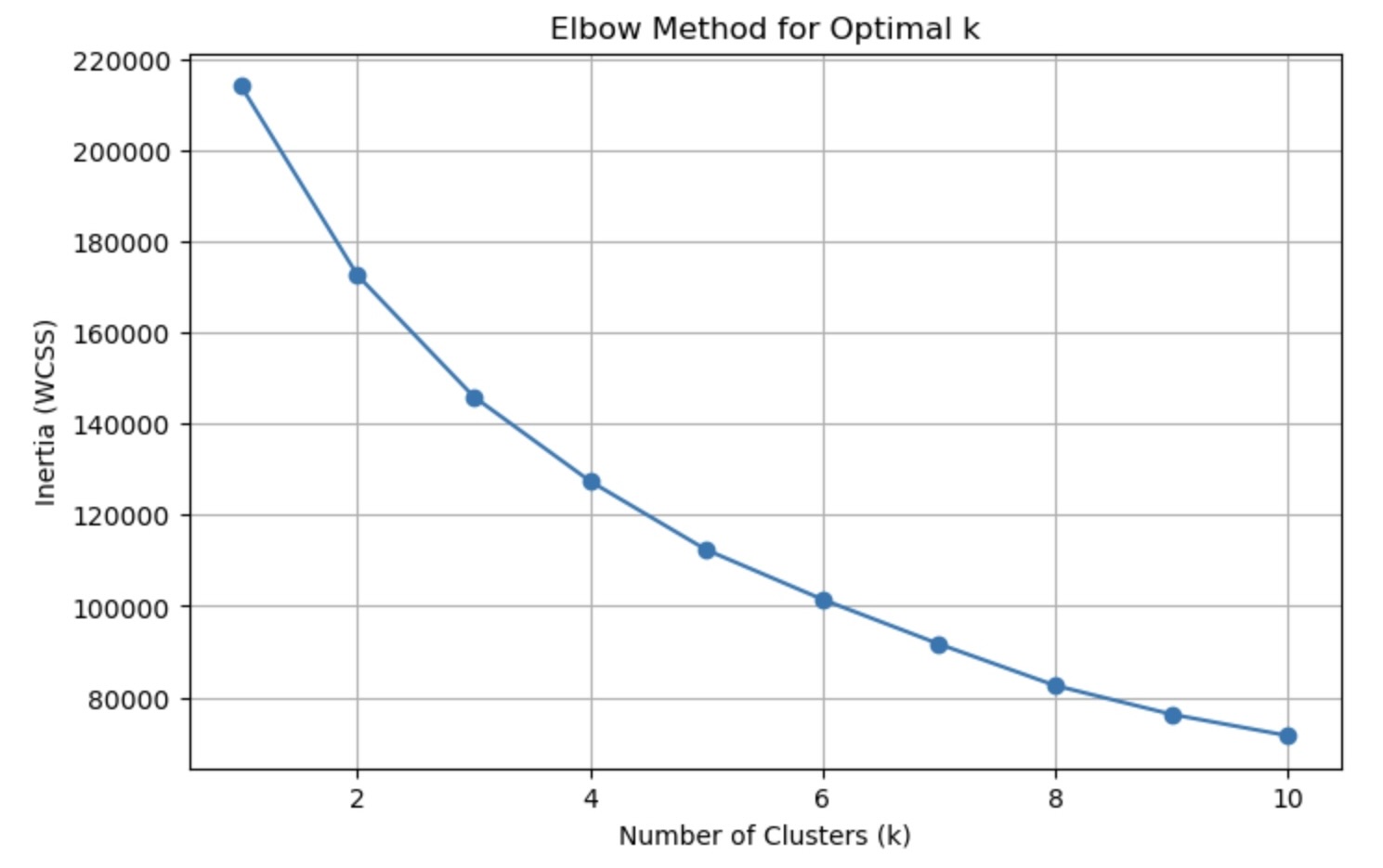
Customer Segmentation Using K-Means Clustering & PCA

# 🧠 Objective

This report explores customer segmentation using the K-Means clustering algorithm on a real-world dataset. We use both the Elbow Method and Silhouette Score to identify optimal cluster count, and visualize results using PCA.

