1. Clustering Methods

Method	General Characteristics
Partitioning methods	• Find mutually exclusive clusters of spherical shape • Distance-based • May use mean or medoid(etc.) to represent cluster center–Effective for small- to medium-size data sets
Hierarchical methods	• Clustering is a hierarchical decomposition (i.e., multiple levels) • Cannot correct erroneous merges or splits • May incorporate other techniques like micro clustering or consider object "linkages"
Density-based methods	• Can find arbitrarily shaped clusters • Clusters are dense regions of objects in space that are separated by low-density regions • Cluster density: Each point must have a minimum number of points within its "neighborhood" • May filter out outliers
Grid-based methods	• Use a multiresolution grid data structure • Fast processing time (typically independent of the number of data objects, yet dependent on grid size)

2. Typical Requirements of Clustering

The following are typical requirements of clustering in data mining

1. Scalability
2. Ability to deal with different types of attributes:
3. Discovery of clusters with arbitrary shape:
4. Requirements for domain knowledge to determine input parameters
5. Ability to deal with noisy data
6. Incremental clustering and insensitivity to input order:
7. Capability of clustering high-dimensionality data
8. Constraint-based clustering
9. Interpretability and usability

3. Cluster and Clustering Evaluation

A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. The process of grouping a set of physical or abstract objects into classes of similar objects is called clustering.

Clustering evaluation assesses the feasibility of clustering analysis on a data set and the quality of the results generated by a clustering method. The tasks include assessing clustering tendency, determining the number of clusters, and measuring clustering quality.