Analysing Social Science Data Using R

Class 1: Introduction

Michał Bojanowski[1] mbojan@icm.edu.pl
Zbigniew Karpiński[2] zkarpinski@ifispan.waw.pl

[1]ICM University of Warsaw

[2]IFiS PAN

February 13, 2012
SNS, Warsaw

Objectives

Main objective:

Educate the participants to use R for a typical data analysis in sociology

Including:

- Loading data
- Basic data manipulation: recoding, aggregating, variable transformation
- Descriptive statistics
- Data visualization
- Regression modeling (linear and logistic)

Work flow

Every meeting will consist of:

- Presentation
- Tutorial exercises
- Practical exercises to be completed in class

You will receive short exercises to complete at home.

Materials

Course materials will consist of:

- Slides
- Data files
- R scripts

All slides, data files, and scripts will be progressively available on-line on course webpage:
https://sites.google.com/site/r4sns2012
The course is rather self-contained: no extra obligatory reading materials etc. are planned. We will provide pointers to literature should anybody feel a need to refresh his/her statistical knowledge.

Grading

Final grade will be computed using:

- Attendance (20%)
- Homeworks (40%)
- Final test (40%)

Schedule

Updated syllabus is available on
https://sites.google.com/site/r4sns2012
Pleas note:

- Dates are provisional. It is likely that we will skip 12.03 and 19.03 and resume on 26.03.

What is R?

Program for data analysis For many people **R** is a program of choice for statistical analysis, visualization.

Programming language Data analysis with functions and scripting. Interactive programming language.

Environment for statistical analysis Access to standard models and cutting edge methods.

Open Source project Open source code. Free. Over 15 years of peer review. Integration with other tools/systems.

Community 20 people in R Core. Approximately 2 mln users worldwide: forums, mailing lists, blogs.

CRAN and addon packages

- **R**'s functionality is contained in packages.
  - Base distribution contains 27 packages.
  - Developers and users created the next 3604 packages (as on 2012-02-13) available on the Internet on the CRAN (Comprehensive **R** Archive Network) servers.
- Other things on CRANs
  - Official documentation
  - User-contributed documentation
  - R Journal

CRAN server in Wrocław http://r.meteo.uni.wroc.pl/

RStudio

http://www.rstudio.org