```
## Credit scoring: a case study
##
## When issuing credit, banks check the "solvency" or "creditworthiness" of the
client. "Solvency" can be evaluated by a statistical model that tries to predict
the probability that a credit is payed back from knowing some economic or
personal characteristics of the borrower. Binary regression models are suited
for this situation.

## Data (in credit1.txt)
# (y=1, the client did not pay back his loan)

credit <- read.table("credit1.txt", header=TRUE)

library(dplyr)

# look at structure

credit %>% ...
credit %>% ...

## To make this variable directly available in the sequel

credit %>% ...

## Exploration

# How many clients did not repaid the loan?

y %>% ...
y %>% table %>% barplot

## Exercise
##
## Carry out an exploratory analysis of the covariates.

...
...
...
...
...
...
...
...
...
...
...
...
...
...

## Logistic Regression

mod1 <- glm(y ~ acc+duration+amount+moral+intuse,family=binomial(link=logit))
summary(mod1)
# comment on the above results

exp(mod1$coefficients[2:3])
# what is the interpretation of the above result?

## A more complicated model
```

```
# Consider a parabolic effect in the new model (continuous covariates squared)

mod2 <- ...


## Exercise (later)
##
## Test if the two new variables, i.e.\ adding a parabolic effect of the two
quantitative covariate, have significantly improved the model (hint: compare the
deviances of the two models).


## Prediction

## Consider a client with the following characteristics: It has no running
account, the term for repaying the debt is 36 months, the amount is 10000 euros,
previous payment beahviour was bad and the money are intended to be used for
business. How risky is to give a loan to him? The predicted probability of
insolvency is:

newclient <- data.frame(...,...,...,...,...)
predict(mod2,newdata=newclient,type="...")


# When you finish your R session Remember to detach the dataframe.
```