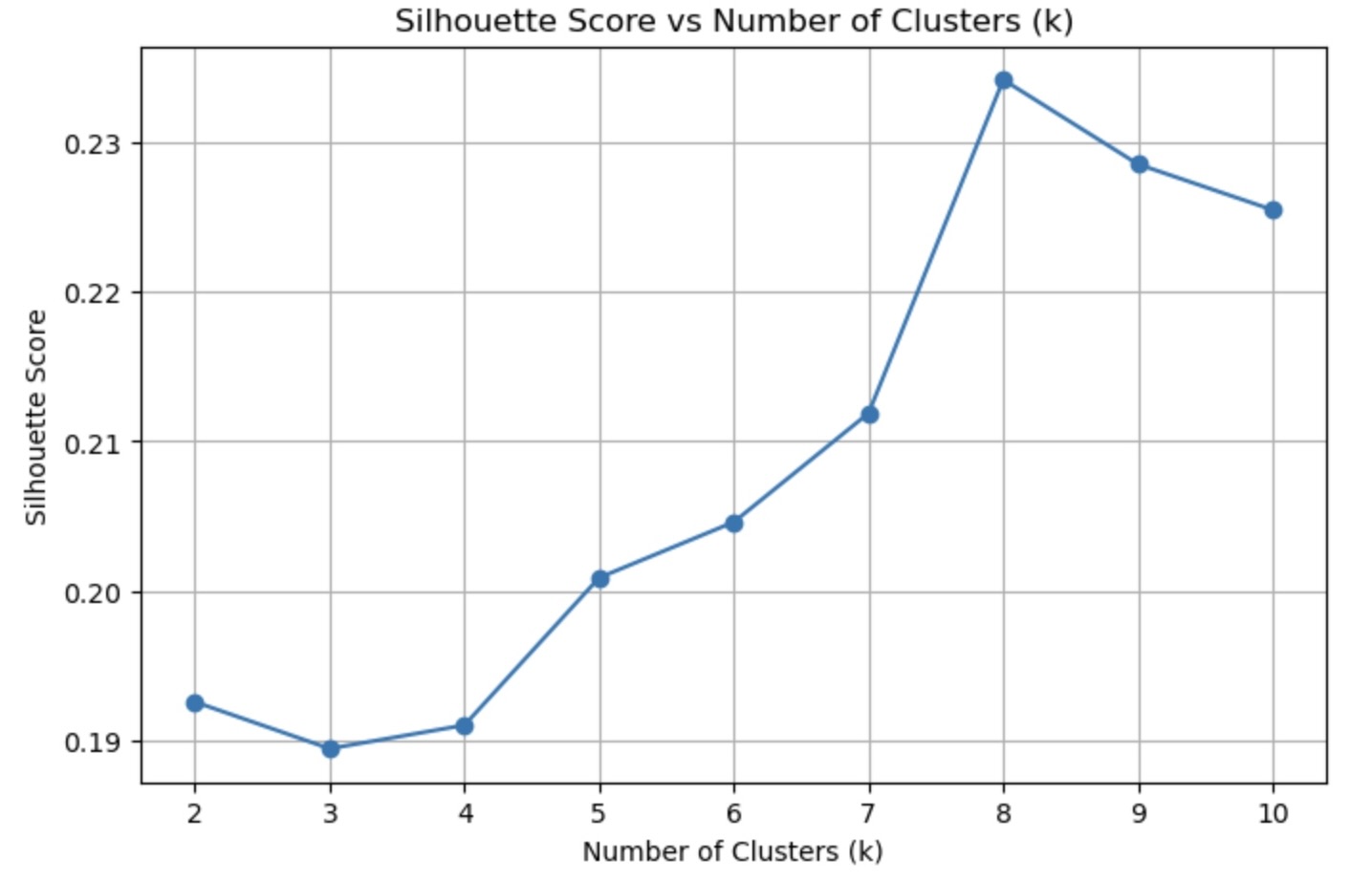
# 📉 Elbow Method

The Elbow Method helps determine the ideal number of clusters (k) by plotting inertia (within-cluster sum of squares). The 'elbow point' is where additional clusters no longer significantly reduce inertia.



# 📏 Silhouette Score

The Silhouette Score quantifies how well-separated the clusters are. Scores near 1 indicate distinct clusters, while scores near 0 indicate overlapping clusters. The highest silhouette score was observed at k = 8.



# 📊 Cluster Profiles

The mean values of Age, Income, Coverage, and Premium for different k values:

