STA 3180: Statistical Modeling
Fall 2024

UF Course Catalog: Overview of modern statistical modeling. Topics include linear regression, binary regression and classification, cross-validation, nonlinear regression and smoothing, tree-based methods, the bootstrap, and causal inference. Approaches will be illustrated in R. **Prerequisite:** STA 3100

Instructor: Beth Johnson
Email: ejohnson5@ufl.edu
Office: 116-A Griffin-Floyd Hall
Telephone: 352-273-1897

Instructor Office Hours: TBD
Graduate Teaching Assistants: TBD
GTA Office Hours/Location: TBD

Class Meeting Times and Locations

Day: Tuesday
Time: Period 2 (8:30-9:20)
Place: TUR 2306

Day: Thursday
Time: Periods 2-3 (8:30-10:25)
Place: TUR 2350

Course Description and Objectives

Course Description: This course will introduce students to modern statistical methods essential for understanding large and complex data that arise in fields from biology to astronomy to the social sciences. This course emphasizes the practical application of these methods and their proper use and interpretation. This course is designed for students who are not statistics or data science majors and is the third core course in the data analytics certificate program.

Course Objectives:

Upon completion of this course,

Students will be able to use the statistical program R to:

Analyze data using methods for linear regression

Analyze data using methods for binary regression and classification

Analyze data using methods for nonlinear regression and smoothing

Analyze data using tree-based methods

Students will be able to explain what is necessary to make proper causal inferences.

Students will build collaborative skills by working with a group to complete a project.

Course Materials

Required Textbook: *An Introduction to Statistical Learning with Applications in R*, second edition, by James, Witten, Hastie, and Tibshirani (Springer, 2013), which can be downloaded at no cost from the website for the book (which is maintained by the authors).

Recommended Textbooks:

Hadley Wickham and Garrett Grolemund , 2017. R for Data Science, O'Reilly, Addison Wesley, download free pdf.

Jared P. Lander, 2017. R for Everyone: Advanced Analytics and Graphics, Second Editions, Addison Wesley Data and Analytics Series

Scientific Calculator (around $10 to $15): You will need a calculator capable of basic arithmetic operations and taking square roots will be needed for in-class exams. Internet-enabled electronic devices, such as cell phones or tablets, cannot be used as calculators during exams.

Web-enabled device: You will need some type of web-enabled device such as a laptop, smartphone, or tablet to use in-class to access Canvas as needed.

Course Resources

The Canvas course website will be used extensively throughout the semester to post notes and make course announcements. You must log on using your gatorlink username and password and access the course webpages from there. Important information about the course will be posted here including this syllabus, announcements, notes, assignments and your grades throughout the semester and computer output to supplement the examples done in class. Please check this site often.

Course Computer Software

Some assignments will require you to use the statistical software package, R, to analyze and visualize data. R is free and used around the world. There are now over 13,000 R packages.

The Comprehensive R Archive Network (CRAN) is the primary place to download R. The Lander and Wickham texts above describe obtaining R and RStudio. The free desktop version of RStudio is fine.

Course Approach

In this third core course of the data analytics certificate, we will focus on developing the following skills: using statistical software to analyze data, interpreting results from a statistical analysis, and stating clear conclusions in context for a lay audience.

These skills will all be assessed through various modes such as homework-lab assignments, a class project, and in-class exams.

Help

Remember to ask for help! You can come by during my scheduled office hours or make an appointment to see me. I can also answer some questions via email. *Emails received during the working week will be answered within 24 hours however emails received over the weekend may not be answered until Monday morning.*

- Always use GatorMail for email. I do not check Canvas inbox regularly.
- Always put STA 3180 in the subject line of your email. I teach multiple courses and use course numbers to search emails from students.

Course Assignments

Your final course grade will be based on a combination of assessment types including computer lab activities, in-class exams, and a final group project. Due dates will be posted on the course schedule on the Canvas course page and announced in class.

Computer Lab activities: Participation is an important component of this course as we learn by doing, so **non-attendance on lab days will result in a 10% penalty on the individual lab score**. There will be a lab activity after most course topics to provide you with experience working with data on a small, guided activity. You will work with an assigned group to complete the assignment, but **everyone is required to submit their own report to be graded.** Groups will be rotated during the semester.

Due dates for the lab activities will be posted on Canvas. There will be a 10% penalty per day for late work and no assignments will be accepted more than one day after the due date. Your lowest lab score will be dropped provided that you complete all seven assignments.

