

DATA SCIENCES AND OPERATIONS

Spring 2024

DSO 499 Section 16248D

Statistical Computing and Data Visualization in R, Python, and Tableau

Professor: Lorena Martin, PhD

Email

lmartin1@marshall.usc.edu

When

Tues and Thurs 4-5:50 pm

Office

BRI 202

Units

4



WHY TAKE THIS COURSE?

Want to learn how to code in R and Python? Want to learn Tableau for Data Visualization?
 Want to gain proficiency in the latest programming languages and software platforms for Data Science?
 Want to analyze and visualize data the correct way?
 This is the course to take!

COURSE OBJECTIVES

1. Successfully use programming languages and Tableau Software to make sense of data numerically and visually
2. Extract, transform, and load data
3. Perform A/B testing and compare group differences
4. Choose the appropriate data visualization based on requirement of data type
5. Create high quality data visualizations that convey business insights
6. Work collaboratively in a team setting

KEY CONCEPTS

Statistical Computing
 Data Wrangling
 Programming
 Data Visualization
 Spatial Analysis

COURSE DESCRIPTION

This class is taught through a combination of lectures, computer labs, hands-on computing tasks in homework, and group project and presentation. The project is key because it integrates multiple computational topics of a modern data problem. Students gain hands-on experience with statistical concepts flowing from contextual problem solving with data, and they make their own discoveries by posing and answering questions

Tentative Syllabus - Subject to Modifications

DSO 499 –Statistical Computing and Data Visualization in R and Tableau

Syllabus – Spring 2024 – 4 Units – Tues/Thurs – Times: 4:00 PM - 6:00 PM (2 hrs/class session)



Professor: Lorena Martin, PhD
Office: TBD
Office Phone: TBD
Email: Lmartin1@marshall.usc.edu

Office Hours: TBD

Course Description

In this course, students will learn how to make sense of data numerically and visually. The course starts with statistical computing, and students will gain experience with a programming language called R and Rstudio. In addition, students will learn Tableau and be able to integrate R with Tableau to maximize these tools' capabilities for business. They will learn the practice of data cleaning, reshaping of data, basic tabulations, and aggregations in order to be able to produce high quality visualizations. This class is taught through a combination of lectures, computer labs, hands-on computing tasks in homework, and group project and presentation. The project is key because it integrates multiple computational topics of a modern data problem. Students gain hands-on experience with statistical concepts flowing from contextual problem solving with data, and they make their own discoveries by posing and answering questions rather than solely fitting models or using "this week's lecture's methodology" as a computing exercise.

Learning Objectives

Upon successful completion of this course, students will be able to:

- 1) Successfully use R and Rstudio programming and Tableau Software to make sense of data numerically and visually
- 2) Extract, transform, and load data
- 3) Perform A/B testing and compare group differences
- 4) Choose the appropriate data visualization based on requirement of data type
- 5) Create high quality data visualizations that convey business insights
- 6) Work collaboratively in a team setting

Required Materials

This is a hands-on course and it is computationally intensive. We will primarily be using **R** (<http://cran.r-project.org>), a GNU-license statistical package and Rstudio. R is available on the Web for free download. Rstudio is a recommended interface for the R software. It is also free, and it runs on Windows, Mac, and Linux operating systems. <http://www.rstudio.org>

There are many reasons for us to focus on R:

- The cost: While commercial distributions exist, open-source R is free.
- The rich features: R has an estimated user community of 2 million, which includes thousands of contributors from different domains expanding the language's capabilities through new libraries.
- The quality: R libraries are enhanced by domain experts and field-tested by the large user base including other experts with real datasets in real analysis scenarios.
- The learning resources: Thanks to the active user community, plenty of tutorials and sample code are readily available.

Tableau Public will also be implemented for Data Visualizations, (<https://public.tableau.com/en-us/s/about>). Tableau Public is a free platform to publicly share and explore data visualizations online.

Wickham, H., & Grolemund, G. (2016). *R for data science: import, tidy, transform, visualize, and model data*. " O'Reilly Media, Inc."

