Analysis of the Clustering Results

1. Elbow Method for Optimal k in K-Means

- The **Elbow Method** was used to determine the optimal number of clusters for K-Means.
- The inertia (sum of squared distances to the nearest cluster center) was plotted against different values of kk (from 1 to 9).
- Based on the elbow method, the optimal number of clusters was identified as **3**, since the inertia shows a noticeable drop at k=3k=3, after which the rate of decrease slows down.

2. K-Means Clustering Results

- After determining k=3k=3, K-Means was applied to the dataset.
- The scatter plot shows three distinct clusters, each represented by a different color.
- The **red 'X' markers** represent the centroids of each cluster.
- K-Means successfully grouped data points into well-defined clusters, aligning with the original data distribution.

3. DBSCAN Clustering Results

- **DBSCAN (Density-Based Spatial Clustering of Applications with Noise)** was applied with parameters eps=0.7 and min_samples=5.
- Unlike K-Means, DBSCAN does not require specifying the number of clusters beforehand and instead groups points based on density.
- The results show clusters of different shapes and sizes.
- Any points that did not fit into a dense cluster were labeled as **outliers (typically assigned -1)**.

4. Comparison of K-Means and DBSCAN

Aspect	**K-Means**	**DBSCAN**
Cluster Shape	Works well for spherical clusters	Handles arbitrary shapes well
Noise Handling	No explicit noise handling	Identifies outliers (labeled -1)
Requires k?	Yes, must specify kk	No, determines clusters based on density
Centroids	Yes, finds cluster centers	No centroids, density-based

5. Key Observations

- **K-Means performed well** because the dataset consists of well-separated clusters.

- **DBSCAN detected some noise** or smaller, less-dense clusters depending on eps and min_samples.
- If the dataset had more irregularly shaped clusters or varying densities, DBSCAN would likely perform better than K-Means.