LOYOLA UNIVERSITY CHICAGO

*Preparing people to lead extraordinary lives*

**FUNDAMENTALS OF MODERN DATA SCIENCE WITH R**
**DSCI 101 - 001**
**MWF 2:45 pm – 3:35 pm**
**Dumbach Hall Rm. 233**

## Instructor Information

***Instructor***: Widad Abdalla Mukhaimer, PhD

***Campus Office***: 301 Loyola Hall

***E-mail***: wabdalla@luc.edu

***Office Hours***:

- *In person*: MWF – 10:20 – 11:50 am
- *Online*: Th – 9-10 am

***Note**: in person office hours will be in my office (301 Loyola Hall). However, in the case that a large number of students come to the office at the same time, I will be holding office hours in one of two common areas on the first floor of Loyola Hall. In this case, I will leave a sign on the door of my office.

## Essential Course Information

***Course Description:*** This course provides students with an introduction to data science using the R programming languages covering such topics as data wrangling, data visualization, principles of reproducible research, building simple statistical models/machine learning and data science ethics.

***Required Textbook***: Modern Data Science with R (2nd edition). Baumer, Kaplan, and Horton
PDF of Book: https://beanumber.github.io/mdsr2e/index.html

***Recommended Textbook:*** Grolemund, G., & Wickham, H. (2017). R for Data Science. O'Reilly Media.
PDF of Book: https://r4ds.had.co.nz/

### *R and Posit (formerly RStudio)*

We will be using/introducing the free statistical software R. While R is the engine, we will use the free and open source IDE (Integrated Development Environment) RStudio to run it. R and RStudio are set up and available on all library computers.

### *Course Structure*

- ***Reading:*** Readings are posted in the course schedule that should be read *before class*. Come to class prepared to solidify the readings through lecturing and group activities.
- ***Class Time***: Class will be composed of lecturing, discussions, collaborative activities, and R practice. Please come to class having done the reading, a charged computer, and ready to discuss and learn in a collaborative manner.
- ***Assessments***: Students will complete homework assignments, take a midterm, 1 project, and a final.
- ***Participation, Discussion, and Group Work***: One important aspect of a Jesuit education is learning to respect the rights and opinions of others. Please respect others by (1) allowing all classmates the right to voice their opinions without fear of ridicule, and (2) not making objectionable (gendered, racial or ethnic) comments, especially comments directed at a classmate. Group work and discussion are vital to this class since no one student will understand everything, please lean on each other for help and learn to hear concepts and ideas from another perspective.

## Grading Policy

Students will be evaluated as follows:

| Grading Components | Percentage of Total Grade |
|---|---|
| Homework | 25% |
| Midterm | 25% |
| Project | 25% |
| Cumulative Final Exam | 25% |

The grading scale is as follows:

| A | A- | B+ | B | B- | C+ | C | C- | D+ | D | F |
|---|---|---|---|---|---|---|---|---|---|---|
| 93-100% | 90-92% | 87-89% | 83-86% | 80-82% | 77-79% | 73-76% | 70-72% | 67-69% | 60-66% | 0-59% |

**Homework**: Homework is due approximately every other week. Discussion between classmates is encouraged; however, the final work should be independent. Homework must be submitted through Sakai. Homework turned in after the due date will receive no credit. To help your final grade, please avoid late homework.

**Exams**: There will be one midterm exam and one final for the semester. Both exams may not be made up unless there is a serious reason for missing and arrangements are made prior to the test. The midterm will be an in-class exam that is cumulative up to that point in class. Final will be cumulative. Both exams would incorporate interpretations and outputs from R. **Final exam is on Friday, May 3rd at 4:15 pm.**

**Project**: The individual project will require students to find a raw dataset, wrangle the data into a useful format, perform some interesting analysis, and present results in a written report following the principles of reproducible research. All code must be version controlled through github (or repository of your choice) and a link to the repository must be submitted along with the final report. More specific details on the project presentations and reports will be given at a later date but note that the project will have multiple due dates throughout the entire semester.

## Additional Course Policy

**Academic Integrity**: Academic dishonesty can take several forms, including, but not limited to cheating, plagiarism, copying another student's work, and submitting false documents. Academic cheating is a serious act that violates academic integrity. Cheating includes, but is not limited to, such acts as:

- Obtaining, distributing, or communicating examination materials prior to the scheduled examination without the consent of the teacher.
- Providing information to another student during an examination.
- Obtaining information from another student or any other person during an examination.
- Using any material or equipment during an examination without consent of the instructor, or in a manner which is not authorized by the instructor.
- Attempting to change answers after the examination has been submitted.
- Unauthorized collaboration, or the use in whole or part of another student's work, on homework, lab reports, - programming assignments, and any other course work which is completed outside of the classroom.
- Falsifying medical or other documents to petition for excused absences or extensions of deadlines.
- Any other action that, by omission or commission, compromises the integrity of the academic evaluation process.
- For more details on Loyola's Academic Integrity Statement please see here.

