

Humboldt University Berlin

Institute of Marketing

Prof. Dr. Daniel Klapper

Advanced Marketing Modeling

SS 2018

Special Work Performance 1: Working with R

This is group work. Each group consists of up to 4 students.

Your answers including all tables and graphs must not exceed 5 pages. Please start a new page when providing your report to a new subtask. Please use typeface Times Roman in 12pt with 1.15 line spacing (in tables and graphs you may use 10pt and 1.0 line spacing) and 1 inch space on all sides. Do not forget to report your names, group number, and student numbers and a page number on each page starting with number one on the first answering page.

Do not include a title page or content page.

Send your team report as pdf to my email address daniel.klapper@hu-berlin.de not later than May 04, 2018, 9am.

Download the Salaries data set (library(car)). Do a descriptive analysis of the Salaries data. Document your results and interpret your findings. In particular focus on the following points:

- How many variables and observations are there in the data set?
- How many professors have more than 40 years of service? (hint: you can sum() a logical vector)
- How many have salary larger than \$150000?
- What is the mean salary for professors with > 20 years service?
- **How do you find out more about the data set?**
- What are the counts of men and women by rank? The proportions?
- Draw a histogram for years of service. Add a density line in red. (Hint: plot proportions, not frequency.)
- Draw a box plot for salary.
- Draw a box plot for salary by rank. Make it horizontal.
- Plot salary vs years since PhD.
- **What is the correlation for salary vs. years since PhD and salary vs. years of service? Are they statistically significant?**
- Draw a visualization of all bivariate relationships.
- What are the mean salaries, by rank and sex?
- Plot those with a horizontal boxplot (conditioned on sex).
- **Does the proportion of women differ by discipline?**