

Intro to R — Stat 5050, Section MW1, Fall 2020 (1 Credit)

Instructor: Dr. Jürgen Symanzik
Office: Virtual
Phone: 435-797-0696
e-mail: symanzik@math.usu.edu
Web: <http://www.math.usu.edu/~symanzik/>
http://www.math.usu.edu/~symanzik/teaching/2020_stat5050/stat5050.html

Office Hours: Monday (M), 9:00am – 10:00am; Wednesday (W) & Friday (F), 8:00am – 9:00am; and by appointment. Office hours have to be scheduled via the Zoom menu in Canvas (or via e-mail) and will be held virtually via Zoom.

Classes & Rooms:

MWF 12:30pm – 1:20pm, M 8/31/20 – M 10/5/20: Virtual Meetings via Zoom.

Please visit the course Web page listed above for emergency announcements, e.g., when Canvas is unavailable. Otherwise, visit Canvas frequently for lecture notes, data sets, R code, etc. — in particular if you miss our lecture periods for any reason. All (additional and updated) materials, announcements, discussions, recordings, etc. from Canvas are part of the course materials. Not seeing one of these in time does not serve as an excuse for not getting point deductions for the course. Deadlines may change or *Coronavirus/Covid-19* regulations and requirements may be updated. It is your responsibility to make sure to receive all announcements in time.

Detailed Class Schedule:

For a 1-credit course, we need 15 lectures/lecture days (in contrast to 43 or 44 lectures/lecture days for a 3-credit course). Those days are marked as “Lecture 01” to “Lecture 15” in the overview below:

Week	Monday	Wednesday	Friday
1	8/31: Lecture 01	9/2: Lecture 02	9/4: Lecture 03
2	9/7: Labor Day	9/9: Lecture 04	9/11: Lecture 05
3	9/14: Lecture 06	9/16: Lecture 07	9/18: Lecture 08
4	9/21: Lecture 09	9/23: Lecture 10	9/25: Lecture 11
5	9/28: Lecture 12	9/30: Lecture 13	10/2: Lecture 14
6	10/5: Lecture 15	10/7: Backup	10/9: Backup

Note: I have marked a few days as “Backup”, e.g., in case we miss lectures because of network problems, power failures, etc. on my side. But, hopefully, this won’t happen. If nothing goes wrong, our tentative last lecture date will be on M 10/5/20.

If a Zoom meeting does not start on time, please wait for 10min so I can try to get it started differently. Similarly, if we get disconnected during a Zoom meeting, also wait for 10min and watch for announcements in Canvas and/or via e-mail. In case you lose the connection on your side, please try to reconnect as quickly as possible.

Course Objectives:

The purpose of this course is to provide students from statistics and other departments with a sound foundation of the R software environment. The course is not designed to

teach students to use R for data analysis or visualizations, but to prepare students to use R *effectively* in other statistics and data science courses.

Students will learn how to read, write, and understand base R code. They will learn about R's data structures, function language, and documentation. They will learn simple programming skills, including how to write their own functions. At the end of the course, a student should be able to “use” an R function and extract the results, read and write data files, and produce an R Markdown document combining text and analysis.

Prerequisites:

I do not expect any prior knowledge of R. Knowledge from an introductory statistics course such as Stat 2000, Stat 3000, or higher is required (Stat 1040 or Stat 1045 are not enough). You should, for example, know basic probability distributions (Bernoulli, Binomial, Gaussian) and know what the Central Limit Theorem (CLT) says.

No programming experience is required, although you must be familiar with your preferred operating system (Windows, Mac, or Linux) and understand how to download files from the web, manage folders, install software, etc.

IDEA Center Learning Objectives:

Objective 1) Gaining factual knowledge (terminology, classifications, methods, trends).

Objective 2) Learning fundamental principles, generalizations, or theories.

Objective 3) Learning to apply course material (to improve thinking, problem solving, and decisions).

Topics: (subject to change)

1. Introduction
2. Getting started with R, RStudio, R Markdown, and \LaTeX
 - (a) Installing and running R, RStudio, and \LaTeX
 - (b) Using R Markdown, knitr, and Sweave to write reports (and more!)
 - (c) Finding functions, objects, and documentation
 - (d) Issuing commands and saving results
3. The R Language
 - (a) Data types
 - (b) Subsetting
 - (c) Vectors
 - (d) Matrices
 - (e) Data frames
 - (f) Lists
 - (g) Apply
 - (h) Functions
4. Programming
 - (a) Control flow: if, for, while
 - (b) Vectorization and efficiency
 - (c) Writing your own functions

Note that this course does not introduce any methods how to visualize and summarize statistical data and results via statistical graphics (such as bar charts, histograms, dot plots, box plots, etc.). To learn how to do so, you should attend “Statistical Visualization I” that usually is offered in the Fall semester. Concurrent enrollment is possible.