Sleeper, R. (2018). *Practical Tableau: 100 Tips, Tutorials, and Strategies from a Tableau Zen Master*. " O'Reilly Media, Inc."

Jones, B. (2014). *Communicating Data with Tableau: Designing, Developing, and Delivering Data Visualizations*. " O'Reilly Media, Inc."

Supplementary Materials

Grolemund, G. (2014). *Hands-on programming with r: write your own functions and simulations*. " O'Reilly Media, Inc."

Prerequisites and/or Recommended Preparation:

None, this course is designed to have the student learn R and Tableau in class.

GRADING DETAIL

Final grades represent how you perform in the class relative to other students. The average grade for this class is expected to average about 3.3. Three items are considered when assigning final grades:

1. Your average weighted score as a percentage of the available points for all assignments (the points you receive divided by the number of points possible).
2. The overall average percentage score within the class.
3. Your ranking among all students in the class.

Assignment Submission Policy:

Assignments must be turned in on the due date/time electronically via Blackboard. Any assignment turned in late, even if by only a few minutes, will receive a grade deduction of 10% per day. Late or not, however, you must complete all required assignments to pass this course.

Homework - There are six homework assignments each worth 5%. The nature of the homework assignments is to provide students with questions that will enable them to practice coding and reinforce what was learned in class.

Participation:

In-class participation is 5% of the total grade and evaluated based on your level of involvement in class discussions and in-class exercises.

- One of the primary goals of this course is to help you develop the ability both to clarify your own position on an issue and to be able to articulate and defend it clearly. Sharing your perceptions and ideas with others is crucial for learning and for understanding how the diverse opinions that you are likely to encounter in an organization are debated. You will find yourself presenting and testing new ideas that are not wholly formulated and assisting others in shaping their ideas as well. You should be prepared to take some risks and be supportive of the efforts of others.
- Effective class participation consists of analyzing, commenting, questioning, discussing, and building on others' contributions; it is not repeating facts or monopolizing class time.

The ability to present one's ideas concisely and persuasively and to respond effectively to the ideas of others is a key business skill. One of the goals of this course is to help you sharpen that ability.

- **Outstanding Contribution:** Your contributions reflect considerable preparation; they are substantive and supported by evidence from the case, readings, and logic. Your comments or questions create a springboard for discussion by making a critical insight. You synthesize and build upon what has already been said in the discussion. The class learns from you when you speak; in your absence, the discussions would suffer.
- **Good Contribution.** You come prepared with substantiated comments. You demonstrate good insight and clear thinking. You are able to make some connection to what has been said in prior discussion. The class notices when you're not part of the discussion.
- **Minimal Contribution.** You participate but are unprepared. You rarely offer interesting insights into the discussion. It appears that you are not listening to what others are saying during discussion.
- **No Contribution.** You say little or nothing in class. If you were not in the class, the discussion would not suffer.

Final Project

The objective of the project is to identify a business problem, apply the R coding programming skills, statistical computing and data visualization techniques to problem solve and facilitate data driven decision making. Each group is expected to find their own dataset, write a report, analyze the data in R and provide relevant data visualizations with R and Tableau. The project consists of a project proposal, a final written report, and a presentation worth 15%

Teams will meet outside of class and provide a Team Project Proposal that will identify the sections responsible by each team member. Instructor will randomize students to teams. Peer review for the team project will be collected at the end of the team presentation (see Appendix II).

R software and Tableau are required. Deliverables will include a project proposal, an abstract, a final written APA report, and a slide presentation.

Exams:

There is a midterm and final exam in this class each worth 25%. The nature of the midterm and final exam are practical and include the material of the chapters covered in each class session prior to the exams.

