Exercise Sheet 3

Problem 1. Factors
Consider the data set `data7.csv` which contains the following variables

`PAT` Patient id number

`ENTRY_D` Date of entry into the study

`Birth` Date of birth

`DIAS` Diastolic blood pressure

`SYS` Systolic blood pressure

`CO_DIABETES` Diabetes status: 1 for diabetes, 0 otherwise

`CO_LIVER` Liver disease status that equals 1 for mild liver disease, 2 for severe liver disease and 0 for no liver disease

`SEX` Gender; 0 for female and 1 for male

(a) Create a factor `Diabetes` with levels `Yes` and `No` with reference value `No`

(b) Create a factor `Liver_Disease` with levels `Mild`, `Severe` and `No` with reference value `No`

(c) Create a factor Gender with levels Female (sex == 0) and Male (sex == 1).

(d) Fit a linear model (function `lm`) with diastolic blood pressure as outcome variable and gender as explanatory variable. Display the results in a nice table.

(e) Change variable `Gender`'s reference level to Male and refit the model. What has changed?