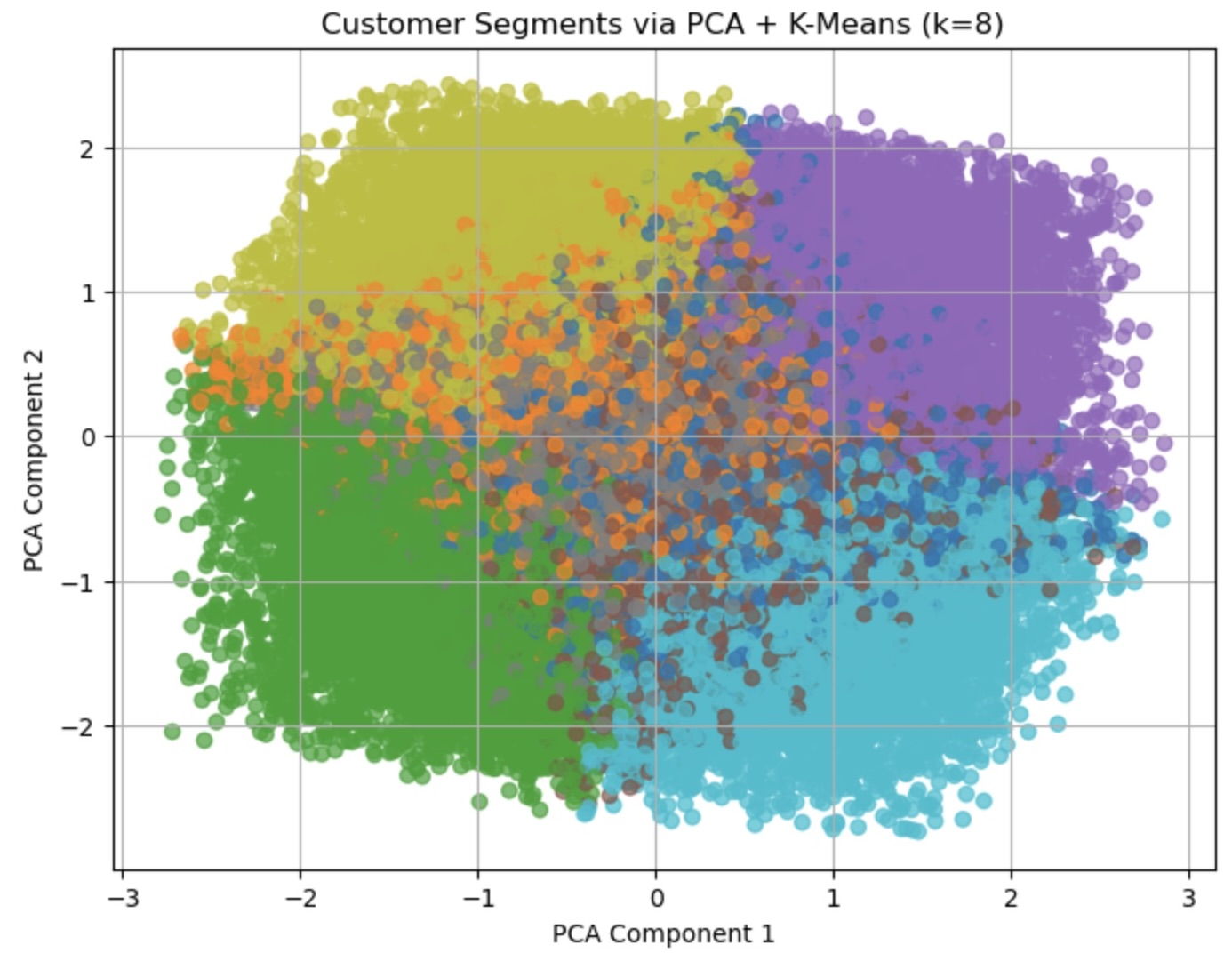
📌 Cluster Means for k=4:  
 Age Income Level Coverage Amount Premium Amount  
Cluster\_k4   
0 33.60 121729.02 462271.73 3151.57  
1 35.31 52435.53 498413.00 3887.95  
2 45.48 66567.82 515044.48 1443.02  
3 61.68 90776.29 493821.18 3640.26

📌 Cluster Means for k=5:  
 Age Income Level Coverage Amount Premium Amount  
Cluster\_k5   
0 44.96 79841.22 785661.55 1780.72  
1 33.46 50346.87 477680.22 3866.99  
2 44.62 82097.21 260623.08 1590.68  
3 34.57 121479.80 471304.93 3777.04  
4 61.90 81235.16 482280.00 3868.27

📌 Cluster Means for k=8:  
 Age Income Level Coverage Amount Premium Amount  
Cluster\_k8   
0 56.89 113581.45 714942.62 1880.12  
1 32.92 114134.13 685197.04 3973.14  
2 32.97 113908.23 284255.05 1939.03  
3 55.76 55161.27 720687.52 3994.94  
4 54.95 52494.02 295560.04 1866.05  
5 33.48 55938.18 302435.85 3940.52  
6 57.46 110286.10 290355.05 3908.30  
7 33.33 54989.34 732120.81 1925.64

# 📌 Visualizing Clusters via PCA

To visualize high-dimensional clustering results, we used PCA to reduce the data to 2 dimensions and color-coded the clusters formed with k=8.



# ✅ Conclusion

Both Silhouette Score and Elbow Method provided insights into the optimal number of clusters. K=8 offered the best separation but depending on business context, K=4 or K=5 may also provide actionable groupings. PCA helped visualize the results effectively.