PUBH 7461, SECTION 001

Exploring and Visualizing Data in R
Fall 2018

COURSE & CONTACT INFORMATION

Credits: 2
Meeting Days: Friday
Meeting Time: 10:10 a.m. - 12:05 p.m.
Meeting Place: Bruininks Hall 312

Instructor: Julian Wolfson
Office Address: Mayo A453
Office Phone: (612) 625-9514
Fax (612) 626-0660
E-mail: julianw@umn.edu
Office Hours: TBD

COURSE DESCRIPTION

This course is intended for students, both within and outside the School of Public Health, who want to learn how to manipulate data, perform simple statistical analyses, and prepare basic visualizations using the statistical software R. While the tools and techniques taught will be generic, many of the examples will be drawn from biomedicine and public health.

COURSE PREREQUISITES

Familiarity with concepts from introductory statistics: mean/median, continuous vs. categorical variables, basic probability, Normal and binomial distribution. While the course is not designed to teach basic programming concepts, no prior programming experience is necessary.

COURSE GOALS & OBJECTIVES

Upon completing this course, students should be able to:
•   Use RStudio and its features efficiently and appropriately
•   Read in and manipulate raw data into R using multiple file types
•   Create and implement functions in R
•   Understand and implement approaches in R to reshaping and restructuring data for analysis
•   Understand and implement approaches in R to visualize data both for exploratory analyses and reports
•   Understand and implement approaches in R to process dates and strings

METHODS OF INSTRUCTION AND WORK EXPECTATIONS

This course is a mixture of lecture, demonstration and hands on activities, with time spent outside of class on programming activities. Students are expected to attend class, participate in class discussions, and complete the assigned homework and project.

Course Workload Expectations
Exploring and Visualizing Data in R is a 2-credit course. The University expects that for each credit, you will spend a minimum of three hours per week attending class or comparable online activity, reading, studying, completing assignments, etc. over the course of a 15-week term. Thus, this course requires approximately 90 hours of effort spread over the course of the term to complete the required material.

Learning Community
Many School of Public Health courses ask students to work in teams or discussion groups. We do not come to our courses with identical backgrounds and experiences and building on what we already know about collaborating, listening, and engaging is critical to successful professional, academic, and scientific engagement with topics.

In this course, students are expected to engage with each other in respectful and thoughtful ways.

In group work, this can mean:
- Setting expectations with your groups about communication and response time during the first week of the semester (or as soon as groups are assigned) and contacting the TA or instructor if scheduling problems cannot be overcome.
- Setting clear deadlines and holding yourself and each other accountable.
- Determining the roles group members need to fulfill to successfully complete the project on time.
- Developing a rapport prior to beginning the project (what prior experience are you bringing to the project, what are your strengths as they apply to the project, what do you like to work on?)

In group discussion, this can mean:
- Respecting the identities and experiences of your classmates.
- Avoid broad statements and generalizations. Group discussions are another form of academic communication and responses to instructor questions in a group discussion are evaluated. Apply the same rigor to crafting discussion posts as you would for a paper.
- Consider your tone and language, especially when communicating in text format, as the lack of other cues can lead to misinterpretation.

Like other work in the course, all student to student communication is covered by the Student Conduct Code (https://z.umn.edu/studentconduct).

COURSE TEXT & READINGS

There is no required text and no formal readings for this course; students will learn from hands-on in-class exercises and by consulting help documentation and online sources (e.g., DataCamp, StackOverflow). However, you may find the following additional references useful:
- Hadley Wickham. "ggplot2: Elegant graphics for data analysis". Available through SpringerLink on campus (http://link.springer.com.ezp3.lib.umn.edu/book/10.1007%2F978-0-387-98141-3)
- Hadley Wickham. "R for Data Science". (http://r4ds.had.co.nz/)
- Garrett Grolemund. "Hands-On Programming with R". (http://shop.oreilly.com/product/0636920028574.do#)
- Norman Matloff. "The Art of R Programming". (https://www.nostarch.com/artofr.htm)

Week	Topic	Activities/Quizzes (in-class)	Assignments (due following week)
Week 1 Sept. 7th	• **Introduction to R**	• **Start on this week's assignments**	• **(DataCamp) Complete *Introduction to R*** • **(DataCamp) Complete chapter 1 and 3 of *Importing Data in R (Part I)*** • **(DataCamp) Complete the *Orientation* chapter of *Working with the RStudio IDE (Part I)***
Week 2 Sept. 14th	• **Getting comfortable with R and RStudio**	• **Programming Quiz #1** • **Exploring data frames, identifying and manipulating variable types**	• **(DataCamp) Complete the remaining chapters of *Working with the RStudio IDE (Part I)*** • **(DataCamp) Complete Chapters 1-3 of *Intermediate R***
Week 3 Sept. 21st	• **Basic programming structures**	• **Programming Quiz #2** • **(DataCamp) Complete Chapters 1 and 2 of *Intermediate R - Practice***	• **(DataCamp) Complete Chapters 4-5 of *Intermediate R*** • **(DataCamp) Complete Chapters 1-2 of *Writing Functions in R***
Week 4 Sept. 28th	• **Functions**	• **Programming Quiz #3** • **(DataCamp) Complete Chapters 3 and 4 of *Intermediate R – Practice***	• **(DataCamp) Complete Chapters 1-2 of *Cleaning Data in R*** • **(DataCamp) Complete Chapters 1-2 of *Categorical Data in the tidyverse***
Week 5 Oct. 5th	• **Data cleaning**	• **Programming Quiz #4** • **(DataCamp) Complete Chapter 3 of *Cleaning Data in R***	• **(DataCamp) Complete Chapters 1-4 of *Data Manipulation in R with dplyr***
Week 6 Oct. 12th	• **Data manipulation**	• **Programming Quiz #4** • **ATUS data exercise (DataCamp) Complete Chapters 1 and 3 of *Introduction to the tidyverse***	• **(DataCamp) Complete Chapter 5 of *Data Manipulation in R with dplyr*** • **(DataCamp) Complete Chapters 1-3 of *Joining Data in R with dplyr***
Week 7 Oct. 19th	• **Data manipulation**	• **Programming Quiz #5** • **ATUS data exercise** •	• **(DataCamp) Complete *Data Visualization in R*** •
Week 8 Oct. 26th	• **Data visualization**	• **Programming Quiz #6** • **Visualizing ATUS data**	• **(DataCamp) Complete *Data Visualization with ggplot2 (Part I)***
Week 9 Nov. 2nd	• **Data visualization**	• **Programming Quiz #7** • **Visualizing ATUS data** • **(DataCamp) Complete Chapter 2 of *Introduction to the tidyverse***	• **(DataCamp) Complete Chapters 1-3 of *Data Visualization with ggplot2 (Part II)***

