What is Inertia?

Inertia, also known as within-cluster sum-of-squares, measures the compactness of clusters. It calculates the total variance within the clusters. In simpler terms, it's the sum of the distances of each data point in a cluster to the centroid of that cluster, squared and summed up for all clusters.

Key Points:

- A lower inertia value implies a better model, as it indicates tighter clustering.
- However, the inertia metric has a drawback: it keeps decreasing with an increase in the number of clusters (k). This is where the "elbow method" is often used to find the optimal (k).

Understanding the Silhouette Coefficient

The **Silhouette Coefficient** is a measure of how similar an object is to its own cluster (cohesion) compared to other clusters (separation). The silhouette ranges from -1 to +1, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighboring clusters.

Key Points:

- A high silhouette score indicates well-clustered data.
- Unlike inertia, the silhouette score provides more nuanced insight into the separation distance between the resulting clusters.

When to Use Each Metric

1. **Inertia**:
    - Good for assessing the compactness of clusters.
    - Best when used with the elbow method to determine the optimal number of clusters.
    - More sensitive to the scale of the data, so normalization or standardization might be necessary.
2. **Silhouette Coefficient**:
    - Ideal for validating the consistency within clusters of data.
    - Useful when the number of clusters is not known.
    - Offers a more balanced view, incorporating both cohesion and separation.