Math 478: Statistical Methods in Data Science
Fall 2024 Course Syllabus

NJIT Academic Integrity Code: All Students should be aware that the Department of Mathematical Sciences takes the University Code on Academic Integrity at NJIT very seriously and enforces it strictly. This means that there must not be any forms of plagiarism, i.e., copying of homework, class projects, or lab assignments, or any form of cheating in quizzes and exams. Under the University Code on Academic Integrity, students are obligated to report any such activities to the instructor.

COURSE INFORMATION

Course Description: This course introduces to students concepts in statistical methods used in data science, including data collection, data visualization and data analysis. Emphasis is on model building and statistical concepts related to data analysis methods. The course provides the basic foundational tools on which to pursue statistics, data analysis and data science in greater depth. Topics include sampling and experimental design, understanding the aims of a study, principles of data analysis, linear and logistic regression, resampling methods, and statistical learning methods. Students will use the R statistical software.

Number of Credits: 3

Prerequisites: Math 333 with a grade of C or better or Math 341 with a grade of C or better.

Course-Section and Instructors:

Course-Section	Instructor
Math 478-001	Professor C. Shi

Office Hours for All Math Instructors: Fall 2024 Office Hours

Required Textbook:

Title	*An Introduction to Statistical Learning: with Applications in R*
Author	Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani
Edition	2nd ed. 2021 edition
Publisher	Springer
ISBN #	978-1071614174

University-wide Withdrawal Date: The last day to withdraw with a **W** is **Monday, November 11, 2024**. It will be strictly enforced.

Homework: There will be six or seven homework assignments. Homework must be submitted through Canvas and needs to be uploaded as ONE PDF file. Late homework is not accepted. For the homework problems that require the use of R, please submit the code and figures in addition to the answers.

Midterm Exam: There will be an in-class midterm exam during the semester, which will be held on the following day:

Midterm Exam	Oct 22, 2024

Final Exam: There will be a comprehensive in-class exam, which will be held during the Finals week:

Final Exam Period	Dec 15 – 21, 2024

The final exam will test your knowledge of all the course material taught in the entire course. Make sure you read and fully understand the Math Department's Examination Policy. This policy will be strictly enforced.

POLICIES

DMS Course Policies: All DMS students must familiarize themselves with, and adhere to, the Department of Mathematical Sciences Course Policies, in addition to official university-wide policies. DMS takes these policies very seriously and enforces them strictly.

Grading Policy: The final grade in this course will be determined as follows:

Homework	30%
Midterm Exam	30%
Final Exam	40%

Your final letter grade will be based on the following tentative curve.

A	90 - 100	**D**	60 - 70
B	80 - 90	**F**	0 - 60
C	70 - 80		

Attendance Policy: Attendance at all classes will be recorded and is **mandatory**. Please make sure you read and fully understand the Math Department's Attendance Policy. This policy will be strictly enforced.

Cheating in Exams: Once caught, the exam will be assigned zero points. To prevent cheating, please leave at least one seat empty between you and your neighbors.

Makeup Exam Policy: There will be **NO MAKE-UP EXAMS** during the semester. In the event an exam is not taken

under rare circumstances where the student has a legitimate reason for missing the exam, the student should contact the Dean of Students office and present written verifiable proof of the reason for missing the exam, e.g., a doctor's note, police report, court notice, etc. clearly stating the date AND time of the mitigating problem. The student must also notify the Math Department Office/Instructor that the exam will be missed.

Cellular Phones: All cellular phones and other electronic devices must be switched off during all class times.

ADDITIONAL RESOURCES

Math Tutoring Center: Located in the Central King Building, Lower Level, Rm. G11 (See: Fall 2024 Hours)

Further Assistance: For further questions, students should contact their instructor. All instructors have regular office hours during the week. These office hours are listed on the Math Department's webpage for Instructor Office Hours and Emails.

Accommodation of Disabilities: The Office of Accessibility Resources and Services (OARS) offers long term and temporary accommodations for undergraduate, graduate and visiting students at NJIT.

If you are in need of accommodations due to a disability, please contact the Office of Accessibility Resources and Services at oars@njit.edu. The office is in Kupfrian Hall, Room 201. A Letter of Accommodation Eligibility from the Office of Accessibility Resources and Services office authorizing your accommodations will be required.

For further information regarding self-identification, the submission of medical documentation and additional support services provided please visit the Office of Accessibility Resources and Services (OARS) website at: https://www.njit.edu/accessibility/

Important Dates (See: Fall 2024 Academic Calendar, Registrar)

Date	Day	Event
September 2, 2024	Monday	Labor Day
September 3, 2024	Tuesday	First Day of Classes
September 9, 2024	Monday	Last Day to Add/Drop Classes
November 11, 2024	Monday	Last Day to Withdraw
November 26, 2024	Tuesday	Thursday Classes Meet
November 27, 2024	Wednesday	Friday Classes Meet
November 28 to December 1, 2024	Thursday and Sunday	Thanksgiving Recess - Closed
December 11, 2024	Wednesday	Last Day of Classes
December 12, 2024	Thursday	Reading Day 1
December 13, 2024	Friday	Reading Day 2
December 15 to December 21, 2024	Sunday to Saturday	Final Exam Period

Tentative Course Outlines

Weeks	Chapters	Topics
Week 1 09/02	Chapter 1	Introduction to Data Science
Week 2 09/09	Chapter 2	Statistical Learning; kNN
Week 3 09/16	Chapter 3	Linear Regression
Week 4 09/23	Chapter 4	Logistic Regression
Week 5 09/30	Chapter 4	LDA; QDA; Naive Bayes; Generalized Linear Models
Week 6 10/07	Chapter 5	Cross-Validation and Bootstrap
Week 7 10/14	Chapter 6	Linear Model Selection
Week 8 10/21	Chapter 6	Shrinkage Methods and Dimension Reduction Methods Review for Midterm Exam **Midterm Exam:** 10/22
Week 9 10/28	Chapter 7	Nonlinear Modeling
Week 10 11/04	Chapter 7	Smoothing Splines, Generalized Additive models
Week 11 11/11	Chapter 8	Tree-Based Methods
Week 12 11/18	Chapter 8	Bagging, Random Forests, Boosting
Week 13 11/25	Chapter 9	Maximal Vector Classifier and Support Vector Classifiers
Week 14 12/02	Chapter 9	Support Vector Machines
Week 15 12/09	Chapter 10	Unsupervised Learning (If time permits)
Week 16 12/16		**Final Exam**

Updated by Professor C. Shi - 08/27/2024
Department of Mathematical Sciences Course Syllabus, Fall 2024