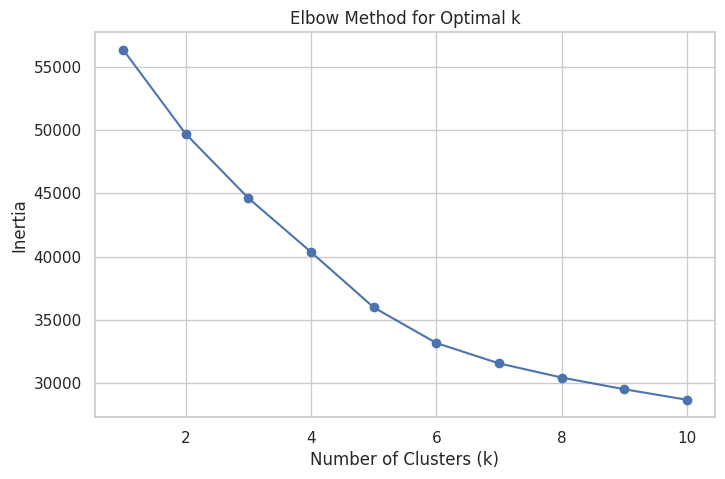
## Optimal Number of Clusters

The analysis to determine the optimal number of clusters in K-Means clustering used the Elbow Method on the scaled training data set. This technique entails computing the inertia (within-cluster sum of squared distances) over an array of cluster values (k) and graphing the results to determine where there is diminishing returns in the number of clusters.

Computations of the inertia values were done with k values of 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 and the resulting plot showed that there was a clear elbow at k = 4. Here is the maximum of the trade-off between the complexity of the model and the amount of variance explained, and additional increment in the number of clusters can deliver only slight improvements in cohesion.



The graph above shows the elbow shape plot which validates the fact that four clusters are the most meaningful structure in the data, which is not overfitted. This was the optimum value and was applied to fit the final K-Means model.