CUSTOMER SEGMENTATION

VATSALYA MANDAGIRI

SANKEERTHANA CHALLA

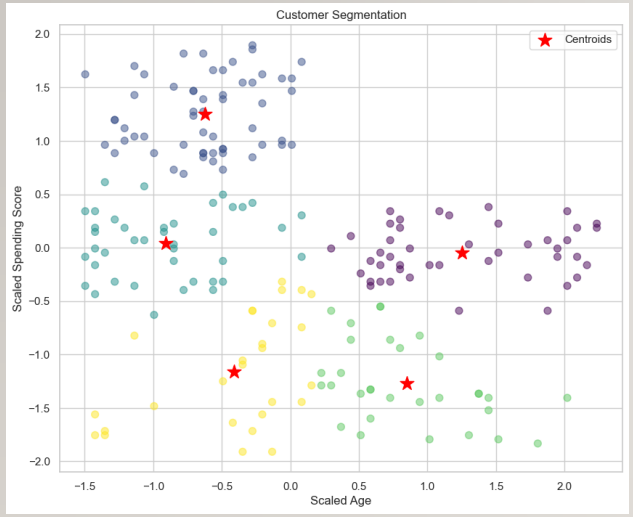
K-MEANS CLUSTERING

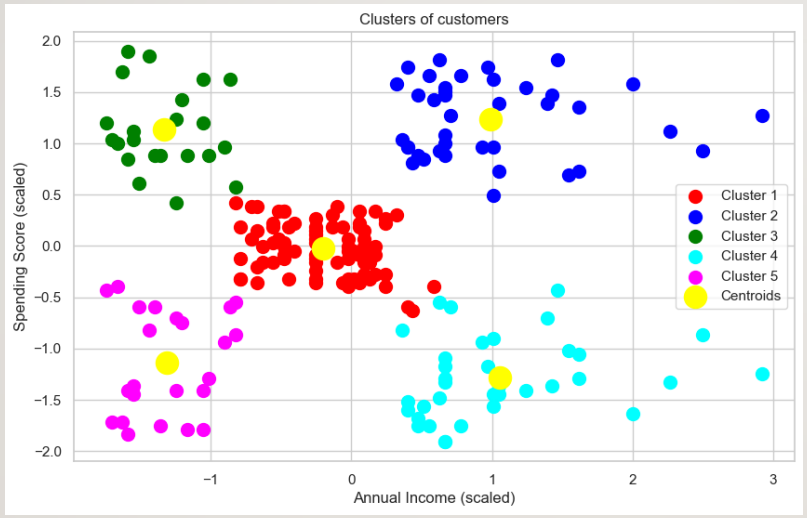
- K-Means Clustering is used to partition data into a set number of clusters.

How It works:

- It starts by selecting random points as initial cluster centers. Each data point is then assigned to the nearest center, and the centers are recalculated. This process repeats until the centers stabilize.
- K-Means is simple and fast, making it effective for clusters that are spherical in shape.
- However, it requires you to specify the number of clusters in advance and is sensitive to the initial choice of centers. It also doesn't work well for clusters of different shapes and densities.

VISUALIZATION





Left plot is a visualization of scaled age and scaled spending score and right plot is visualization of scaled annual income and scaled spending score

DB SCAN

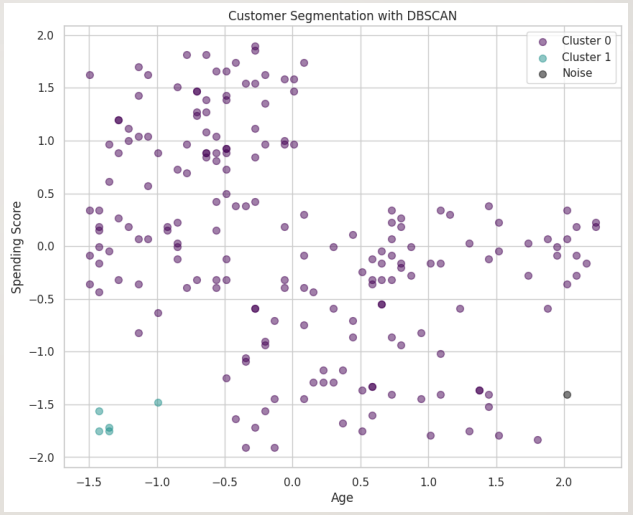
- DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a clustering method that can find clusters of arbitrary shapes and handle noise.

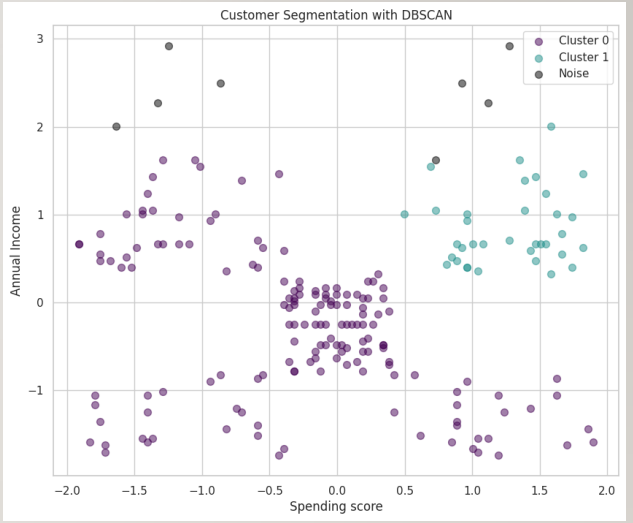
How it works:

- It uses two parameters: Epsilon (the maximum distance between two points to be considered neighbors) and MinPts (the minimum number of points to form a cluster). Points with enough neighbors are core points, forming clusters by connecting to other core points and their neighbors. Points that do not meet these criteria are marked as noise.
- DBSCAN does not require specifying the number of clusters and is robust to noise.
- However, its effectiveness depends on the chosen parameters, and it struggles with clusters of varying densities.

VISUALIZATION

Left plot is a visualization of scaled age and scaled spending score and right plot is visualization of scaled annual income and scaled spending score





THANK YOU