湖北工業大學
HUBEI UNIVERSITY OF TECHNOLOGY

Course Title	R for Data Science
Course Code	STAT 3613
Semester	Summer 2026
Course Length	8 Weeks, 60 Contact Hours
Credits	4
Instructor	TBA
Office	TBA
Email	TBA
Prerequisite	N/A
Antirequisite	STAT 3611 R for Data Science (4 Weeks)

Course Description:

This course introduces the foundational techniques in data science using R. Topics include data manipulation and cleaning, exploratory data analysis, statistical inference, linear and nonlinear regression, classification, clustering, resampling methods, and model evaluation. Students will also explore advanced methods such as tree-based models, support vector machines, and ensemble techniques. Emphasis is placed on both the theoretical principles behind these techniques and their practical application.

Course Goals:

Students who successfully complete this course will demonstrate competency in the following general education core goals:

- **Critical Thinking Skills** – Students will engage in analytical thinking, demonstrating the ability to critically evaluate, synthesize, and apply knowledge to complex problems, and construct well-reasoned solutions and arguments.
- **Independent Research and Inquiry** – Students will conduct independent research, utilizing academic resources to explore relevant topics, formulating research questions, analyzing data, and presenting findings in a coherent, scholarly manner.
- **Problem-Solving and Application** – Students will apply theoretical concepts and methodologies learned in the course to real-world problems, demonstrating the ability to develop practical solutions informed by academic inquiry.
- **Collaboration and Teamwork** – Students will collaborate effectively with peers in group projects and discussions, contributing to a collective goal while respecting different perspectives and working towards a shared purpose.

Student Learning Outcomes:

Upon completion of this course, students will be able to:

- import, munge, summarize, and visualize data in the R statistical software environment;
- interpret the mathematical underpinnings of regression models, including model assumptions and inferences;
- apply regression models to data in the R statistical software environment;
- approach scientific problems and communicate statistical results from a data scientist's perspective.

Textbooks/Supplies/Materials/Equipment/ Technology or Technical Requirements:

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2021) *An Introduction to Statistical Learning with Applications in R,* Second Edition. Springer.
Hadley Wickham, Mine Çetinkaya-Rundel, Garrett Grolemund (2023). *R for Data Science.* O'Reilly Media.

Course Requirements:

In-Class Quizzes

A very short quiz will be given during the first 5-10 minutes of a randomly picked class meeting. There will be altogether 7 quizzes throughout the semester and the lowest two scores will be dropped. All quizzes will contribute equally.

Homework

There will be 5 graded homework assignments using R.

Project

Students will be expected to complete a written project at the end of the semester and present their findings to the class. The grade for the final project will consist of three main components: statistical modeling decisions and code, a written report, and an oral presentation. The first component is based on whether the student used appropriate statistical tools for the setting and objectives, whether the coding done was efficient and correct, and whether the conclusions are consistent with their analysis. The written report will state the objectives of the study, describe data collection, describe the statistical model used, explain any assumptions required by the analysis, and provide conclusions for the main study questions. The oral presentation will be a ten-minute presentation that should cover the key components of the written report. Presentations should clearly state the objectives of the project while using visualizations to illustrate the main results and findings of the project.

Exams

You will have two take-home exams that are to be assigned roughly in the middle of and near the end of the semester.

Assessments: Activity	**Percent Contribution**
In-Class Quizzes	15%
Homework	15%
Project	30%
Take-Home Exam 1	15%
Take-Home Exam 2	25%

Grading:

Final grades will be based on the sum of all possible course points as noted above.

Grade	Percentage of available points
A	94-100
A-	90-93
B+	87-89
B	84-86
B-	80-83
C+	77-79
C	74-76
C-	70-73
D	64-69
D-	60-63
F	0-59

Course Schedule:
The schedule of activities is subject to change at the reasonable discretion of the instructor. Minor changes will be announced in class, major ones provided in writing.

