Quick cluster analysis of Thomas Weber's data set
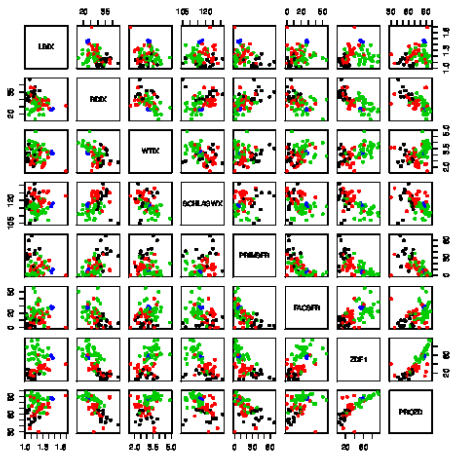
Christian Hennig

October 31, 2012
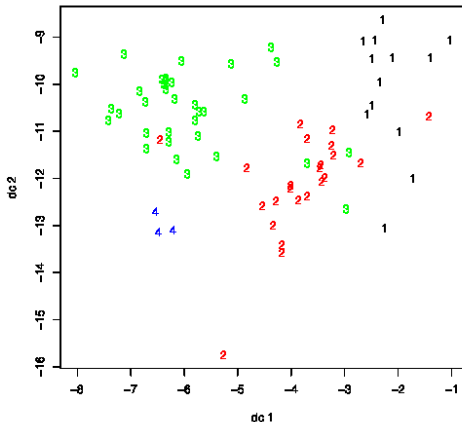
Variables used:
LBIX, RDIX, WTIX, SCHLAGWX, PRIMSFR, FACSFR, ZDF1, PROZD
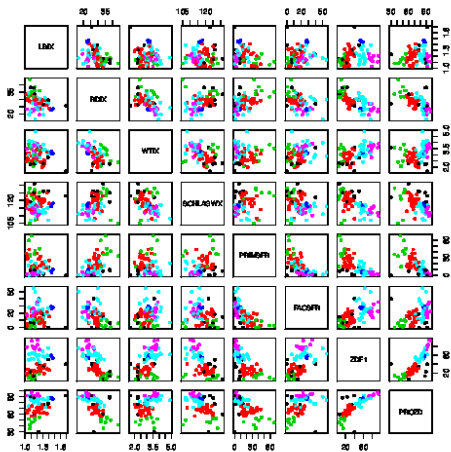
Matrix plot of original groups:

Discriminant coordinates of original groups
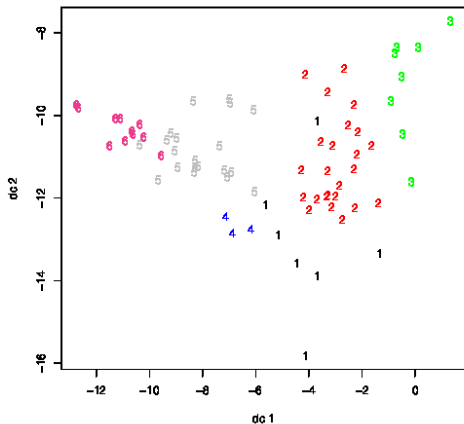(plot that shows group means optimally separated)

Clustering by R-package mclust,
fitting a Gaussian mixture model
optimising over all models by BIC.

Optimal model 6 clusters, VEI
(covariance matrix diagonal,
varying volume, equal shape)

Matrix plot of mclust clustering:

Discriminant coordinates of mclust clustering:
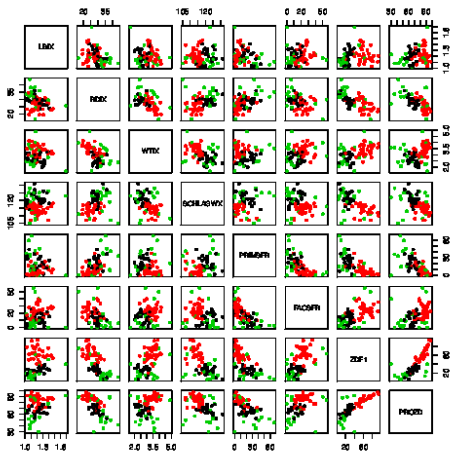
Original groups vs. clusters

```
table(groups,smpoprc$classification)
groups  1  2  3  4  5  6
     1  1  5  7  0  0  0
     2  6 15  1  0  1  0
     3  0  4  0  0 19 11
     4  0  0  0  3  0  0
```

mclust clustering with "noise component",
classifying some points as "not belonging to any cluster".

Optimal 2 clusters plus noise, EEE
(ellipsoidal, equal volume, shape, and orientation)

Matrix plot of mclust clustering with noise (noise is green):

Original groups vs. clusters (noise is 3)

```
> table(groups,npopm$partition)

groups   1   2   3
     1   6   0   7
     2  14   2   7
     3   2  29   3
     4   0   3   0
```