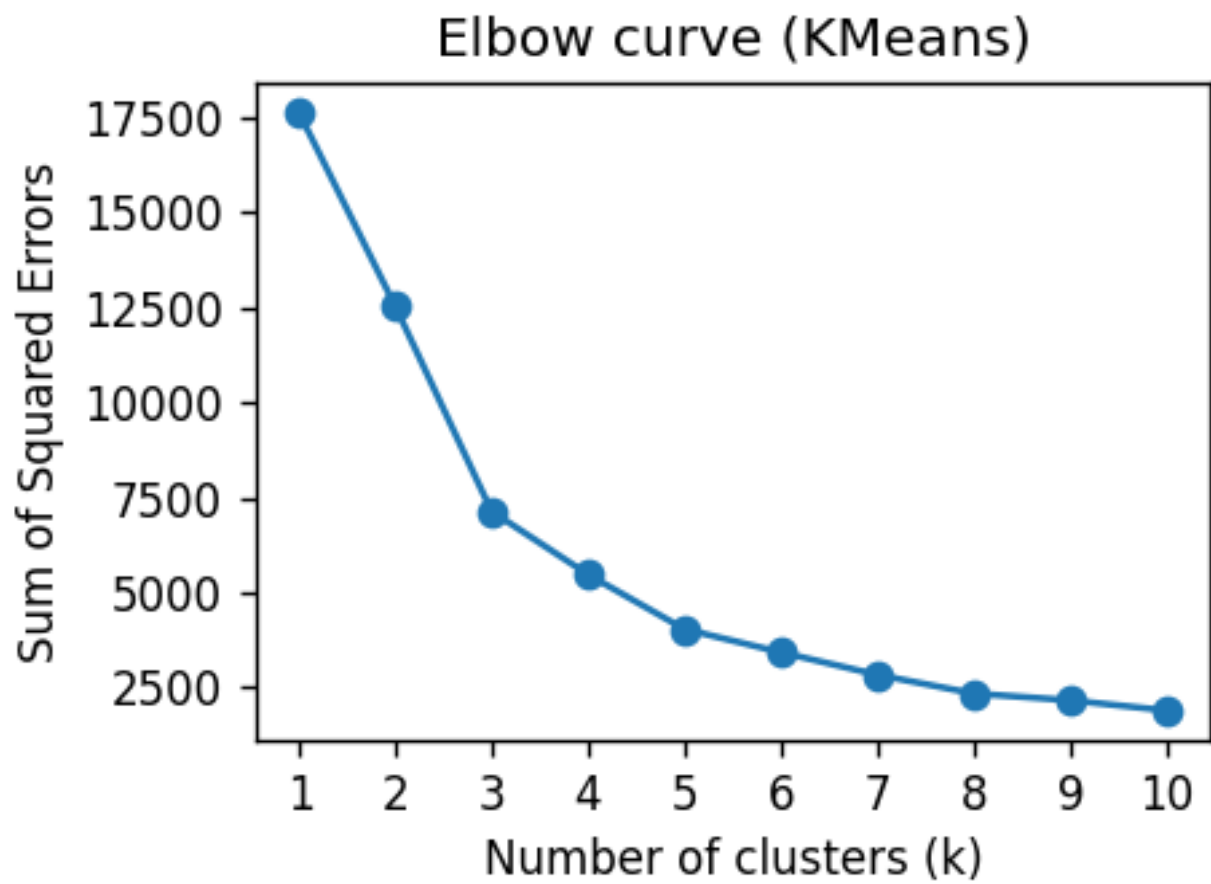


Figure 6: Relative Importance Heat-Map

