# Multiple Linear Regression in R

# (predicting income)

This exercise uses file Income\_data\_categ.csv.

## Tasks to be performed:

1. Find out how income depends on age, education, smoking and residence.
2. How is the categorical variable Residence represented in the model?
3. Find out if one of the independent variables is insignificant and rebuild the model without it.
4. Compare the quality of the model (explanatory power) and see if removing that variable made the model any worse.
5. Initialise the input for several people and compute the predictions for their income.
6. Compute the income for a new-born baby (age = 0, education = 0, choose any residence). Are you surprised by the result? Can you explain it?