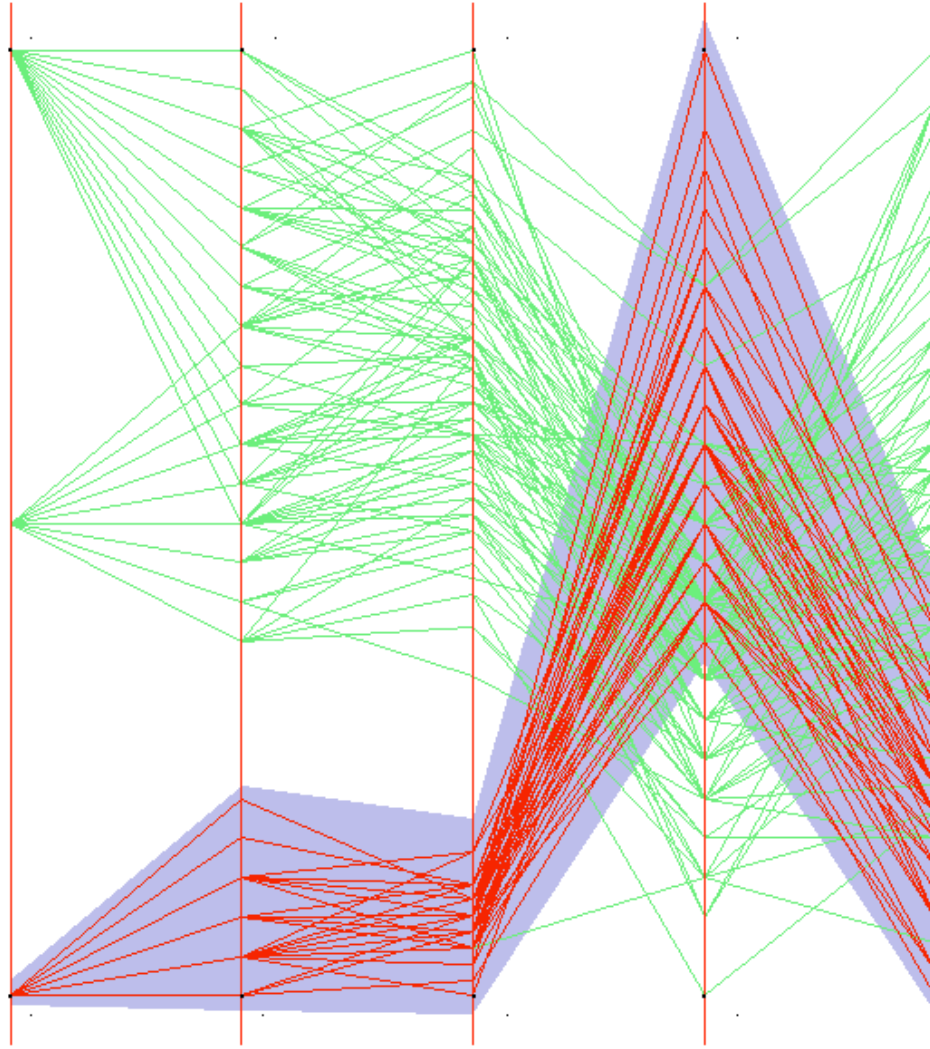