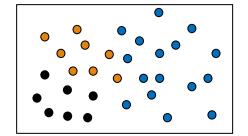
Clustering Tendency: Whether the Data Contains Inherent Grouping Structure

- Assessing the suitability of clustering
 - ☐ (i.e., whether the data has any inherent grouping structure)
- Determining clustering tendency or clusterability



- □ A hard task because there are so many different definitions of clusters
 - E.g., partitioning, hierarchical, density-based, graph-based, etc.
- □ Even fixing cluster type, still hard to define an appropriate null model for a data set
- □ Still, there are some **clusterability assessment methods**, such as
 - Spatial histogram: Contrast the histogram of the data with that generated from random samples
 To be covered here
 - □ **Distance distribution**: Compare the pairwise point distance from the data with those from the randomly generated samples
 - Hopkins Statistic: A sparse sampling test for spatial randomness