R Programming Course

Data Analysis

UF INFORMATION TECHNOLOGY

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

R Programming Course

Data 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 course. 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 June 29 (8:00 am) to August 7 (5:00 pm), 2026. **No extension will be granted.**

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

Goal and Objectives

This course is an introduction to the program language R specifically designed for faculty, staff, postdocs, graduate students, and teaching assistants. The purpose is to become familiar with R commands for statistical analysis so that participants feel confident continuing on their own. After completing this course, participants will be able to:

- Perform elementary statistical analyses such as two or more mean comparison, one-way and two-way ANOVA, correlation, bivariate and multiple linear regression, proportion comparisons, and logistic regression.

Course Description

This **free** introductory course was designed for participants with graduate-level knowledge of basic statistics. It is oriented towards applications in the Natural and Social Sciences. Participants are expected to be familiar with the fundamental concepts and statistical procedures covered in the course (see next page). Therefore, **prior completion of at least one graduate-level statistics course is required**, as this course synthesizes key analyses typically taught in three graduate courses in the Statistics Department: *STA 6166 Statistical Methods in Research I*, *STA 6126 Statistical Methods in Social Research I*, and *STA 6127 Statistical Methods in Social Research II*.

It is important to emphasize that, although this training series covers several statistical topics, it is not a substitute for any of the courses listed above. R is a statistical programming language, and the primary focus of this training is to demonstrate how to use R to perform statistical analyses.

Participants are expected to prepare their own data and conduct exploratory data analysis (EDA) prior to performing statistical procedures. Accordingly, this course assumes prior familiarity with data preparation and EDA. If you do not have this background, it is strongly recommended that you first complete the training course *"Data Wrangling and EDA."*

The course begins with hypothesis formulation for comparing means and medians, including techniques such as the two-sample t-test and ANOVA. It then introduces regression and correlation, which are essential components of classical statistical analysis. Next, cross-tabulation methods are used to examine goodness-of-fit and independence. Finally, the course concludes with logistic regression, providing a more advanced approach to statistical modeling. Overall, this course introduces R as a free and powerful tool for conducting commonly used statistical analyses.

The instructional approach follows a traditional format that combines lectures with hands-on activities. Participants engage actively by watching instructional videos, asking questions, and completing practical exercises. Lectures and activities are integrated to reinforce both conceptual understanding and applied skills. **The course is delivered entirely online and is 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. Because learning R requires consistent practice, participants are encouraged to master the content of each workshop before attempting the corresponding activities. This approach allows participants to evaluate their own progress and understanding.

Course Content

1. **Mean and Median Comparison**:
    - Making inferential statistics about the mean and the median

- Appling the parametric t-test:
  * One-sample t-test
  * Two-independent samples t-test
  * Paired difference t-test
- Applying non-parametric t-test:
  * Sign test
  * Mann-Whitney U Test
  * Wilcoxon Signed Rank Test
- Build charts for two mean comparison
- Carry out t-test involving violations of normality

2. **ANOVA and Correlation**:
   - Making inferential statistics by using one-way ANOVA and two-way ANOVA
   - Executing the Kruskal-Wallis test (non-parametric of one-way ANOVA)
   - Performing bivariate correlation parametric and non-parametric
3. **Bivariate Linear Regression and Multiple Linear Regression**:
   - Making EDA and charts for more than 2 variables
   - Bivariate linear regression parametric and non-parametric
   - Multiple linear regression and correlation
   - Model building and selection
   - Quantile regression
4. **Proportion Comparison:**
   - Two proportions comparison
   - Chi-squared goodness of fit tests
   - Chi-squared tests of independence
5. **Logistic Regression:**
   - Demonstrating the general procedure to carry out the logistic regression with scale, binary and categorical predictor variables:
     * Fit the model to the data
     * Estimate and interpret regression coefficients, the odds ratio, and confidence intervals
     * Estimate the probability of occurrence
   - Carrying out exercises with the combination of binary, categorical, and scale predictor variables.
   - Testing the linearity assumption

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 several forms, and one of them must be turned into the instructor. Since some of this accommodation requires time to be in place, I would appreciate it if the form were given to me two weeks in advance.

Course Policies

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

Software Use

All faculty, staff, and students at 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