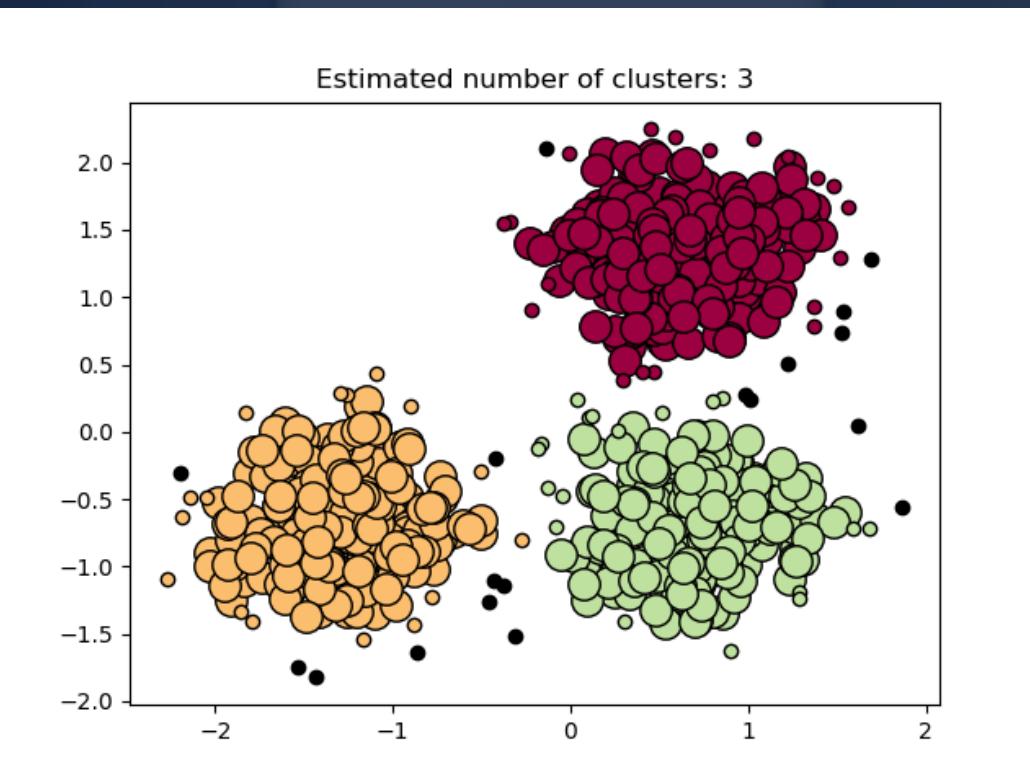


# DBSCAN

PAGE 01



DBSCAN is a clustering algorithm that groups points based on density. It forms clusters where data points are closely packed and marks isolated points as noise

# Advantages

PAGE 02

- No need to specify number of clusters
  - Unlike K-Means, you don't need to give K in advance.
- Finds arbitrary shaped clusters
  - Can detect non-spherical shapes
- Handles noise and outliers
  - Automatically labels sparse points as noise.
- Robust to small variations
  - Small movement in points won't break clusters.

# Disadvantages

- Struggles with varying density
  - If one cluster is dense and another is sparse, DBSCAN fails.
- High-dimensional data problem
  - Distance becomes meaningless in many dimensions (curse of dimensionality).
- Not good for very large datasets without indexing
  - Can be slow if data is huge.
- Border points ambiguity
  - A border point may belong to more than one cluster.