STAT 4131
Regression Analysis and Statistical Learning
Fall 2025

Instructor:	Dr. Emily L. Kang		
Office:	4428B French Hall West	**Office Hours:**	T R 2:00 PM – 3:00 PM Office
Email:	kangel@ucmail.uc.edu		W 3:30 PM – 4:30 PM Zoom
When:	TR 9:30 AM – 10:50 AM		(or by appointment)
Where:	Room 270 60WCharl		

Course Objective

This 3-credit undergraduate course introduces classical statistical inference and data analysis through simple and multiple linear regression models, and extends to modern statistical learning techniques such as ridge regression, the LASSO, decision trees, and random forests. Additional topics, including logistic regression, linear discriminant analysis, nonparametric regression, and neural networks, will be covered as time permits. Students will gain hands-on experience implementing methods that have had significant impact in science and industry. The course also provides preparation for advanced undergraduate and graduate-level courses (5000/6000 level) in Statistics.

Learning Outcome & Transferable Skills

In this course, students will develop a set of transferable skills that are valuable across a wide range of career paths:

- Critical Thinking: Develop a solid understanding of basic concepts in statistics and probability, evaluate statistical arguments critically, and recognize the broader importance of statistical reasoning.
- Quantitative Reasoning: Apply statistical methods to analyze data, interpret results, and solve real-world problems.
- Scientific Literacy: Use statistical techniques such as hypothesis testing and regression appropriately to analyze scientific data and draw valid conclusions.
- Teamwork and Collaboration: Work effectively in groups to complete a project, share ideas, and contribute to collective problem-solving.

Assumed Background Knowledge

Prerequisite: Probability & Statistics I & II (15 STAT 2037, 3038). Students are assumed to have background knowledge of basic concepts of probability and distribution theory and inferences about population means and variance.

Textbook

- *Applied Linear Statistical Models, 5th edition*, by Kutner, Nachtsheim, Neter, & Li. (KNN)
- *An Introduction to Statistical Learning with Applications in R*, 2nd edition, by James, Witten, Hastie & Tibshirani. (ISLR) `https://www.statlearning.com/`

Software
`R` will be used throughout this course. You can download it for free at `R website`. I also encourage you to install RStudio, a user-friendly interface for R. All course data sets and R code will be posted on Canvas.

Homework
Homework will be assigned and announced in class as well as on Canvas throughout the semester. While the instructor encourages students to work together and learn from each other, all submitted work must **be your own**. You may discuss homework problems with classmates, but each student is responsible for completing all work and writing up solutions independently. Submitting work that is not your own will be considered a violation of the University's academic integrity policy.
Homework Submission: All homework must be submitted via Gradescope by the stated deadline. This is the only acceptable method of submission; please do not leave hard copies under the instructor's office door or email electronic files. Students can access Gradescope directly through Canvas.
Late work will not be accepted, except in cases of extreme and well-documented circumstances that are approved by the instructor. To accommodate unexpected emergencies, your *lowest homework score will be dropped* when calculating the final grade. Please reserve this "freebie" for when you truly need it.
Communication of Results: The ability to interpret statistical analyses and clearly explain the results to a non-statistician is an essential skill, and we will emphasize this throughout the course. Therefore, *solutions to applied problems must always be written in proper English and framed in the context of the problem. Full credit will not be awarded otherwise.* **For problems involving coding and data analysis, you must submit both the code and its output, along with a clear interpretation of the results as required by the problem.**

Group Final Project
A group project will be assigned using a dataset posted on Kaggle. Each group will submit a written report (maximum 5 pages, single-spaced, 11pt font) and give a class presentation. The project requirements are as follows:

1. Form a group of 2 – 3 students.
2. Provide a descriptive analysis of the dataset using appropriate statistical tools (e.g., summary statistics, scatter plots, histograms, boxplots).
3. Analyze the dataset using linear regression models. Conduct model diagnostics, identify any violations of assumptions, and propose appropriate remedies. Select the "best" model for your dataset. Upload predictions to Kaggle, which will automatically rank models from all groups using specified criteria.
4. Apply at least two additional methods introduced in this course and compare results to those from the regression models.
5. Summarize findings in the written report and present them as a group in class.
6. Submit the final report, code, and presentation slides by the deadline (TBD). The instructor may also assign intermediate milestones, such as an initial proposal, a midterm progress report, and a draft presentation.
7. On the cover page of the final report, clearly describe each group member's contributions. All members must sign this cover page. The project will not be graded without these contribution statements and signatures.

Exams
There will be **one in-class written exam** and **one take-home data analysis exam** during the semester. The take-home exam is tentatively scheduled for Weeks 7–8, and the in-class written exam for Weeks 10–12.
Full credit for exam problems will only be awarded when solutions include proper justification written in clear English and framed in the context of the problem.
For the in-class written exam, you are permitted to bring to the exam *one cheat sheet (8.5" × 11" sheet of paper, both-sided, any content, but nothing photocopied).* A calculator (no cell phone calculators) needs to be brought to the written exam. *Sharing the cheat sheets and/or calculators with other students during the exams is prohibited.*
More specifics regarding these exams will be provided later in the semester. Make-up examination will be given only if a student has a written medical excuse signed by a doctor or other health professional and if he/she contacts the instructor prior to the exam.

