Sheet 1

Hand in solution by February 28 in the lecture room

The goal is to get an overview over graphical methods in R.
To make sure you reach this goal, defer solving the problems marked with (*) until the end.

In the exercises for this course we will often come back to a basic dataset. It consists of a data collection on a mountain meadow called Alp Flix in Ticino. For 82 "plots", there are

- the abudances for 64 plant species,
- 4 measurements of soil properties for each of 3 depths,
- 2 variables describing the intensity of grazing,
- some additional variables.

A detailed description can be found on `http://stat.ethz.ch/~stahel/courses/multivariate/datasets/vegenv.html`

The data are made available to R by typing

```
d.vegenv ←read.table(
"http://stat.ethz.ch/∼stahel/courses/multivariate/datasets/vegenv.dat",
header=TRUE)
```

1. Try the possibilities of the R function `pairs` for the soil variables `pH10, P10, N10, C10` in the dataset `vegenv`!

a) Generate a scatterplot matrix. What do you spot?

b) You can display the relationship with grazing by coloring points: Set the argument `col = ifelse(d.vegenv$Dungdensity>10, "red", "blue")`. This relates red to intensive, blue to extensive grazing. Which variables show noticeable differences?

c) (*) Examine the help file for `pairs`. You will find the functions `panel.hist` und `panel.cor`. Alternative versions which allow for using the argument `col` from b) are available in the file `panel-fn.R`. Copy these function into your R script file and then into the workspace. `panel.hist` is used to display a histogram of each variable in the diagonal of the scatterplot matrix. `panel.cor` calculates the (Pearson) correlation for each pair of variables and displays this numerical result. Use these functions in your call to `pairs`. The examples in the help file are informative!

2. **a)** Display a coplot of `P10` against `pH10`, given `Dungdensity` and `Slope`. Anything worth mentioning?
R Hint: `coplot(P10 ~ pH10 | Dungdensity+Slope, data=na.omit(d.vegenv))`

b) Add a smooth line by using the argument `panel = panel.smooth` .

3. (*) Study the help file of the function `symbols`.

Generate a star plot of some of the abundancies of the species, using `P10` and `N10` as the axes of the plot.
Are there relations between the two soil variables and any species abundancies?
R Hints:

```
symbols(log10(d.vegenv[,"P10"]),d.vegenv[,"N10"],
        stars=as.matrix(d.vegenv[,c("Nardstri","Caluvulg","Festrubr")]),
        xlab="log10(P)",ylab="N", inches=0.5)
stars(as.matrix(d.vegenv[,c("Nardstri","Caluvulg","Festrubr")]),
      location=cbind(log10(d.vegenv[,"P10"]),d.vegenv[,"N10"]), axes=TRUE,
      key.loc=c(-0.5,0.3),len=0.07,full=TRUE, xlab="log10(P)",
      ylab="N", labels=NULL, xlim=c(-1.5,-0.5))
```