R Programming Training

Data Preparation for Analysis

UF INFORMATION TECHNOLOGY

Spring 2026
Authored by: Jose Lorenzo Silva-Lugo, Ph.D.

R Programming Training

Data Preparation for Analysis

Instructor

Dr. Jose Lorenzo Silva-Lugo, HUB 285, Phone: (352) 273-1579. Email: joselugo@ufl.edu.
Office Hours: TBA.

Textbooks

- Field, A, J. Miles, and Z. Field. 2012. Discovering Statistics using R. Sage Publications Ltd, London.
- Michael, W. Trosset. 2009. An Introduction to Statistical Inference and its Application with R. A Chapman & Hall Book, CRC Press. Boca Raton, Florida.
- Muenchen, R. A. 2011. R for SAS and SPSS Users. 2nd Edition. Springers, New York.
- William N. Venables, D. M. Smith, and the R Development Core Team. 2009. An Introduction to R. 2nd edition. Network Theory, Ltd.
- Hothorn, T. and B. S. Everitt. 2014. A handbook of statistical analysis using R. 3rd edition. CRC Press, Boca Raton, Florida.

These textbooks are not required for the training series. But if you are interested in learning more about R, I suggest purchasing any of them. Particularly, I like Field's book because his writing style is friendly and very easy to read.

Training Schedule

From January 12 (8:00 am) to February 6 (5:00 pm), 2026. The course may be extended until April 24 (5:00 pm) for participants who do not complete it by the end date.

Location: **online (self-paced)**

Goal and Objectives

These training series are an introduction to the program language R specifically designed for faculty, staff, postdocs, graduate students, and teaching assistants. The purpose is to get familiar with the R commands in such a way that participants will feel self-prepared to continue by their own. After completing these trainings, participants will be able to:

- Install R, install and load packets and entering data
- Manage the data set and prepare it for analysis
- Explore the data set with descriptive statistics, normality test and graphics

Training Series Description

This free introductory R training is designed for participants with a graduate-level background in statistics. It is tailored for applications in the Natural and Social Sciences. Participants are expected to be familiar with fundamental statistical concepts and procedures covered in the training. Therefore, enrollment requires prior completion of at least one graduate-level statistics course.

This training course provides a hands-on introduction to R programming, covering installation, packages, data structures, functions, and visualization tools. It guides learners in data preparation, including handling missing values, outliers, duplicates, normalization, and file merging. Participants also practice exploratory data analysis (EDA), learning to detect patterns, create charts, and apply mathematical or statistical transformations. The course emphasizes building new variables, converting scales, and making informed decisions on suitable statistical analyses based on EDA results.

The teaching technique will follow the classical training session composed by lecturing and activities, in which students take the dynamic role of watching videos, asking questions, and working on hands-on activities. Lectures and activities will always be together to provide knowledge and to practice the skill. **Class is 100% online and self-paced**.

There is no formal evaluation because this is not a registrar course. **The hands-on activities were designed for the participants to evaluate themselves**. R requires a lot of practice, and the idea is that participants learn the content of each workshop before doing the hands-on activity. Participants will assess whether they learn the content in this way.

Training Content

1. **Introduction to R**:
    - What is R?
    - Installing R, R Studio and R Commander
    - Installing and loading packages
    - Loading data sets from a computer
    - Mathematical operations with objects
    - Working with vectors
    - Data structures
    - Working with functions
    - Working with data frames and matrices
    - Adding and removing columns/rows
    - Renaming/reordering columns
    - Working with graphics

- Saving, cleaning up, and quitting
- Getting help with R

2. **Data Preparation**:
   - Detecting anomalies
   - Working with missing values
   - Removing duplicate cases
   - Dealing with outliers
   - Aggregating cases
   - Normalizing variables
   - Merging files
3. **Exploratory data analysis**:
   - Performing exploratory data analysis for one or more variables
   - Creating charts for one or more variables
   - Transforming data
   - Computing a new variable:
     * SqRt, Log10/Ln, and 1/X
     * Mathematical Operations (+, -, *, /)
     * Raising to the power
     * Trigonometric transformations
   - Transforming scale data into an ordinal/binary one
   - Making decisions about the statistical analysis to follow based on the exploratory data analysis.

GENERAL NOTICE TO STUDENTS

Students with Disabilities

If you need classroom accommodation because of a disability, you must register with the Disability Resource Center. This office will provide you with several forms, and one of them must be turned in to the instructor. Since some of these accommodations require time to be in place, I will appreciate that the form is given to me with two weeks in advanced.

Training Policies

- You are responsible for following the workflow during training and for studying all materials and resources posted in Canvas
- You are responsible for carrying out all activities during the training sessions
- If you have questions, please contact me at joselugo@ufl.edu

Software Use

All faculty, staff, and students of the University are required to obey the laws and legal agreements regarding software use. It is Illegal to copy licensed and/or copy written materials. This is a third-degree felony under Florida law. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. The Office of Academic Technology and the members of the University of Florida community, pledge to hold our peers and ourselves to the highest standards of honesty and integrity.

UF Counseling Services

Resources are available on campus for students having personal problems or lacking clear career and academic goals that interfere with their academic performance. These resources include:

- U Matter We Care, 352-294-2273 | umatter@ufl.edu, help for students in distress
- Counseling and Wellness Center, 3190 Radio Road, 392-1575, personal, sexual assault, and career counseling
- Career Resources Center, Reitz Union, 392-1601, career development assistance and counseling