STAT 3613 Schedule		
Lecture	Topic	Reading Chapters
L1	Course overview Introduction to statistical learning	*James 1-2*
L2	Install and configuration of R programming environment	--
L3	Basic language elements and data structures	*James 2*
L4	Data visualization: The grammar of graphics and ggplot2 API	*Hadley* 1-2
L5	Data visualization: Designing effective plots	*Hadley* 1-2
L6	Data visualization: Small multiples and patchwork **HW1 due**	*Hadley* 1-2
L7	Data wrangling: Data import, Data transformation	*Hadley* 3, 7
L8	Data wrangling: Tidy data, Strings, Factors	*Hadley* 5, 14-16
L9	Data wrangling: Dates and times, Iteration, Skills assessment	*Hadley* 17, 26
L10	Models: Data science ethics, Building models	*Hadley* 10, 20
L11	Models: Tidy models introduction **HW2 due**	*Hadley* 22-24
L12	Communicating data science	*Hadley* 28-29
L13	Communicating data science (cont.)	*Hadley* 28-29
L14	Linear Regression: Model assessment, Some practical issues	*James 3*
L15	Classification: Classification theory and evaluation **Take-Home Exam 1 Submission**	*James 4*
L16	Resampling Methods **HW3 due**	*James 5*
L17	Linear Model Selection and Regularization: Subset selection, Ridge regression, Lasso	*James 6*
L18	Tree-Based Methods: Regression tree, Pruning, Tree ensembles	*James 8*
L19	Classification Tree and Boosting	*James 8*
L20	Nonlinear Regression	*James 9*
L21	Linear Classifier: Logistic regression with regularization, Linear discriminant analysis	*James 9*

	HW4 due	
L22	Clustering	*James 12*
L23	Clustering (Cont.)	*James 12*
L24	Projects: Work session	--
	HW5 due	
L25	Projects: Presentations day	--
	Take-Home Exam 2 Submission	

Accommodation Statement:

Academic accommodations may be made for any student who notifies the instructor of the need for an accommodation. It is imperative that you take the initiative to bring such needs to the instructor's attention, as he/she is not legally permitted to inquire. Students who may require assistance in emergency evacuations should contact the instructor as to the most appropriate procedures to follow.

Academic Integrity Statement

Each student is expected to maintain the highest standards of honesty and integrity in academic and professional matters. The University reserves the right to take disciplinary action, up to and including dismissal, against any student who is found guilty of academic dishonesty or otherwise fails to meet the standards. Any student judged to have engaged in academic dishonesty in coursework may receive a reduced or failing grade for the work in question and/or for the course.

Academic dishonesty includes, but is not limited to, dishonesty in quizzes, tests, or assignments; claiming credit for work not done or done by others; hindering the academic work of other students; misrepresenting academic or professional qualifications within or without the University; and nondisclosure or misrepresentation in filling out applications or other University records.

Other Items:

Attendance and Expectations

All students are required to attend every class, except in cases of illness, serious family concerns, or other major problems. We expect that students will arrive on time, be prepared to listen and participate as appropriate, and stay for the duration of a meeting rather than drift in or out casually. In short, we anticipate that students will show professors and fellow students maximum consideration by minimizing the disturbances that cause interruptions in the learning process. This means that punctuality is a must, that cellular phones be turned off, and that courtesy is the guiding principle in all exchanges among students and faculty. You will be responsible for the materials and ideas presented in the lecture.

Assignment Due Dates

All written assignments must be turned in at the time specified. Late assignments will not be accepted unless prior information has been obtained from the instructor. If you believe you have extenuating circumstances, please contact the instructor as soon as possible.

Make-Up Work
The instructor will not provide students with class information or make-up assignments/quizzes/exams missed due to an unexcused absence. Absences will be excused and assignments/quizzes/exams may be made up only with written documentation of an authorized absence. Every effort should be made to avoid scheduling appointments during class. An excused student is responsible for requesting any missed information from the instructor and setting up any necessary appointments outside of class.

Access, Special Needs, and Disabilities
Please notify the instructor at the start of the semester if you have any documented disabilities, a medical issue, or any special circumstances that require attention, and the school will be happy to assist.