Introduction to Statistics with R

Christopher Maronga

September 28, 2021

Session outline

- External packages.

- Data transformations - basics.

- Steps in Exploratory Data Analysis (EDA)
  - Numeric summaries
  - Visualizations.

- Robust ways of creating contingency tables

- Exporting/saving outputs for later use.

External packages

- R packages are contributed libraries for increased efficiency.

- There are currently over 18,200 packages in CRAN repository.

- Install package - `install.packages("name")`.

- Load package `library("name")` or `require("name")`.

Data transformation - Basics

- Critical step in analysing and summarizing data .

- Common forms of data.
  - Long format - observation appears multiple times (longitudinal data)
  - Wide format - observation appears ONLY once

- Variable transformations.
  - For computing necessary statistics
  - Outputting correct summaries
  - Effective visualization

- Saving objects (datasets, summary tables and plots).

Reference material

- An introduction to R hand-out

- Chapter 7 of R for Data Science book

- Introduction to R slides by Dr. Norma Coffey

- The book R in Action for data analysis and graphics