Series 1

1. We investigate graphically the R internal dataset `swiss` which you can load by `data(swiss)`. The data contains the variables

`Fertility`	common standardized fertility measure
`Catholic`	number of catholics
`Agriculture`	% of men working in agriculture enviroment
`Examination`	% draftees receiving highest mark on army examination
`Education`	% education beyond primary school for draftees
`Infant.Mortality`	live births who live less than 1 year

of 47 counties in the west of switzerland dated at 1888. With `?swiss` you get more information on the meaning of the variables.

a) Read the help file of `stars()`.

b) Make a star plot of all variables. What can you say about Sierre?

R-Hint:
```
data(swiss)
stars(as.matrix(swiss), ...)
```

c) We are interested in the relation between `Fertility` and `Education`. Therefore we would like to make a scatter-plot of `Fertility` against `Education` whose points are stars with the information of the other variables. In addition we need the argument `location`.
R-Hint:
```
stars(as.matrix(swiss[, c(2,3,5,6)]),
      location = as.matrix(swiss[, c(4,1)]),
      axes = T, ...)
```

d) (*) Set the argument `draw.segments` to `TRUE` to get segments instead of stars. Place a legend with `key.loc`.

e) Which relation do you get from the plots?

2. The data set gives the latitude, longitude, depth, magnitude and number of reporting stations of 1000 seismic events of Mb > 4.0 (body wave magnitude). The events occurred in a cube near Fiji since 1964.

a) Load the data set saved in quakes.csv
R-Hint:
```
url <- "http://stat.ethz.ch/Teaching/Datasets/WBL/quakes.csv"
x <- read.csv(url, header = TRUE, row.names = 1)
```

b) Does the magnitude of the earthquake depend on the depth? (Make a scatterplot)

c) Does the number of reporting stations depend on the magnitude? (Make a scatterplot)

d) Investigate the relationships between all variables in the data set using a parallel coordinate plot and a scatter plot matrix. Which method do you find more useful?

e) How does the depth depend on longitude and latitude? (Plot a point (pch = 20) at the position of the earthquake; the color should be green, orange or red according to the depth)
R-Hint: `deepVec <- cut(x$depth, breaks = c(0, 250, 450, 700), labels = c("green", "orange", "red"))`
`deepVecString <- as.character(deepVec)`

3. The data set `Titanic` provides information on the fate of passengers on the fatal maiden voyage of the ocean liner 'Titanic', summarized according to economic status (class), sex, age and survival. The data set is available in R.
R-Hint: `dfTitanic <- as.data.frame(Titanic)`.
Load the package vcd: `library(vcd)`

a) Is there a signigficant connection between survival and class?
R-Hint: `?mosaic`, `?structable`

b) In a sinking ship women should leave the ship first. Was this the case on the Titanic?

c) In a sinking ship children should leave the ship first. Was this the case on the Titanic?

d) Do you think, that the "women-first policy" also holds true in all classes?
R-Hint: `?cotabplot`

Preliminary discussion: 25.02.13.

Deadline: No hand-in.