STATISTICS 3
WS 2019 (Mag. Thomas Forstner)

Course-Number: 366.542

21) Fit an appropriate regression model to predict the infant mortality rate based on the GDP per capita.

Use the built in R data frame "UN" from the package "car". The observations are nations of the world and the variables are infant mortality rate (infant deaths per 1000 live births) and GDP per capita (U.S. dollars).

22) Estimate an appropriate regression model for "Had an affair [no/yes]" based on the so called "Fair's Affairs Data". The data is based on a survey conducted by Psychology Today in 1969. A shortened version of this data is available on the course homepage ("affairs.txt").

Create the binary dependent variable based on the variable y "number of affairs in the last year"

Description of the independent variables in the shortened dataset:

z1 … gender (0 = female, 1 = male)
z2 … Age in years
z3 … Number of years married,
z4 … Children (0 = no, 1 = yes)
z8 … Self rating of marriage (1 = very unhappy, 2 = somewhat unhappy, 3 = average, 4 = happier than average, 5 = very happy).

Fair's Affairs Data
The complete unmodified data set is available under the following address:
http://pages.stern.nyu.edu/~wgreene/Text/tables/tablelist5.htm (Table F22.2).

In the unmodified dataset age and years married are represented as categorical variables but for the above examples both variables are treated as numeric.

Sources for the dataset:

- Greene, W.H. (2003). Econometric Analysis, 5th edition. Upper Saddle River, NJ: Prentice Hall.
- Fair, R.C. (1978). A Theory of Extramarital Affairs. Journal of Political Economy, 86, 45–61.