Exams: There will be four in-class exams which will require students to read and interpret statistical output and answer conceptual questions. In case of conflict or illness, if a student is unable to take an exam at the scheduled time, they must get in touch with the instructor prior to the exam time for any arrangements to be made for a makeup. Each case will be reviewed individually. Valid and detailed documentation is a prerequisite under such extenuating circumstances. A grade of zero is the minimum punishment of any type of dishonesty on an exam.

Group Project: Students will work in assigned or self-selected teams of no more than four people to complete a research project based on a topic of interest. Each team will: 1) write a statistical paper, of approximately 5-7 pages, and 2) prepare two 10-minute video presentations of the project. The first video will cover the introduction, research questions and the data collection process while the second video will contain the methods, results and conclusions for your project. Rubrics will be provided.

Each group must obtain approval of their data set and research questions before data analysis begins. Each group must submit a statement of each member's duties in the project, i.e., data cleaner, literature reviewer, etc. Each member of a team will be required to utilize a different statistical technique or type of analysis to analyze the common data set and include their own contribution in the final common report and video. If a student fails to perform their duties as part of the group a point deduction may be applied to that student's individual project grade.

It is very important to review the work of your peers so each student will be assigned to view two other group project videos and to provide constructive comments on each video. Discussion topics will be provided.

Grading Scheme:

Exams (4 at 12% each)	48%
Computer Labs (best 6* at 5% each)	30%
Group Research Project:	
Video Presentations	10%
Written Report (rubric will be provided)	9%
Group Critiques	3%

The instructor reserves the right to adjust the percentages if needed.

Your final overall numeric score is rounded to the nearest integer.

So, for example, if your average is 76.4 your grade will be 76.
If your grade is 76.5, your grade will be 77.

Letter grades will be assigned according to the table shown.
Note: No D+ or D- will be assigned.

Numeric Score	Letter Grade
91-100	A
88-90	A-
85-87	B+
81-84	B
78-80	B-
75-77	C+
68-74	C
65-67	C-
60-64	D
0-59	E

Course Policies

Grading Policies:

Requirements for class attendance and make-up exams, assignments, and other work in this course as well as policies regarding absences, religious holidays, illness, and student athletes are consistent with UF Attendance Policies.

Additional make-up policy requirements:

- Every effort should be made to complete the assignment/exam during the open period. Only extreme situations will warrant a makeup. Contact the instructor prior to the exam - as soon as you realize you will be unable to take the assignment/exam at the scheduled time. Each case will be reviewed individually. Valid and detailed documentation is a prerequisite for scheduling a makeup under such extenuating circumstances.
- If you have an emergency on the day of the assignment/exam, the instructor must be contacted by midnight of the day of the assignment/exam.
- Make-up exams will be scheduled within a week from the assignment deadline. Student is responsible for attending scheduled make-up. Instructor reserves the right to utilize the UF posted final exam day as a make-up date.
- The UF Religious Holidays Policy is available at: https://catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/#religiousholidaystext.
- Additional Note: Being on vacation or booking a trip prior to the completion of the semester is not a valid reason to request a makeup. Please reference the most recent Academic Calendar, https://catalog.ufl.edu/UGRD/dates-deadlines/pdfs/.

*If you have a disability that requires academic accommodation, contact the Disability Resource Center (DRC). The DRC will provide documentation to the students who must then provide this documentation to the instructor when requesting information. You must submit this documentation prior to submitting any assignments for which you are requesting accommodation.

* Incompletes are only assigned when extraordinary circumstances (such as an accident, or extended hospitalization), arising after the date for dropping the course, prevent the student from completing the course requirements. Having a failing grade in the course is not a valid reason for requesting an Incomplete. Information on Medical Withdrawal can be found at https://umatter.ufl.edu/ . Information on how to Drop a class can be found in UF's Academic Catalog https://catalog.ufl.edu/ and https://catalog.ufl.edu/UGRD/academic-regulations/dropping-courses-withdrawals/

*There is no "extra credit" or forgiven grades – you are responsible for all your work done (or left undone).

*If you have a question concerning a graded assignment, you should notify me within seven days after a graded assignment is posted to schedule a meeting.

Classroom Behavior: During class students should silence their cellular phones and refrain from eating, drinking, reading newspapers, doing homework, listening to music, excessive talking and all other behaviors that are distracting and disrespectful to the instructor and their fellow students.

Privacy Policy: Student records are confidential. Only information designated "UF directory information" may be released without your written consent. This applies to parents or anyone else who contacts me about your grades.

Faculty Course Evaluations: Student feedback is welcomed by the instructor and beneficial to future students in the course. Students are requested to provide feedback on the quality of instruction in this course by completing a brief confidential evaluation towards the end of the semester at https://evaluations.ufl.edu. Summaries of the evaluation results can be found at https://evaluations.ufl.edu/results.

