QMB: Exploring Housing Rents

New Real Estate attacked by local newspaper

In Seldwyla, a city located in Northwestern Switzerland, the local newspaper Seldwyla Herald attacked the property holder **New Real Estate** (NRE) for overpricing housing rents in a difficult period for tenants. Seldwyla Herald had investigated the housing rents of a sample of appartments in Seldwyla and claims that the rents for NRE appartments cost 300 CHF more than other appartments.

You are the Chief Analytics Officer (CAO) of NRE and your CEO has asked you to react in order to avoid a bad reputation of NRE. You contact the newspaper and succeed, after hard negotiations, to obtain the data of the survey. You get an Excel file with the observations on housing rents collected by the Seldwyla Herald. The data set “housingrents.csv" contains the following variables

variable name	variable label
id	Identification
rooms	Number of rooms (1,..., 6 or more)
area	area (m^2)
rent	Net housing rent (CHF/month)
nre	Property of NRE (1 for yes, 0 for no)
econage	Economic age of the appartment
balcony	Balcony (yes, no)

- Use Rcmdr to solve the exercise.
- Download the data set "housingrents.csv" from moodle and store it in a folder on your notebook.
- Start `Rgui.exe`, type `library(Rcmdr)`.
- Solve the problem set below and write the code, the results and your comments (!!!) into a solution report. Give arguments for your choices!
- You may establish the solution report with the Markdown package (see the `R Markdown` tab in Rcmdr). This will help you 1) to find and identify possible errors later on, and 2) may serve as a guide for future applications. You can write your comments on the output into the Markdown file (always outside the code chunks delimited with lines that start with ```` ```{r} ````) . You can create an output by hitting the **Generate HTML output** button.
- Clean out unnecessary stuff from the markdown file. Either print it as pdf or tidy it in a text processor first. Create a pdf and upload the solution report as a pdf-file to moodle.

Hint: To load the data set into Rcmdr use **Data/Import...** in the Rcmdr Commander menu and use the text file import. Depending on your system you may have to specify the field separator as a semicolon (**;**). Give the data set a name (`housingrents`).

Task 1. 1. How many variables and observations are in the data set?

2. Of what type (binary, categoric, numeric, character) are the respective variables?

3. Are there missing items? (Hint: Use **Statistics/Summaries/Active data set**)

Task 2. Create a bar chart of the frequency of appartments according to their number of `rooms`. Colour the bars in blue.

Hint: Before you can create a barplot with Rcmdr you have to convert the variable `rooms` to a `factor`. Use **Data/Manage.../Convert...**. Then use **Graphs/Bar...**. Add `... , col="blue"` to the arguments of `barplot` in the script-window and hit CTRL-R.

Task 3. Create a contingency table of `rooms` (rows) by `nre` (columns) (Use **Statistics/Contingency...**). Compute row percentages (percentages sum to 100% across rows) of the contingency table. Make a plot of the contingency table. Comment!(Hint: Store the outcome of `xtabs` assigning it to a variable: `rooms2nre<-xtabs(...)`. You may view your objects with command `ls.str()`. Add margins to the table with `addmargins(rooms2nre)` . Plot the table with `plot(rooms2nre)`) .

Task 4. Create a histogram of the distribution of the variable `area` (Use **Graphs/ Histogram...**). Do you think the variable area follows a normal distribution?

Task 5. Draw boxplots of the `rent` for NRE appartments and for non-NRE appartments (Use **Graphs/Boxplot...** and plot by groups). Calculate the mean and median rent for NRE and non-NRE appartments (Use **Statistics/Summaries/Numerical...**). Comment! (Hint: You may have to convert `nre` to a factor first.)

Task 6. Draw a simple scatter plot of `rent` vs. `area`, i.e. rent on the vertical axis and area on the horizontal axis (Use **Graphics/Scatterplot...** with option **Identify points/Interactively...**). Comment the form of the point cloud. Click on outliers to identify them.

Task 7. Create a new variable `rps=rent/area`, the rent per square meter (Use **Data/Manage.../Compute...**). Draw a Boxplot of `rps` for NRE and non-NRE appartments. Compare means and medians of `rps` for the two groups. Comment!