Week 10 Nov. 9th	• Data visualization	• Programming Quiz #8 • (DataCamp) Complete Chapter 4 of *Introduction to the tidyverse* • (DataCamp) Complete Chapter 5 of *Data Visualization with ggplot2 (Part II)*	• (DataCamp) Complete Chapters 1-3 of *Data Visualization with ggplot2 (Part III)*
Week 11 Nov. 16th	• Synthesis	• Programming Quiz #9 • (DataCamp) Complete the *Dr. Semmelweis and the discovery of handwashing* project	• (DataCamp) Complete the *Visualizing Inequalities in Life Expectancy* project
Week 12 Nov. 23rd	THANKSGIVING		
Week 13 Nov. 30th	• Exam Prep	• Programming exam review and practice	• Study for in-class programming exam
Week 14 Dec. 7th	IN-CLASS PROGRAMMING EXAM		

SPH AND UNIVERSITY POLICIES & RESOURCES

The School of Public Health maintains up-to-date information about resources available to students, as well as formal course policies, on our website at www.sph.umn.edu/student-policies/. Students are expected to read and understand all policy information available at this link and are encouraged to make use of the resources available.

The University of Minnesota has official policies, including but not limited to the following:
- Grade definitions
- Scholastic dishonesty
- Makeup work for legitimate absences
- Student conduct code
- Sexual harassment, sexual assault, stalking and relationship violence
- Equity, diversity, equal employment opportunity, and affirmative action
- Disability services
- Academic freedom and responsibility

Resources available for students include:
- Confidential mental health services
- Disability accommodations
- Housing and financial instability resources
- Technology help
- Academic support

EVALUATION & GRADING

The final grade will be determined as follows:

Completion of assignments: 15%
Completion of in-class activities: 10%
In-class quizzes: 50%
Final programming exam: 25%

Grading Scale
The University uses plus and minus grading on a 4.000 cumulative grade point scale in accordance with the following, and you can expect the grade lines to be drawn as follows:

% In Class	Grade	GPA
93 - 100%	A	4.000
90 - 92%	A-	3.667
87 - 89%	B+	3.333
83 - 86%	B	3.000
80 - 82%	B-	2.667
77 - 79%	C+	2.333
73 - 76%	C	2.000
70 - 72%	C-	1.667
67 - 69%	D+	1.333
63 - 66%	D	1.000
< 62%	F	

- A = achievement that is outstanding relative to the level necessary to meet course requirements.
- B = achievement that is significantly above the level necessary to meet course requirements.
- C = achievement that meets the course requirements in every respect.
- D = achievement that is worthy of credit even though it fails to meet fully the course requirements.
- F = failure because work was either (1) completed but at a level of achievement that is not worthy of credit or (2) was not completed and there was no agreement between the instructor and the student that the student would be awarded an I (Incomplete).
- S = achievement that is satisfactory, which is equivalent to a C- or better
- N = achievement that is not satisfactory and signifies that the work was either 1) completed but at a level that is not worthy of credit, or 2) not completed and there was no agreement between the instructor and student that the student would receive an I (Incomplete).

Evaluation/Grading Policy	Evaluation/Grading Policy Description
Scholastic Dishonesty, Plagiarism, Cheating, etc.	You are expected to do your own academic work and cite sources as necessary. Failing to do so is scholastic dishonesty. Scholastic dishonesty means plagiarizing; cheating on assignments or examinations; engaging in unauthorized collaboration on academic work; taking, acquiring, or using test materials without faculty permission; submitting false or incomplete records of academic achievement; acting alone or in cooperation with another to falsify records or to obtain dishonestly grades, honors, awards, or professional endorsement; altering, forging, or misusing a University academic record; or fabricating or falsifying data, research procedures, or data analysis (As defined in the Student Conduct Code). For additional information, please see https://z.umn.edu/dishonesty The Office for Student Conduct and Academic Integrity has compiled a useful list of Frequently Asked Questions pertaining to scholastic dishonesty: https://z.umn.edu/integrity. If you have additional questions, please clarify with your instructor. Your instructor can respond to your specific questions regarding what would constitute scholastic dishonesty in the context of a particular class-e.g., whether collaboration on assignments is permitted, requirements and methods for citing sources, if electronic aids are permitted or prohibited during an exam. Indiana University offers a clear description of plagiarism and an online quiz to check your understanding (http://z.umn.edu/iuplagiarism).
Late Assignments	Assignment completion will be tracked via DataCamp.
Attendance Requirements	
Extra Credit	Extra credit opportunities may be offered on in-class activities or assignments. For reasons of fairness, all extra credit opportunities will be made available to the whole class as part of the planned curriculum.