Q1. Difference between K-Means and Hierarchical Clustering

K-Means partitions data into a fixed number of clusters using centroids and requires predefined k. Hierarchical clustering builds a tree of clusters without predefining k. Use case: K-Means for customer segmentation, Hierarchical for gene analysis.

Q2. Purpose of Silhouette Score

It measures how well a data point fits within its cluster versus others. Score ranges from -1 to 1. Higher is better.

Q3. Core parameters of DBSCAN

eps defines neighborhood radius, min_samples defines minimum points to form a dense region. They control cluster density and noise.

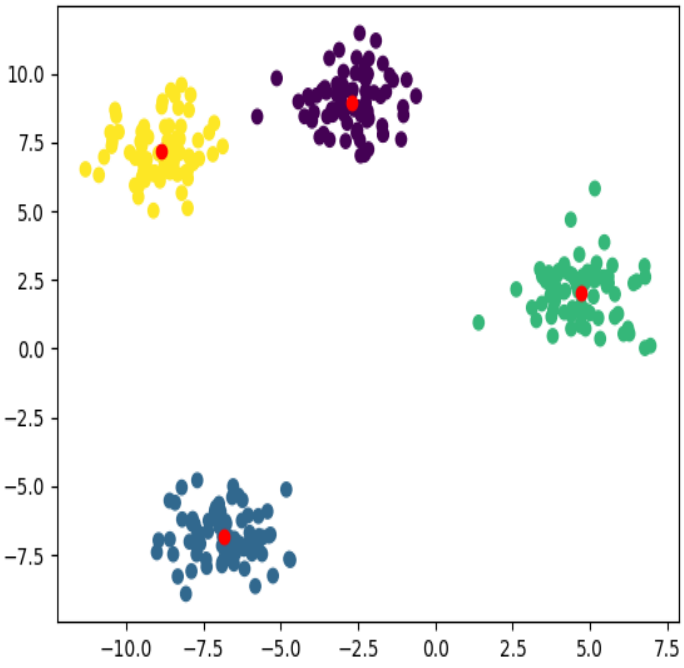
Q4. Importance of feature scaling

Clustering relies on distance. Scaling ensures all features contribute equally.

Q5. Elbow Method

It plots inertia vs k and identifies a point where improvement slows, indicating optimal clusters.

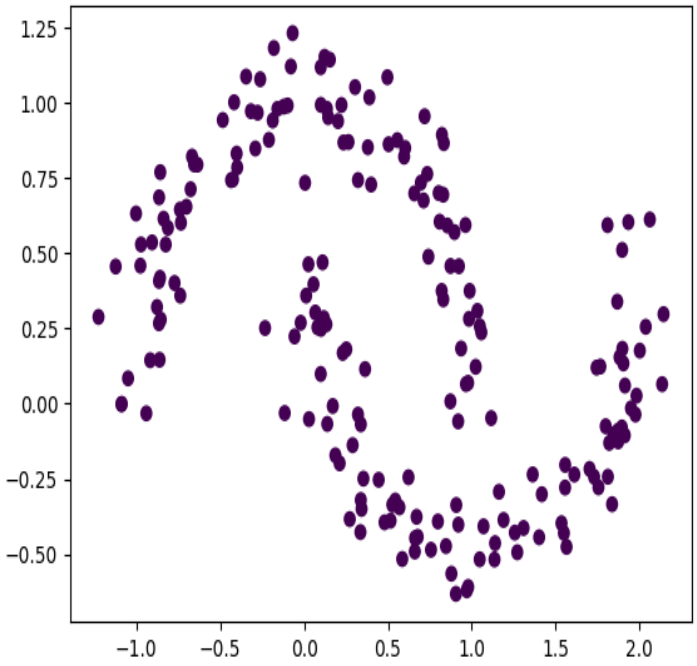
Q6. KMeans on make_blobs



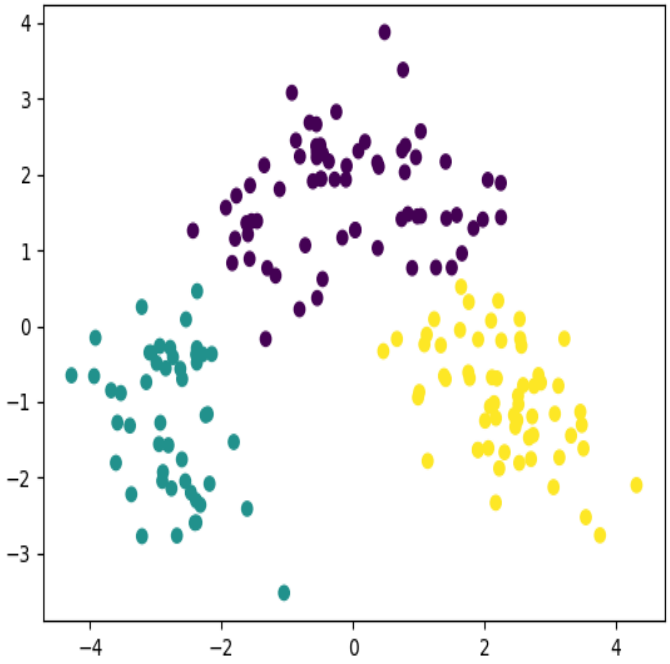
Q7. DBSCAN on Wine Dataset

Number of clusters found (excluding noise): 0

Q8. DBSCAN on make_moons



Q9. Agglomerative Clustering with PCA



Q10. Real-world clustering workflow

I would clean data, scale features, apply KMeans or DBSCAN, use Elbow/Silhouette to pick clusters, and help marketing target customer groups effectively.