

Multivariate Analysis: Final Exam

Due on 6/21/2018 in class

Announcement: Please follow the guidelines when writing your report:

- The presentation should be as formal as possible (with a title page, names and IDs, and necessary outputs and plots);
 - Please (i) clearly address the problem; (ii) clearly describe the basic ideas of each method you use; (iii) clearly introduce every step required to approach the answer; (iv) carefully interpret the results.
 - **Do not** attach any R code in your report;
 - Writing either in English or Chinese is fine.
- 1) The data set “[Seeds.txt](#)” describes seven measured geometric parameters of 205 wheat kernels which belong to three different types of wheat: Kama (Y=1), Rosa (Y=2) and Canadian (Y=3). Please download the file “[Seeds_Variables.txt](#)” for the detailed description.

Q1: Perform clustering analysis based on the attributes X_1, \dots, X_7 by using the following methods:

- (1) The hierarchical tree of Agglomerative Nesting by choosing the most adequate linkage;
- (2) The K-medoids;
- (3) The self-organizing maps (SOM).

Q2: Does the clustering result recover mostly the original wheat type for each of the above method?