Other University Services

U Matter, We Care: Your well-being is important to the University of Florida. The U Matter, We Care initiative is committed to creating a culture of care on our campus by encouraging members of our community to look out for one another and to reach out for help if a member of our community is in need. If you or a friend is in distress, please contact umatter@ufl.edu so that the U Matter, We Care Team can reach out to the student in distress. A nighttime and weekend crisis counselor is available by phone at 352-392-1575. The U Matter, We Care Team can help connect students to the many other helping resources available including, but not limited
to, Victim Advocates, Housing staff, and the Counseling and Wellness Center. Please remember that asking for help is a sign of strength. In case of emergency, call 9-1-1.

***Sexual Assault Recovery Services (SARS): Student Health Center, 392-1161**
***University Police Department, 392-1111 (or 9-1-1 for emergencies), http://www.police.ufl.edu**

***Student Health Care Center:** Call 352-392-1161 for 24/7 information to help you find the care you need or visit the Student Health Care Center website.

*__GatorWell Health Promotion Services:__ For prevention services focused on optimal wellbeing, including Wellness Coaching for Academic Success, visit the GatorWell website or call 352-273-4450.

Academic Resources

E-learning technical support: Contact the UF Computing Help Desk at 352-392-4357 or via e-mail at helpdesk@ufl.edu.

Writing Studio: 2215 Turlington Hall, 352-846-1138. Help brainstorming, formatting, and writing papers.

Student Complaints On-Campus: Visit the Student Honor Code and Student Conduct Code webpage for more information.

Tentative Schedule Fall 2024					
Date	**Day**	**Topic**	**Topic**	**Textbook Chapter**	**Assignment Due Dates**
22-Aug	R	1	Introduction to Statistical Learning	1 and 2	
		1	Simple Linear Regression with Inference and Residual Plots	3.1	
27-Aug	T	2	Transformations,Outliers,and Influential Points	3.2	
29-Aug	R	2	Multiple Linear Regression with Assumption Checks	6.1	
		3	Lab1: Simple Linear and Multiple Regression		Lab 1: Tuesday 9/3
3-Sep	T	3	Model Selection		
5-Sep	R		Exam 1 - Simple and Multiple Regession		
10-Sep	T	5	Logistic Regression	4.1-4.3	
12-Sep	R	5	Multiple Logistic Regression	4.1-4.3	
		5	Lab 2: Logistic Regression	4.1-4.3	Lab 2: Tuesday 9/17
17-Sep	T	6	Linear Discriminant Analysis	4.4	
19-Sep	R	6	Linear Discriminant Analysis		
		6	Lab3: Linear Discriminant Analysis		Lab 3: Tuesday 9/24
24-Sep	T		Work on Projects/Review		Project teams are formed
26-Sep	R		Exam 2- Logistic and Linear Discriminant		
1-Oct	T	7	Cross-Validation and the Bootstrap	5.1 and 5.2	
3-Oct	R	8	Lab 4: CV and Bootstrap		Lab 4: Tuesday 10/8
		8	Lab 4: CV and Bootstrap		
8-Oct	T	9	Polynomial Regression	7.1	
Oct-24	R	9	Polynomial Regression	7.1	
		10	Smoothing/Transformations	7.2-7.4	
15-Oct	T	10	Work on Projects		ALL data sets and research questions must be approved
17-Oct	R	10	Lab5: Regression and Smoothing		Lab 5: Tuesday 10/22
		11	Lab5: Regression and Smoothing		
22-Oct	T	11	Work on Projects		Project background and data cleaning
22-Oct	R	11	Exam 3: Polynomial Regression and Resampling Methods		
29-Oct	T	13	Cluster Analysis	12.1, 12.4	
31-Oct	R	13	Cluster Analysis	12.4	
			Cluster Analysis	12.4	
5-Nov	T		Lab 6: Cluster Analysis		Lab 6: Tuesday 11/12
7-Nov	R	14	Principal Components Analysis (PCA)	12.2	
		14	Principal Components Analysis (PCA)	12.2	Project draft analysis and preliminary reults
12-Nov	T	13-14	Lab 7: PCA	12.2	Lab 7: Tuesday 11/19
14-Nov	R	13-14	Lab 7: PCA	12.2	Lab 7: Tuesday 11/19
		15	Work on Projects		
19-Nov	T	12	Tree Based Methods and Random Forests	8.1,8.2	
21-Nov	R	15	Exam 4: Cluster and PCA		
		15	Exam 4: Cluster and PCA		
3-Dec	T		Projects		Videos and Written Report Due Wednesday 12/4