Instructions to the project work - SF2930 Regression Analysis

Tatjana Pavlenko and Felix Rios

January 18th, 2017

Instructions

General project work objectives:

Regression analysis is one of the most widely used methodology within the framework of applied statistics. The main goal of this part of the course is to provide students with practically oriented introduction to regression modeling techniques. To achieve this goal, the two compulsory course projects based om real life data are designed. Specifically, when working on the projects the students will

- acquire practical skills in applying regression analysis to solve real life problems;
- get hands-on experience in using the statistical software R for the regression model development, diagnostics and validation;
- learn to correctly interpret the results of the regression analysis, make quantitative and qualitative statements about the regression model-building process;
- summarize the results of the projects in terms of the written report containing the problem statement, description and motivation of the regression techniques used, interpretation of the resulting models and conclusions.

Project performance:

Students will work on the projects in groups of at most two members. The results of the project work will be presented in written form by handing in two written reports.

Project evaluation:

Each project report will be evaluated and graded as *Pass* or *Fail*. To pass the assignments part of the course both projects must be graded as *Pass*. Grading of the reports will be based on the following set of criteria.

- **Introduction to the data set and purpose of the project:** Clear and concise explanation of context, statement of the problem in terms of the regression analysis and clearly specified goals of the projects.
- **Analysis and model development:** Clear and concise description of the regression modeling strategy for the model building and validation (e.g. exploration of data to find an initial reasonable model, initial testing, treatment of outliers checking and modifying the model using correctly motivated techniques).
- When writing down the report **you do not have to insert all of the code into the report**, instead describe the process of solving the exercise and also why you do things. For example, when creating a new variable, instead of just giving the code for how this variable is created, mention why you are doing this and why it is impossible use the original variable?
- **Results:** Clear presentation and interpretation of the resulting model(s), their properties, advantages and limitations.
- **Conclusion:** Clear and supported conclusions.
- Report quality: organization (format/structure), quality of discussion and logic.

Report template

To get started with the report there is a report template in LaTeX available for download on the course webpage. Please use the sections outlined in the report template. If you are not able to use LaTeX to typeset your report you may use other software (e.g. Word, Pages, etc.)