Introduction to Data Analysis in R

Course Syllabus - Fall 2019

Instructor: Dr. Geoff Zahn

Office: SB243c – Hours: Mon 11:00 - 11:50 ; Tue 8:00 - 8:50 ; or by appointment

Email: zahn.geoff@gmail.com

Course GitHub Repository: https://github.com/gzahn/Data_Course

You need to bring your own laptop to class every day. That way you have installation priveleges and you will be set up for the future as you continue to analyze data.

Course Description

This course provides an introduction to analyzing data in the R software environment, assuming no previous R experience. We cover:

- the justification of using code to explore and analyze data
- best practices for dealing with data
- experimental design, modeling, and hypothesis testing
- how to create publication-quality figures that show interesting relationships in our data sets
- how to deal with common genetic sequence data types
- command-line computing skills and proper reporting

Pre-requisites: Advanced standing or instructor permission

Learning outcomes Students completing this course should be able to:

1. Demonstrate proficiency in proper data entry, management, and storage with an emphasis on reproducibility.
2. Convert untidy data to "tidy data" for analyses.
3. Understand the basic principles of exploratory data analyses within a computational software environment.
4. Evaluate the rationale behind using code to analyze data and present results.
5. Develop computational skills for processing common biological data formats, such as DNA sequences.
6. Create appropriate and meaningful data visualizations using the R software environment.
7. Integrate principles of experimental design, statistical modeling, hypothesis testing, and data visualization to critically analyze a unique data set and present a fully-reproducible report.

. . .
. . .
. . .

Advice

This course requires significant outside work... meaning that if you aren't practicing at least 6 hours a week *outside of class*, then you are going to have a really rough time of it. Like, you'll fall behind and will be frustrated and scared and sad. Don't let this happen. Practice! Come to office hours! Ask questions! And spend time **every day** using R. Seriously. Every day.

Grading

We will use a point system. The points you have accumulated by the end of the course determine your grade as follows:

Points	Letter Grade
700-800	A
640-699	B
560-639	C
480-559	D
<480	E

Points are based on:

- 10 Assignments - 20 pts each
- 4 Skills Tests - 100 pts each
- 1 Final Project - 200 pts

Assignments

These will be announced in class and detailed instructions will be posted on the course GitHub repository. They will generally consist of requiring you to complete a task using R and to upload your code as an R script to Canvas and/or GitHub. Some may vary. Any data sets required for the assignment will be on the GitHub repository as well. These assignments will not be accepted for credit after the due date.

Skills Tests

These will take place during class time. They will be similar to the sorts of tasks on the precedent assignments, and they will be open-source (you can use notes, internet, etc.) but they will be timed.

Final Project

Beginning early in the semester, we will decide on individual projects based on personal interest. Working with instructor feedback, you will come up with a question that interests you and will identify a data set that can address that question. You will then apply the data exploration and visualization skills you learn in class to prepare a well-formatted report that contains all the code and results of your analyses. You are encouraged to use your own data if you have any. **Most importantly this project will require you to teach yourself new R skills using the resources we learn about (or any others). You will be required to learn analyses explicitly not covered in this course (Check the course schedule to see what we are covering), and to apply them correctly to your analysis and report.**

Examples of unique topics could include, but are not limited to:

- Time-series analyses
- Genomic Profiling
- New statistical testing methods
- Metagenomic assembly
- Learning a new R package
- Unique visualization methods
- Etc.

Your topic/question is due before Skills Test 2

Soon after, you will work with the instructor to identify a suitable publicly-available data set (or use your own)

An initial draft of your exploratory code is due before Skills Test 3

Then you will work with the instructor to identify a new skill to be applied for your project

Your project will culminate in a "markdown" report that integrates your background research, code, results, figures, and discussion into a single reproducible document

---

The tentative course schedule can be found at:

https://github.com/gzahn/Data_Course/blob/master/Rough_Course_Outline.pdf

This course schedule document also serves as a template for proper reporting (for final project).

It (along with this syllabus) was created in R, and it integrates explanatory text, code, and results into a single document.

You can see the raw R code used to generate the document at: https://raw.githubusercontent.com/gzahn/Data_Course/master/Rough_Course_Outline.rmd

---

Important legal stuff, etc.:

Americans with Disabilities Act: Students with disabilities that may negatively impact their ability to successfully complete this course should let the instructor know as soon as possible, with documentation from the ASD. "Students who need accommodations because of a disability should contact the UVU Accessibility Services Department (ASD), located on the Orem Campus, in LC 312. To schedule an appointment, or speak with a counselor, call the ASD office at 801-863-8747, or for Deaf/Hard of Hearing individuals, use the video phone number, 886-760-1819."

