Data Analysis with R, fall 2010

Demonstrations 1

1. Read in .txt-data *fuel.txt* from the course web page to R. You'll have to end up with a data-frame, where *year* variable is consider as a factor and *carbon* variable an integer.

2. Calculate mean, standard deviation (SD), minimum and maximum for the following variables (the datasets here are attached by default when R is started)

    a. Column *conc* (concentration) from the data frame called *Indometh*
    b. A Random sample of 100 from the standard Normal-distribution [rnorm(100)]
    c. The sum of columns *Vic* and *Qld* from the data frame called *austpop*

3. Create a vector x, which contains 100 random values drawn from the standard normal distribution. Code for this is given for you below.

```
x <- rnorm(100)
```

    a. Form a vector which contains the entries of *x* at the positions 3, 46 and 79?
    b. Form a vector which contains all the entries of *x* except the last and the second last?
    c. Create a logical vector *n*, whose i'th entry is TRUE if and only if the i'th entry of *x* is greater than -1.5.
    d. Select those entries of x which are greater than -1.0 and less than 1.5?

4. Create a categorical variable *warm* to dataset *airquality,* which gets values based on following:

    - 1, if Temp < 76
    - 2, if 76 <= Temp < 81
    - 3, if Temp >= 81

    Now calculate mean of *Wind* in the groups defined by the created variable *warm*.

5. Extract the following subsets from the data frame *ais* (*library(DAAG)*):
    a) Extract the data for the rowers.
    b) Extract the data for the rowers, the netballers and the tennis players.
    c) Extract the data for the female basketballers and rowers.