STA 4210 (Class Number: 17785)

Regression Analysis

Fall 2021

MWF 1:55-2:45 p.m. (Room: FLG 0230)

The course is offered fully in person. You are expected to wear approved face coverings at all times during class and within buildings even if you are vaccinated.

The course syllabus is a general plan for the course; deviations announced to the class by the instructor may be necessary.

Instructor: Juhyung Lee

Office:	Griffin-Floyd 103B
Office Hours:	TR 11:00 a.m.-12:00 p.m. via Zoom (link: TBA)
Email:	juhyunglee@ufl.edu

Teaching Assistant: Cheng Zeng

Office:	TBA
Office Hours:	TBA
Email:	czeng1@ufl.edu

Course Description: The course is primarily on the linear regression model, for which the main techniques are rooted in the method of least squares. Procedures are motivated by applications. Mathematical results are stated and explained, and occasionally derived. A course in mathematical statistics is helpful but not strictly required. The focus is on carrying out and explaining the methods. After some review of basic statistics, the course proceeds systematically through the simple regression model, the multiple regression model, the matrix formulations of both these models, and a number of related tools such as model diagnostic measures, collinearity statistics, and variable selection procedures.

Prerequisites: STA 3024 or STA 3032 or (STA 4321 and STA 2023) or (MAS 3114 and STA 2023) or (MAS 4105 and STA 2023). STA 3100 will be an added prerequisite to this course, so it is highly recommended you have some programming experience.

Course Website: e-Learning. Check course website regularly for any updates.

Course Material: Lecture notes posted on course website.

Required Text:

1. *Applied Linear Statistical Models* (5th ed.) by Kutner, Nachtsheim, Neter, and Li.

Software: We will use the free statistical software R.

Course Communication: Office hours and email.

- Always wear a mask if you plan to visit me or TA in person.

- Put **4210** in the subject line of your email. I use course numbers to search emails from students.

Course Objectives:

1. With two quantitative variables, be able to carry out simple regression analysis and to correctly interpret such an analysis.

2. Know the multiple linear regression model in its matrix form, including all the common variations on this model (e.g., continuous predictors, categorical predictors, square and interaction terms).

3. Be able to carry out and interpret inference procedures for the various types of multiple regression model, including the quadratic regression model.

4. Know what the purposes of diagnostic methods in simple and multiple regression are; be able to carry out several common diagnostic procedures and interpret them.

5. Know what multicollinearity is, why it is an issue in multiple regression, and how to analyze and deal with its presence.

6. Know several measures of model performance, how to compute and interpret them for a multiple regression model.

Assignments and Grading:

- Grades will be based on the following components.

Homework	25%	
Exam 1	25%	in-class, September 27 (tentative)
Exam 2	25%	in-class, November 1 (tentative)
Exam 3	25%	in-class, December 6 (tentative)
Total	100%	

- Homework should be submitted to the relevant dropbox on course website in the form of **one single pdf file**.

- **No late homework will be accepted.** However, the lowest homework score will be dropped at the end of the semester.

- **All exams will be closed-book.** However, you may bring one letter-size (8.5 by 11 inches) cheat sheet (i.e., two pages) to each exam. You may bring your old cheat sheet(s) in addition to the current one.

- **Bring a scientific calculator (no cell phones!) to each exam.**

- **If you miss an exam, you must contact me by the end of the day.** Your case will be reviewed by me and if approved, (i) the other two exams will be reweighted if you missed Exam 1 or Exam 2; (ii) a take-home exam will be given if you missed Exam 3. This policy cannot be applied to more than one exam.

- I will strictly follow the cutoffs below for letter grades.

		A	92 to 100	A-	90 to < 92
B+	87 to < 90	B	82 to < 87	B-	80 to < 82
C+	77 to < 80	C	72 to < 77	C-	70 to < 72
D+	67 to < 70	D	62 to < 67	D-	60 to < 62
E	< 60				

Academic Integrity:
You may discuss homework with each other, but you must write up your solutions independently. You may never discuss exams with each other. You are held accountable to the UF Student Honor and Conduct Code.

Students with Disabilities:
Students requesting accommodation for disabilities must first register with 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.

Course Evaluations:
Students are expected to provide feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at `https://gatorevals.aa.ufl.edu/students/`. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via `https://ufl.bluera.com/ufl/`. Summaries of course evaluation results are available to students at `https://gatorevals.aa.ufl.edu/public-results/`.

Tentative Course Outline

Text	Content
Appendix A	Review of Basic Statistical Results
Chapter 1	Simple Linear Regression (SLR)
Chapter 2	Inferences in SLR
Chapter 3	Diagnostics and Remedial Measures
Exam 1	
Chapter 4	Simultaneous Inferences
Chapter 5	Matrix Approach
Chapter 6	Multiple Linear Regression I
Chapter 7	Multiple Linear Regression II
Exam 2	
Chapter 8	Qualitative Predictors
Chapter 9	Model Selection and Validation
Chapter 10	Diagnostics
Exam 3	