


STATISTIK 2 (107.325) WS 2014
COMPUTERSTATISTIK (107.258) WS 2014

Übung 1

5. Oktober 2014
Dutter

Analysieren Sie die Daten “Werner Chemistry Data“ mit dem Computer Program System . Das Originalfile *b.2d.inp* (siehe <http://www.statistik.tuwien.ac.at/public/dutt/vorles/>) enthält als Kopfzeilen die Steuerbefehle für das Programm BMDP, was nicht weiter wesentlich ist, aber die Bezeichnung der Variablen sind enthalten. Das File *werner_bcd.txt* enthält nur die nackten Zahlen (allerdings mit fehlenden Werten kodiert durch „*“).

Das Importieren der Daten könnte mit den R-Befehlen innerhalb einer Funktion so ausschauen:

```
"f.ueb1" <-
function ()
{
# werner_bcd <- read.table("werner_bcd.txt")
werner_bcd <- matrix(scan("werner_bcd.txt",na.strings="*"),ncol=9,byrow=T)
dimnames(werner_bcd) <- list(werner_bcd[,1],
  c("ID", "AGE", "HEIGHT", "WEIGHT", "BRTHPILL", "CHOLSTRL", "ALBUMIN",
    "CALCIUM", "URICACID"))
}
```

Dabei werden schon Namen der Variablen (Spalten) und Zeilen (Patientennummern) vergeben.

- 1 Wählen Sie eine Variable aus den “Blutkenngößen“ und zeichnen Sie ein Histogramm mit einem “eindimensionalen Streuungsdiagramm“.
- 2 Berechnen Sie Schätzungen für Lokation (Ort), Streuung, Form (Schiefe und Kurtosis) und beschreiben Sie verbal die Datenverteilung.
- 3 Berechnen Sie den Korrelationskoeffizient mit den anderen Variablen. Wählen Sie eine (sinnvolle) zweite Variable, zeichnen Sie ein Streuungsdiagramm (xy-plot), und schätzen Sie eine Regressionsgerade. Kommentar!
- 4 Machen Sie eine Boxplotdarstellung über möglichst alle Variablen, diskutieren Sie das Resultat.

Bitte, stellen Sie die Ausarbeitung in Form eines pdf-Files (nicht mehr als 3 Seiten) mit den Resultaten (Outputs plus textliche Kommentare) und Kurz-Listing des Programmcodes (Funktion) in die TUWEL-Seite

<https://tuwel.tuwien.ac.at/course/view.php?idnumber=107258-2014W>
oder schicken Sie es einfach an

R.Dutter@tuwien.ac.at

bis incl. 17. Oktober 2014.

Bitte den Namen des pdf-File folgendermaßen:

name_exer_1.pdf

wobei 'name' für den Familiennamen steht. Bitte im pdf-File Name, Datum und Seitennummerierung nicht vergessen!

Empfehlenswert ist es, den Bericht mit „Sweave“ zu erstellen. (Vorlage siehe <http://www.statistik.tuwien.ac.at/public/dutt/vorles/> .)