

## Lecture “Introduction to Stata”

### Problem Set 1

This problem set should help to practice what you have learnt so far. You analyze data on characteristics on cell phones. The data set is provided in the Internet as file “handys.dta”. The variables in this data set are the following:

- preis: price of the mobile in €
- gewicht: weight of mobile in kg
- strahl: radiation pollution of the mobile in SAR (amount of absorbed radio signals when using the mobile)
- ruftoene: number of ringing tones that can be used
- funkakku: duration of maximal conversation time provided by the mobile’s battery in minutes
- lnpreis: logarithm of the price in €

Solve the following problems and answer the questions. Use a do file and log file for this purpose.

1. Open the data handys.dta in Stata.
2. Calculate the descriptive statistics of all variables in the data. What is the average and what is the median price of the mobiles?
3. How many observations on cell phones do the data provide?
4. Generate a variable obs with consecutive numbering from 1 to the maximal number of observations (hint: type `help _n`)
5. Change the name of variable obs to id.
6. Erase variable lnpreis. Then generate a variable price\_ln which represents the logarithm of the respective prices (as lnpreis before).
7. How many mobiles have a price above €250?
8. Calculate the descriptive statistics for the group of mobile phones with a price not higher than €250 and the group of mobiles with a price above €250, respectively.

9. Produce a new variable ring that equals 1 if the mobile provides the user with at most 15 ringing tones, 2 if it is 16 to at most 40 ringing tones, and 3 if it is more than 40. What is the mean price and the minimum as well as maximum price in the respective ring-group?
10. Now generate a new variable price\_segment with the price groups 1 = between 0 to 125 €, 2 = 125 to 250 €, 3 = 250 to 500 €, and 4 = more than 500 €. Then produce a cross table for price\_segment and ring.
11. Produce a variable that covers the information weight squared.
12. Are there any missings?
13. Generate a figure with a scatterplot for the variables logarithmized price and weight: log price at X axis and weight at Y axis.
14. Reproduce the figure and add into the same figure the estimated bivariate regression line (regression of log price on weight).
15. Finally, estimate a hedonic price function for mobile phones, that is, regress log price on weight, weight squared, radiation pollution, number of provided ringing tones, and battery performance. Interpret and conclude. Drop some regressor(s) and observe the effect. Interpret!