Academic Honesty and Integrity

Utah Valley University expects all students to maintain integrity and high standards of individual honesty in academic work, to obey the law, and to show respect for others. Students of this class are expected to support an environment of academic integrity, have the right to such an environment, and should avoid all aspects of academic dishonesty. Examples of academic dishonesty include plagiarizing, faking of data, sharing information during an exam, discussing an exam with another student who has not taken the exam, consulting reference material during an exam, submitting a written assignment which was authored by someone other

than you, and/or cheating in any form. In keeping with UVU policy, evidence of academic dishonesty may result in a failing grade in the course and disciplinary review by the college. Any student caught cheating will receive, at minimum, zero points on that particular assignment for the first offense. A second offense can result in failing the course and will entail being reported to Student Advising. Academic dishonesty includes, in part, using materials obtained from another student, published literature, and the Internet without proper acknowledgment of the source. Additional information on this topic is published in the student handbook and is available on the UVU website. Situation of academic dishonesty will be reported to the Student Conduct and Conflict Resolution office as, at a minimum, an informational item and potentially as an item for investigation, resolution, or other.

Student Code of Conduct

All UVU students are expected to conduct themselves in an appropriate manner acceptable at an institution of higher learning. All students are expected to obey the law, to perform contracted obligations, to maintain absolute integrity and high standards of individual honesty in academic work, and to observe a high standard of conduct for the academic environment. The Student Rights and Responsibilities Code, or Code of Conduct, outlines for students what they can expect from the University and what the University expects of them. It can be found at http://www.uvu.edu/studentconduct/students/

Student Timelines

You are responsible for being aware of and abiding by the dates and deadlines on the student timelines for this term found at UVU Schedule (http://www.uvu.edu/schedule/)

Understand When You May Drop This Course

It is the student's responsibility to understand when they need to consider dropping from a course. Dates and deadlines for dropping a course can be found in the student timetables located at http://www.uvu.edu/schedule/. After the period specified, a serious and compelling reason is required to drop from the course and would need approval from not just the faculty but also the department chair as well as the Dean's Office.

Incomplete Policy

An Incomplete (I) Grade is only given due to an extenuating circumstance. The final decision to issue an Incomplete grade is up to the instructor. Review the following criteria taken from Policy 523 at for information on when an incomplete grade can be given: The student was unable to complete the course due to extenuating circumstances:

- Incapacitating Illness
- Death in the immediate family
- Change in the student's work schedule by the employer
- Other deemed acceptable by the instructor The student must have completed a minimum of 70% of the required work An Incomplete Grade form indicating work completed and work to be completed must be signed by the department chairperson, and turned into the Registrar's Office at the end of the semester. Incomplete Grade forms can be obtained from the Registrar's Office.

Unofficial Withdrawals

UW will be given only when a student stops attending and completing work before the regularly scheduled date to officially withdraw as found in the Student Timetable but did not actually officially withdraw. Rationale: The assumption here is that the student did not realize the necessity to physically withdraw or

made an error in this process, but all intentions of the student were to withdraw from the class prior to the withdrawal deadline See also Academic Policies and Standards in the UVU Catalog which notes, "If a student stops attending (but does not officially withdraw) before the last day to withdraw, he/she should receive a UW."

Late Work

Late work will not be accepted without a doctors or lawyers note unless previously discussed with and approved by your instructor. Computer problems, work schedule, lack of internet access, traffic, travel, or similar do not constitute rationale for late homework. Your instructor has the right to modify their handling of late work.

Individual versus Group Work

All work is to be done individually unless it is clearly identified as group work. By default all assignments are not group work. This means that you cannot work on assignments with others, even if your spouse, friends or family, and turn them in as your own. A "group" in this context means more persons than yourself. You also cannot have others complete the work for you or in part on your behalf.

Inclusivity Statement

We understand that our members represent a rich variety of backgrounds and perspectives. The Technology Management department is committed to providing an atmosphere for learning that respects diversity. While working together to build this community we ask all members to:

- share their unique experiences, values and beliefs
- be open to the views of others
- honor the uniqueness of their colleagues
- appreciate the opportunity that we have to learn from each other in this community
- value each other's opinions and communicate in a respectful manner

Academic Freedom

For the purposes of this class, academic freedom is generally defined as follow: For faculty and students alike, you have the freedom to disagree with the views and statements of others, but academic freedom should entail tolerance and respect for others, professionalism, and, to the degree possible, fact-based information. Intellectual vigor is best sustained when the free exchange of ideas is carried on within an environment supportive of human dignity and self-esteem. For full guidelines see http://www.uvu.edu/tm/pages/department/acad-freedom.html