Course Format and Lecture Attendance Points:

The course will be offered in a blended web broadcast format. See <https://www.usu.edu/ais/scheduling/deliverymethods> for requirements on your side.

Under this setup, we will basically use a flipped classroom approach. You will have to watch recordings of past classroom-based lectures for this course by yourself and you have to work through the lecture slides and R code also by yourself. The scheduled lecture periods will be used to discuss your questions related to the lecture recordings and other course materials and for help with the next homework assignment. I also plan to summarize the most important parts of each lecture at the start of each lecture period. Lecture periods will differ in lengths, depending on your questions. Some may be as short as 10min, while others may take up the entire 50min.

You will be awarded up to three lecture attendance points (LAPs) for each lecture, i.e., up to 45 LAPs in total. You will be asked early, in the middle, and towards the end of each lecture to type a short confirmation into the Chat box in Zoom so that I can see who was present at that time. LAPs will contribute to 10% of your course grade. You will obtain 10 points for 90% or more (41 to 45) of all possible LAPs, 9 points for 80% up to 90% (36 to 40) of all possible LAPs, 8 points for 70% up to 80% (32 to 35) of all possible LAPs, 7 points for 60% up to 70% (27 to 31) of all possible LAPs, 6 points for 50% up to 60% (23 to 26) of all possible LAPs, and 0 points for less than 50% (0 to 22) of all possible LAPs.

The lecture periods will be recorded. If you really cannot participate at a lecture, at least watch the recording before the next lecture. In case of an excused absence, e.g., for medical reasons, family emergencies or funerals, court appointments, university-approved travel, etc., please provide some supporting information and your LAP score will be adjusted according to the number of lecture periods you could attend. Private reasons such as travel, most family events (such as weddings), etc. do not count as an excused absence.

Homework Assignments:

There will be 5 HW assignments for this course, roughly one per week. Each HW assignment will include a value (typically 20–100 points) that it will be scored out of. HW assignments will contribute to 90% of your course grade. The value of each HW assignment will be roughly proportional to its importance and the amount of work involved. Your final course grade will be determined as the weighted average of your LAPs and the sum of your points in all HW assignments.

You will be allowed to discuss general approaches to questions on the HW assignments with other students, but each student must write and submit their own R code and comments. Any students caught sharing R code or other parts of their homework submissions will fail the class.

Unless otherwise stated on the HW assignment sheet, all homework assignments have to be submitted electronically via Canvas.

The following deductions will be applied to late homework submissions: 1 min – 24 hours late: 10% off; > 24 hours – 48 hours late: 25% off; > 48 hours – 72 hours late: 50%

off. Homeworks won't be accepted later than 72 hours (i.e., 3 days) after the submission deadline.

There will be no (in-class or take-home) quizzes, midterm exams, or final exams. We will have a few worksheets for training purposes only.

Textbooks:

There is no formal textbook for this course, but the instructor will provide links to suitable online resources.

Software:

We will primarily be using R (<http://cran.r-project.org/>), a free software environment for statistical computing and graphics. Please install the most recent version of R, i.e., 4.0.2, on your own computer so we can exchange code. Also install RStudio (<https://www.rstudio.com/>) as a front end to R and MiKTeX (<https://miktex.org/>) that will allow us to combine code and results from R into text documents. We will jointly look into the software installation in our first two lectures.

Courtesy:

One of the intents of the blended web broadcast format is to make our courses somewhat more personal again, compared to the pure online format during the second half of the Spring 2020 semester. For this reason, please activate your webcams during all of our Zoom meetings so we each can see each other. However, to avoid interference with the audio, please mute your microphone unless you want to speak yourself. Let me know that you want to speak by indicating so in the Chat box in Zoom or by unmuting your microphone, but wait until the previous speaker has ended. When you are done speaking, please mute your microphone again.

Please be aware of the *Code of Policies and Procedures for Students at Utah State University* (<https://studentconduct.usu.edu/studentcode/>) and follow the code accordingly. Also be aware of the USU *Coronavirus/Covid-19* regulations and requirements (<https://www.usu.edu/covid-19/>). The latter ones may change with little advance notification. Useful resources for students can be found at <https://www.usu.edu/ready/>. Also, as a reminder, all of our lecture periods, including audio, video, and chats, will be recorded and might be used as evidence in case of any student code violations during these lecture periods, in particular those listed in Article 5 of the student code (<https://studentconduct.usu.edu/studentcode/article5>).

Americans with Disabilities Act:

If a student has a disability that will likely require some accommodation by the instructor, the student must contact the instructor and document the disability through the Disability Resource Center (DRC – <https://www.usu.edu/drc/>), preferably during the first week of the course. Any requests for special considerations relating to attendance, pedagogy, taking of examination, etc. must be discussed with and approved by the instructor. In cooperation with the Disability Resource Center, course materials can be provided in alternative formats — large print, audio, or Braille.

Note:

The above schedule and procedures in this course are subject to change in the event of extenuating circumstances.