**Tutoring Help:** The www.luc.edu/tutoring embodies the mission of Loyola University Chicago by providing academic services and resources which foster development of skills and attitudes necessary to increase the knowledge and academic independence of all students. Through multiple learning services, the Tutoring Center helps to contribute towards student success and growth efforts that are made by Loyola University Chicago. ***Get help early if you are having difficulty.***

**Diversity Equity and Inclusion Statement:** The diversity that students bring to this class, in all its forms, is viewed as a resource, a strength, and a benefit. It is my intent to invest in each student's success and attend to each student's learning needs, both in and out of class. It is my intent to present materials and activities that are respectful of diversity, equity and inclusion, and that students from all diverse backgrounds and perspectives be well-served by this course.

**Accommodations for Disabilities or Conditions:** Loyola University provides reasonable accommodations for students with disabilities. Any student requesting accommodations related to a disability or other condition is required to register with Student Accessibility Center (SAC), located in Sullivan Center, Suite 117. Students will provide professors with an accommodation notification from SAC,

preferably within the first two weeks of class. Students are encouraged to meet with their professor individually in order to discuss their accommodations. All information will remain confidential. For more information or further assistance, please call 773.508.3700.

**Intellectual Property:** All lectures, notes, PowerPoints, and other instructional materials in this course are the intellectual property of the professor. As a result, they may not be distributed or shared in any manner, either on paper or virtually without my written permission. Lectures may not be recorded without my written consent; when consent is given, those recordings may be used for review only and may not be distributed. Recognizing that your work, too, is your intellectual property, I will not share or distribute your work in any form without your written permission.

## Tips to Succeed in This Course

1. Come to class ready to learn.
2. Do not skip class. You are expected to attend and participate in class. I will be taking attendance every day.
3. Unless otherwise stated, please put cellphones away and make sure that they are in silent mode.
4. Ask questions! There's no such thing as a dumb question. If you need me to review something, please let me know.

## Tentative Course Schedule

| Date | Topics |
|---|---|
| January 17 | **First Day of Class** |
| January 19 | Chapter 1: Why DSCI? |
| January 22 | Chapter 4: Data Wrangling |
| January 24 | Chapter 4: Data Wrangling |
| January 26 | Chapter 4: Data Wrangling |
| January 29 | Chapter 5: Joins |
| January 31 | Chapter 5: Joins |
| February 2 | Chapter 6: Tidy Data |
| February 5 | Chapter 6: Tidy Data |
| February 7 | Chapter 6: Tidy Data |
| February 9 | Chapter 3: GoG |
| February 12 | Chapter 2: Data Visualization |
| February 14 | Chapter 2: Data Visualization |
| February 16 | Chapter 2: Data Visualization |
| February 19 | Chapter 7: Maps |
| February 21 | Chapter 7: Maps |
| February 23 | Chapter 7: Maps |

| Date | Topics |
|---|---|
| February 26 | Chapter 8: Ethics |
| February 28 | Chapter 8: Ethics & Exam Review |
| March 1 | **Midterm Exam** |
| March 4 | **Spring Break – Classes do not meet** |
| March 6 | |
| March 8 | |
| March 11 | Chapter 9: Bootstrap |
| March 13 | Chapter 9: Bootstrap |
| March 15 | Chapter 9: Bootstrap |
| March 18 | Chapter 9: Statistics |
| March 20 | Chapter 9: Statistics |
| March 22 | Chapter 9: Statistics |
| March 25 | Chapter 10: Predictive |
| March 27 | Chapter 10: Predictive |
| March 29 | **Easter Holiday – Classes do not meet** |
| April 1 | |
| April 3 | Chapter 10: Predictive |
| April 5 | Chapter 11: Supervised Learning |
| April 8 | Chapter 12: Unsupervised Learning |
| April 10 | Chapter 12: Unsupervised Learning |
| April 12 | Chapter 12: Unsupervised Learning |
| April 15 | Chapter 14: Data Visualization + |
| April 17 | Chapter 14: Data Visualization + |
| April 19 | Chapter 14: Data Visualization + |
| April 22 | Chapter 15: SQL |
| April 24 | Chapter 15: SQL |
| April 26 | Chapter 15: SQL |
| **May 3** | **Final Exam** |