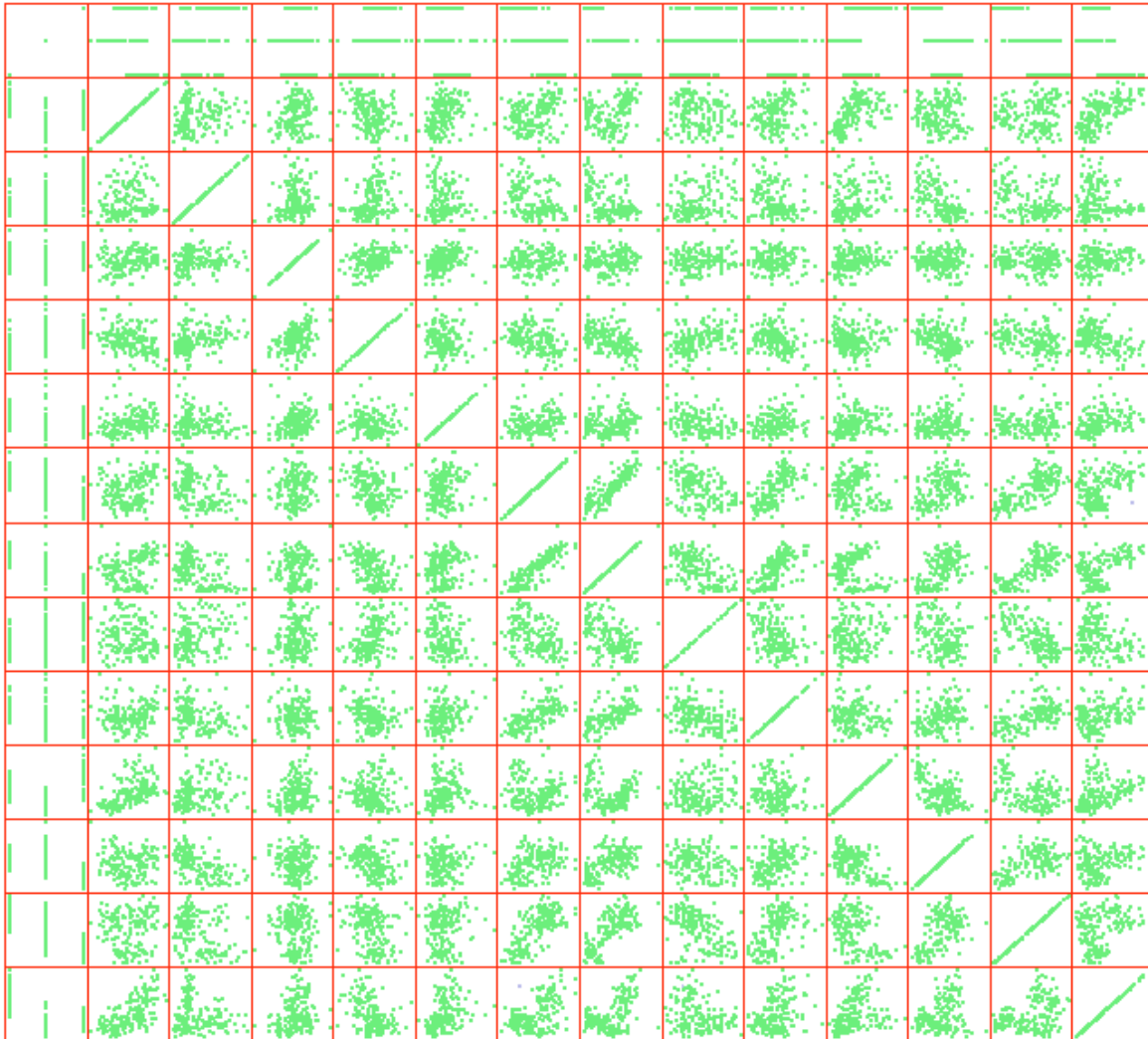**Iris Data**

Method used: Parallel Coordinates

For the Iris data I chose a Parallel Coordinates plot as the groupings of the lines show clear trends in the species data.

For instance, when exploring the data interactively, it quickly becomes apparent that Species #1 plants have a small petal width coupled with small petal length while Species #3 plants are the opposite , with large petal width and length. Conversely, Species #1 tend toward large sepal widths coupled with small sepal lengths while Species #3 tends toward the opposite of small sepal widths with large sepal lengths.

In short, one quickly sees that the species 1, 2, and 3 groupings have radically different trends in petals and sepals, particularly that Species #1 has much smaller petals.

CSCIE 236 – Homework 3, Q4, Information Visualization
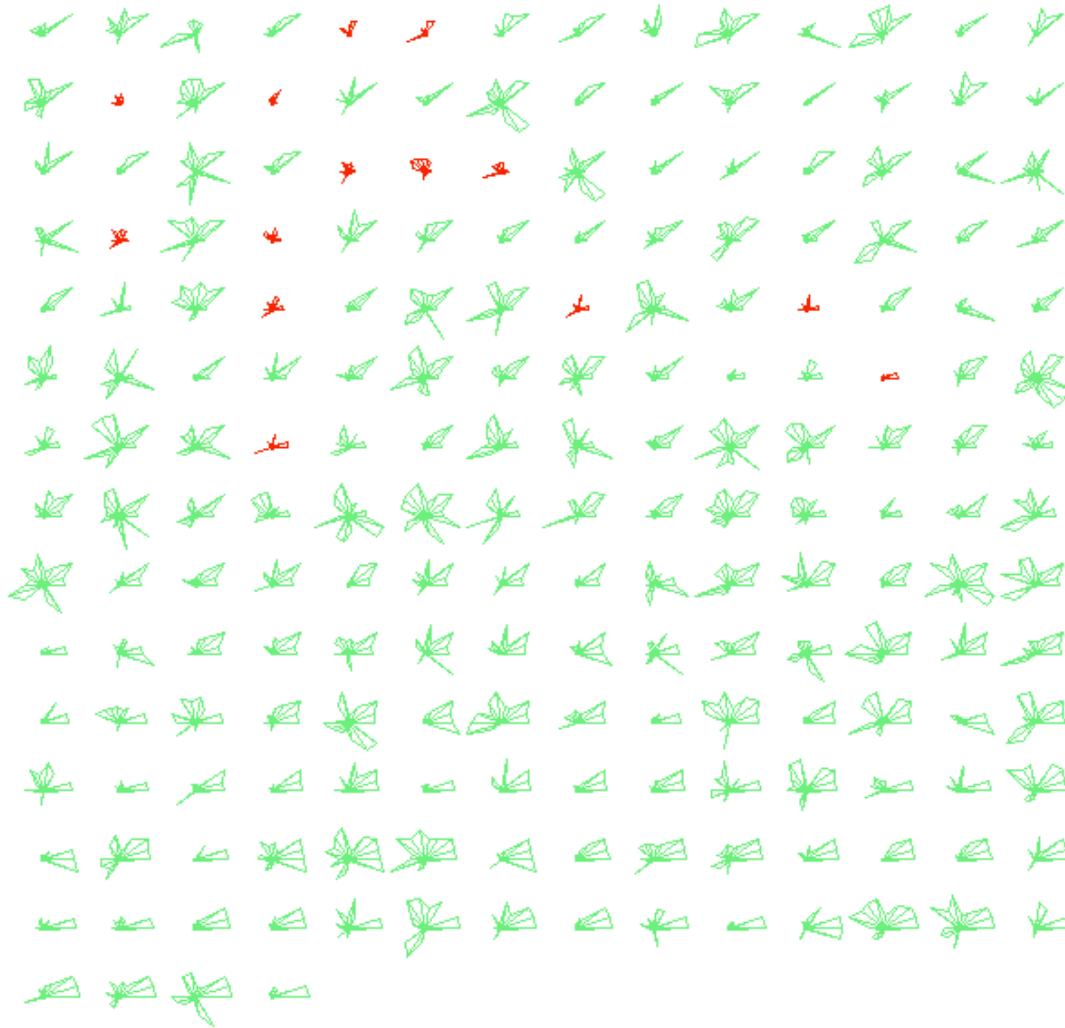Russell Lowke April 19th 2005

Wine Data

Method used: Scatterplots

I have used a Scatterplot in the case of the wine data as it makes it easy to look and see correlation between each of the fourteen elements with each of the other fourteen. For instance, there is a strong correlation between "Malic Acid" and "Hue," a high Malic Acid tending toward a low Hue. Similarly, other correlation/tends can quickly be spotted; such as between "Alcohol" and "Proline"; "Flavanoids" and "Proline"; and "OD280/OD315" and "Flavanoids."

To a wine chemist looking for various trends a Scatterplot such as this is a valuable analyzing tool.
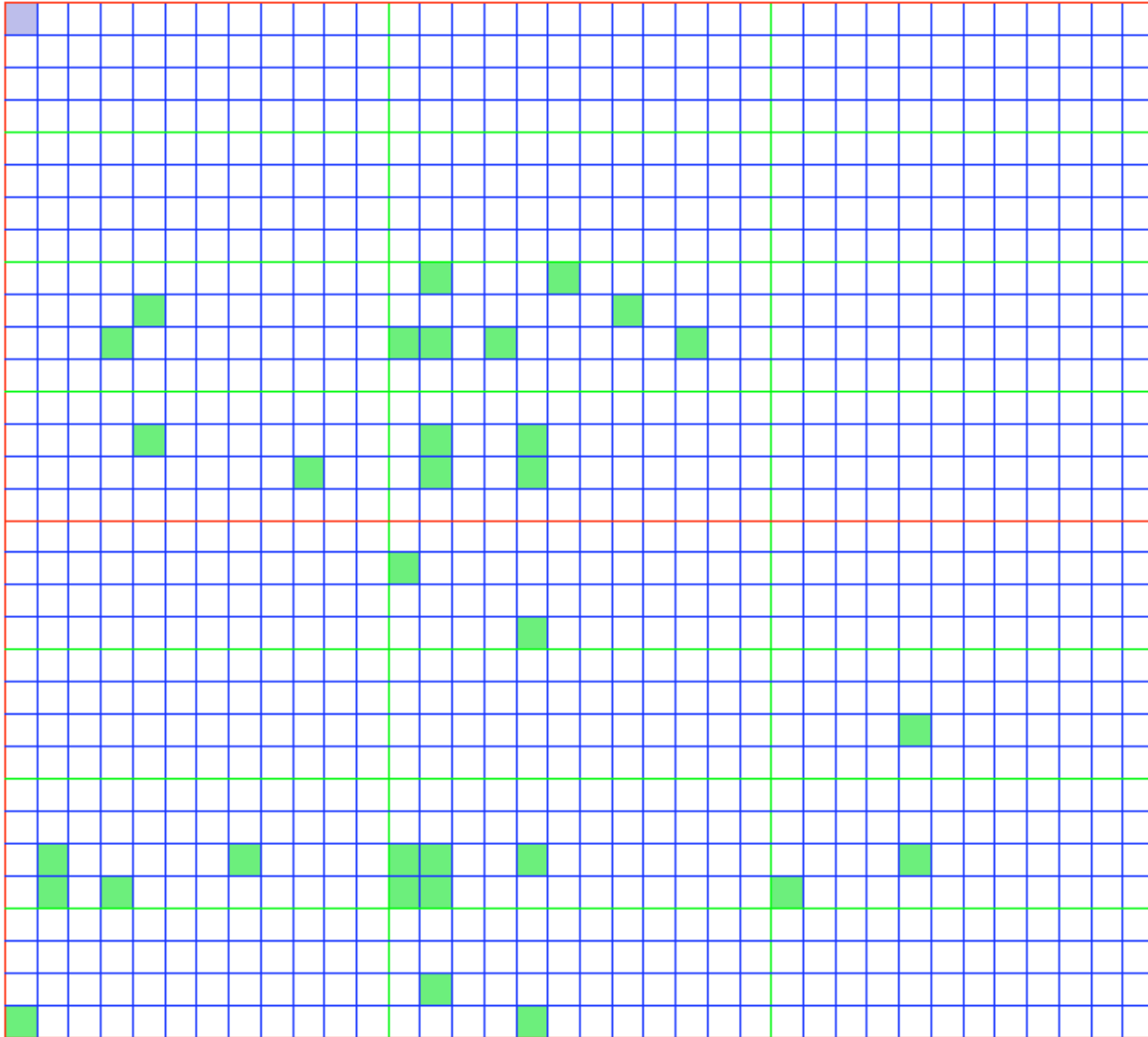
Charles Book Club Data

Method used: Star Glyphs

       When analyzing the Book Club data one is confronted with many categories all with widely differing scopes, making it hard to digest. Regardless, the Star Glyph plot shows trends in the book data that are visible due to size and shape of the glyph. For instance, less popular books may be spotted immediately by the fact that they tend towards small glyphs. Children's books on the other hand bulge down and to the right, while cook books can be seen to bulge down and to the left. Common data attributes may be isolated simply by glyph shape and as such complex data sets such as the book club data become more decipherable. Interactively selecting glyphs and then checking other plot types is also useful to reveal trends.

Flying Fitness Data

Method used: Dimensional Stacking

I chose to use Dimensional Stacking to analyze the Flying Fitness data as the categorical variables used in the data set have a limited number of values, particularly the output variable, "Var 1", which is either 0 [false] or 1 [true], for the crucial information of whether the pilot passed or failed his fitness evaluation.

By setting the cardinality of "Obs" = 1 and "Var 1" = 2, one can immediately see in the graph which pilots passed and which failed, as those who passed are in the top "RED" bin above the red horizontal line and those who failed are below it.

It is also visually evident that a high (> 1.7) "Var 2" score, as indicated by placement in the right hand third, or a high (>2) "Var 3" score, as indicated by placement in the upper third of the red bin is an attribute of a failed pilot.

header used or FlyingFitness.okc

7 40
Obs
Var1
Var2
Var3
Var4
Var5
Var6
1 40 1
0 1 2
0 2 3
0 3 4
0 3 4
0 3 4
0 2 3