KSL R Syllabus

Лето 2025

Sanzhar Kalkanbay
s.kalkanbay@gmail.com

Maqsut Narikbayev University

Last Update: 2025-06-12

General Information

Class hours

Official Class Hours for this class is **2 hours on Tuesdays/Thursdays**. In general, the course contents will be provided in the following way (adjustments may be made in the case the class goes online, subject to change as the semester proceeds):

- Theoretical contents of the course will be offered in a lecture format on first half hour of the class.
- Applied contents (i.e., implementation of methods in R) will be offered in a lab format during rest of the class. Students are expected to bring their PCs to the classroom and work on the problems together.

Office Hours

Office hours are times when you can meet with your teacher to discuss the material being presented in class or other related interests you have. Course-related discussions include asking for extra help, seeking clarification of material presented in class and following up on aspects of the class you find compelling.

Office hours will be held by appointment, 17:00 −18:00 on Tuesdays/Thursdays. Check the availability from the URL below:

https://calendly.com/s-kalkanbay/ofistik-sagattar

Each meeting slot is for 15 minutes. If you want a longer meeting, reserve two slots separately. Make an appointment *at least 4 hours prior to the meeting times*. **If the given slots do not work for you, contact me directly with the proposal of your available times.**

Prerequisites

This course assumes that students have no prior knowledge of statistical software and/or programming. Students are assumed to have basic knowledge of calculus and statistical terms, as well as low to moderate fluency in English. However, during the first day of the class, a short survey will be gathered to learn more about prior knowledge and skills.

Course Objectives

R is a widely used programming language for data analysis, especially in social sciences. It offers a flexible environment for working with data, performing statistical analysis, and creating visualizations. This course

introduces the basics of R, focusing on practical skills for data manipulation, visualization, and simple statistical modeling using the `tidyverse` and base R.

After completing this course, students are expected to acquire the skills to:

- Understand the basic syntax and structure of R.
- Import, clean, and manipulate data using R.
- Create engaging and informative data visualisations.
- Perform basic statistical analysis.
- Create and analyse statistical models.
- Write documents and presentations using R Markdown.

The course emphasizes hands-on practice and provides foundational skills for conducting independent research and quantitative analysis in social sciences.

Statistical Software

In this course we use the open-source statistical software R (http://www.r-project.org), as well as RStudio (http://www.rstudio.com/) — a user interface that simplifies many common operations. Students are expected to have personal computers capable of installing and launching R and RStudio.

Course Outline

Class	Date	Topics and In-Class Content
Class 1	TBD	1. Course introduction and structure 2. Installing R and RStudio, working environment setup 3. R as a calculator: basic syntax, variables, assignment 4. Data types and vectors 5. How to get help
Class 2	TBD	1. Data structures: vectors, matrices, lists, data frames 2. Subsetting and indexing 3. Introduction to ‘tibble‘ and basic wrangling 4. Reading data from CSV/Excel
Class 3	TBD	1. Cleaning data with ‘dplyr‘ 2. Filtering, selecting, mutating, summarizing 3. Grouping and aggregating 4. Writing clean, readable scripts
Class 4	TBD	1. Introduction to ‘ggplot2‘: the grammar of graphics 2. Scatterplots, barplots, histograms 3. Aesthetic mappings and themes 4. Exporting plots
Class 5	TBD	1. Working with categorical variables (factors) 2. Working with dates and time data 3. Reshaping with ‘tidyr‘ (pivoting, joining) 4. Project organization tips
Class 6	TBD	1. Descriptive statistics in R 2. Summarizing categorical and numeric data 3. Interpreting central tendency and spread 4. Presenting summary tables and charts

(continued)

Class	Date	Topics and In-Class Content
Class 7	TBD	1. Intro to linear modeling with ‘lm()‘ 2. Running basic regressions 3. Interpreting model output 4. Visualizing regression results
Class 8	TBD	1. Introduction to R Markdown 2. Creating dynamic documents 3. Tables, plots, and inline stats 4. Final recap: applying R to research