Final Grade
Your final course grade will be determined by the following components:

Class participation	Homework	In-Class Exam	Take-Home Exam	Group Final Project
5%	20%	30%	20%	25%

Your final course letter grade will be assigned according to the following grading scale:

		B+	87 – 89	C+	77 – 79	D+	67 – 69		
A	93 – 100	B	83 – 86	C	73 – 76	D	63 – 66		
A-	90 – 92	B-	80 – 82	C-	70 – 72	D-	60 – 62	F	below 60

Attendance
Beginning Fall 2016, Title IV provisions require undergraduate students to demonstrate participation in each course in order to remain eligible for federal financial aid. To meet this requirement, the University and the College of Arts & Sciences have implemented a simple procedure through Canvas. When you access the Canvas site for each of your courses, you will see a link in the left-hand control panel titled "Attendance Verification." Clicking this link will take you to a short question. Submitting your response will serve as verification of your participation.

Email Correspondence
The best way to contact the instructor is via email at `kangel@ucmail.uc.edu`. All course-related email communication must be conducted through your UC email account or Canvas. The instructor will not send messages to external accounts (e.g., Gmail).

Communication Devices
Personal communication devices (e.g., cell phones) must be **turned off or set to vibrate** during class. Please refrain from texting or using devices in a way that disrupts class.

Virtual Office Hours
In addition to in-person office hours, the instructor will hold virtual office hours via Zoom. Students may reserve a time slot by going to *Canvas → Calendar → Find an Appointment.*

Campus Safety Measures
All faculty, staff, instructors, and students are required to follow campus safety measures, which can be found here: https://www.uc.edu/about/publicsafety.html.

Academic Integrity
Please help maintain an academic environment of mutual respect and fairness. You are expected to produce original and independent work on exams. For homework, discussion is encouraged; however, copying someone else's work and presenting it as your own constitutes plagiarism. All students must submit their own written work in their own words. Academic misconduct **will not** be tolerated. For more information, see: `https://www.artsci.uc.edu/student-experience/academic-forms-and-policies/misconduct-process.html`.

Accessibility Resources
Reasonable accommodations will be provided for students with documented needs. To access these accommodations, students must contact Accessibility Resources as described on their website: `https://www.uc.edu/campus-life/accessibility-resources.html`.

Religious Accommodations
Ohio law and the University's Student Religious Accommodations for Courses Policy 1.3.7 permits a student, upon request, to be absent for reasons of faith or religious or spiritual belief system or participate in organized activities conducted under the auspices of a religious denomination, church, or other religious or spiritual organization and/or to receive alternative accommodations with regard to examinations and other course requirements due to an absence permitted for the above-described reasons. Not later than fourteen days after the first day of instruction in the course, a student should provide the instructor with written notice of the specific dates for which the student requests alternative accommodations. University policy can be found at: `https://www.uc.edu/about/equity-inclusion/equal-opportunity/student-religious-accommodations-for-courses-policy.html`. Additional information about or questions related to the policy can be directed to the Office of Equal Opportunity (OEO).

Drop and Withdraw Dates
The last day to drop without entry to the academic record is Monday, September 8, 2025. The last day to withdraw is Friday, November 21, 2025.

Holidays
There will be no class on the following days:
Labor Day: Monday, September 1, 2025
Reading Days: Thursday, October 9 – Friday, October 10, 2025
Veteran's Day: Tuesday, November 11, 2025
Thanksgiving Weekend: Thursday, November 27 – Sunday, November 30, 2025

Receiving an 'I' for the course:
You cannot receive receive an incomplete for the course unless 70% of the work in the course (especially the attendance) has been completed. Extenuating circumstances will be handled on a case-by-case basis.

Tentative Schedule

Chapter(s)	**Description**	**Time**
KNN 1-4	Overview, simple linear regression	3 wks
KNN 5-6	Multiple regression	3 wks
KNN 7-10	Model selection, polynomial regression, & diagnostics	3 wks
ISLR 6	Shrinkage methods: Ridge regression and LASSO	2 wks
ISLR 8	Tree-based methods	2 wks
ISLR 4	Classification and Generalized Linear Models	1 wks
ISLR 7 & 10	(brief) Nonparametric regression & Deep learning	1 wk as time permits

Note: The instructor reserves the right to change the class syllabus to meet class needs.

STAT 4131 (Fall 2025 TR 9:30AM – 10:50AM)

I read the entire course syllabus thoroughly and will obey the policies for this course through the semester.

Name (Last, First): _____

Major: _____

Signature: _____

Date: _____