

Visualizing Cluster Results Using Package FlexClust and Friends

Friedrich Leisch

Department of Statistics, Ludwig-Maximilians-Universität München, Ludwigstrasse 33, 80539 München, Germany.
Friedrich.Leischr@R-project.org

Keywords: cluster analysis, visualization

Centroid-based partitioning cluster analysis is a popular method for segmenting data into more homogeneous subgroups. Visualization can help tremendously to understand the positions of these subgroups relative to each other in higher dimensional spaces and to assess the quality of partitions. In this talk we present several improvements on existing cluster displays using neighborhood graphs with edge weights based on cluster separation and convex hulls of inner and outer cluster regions. Using symbols or complete high-level plots in the nodes of the graph help to understand the association of background variables and clusters. A new display called shadow-stars can be used to diagnose pairwise cluster separation with respect to the distribution of the original data. Barplots of centroid profiles are improved by shading bars corresponding to statistically significant or user-relevant differences in darker colors. All methods will be demonstrated using real data from market segmentation and microarray data analysis.