STAT 503X Assignment 2
Classifying Flea beetles (10pts)
Due: in class Feb 22

This data is from a paper by A. A. Lubischew, "On the Use of Discriminant Functions in Taxonomy", Biometrics, Dec 1962, pp.455-477.

There are three species of flea-beetles: Ch. concinna, Ch. heptapotamica, and Ch. heikertingeri, and 6 measurements on each:
X1 width of the first joint of the first tarsus in microns (the sum of measurements for both tarsi).
X2 the same for the second joint.
X3 the maximal width of the head between the external edges of the eyes in 0.01 mm.
X4 the maximal width of the aedeagus in the fore-part in microns.
X5 the front angle of the aedeagus (1 unit = 7.5 degrees).
X6 the aedeagus width from the side in microns.

This assignment requires you to construct your own report on this data, using the template problem sheet handed out in the first class.
Specifically, I need you to address the questions below.

Exercise 1  1. Describe how to set up colors (and glyphs) in order to find classification rules for identifying the 3 species based on the 6 measured variables.

2. Explore the data for a strategy for visually constructing the classification rules.

3. Write down your final classification rules.

4. Compare the results with those of linear discriminant analysis.

5. Also pretend you forgot the species identity of the points, and have to start from scratch trying to find the clusters. Compare clustering methods to see which ones do a better job finding the true species identity.

The data is fairly clean, so you should not need to do too much restructuring. You will need to present your results in tabular and graphical form, and concisely summarize the major findings.
The grade for the assignment will depend on neatness of the report, how comprehensive is the analysis (are all possible questions answered), clarity of results presentation, and preciseness of final conclusions.