

Syllabus

Statistics with R, GSERM

Instructor: Oliver Westerwinter
Assistant Professor
Department of Political Science
University of St. Gallen
June 3-5, 2016

Time & room

Class: 4:30-7pm, 03.06, 04.06 in 09-012
05.06 in “Friedegg”, 3rd floor,
hotel Einstein
Office hours: by appointment

Office

Oliver Westerwinter
Room: 33-506, Rosenbergstr. 51, 5th floor
Email: oliver.westerwinter@unisg.ch

This syllabus may be subject to adjustments.

Overview and goals

These workshop lectures introduce participants to R, one of the most flexible and most widely used statistical computing environments in the social sciences. In this course, students will learn how to use R for data analysis. While we will cover a selected range of basic topics (e.g., reading data, data structures, data manipulation, data visualization), the emphasis of the course is on more advanced topics (e.g., writing functions, loops, reshaping data, estimation of statistical models, and analyzing the results of statistical models). We will also learn how to prepare outputs generated in R for the use in L^AT_EX.

By the end of this class, students should be able to manipulate data, run statistical models, write functions and loops, produce high-quality graphics, and export outputs generated in R to L^AT_EX. This course is also helpful as a primer for other GSERM courses that will use R, such as the courses on regression analysis. No prerequisites are required for this course, although familiarity with matrix algebra and statistics will be helpful.

Lectures and lab sessions

Each day consists of a lecture and a lab session. The lectures will be mostly applied and cover several topics. There will be examples and exercises during the lectures. At the end of each day, there will be a lab session during which we will work through a set of exercises and solve them together. The exercises are designed to deepen your understanding of the material covered in the lecture and to practice the use of R. The exercises and the data needed to work through them will be available prior to class.

Office hours

Students with questions and problems related to the materials covered in class as well as their own data and research are invited to ask for individual appointments.

Class materials

Students are strongly recommended to bring their own laptops to class to be able to follow the lecture examples and participate in the lab sessions (class does not take place in a computer lab). Please make sure that you install R on your own computer prior to the start of the class and make sure it runs properly. You can download R for free from www.r-project.org.

Scripts, data, and exercises will be made available prior to the start of each class session via a shared dropbox folder to which all course participants will have access.

Students who have never worked with R before are recommended to work through the introductory script available in the dropbox. The script introduces some fundamental basics of R and guides you in your first few steps in this environment.

In addition to the provided materials, students are encouraged to make extensive use of detailed help files that come with R as well as the vast and oftentimes freely available online sources.

Readings

Much of the material covered in this class is concisely and accessibly summarized in:

Monogan III, James E. 2015. *Political Analysis Using R*. New York: Springer.

Other texts that students may find useful include:

Dalgaard, Peter. 2008. *Introductory Statistics with R*. Second edition. New York: Springer.

Fieller, Nick. 2016. *Basics of Matrix Algebra for Statistics with R*. London: Taylor & Francis.

Robert, Christian P. and George Casella. 2010. *Introducing Monte Carlo Methods with R*. New York: Springer.

Ugarte, Maria Dolores, Ana F. Militino and Alan T. Arnholt. 2016. *Probability and Statistics with R*. Second edition. London: Taylor & Francis.

Script or editor?

There are several ways of interacting with R. You may use R with the RStudio editor to facilitate accessibility. Like R, RStudio is freely available. This class will use the built-in R script editor to have wider appeal. Students are more than welcome to use their editor of choice though the instructor's ability to provide assistance may vary for editors other than the script editor.

Schedule

Date	Topics	Readings
June 3	Topic I: Loops, functions & probability distributions Topic II: Reading & manipulating data	Monogan ch. 1-2, 10-11
June 4	Topic I: Descriptive statistics Topic II: Estimating linear and non-linear models	Monogan ch. 4-7
June 5	Topic I: Data visualization Topic II: From R to L ^A T _E X	Monogan ch. 3