UTRECHT
UNIVERSITY

Agnese Giacomello

Department of Social and Behavioural Sciences

May 15, 2019

In the following, the folder structure of the Research Archive for my Master's thesis is illustrated. I am responsible for this archive and it can be accessed by my supervisor, my mentor and my program-coordinators for one-year time.
The main folder is *Research Archive*. Sub-folders and the corresponding files are listed and described.
All analyses have been performed with the statistical software R.

Research Archive:

- **Data folder**
  This folder has been emptied for privacy reasons. Data are subjected to be kept within the CBS (Centraal Bureau voor de Statistiek) environment. If any information is needed, my supervisor Jeroen Pannekoek can be contacted (j.pannekoek@cbs.nl).

  Two datasets have been analysed:
  - "*Data_wholesalers.csv*" → Dutch Structural Business Dataset
  - "*Healtcare.csv*" → Dutch Healthcare Dataset

- **One predictor models folder**
  - "*R_estimates.R*" → wls, Huber and Tukey ratio *R* estimates
  - *"RatioModels_OnePredictor.R"* → wls, Huber and Tukey ratio models with one predictor, evaluated on RMSE and MAPE
  - *"Means_comparison.R"* → after-imputation means comparison (wls, Huber and Tukey)

- **Multiple predictor models folder**
  In this folder, two files for each considered target variable are present; an additional file *("Boosted Huber and Tukey implementation.R")* is added. The structure of all .R files, except for the additional one, is the same. Further information can be found in the files themselves. The .png files show the accuracy overview of all the fitted models.

  - *"Boosted Huber and Tukey implementation.R"* → implementation of the boosted Huber and Tukey models (not implemented in the package *mboost*) and comparison of tuning parameters;
  - *"Business_othercost.R"* and *"Business_othercost.png"*
  - *"Business_costpurchases.R"* and *"Business_costpurchases.png"*
  - *"Business_costemployees.R"* and *"Business_costemployees.png"*
  - *"Business_costdepreciation.R"* and *"Business_costdepreciation.png"*
  - *"Healthcare_othercost.R"* and *"Healthcare_othercost.png"*
  - *"Healthcare_costpurchases.R"* and *"Healthcare_costpurchases.png"*
  - *"Healthcare_costdepreciation.R"* and *"Healthcare_costdepreciation.png"*
  - *"Healthcare_wagesalaries.R"* and *"Healthcare_wagesalaries.png"*
  - *"Healthcare_totalcosts.R"* and *"Healthcare_totalcosts.png"*