Solutions theoretical exercises for STK4900/9900.

Exercise 5 We see that x_1 (mass type), x_3 (pressure) and their interaction x_1x_3 have large (absolute values of) t-values and large contributions to the R^2. These variables should clearly be kept in the model.

The interaction between x_2 (temperature) and x_3 is also significant and has a fairly large R^2, thus we should likely also keep x_2x_3 in the model. Some statisticians prefer to keep main order variables in the model whenever interactions or second order terms are entered. They would then also include the main effect x_2 although it is insignificant and has a small predictive power.

The second order term x_3^2 (and to smaller degree x_2^2) might also be considered for inclusion in the model since it has a p-value lying between 0.05 and 0.10 (the exact value depends on the number of variables included in the model).

The remaining variables are insignificant and will not improve prediction, and could be kept out of the model.