K-Means Clustering

Cluster Analysis

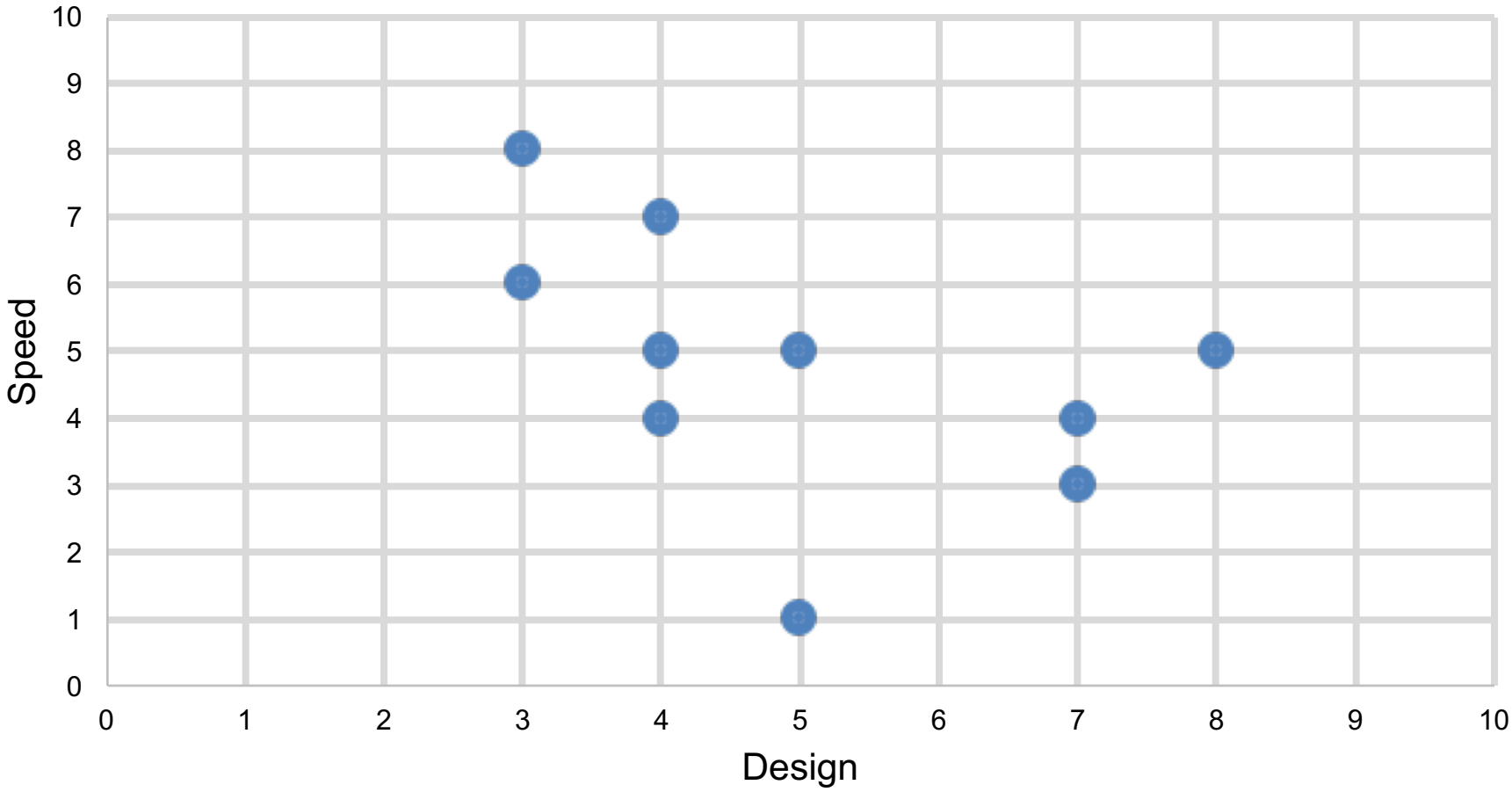
- Deriving Clusters and Assessing Overall Fit
  - Decide on clustering algorithm
    - Many algorithms are available
    - More coming all the time
- **Main goal**: maximize differences **between** clusters relative to variation **within** clusters

Customer Preferences

Customer	Design	Speed
1	5	1
2	7	3
3	4	4
4	7	4
5	4	5
6	5	5
7	8	5
8	3	6
9	4	7
10	3	8

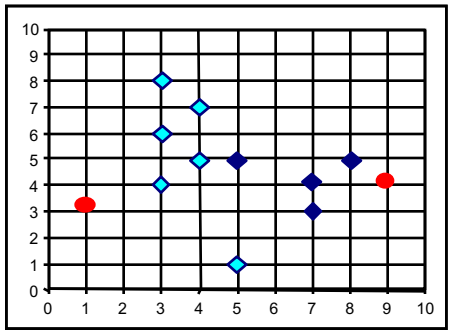
Preference on a 10-point scale, 10 most preferred and 1 least preferred for design and speed of a laptop

Customer Preferences



Nonhierarchical Clustering

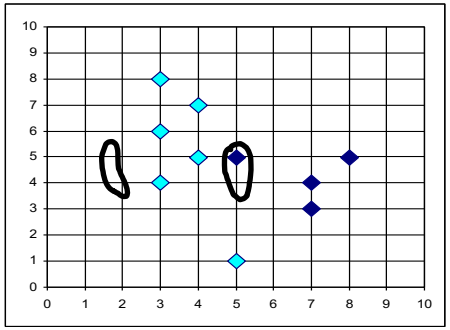
K-Means Clustering



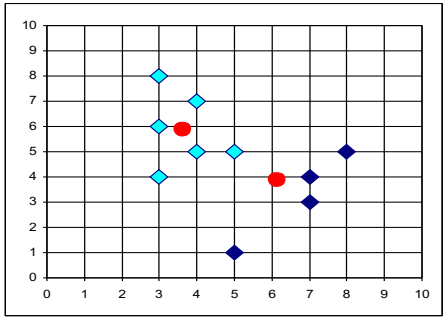
K = 2

Arbitrarily choose K object as initial cluster center

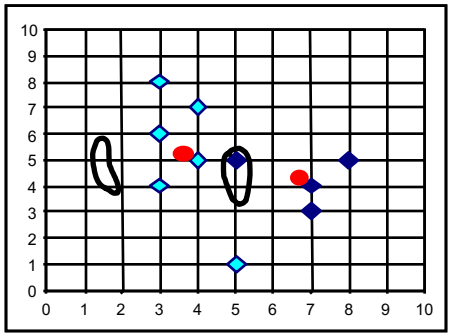
Assign each object to most similar center



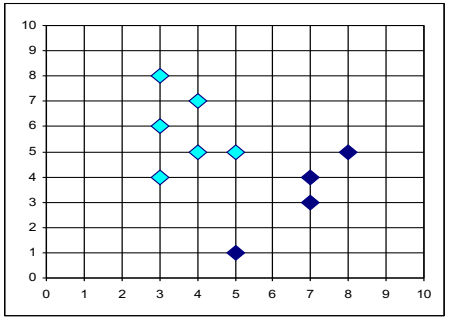
Reassign



Update the cluster means



Reassign



Update the cluster means