

Course syllabus «Introduction into R»

Утверждена
Академическим советом ООП
Протокол № 1 от «28» июня 2018 г.

Author	Rustam Tursun-Zade
Number of credits	2
Contact hours	28
Self-study hours	48
Course	1
Educational format	Without use of online course

I. Aim, Results of Mastering the Discipline and Prerequisites

The present program outlines minimum demands of students' knowledge and skills and the content of the course.

The syllabus is intended for the department teaching this course, its teaching assistants, and students of the degree program 38.04.04. 'Public Administration', master's program 'Population and Development'.

This syllabus meets the standards required by:

- FSES of 38.04.04 «Public Administration»;
- the educational standards of the Higher School of Economics developed for MA program 38.04.04 "Public Administration" (0401000.68 «Social sciences»)
- the curriculum of the MA program 38.04.04 "Public Administration" as of 2018.

The course is taught on the 1st year

Learning Objectives

This course is aimed to provide thorough introduction to R programming language with minimal theory and focus on its practical applications. It is also aimed to elaborate on the essentials of reproducible approaches to the academic workflow. During the course students will be introduced to the essentials of data acquisition, data manipulation and data visualization. We will also cover some applied demographic techniques, elements of calculus and matrix algebra.

Within this course, students will be provided with rigorous introduction to tidyverse – a set of packages designed for data manipulation and visualization. One of the major goals is to provide students with an essential toolbox needed for data manipulations rather than explaining the theoretical aspects of programming or logic behind the R algorithms. This course is not designed for programming per se, but rather for developing skills for applied data analysis.

After the end of this discipline the students get the basic skills of working with data in R – the essentials of data acquisition, processing, manipulation and visualization.

Upon a completion of this course students are expected to:

- Demonstrate an understanding of R programming language: its workflow, advantages and maluses,

- Have essential skills of data acquisition (importing) from various sources including the on-line databases,
- Be able to do the data manipulations and cleaning using the dplyr package,
- Know the basics of data visualization techniques using the ggplot2 package;
- Have a good grasp of R studio allowing for its use while studying in academia.
- Students are also expected to develop an analytical thinking and critical thinking allowing for deep understanding of data. By the end of the course, students are also expected to know how to evaluate the quality of statistical data, and be able to calculate the basic demographic and statistical indicators.

II. Content of the Course

Section 1. Course introduction. Basics of R, R studio and reproducible research.

Section 2. Main data structures, matrix algebra and data importing.

Section 3. Basics of data wrangling and manipulation with dplyr.

Section 4. Basics of data visualization with ggplot2 and demographic data in R.

III. Grading

The grade will be based on participation in coding sessions, completion of daily class tests and exercises, home assignments and written exam.

Final grade will be estimated in the following way:

$$0.2 * \text{hometask}_1 + 0.2 * \text{hometask}_2 + 0.2 * \text{class work} + 0.4 * \text{exam}$$

IV. Grading Tools

Homework:

Students will be given set of exercises designed so that they could put their newly acquired skills in practice. Students will be provided with raw dataset they are expected to manipulate with. They will be asked to add or remove variables and observations, to change the order of variables and observation, to manipulate upon the data structure, to convert the data types, to compute the statistical tests and common arithmetic upon that data, to clean the dataset, to visualize certain parts of it. General understanding of R workflow and logic will also be evaluated.

Exam:

Exam will be held in a form of practical written assignment, similar to that we will have during the lab session. The raw data frames will be distributed and a set of assignments will be offered to the class. Students will be asked to perform general exploratory data analysis followed by data manipulation, some computations and visualization of the results. All aspects learned within the course will be in one way or another included in the exam curriculum. Code clarity and consistency will also be evaluated.

N.B. Only those packages and functions covered within the classes will be included in exam and home task curriculum.

V. Sources

5.1 Main Literature

Tattar, P. N., Ramaiah, S., & Manjunath, B. G. (2016). *A Course in Statistics with R*. John Wiley & Sons.

ProQuest Ebook Central, [<https://ebookcentral.proquest.com/lib/hselibrary-ebooks/detail.action?docID=4452971>]

Additional literature

Chapman, C., & Feit, E. M. (2015). *R for marketing research and analytics*. New York, NY: Springer.

[<https://link.springer.com/book/10.1007/978-3-319-14436-8>]

Zelterman, D. (2015). *Applied multivariate statistics with R* (p. 54). Cham: Springer.

Chris Chapman, Elea McDonnell (2015) *R for Marketing Research and Analytics*

[<https://link.springer.com/book/10.1007/978-3-319-14093-3#about>]

5.2 Software

№ п/п	Name	Access conditions
1.	MicrosoftWindows 7 Professional RUS MicrosoftWindows 10 MicrosoftWindows 8.1 Professional RUS	<i>From the university's internal network (contract)</i>
2.	Microsoft Office Professional Plus 2010	<i>From the university's internal network (contract)</i>

5.3 Material and technical support

Classrooms for lectures on the discipline provide for the use and demonstration of thematic illustrations corresponding to the program of the discipline, consisting of:

- PC with Internet access (operating system, office software, antivirus software);
- multimedia projector with remote control.