STA 4210 (Class Number: 19975)

Regression Analysis

Spring 2022

MWF 10:40-11:30 a.m. (Room: AND 0134)

*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: | TBA |
| Email: | juhyunglee@ufl.edu |

**Teaching Assistant:** Maoran Xu

| | |
|---|---|
| Office: | TBA |
| Office Hours: | TBA |
| Email: | maoranxu@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, explained, and derived. 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 3100 and (STA 3024 or STA 3032 or STA 4321 or MAS 3114 or MAS 4105).

**Course Website:** e-Learning. Check course website at least once a day for updates.

**Course Material:** Lecture notes (posted on course website) and in-class notes (will not be posted).

**Required Text:**

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

**Software:** R (free statistical software).

**Course Communication:**   Office hours and email.

- **Always wear a mask during in-person office hours.**

- **Always use GatorMail for email.** I do not check Canvas inbox.

- **Always put 4210 in the subject line of your email.** I teach multiple courses and 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, February 7 |
| Exam 2 | 25% | in-class, March 16 |
| Exam 3 | 25% | in-class, April 18 |
| Total | 100% | |

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

| | | | | | |
| --- | --- | --- | --- | --- | --- |
| | | A | $\geq 92$ | 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 | | | | |

- **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 Exam $k, k = 1, 2, 3$, you will receive $10(9 - k)\%$ of the minimum of your other two exam scores for Exam $k$.** For example, suppose you missed Exam 3, where your homework, Exam 1, and Exam 2 scores are 100, 90, and 80, respectively. Then your Exam 3 score would be $80 \times 10(9 - 3)\% = 48$ and your final score would be

$$100 \times 25\% + 90 \times 25\% + 80 \times 25\% + 48 \times 25\% = 79.5,$$

  which is a C+. **You will automatically fail the course if you miss more than one exam.**

- **If you take all three in-class exams, there will be 3% extra credit toward your final grade.** For example, if your homework, Exam 1, Exam 2, and Exam 3 scores are 100, 90, 80, and 70, respectively, your final score would be

$$100 \times 25\% + 90 \times 25\% + 80 \times 25\% + 70 \times 25\% + 100 \times 3\% = 85 + 3 = 88,$$

  which is a B+.

**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/`.