Assessment	% of Grade
Homework	30%
Participation	5%
Project	15%
Midterm	25%
Final Exam	25%

COURSE CALENDAR

	Topics/ Daily Activities	Readings and Homework	Deliverables and Due Dates
Week 1	Introduction to R software programming	Chapter 1 from the book: R for Data Science	
Week 2	R functions	Chapter 2 from the book: R for Data Science Homework 1 Basic Calculations and functions	
Week 3	Data Visualization	Chapter 3 from the book: R for Data Science	Deliverable: HW1 consisting of basic calculations and functions and is due Week 3 in Blackboard (TBD)
Week 4	Statistical Modeling	Chapter 4 from the book: R for Data Science Homework 2 Statistical Computing and Modeling (Parametric Tests)	
Week 5	Statistical Transformation and Visualizing Findings	Chapter 5 from the book: R for Data Science	Deliverable: HW2 consisting of statistical computing and modeling is due Week 5 in Blackboard (TBD)
Week 6	Data Manipulation	Chapter 7 from the book: R for Data Science Homework 3 Data Manipulation (Filtering, Selecting, Mutating, Arranging)	Deliverable: Project Proposal
Week 7	Data Wrangling	Chapter 9 from the book: R for Data Science	Deliverable: HW3 consisting of Data Manipulation is due Week 7 in Blackboard (TBD)
Week 8	REVIEW & MIDTERM EXAM		
Week 9	Spatial Analysis in R	Homework 4 Analysis of Countries Types of Business	Deliverable: Abstract
Week 10	Introduction to Tableau	Chapter 1-2 from Book: Communicating Data with Tableau	Deliverable: HW4 consisting of Data Manipulation is due Week 10 in Blackboard (TBD)
Week 11	Spatial mapping in Tableau	Chapter 10 from Book: Communicating Data with Tableau Homework 5 Worksheet Chart Creation in Tableau	

Week 12	Motion Charts in Tableau	Chapter 41 from Book: Practical Tableau	Deliverable: HW5 consisting of a Tableau Data Visualization is due Week 12 in Blackboard (TBD)
Week 13	Dashboarding in Tableau	Chapter 12-14 from Book: Communicating Data with Tableau Homework 6 Dashboard Creation in Tableau	Deliverable: Project Written Report
Week 14	R integration into Tableau	Chapter 79 from Book: Practical Tableau	Deliverable: HW6 Creation of Dashboard in Tableau is due Week 14 in Blackboard (TBD)
Week 15	Project Presentations		PROJECT is due (Tuesday & Thursday)
FINAL	FINAL EXAM		Exact Time and location TBD

STATEMENT ON ACADEMIC CONDUCT AND SUPPORT SYSTEMS

Academic Conduct:

Students are expected to make themselves aware of and abide by the University community’s standards of behavior as articulated in the [Student Conduct Code](#). Plagiarism – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Part B, Section 11, “Behavior Violating University Standards” policy.usc.edu/scampus-part-b. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct at <http://policy.usc.edu/scientific-misconduct>.

Support Systems:

Counseling and Mental Health - (213) 740-9355 – 24/7 on call
studenthealth.usc.edu/counseling

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call
suicidepreventionlifeline.org

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention and Services (RSVP) - (213) 740-9355(WELL),
 press “0” after hours – 24/7 on call
studenthealth.usc.edu/sexual-assault

Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

Office of Equity and Diversity (OED)- (213) 740-5086 / *Title IX* – (213) 821-8298
equity.usc.edu, titleix.usc.edu

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants. The university prohibits discrimination or harassment based on the following *protected characteristics*: race, color, national origin, ancestry, religion, sex, gender, gender identity, gender expression, sexual orientation, age, physical disability, medical condition, mental disability, marital status, pregnancy, veteran status, genetic information, and any other characteristic which may be specified in applicable laws and governmental regulations. The university also prohibits sexual assault, non-consensual sexual contact, sexual misconduct, intimate partner violence, stalking, malicious dissuasion, retaliation, and violation of interim measures.

Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298
usc-advocate.symplicity.com/care_report

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office of Equity and Diversity |Title IX for appropriate investigation, supportive measures, and response.