Discrimination / Harassment / Sexual Misconduct

It's On Us to prevent discrimination/harassment against members of protected classes. If you see something, step up and say something. If you encounter discrimination, harassment, or gender-based discrimination (including discrimination or harassment based on pregnancy, pregnancy related conditions, sexual orientation or gender identity) or sexual violence (dating violence, domestic violence, stalking including unwanted electronic contact), please contact the Office for Equal Opportunity/Affirmative Action and Title IX. Additional information regarding options and services can be found at: http://www.uvu.edu/equalopportunity/ and http://www.uvu.edu/studentconduct/docs/title9optionsbrochure.pdf Consistent with State and Federal Law such as Title VI of the Civil Rights Act of 1964 and Title IX of the Education Amendments of 1972, Utah Valley University is committed to maintaining a respectful and safe environment for its students, faculty,

staff, and visitors. UVU policies 162 and 165 define and prohibit all forms of unlawful discrimination, harassment, sexual misconduct and retaliation in its programs services and activities. Utah Valley University policy 165 prohibits any form of: • (a) Discrimination based on one or more protected classes, (Race, color, religion, national origin, sex, sexual harassment, sexual orientation, gender identity, age (40 and over), disability, veteran status, pregnancy, childbirth, or pregnancy-related conditions, genetic information, or other bases protected by applicable federal, state, or local law) • (b) Harassment based on one or more protected classes, including sexual harassment (policy 162), or • (c) Retaliation(a negative action or discouraging someone from reporting discrimination/harassment or participating in a University internal investigation). If you need accommodations, please see the appropriate contact below: • ADA: Students who need accommodations because of a disability may contact the UVU Accessibility Services Department (ASD), located on the Orem Campus in LC 312. To schedule an appointment or to speak with a counselor, call the ASD office at 801-863-8747. Deaf/Hard of Hearing individuals, email nicole.hemmingsen@uvu.edu or text 385-208-2677. • Pregnancy: Accessibility Services or Office for Equal Opportunity & Affirmative Action (EOAA) - http://www.uvu.edu/equalopportunity/, ext. 5704, BA 203. • Religious: You may begin by asking your professor. If you have questions or concerns, contact the EOAA. • Veterans (Military Leave, Course Withdraw, or Other Assistance): Veteran Services - http://www.uvu.edu/veterans/, ext. 8212, WB100. • As the result of unwanted sexual harassment, contact, stalking, dating or domestic violence: • EOAA or Student Ombuds - ext. 7237, SC 107 • Student Health Services, if Anonymity is desired- ext. 8876, SC 221 • UVUPD, if you wish to press criminal charges - ext. 5555, GT 331

Information about gender neutral facilities or gender personal pronouns:

You may ask your professor/supervisor to use your personal name and pronouns. If you are unclear about pronouns or identifiers, you may request clarification. If you have questions or concerns contact LGBT Student Services - https://www.uvu.edu/multicultural/lgbt/ LA 126, ext. 8885 or the EOAA.

Disability Services

Students who need accommodations because of a disability may contact the UVU Accessibility Services Department (ASD), located on the Orem Campus in LC 312. To schedule an appointment or to speak with a counselor, call the ASD office at 801-863-8747. Deaf/Hard of Hearing individuals, email nicole.hemmingsen@uvu.edu or text 385-208-2677. Veteran's Services UVU is committed to providing a working and learning atmosphere for student veterans and their families. If you are a student veteran or a family member of a student veteran, you are eligible for support services from the Veterans Success Center and may be eligible for Post-9/11 GI Bill benefits. For help receiving your benefits, please contact the Veterans Success Center in the Woodbury Business Building, room 100, via email at veterans@uvu.edu or by calling 801-863-8212. Learn more: http://www.uvu.edu/veterans/

Biology Department Policy Students in this class are expected to understand and use proper English grammar, sentence structure, and spelling. Use of dictionaries during quizzes and exams is NOT allowed. Students are also expected to have basic calculating skills that include fractions, decimals, exponents (e.g., squares & square roots, powers of ten) and the ability to solve simple algebraic expressions. In addition, they must be able to add, subtract, multiply, and divide small numbers without a calculator. Understanding of logarithms (logs) will be helpful. Course rigor level should be such that the average grade is about a C.

Student Responsibility It is up to you to see that the requirements for this course are completed. This syllabus is your guide to those requirements. Make certain that you understand what is expected of you. It is up to you to attend class and arrange for time to complete the readings, assignments, and exams. This course is designed for biology majors. If you are not reading at a college level, you WILL have difficulty with this course. There is only one type of "excused" absence, and that is an absence for an official University activity, such as participation in athletics, not just watching but doing, field trip for another course, etc. Make-up exams will NOT be given except for a situation beyond your control for which you provide verifiable (I will verify), written documentation. Family vacations, Illness, family crises, vehicle problems, legal problems, weddings, not studying, or forgetting to take an exam are NOT valid excuses or excused absences, even if

you tell me beforehand; emergencies are. If such difficulties require asignificant number of absences during the semester, you are advised to drop the course and take it at another time. You do not need to call me if you will not be in class, I don't take roll. Class attendance and obtaining the information for the tests is your responsibility.