Lab 6: ANCOVA

A. Forbes 500 Companies (1986)

The dataset **Forbes.xlxs** holds data about 40 companies selected from the Forbes 500 list for 1986.

Sector is coded as 2 = Energy, 3 = Finance, 4 = HiTech

It is of interest to investigate the relationship between Sales (Dependent) and Assets (Independent) to see if there are differences between the companies in the different Sectors.

(i) Determine the nature of the distribution of the variables Sales and Assets.
(ii) Suggest a transformation to normalise the data
(iii) Use ANCOVA to investigate the relationship between transformed sales, assets and the categorical variable sector. What are your conclusions on the analysis?
(iv) Create a scatterplot to illustrate this relationship.

Source: http://www.statlab.uni-heidelberg.de/data/ancova/Forbes500.html

B. ANCOVA with two factors

The data in file **Clinical Trial.xlsx** results from a Clinical trial with baseline symptoms as a covariate. The trial investigated the relative effectiveness of cognitive therapy (1 = yes, 0 = No) and an antidepressant drug (1 = yes, 0 = no) on the treatment of depression. Baseline and Six Weeks Post treatment scores were recorded.

(i) Analyse this data using R and report on the results.
(ii) What conclusions can you make on the basis of this analysis?

C. Cholesterol in Iowa vs Nebraska

The data in file **Cholesterol.xlsx** are the results from Cholesterol levels [mg/ml] measured for 30 women from two US states, State is coded 1 = Iowa and 2 = Nebraska. Age [years] may be a relevant covariate. Is there evidence of a difference cholesterol levels between the two regions? Check whether the assumptions of ANCOVA are satisfied by the model fitted to the data.