The Office of Disability Services and Programs - (213) 740-0776
dsp.usc.edu

Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC Support and Advocacy - (213) 821-4710
uscsa.usc.edu

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity at USC - (213) 740-2101
diversity.usc.edu

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, *HSC*: (323) 442-1000 – 24/7 on call
dps.usc.edu, emergency.usc.edu

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, *HSC*: (323) 442-120 – 24/7 on call
dps.usc.edu

Non-emergency assistance or information.

Appendix I.

Undergraduate Program Learning Goals and Objectives

Learning goal 1: Our graduates will demonstrate critical thinking skills so as to become future-oriented problem solvers, innovators and decision makers in diverse and rapidly changing business environments.

- Students will demonstrate the ability to anticipate, identify and solve business problems. They will be able to identify and assess central problems, identify and evaluate potential solutions, and translate a chosen solution to an implementation plan
- Students will demonstrate the ability to be accurate, thorough, and detailed thinking
- Students will critically analyze concepts, theories and processes by stating them in their own words, understanding key components, identifying assumptions, indicating how they are similar to and different from others and translating them to the real world
- Students will be effective at gathering, storing, and using qualitative and quantitative data and at using analytical tools and frameworks to understand and solve business problems
- Students will understand the concepts of critical thinking, entrepreneurial thinking and creative thinking as drivers of innovative ideas

Learning Goal 2: Our graduates will develop people and leadership skills to promote their effectiveness as business managers and leaders in the 21st century's evolving work and organizational structures.

- Students will recognize, understand and analyze the roles, responsibilities and behaviors of effective managers and leaders in diverse business contexts e.g., functionally diverse, culturally diverse, geographically diverse, etc.
- Students will understand factors that contribute to effective teamwork including how to elicit, manage and leverage diverse perspectives and competencies.
- Students will recognize, understand, and analyze the motivations and behaviors of stakeholders inside and outside organizations

Learning Goal 3: Our graduates will be effective communicators to facilitate information flow in organizational, social, and intercultural contexts.

- Students will identify and assess diverse personal and organizational communication goals and audience information needs
- Students will demonstrate an ability to gather and disseminate information and communicate it clearly, logically, and persuasively in professional contexts
- Students will understand individual and group communications patterns and dynamics in organizations and other professional contexts

Learning goal 4: Our graduates will demonstrate ethical reasoning skills, understand social, civic, and professional responsibilities and aspire to add value to society.

- Students will recognize ethical challenges in business situations and assess appropriate courses of action
- Students will understand professional codes of conduct

Learning goal 5: Our graduates will develop a global business perspective. They will understand how local, regional, and international markets, and economic, social and cultural issues impact business decisions so as to anticipate new opportunities in any marketplace.

- Students will understand how stakeholders, stakeholder interests, business environments (legal, regulatory, competitor) and business practices vary across regions of the world as well as how to communicate through data visualization
- Students will understand how local, regional and global markets interact and are impacted by economic, social and cultural factors.

Learning goal 6: Our graduates will understand types of markets and key business areas and their interaction to effectively manage different types of enterprises.

- Students will demonstrate foundational knowledge of core business disciplines that include statistical computing and data visualization
- Students will understand the interrelationships between functional areas of business so as to develop a general perspective on business management
- Students will apply theories, models, and frameworks to analyze relevant markets
- Students will be able to competently implement technologies such as R and Tableau relevant to contemporary business practices

Appendix II.

PEER EVALUATION FORM

Please identify your team and team members for the Group Project that you worked on. Then rate all of your team members, *including yourself*, based on the contributions of each team member for the selected assignment according to the criteria listed below. On a scale of 0 – 2 with 0 equal to does not meet expectations, 1 meets expectations and 2 exceeds expectations, rate each person on each of the five criteria, then sum the points, with the maximum points being 10.

Team Members/ Assessment Criteria of Team Contributions	Team Member 1	Team Member 2	Team Member 3	Yourself
1. Role Performance				
2. Assists Team Members				
3. Listening and Discussing				
4. Research and Information Sharing				
5. Time Management				
Total				

Comments: