MATH 678: Stat Methods in Data Science
Spring 2022 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.

Please be sure you read and fully understand our DMS Online Exam Policy.

COURSE INFORMATION

Course Description: This course introduces students to 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 661 or MATH 663, or permission by instructor.

Course-Section and Instructors:

Course-Section	Instructor
Math 678-102	Professor W. Guo

Office Hours for All Math Instructors: Spring 2022 Office Hours and Emails

Required Textbook:

Title	*An Introduction to Statistical Learning: with Applications in R*
Author	Gareth James, et al
Edition	2nd
Publisher	Springer
ISBN #	978-1071614174

Reference Book	The Elements of Statistical Learning: Data Mining, Inference, and Prediction, by Hastie, Tibshirani, and Friedman; Publisher: Springer, 2nd edition (2009); ISBN: 978-0387848570.

University-wide Withdrawal Date: The last day to withdraw with a **W** is **Monday, April 4, 2022**. It 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	25%
Project	20%
Midterm Exam	25%
Final Exam	30%

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

A	90 - 100	C+	75 - 79
B+	85 - 89	C	70 - 74
B	80 - 84	F	0 - 69

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.

Exams: There will be one exam during the semester and a cumulative final exam during the final exam week:

Midterm Exam	March 23, 2022
Final Exam Period	May 6 - May 12, 2022

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.

Makeup Exam Policy: There will be **NO MAKE-UP QUIZZES OR 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

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 Scott Janz, Associate Director of Disability Support Services at 973-596-5417 or via email at scott.p.janz@njit.edu. The office is located 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/studentsuccess/accessibility/

Important Dates (See: Spring 2022 Academic Calendar, Registrar)

Date	Day	Event
January 18, 2022	Tuesday	First Day of Classes
January 22, 2022	Saturday	Saturday Classes Begin
January 24, 2022	Monday	Last Day to Add/Drop Classes
March 14, 2022	Monday	Spring Recess Begins
March 19, 2022	Saturday	Spring Recess Ends
April 4, 2022	Monday	Last Day to Withdraw
April 15, 2022	Friday	Good Friday - No Classes
April 17, 2022	Sunday	Easter Sunday - No Classes
May 3, 2022	Tuesday	Friday Classes Meet
May 3, 2022	Tuesday	Last Day of Classes
May 4 - May 5, 2022	Wednesday and Thursday	Reading Days
May 6 - May 12, 2022	Friday to Thursday	Final Exam Period

Course Outline

Date	Lecture	Sections	Topic	Assignment
Week 1 1/19	1	Chapter 1	**Introduction to Data Science**	
Week 2 1/26	2	Chapter 2	**Statistical Learning; KNN**	Homework 1
Week 3 2/2	3	Chapter 3	**Linear Regression; R Lab**	
Week 4 2/9	4	Chapter 4	**Logistic Regression**	Homework 2
Week 5 2/16	5	Chapter 4	**LDA, QDA; R Lab**	
Week 6 2/23	6	Chapter 5	**Cross-Validation and Bootstrap**	Homework 3
Week 7 3/2	7	Chapter 6	**Linear Model Selection; R Lab**	
Week 8 3/9	8	Chapter 6	**Shrinkage Methods and Dimension Reduction Methods**	Homework 4
Week 9 3/16			**SPRING RECESS (NO CLASSES)**	
Week 10 3/23			**MIDTERM EXAM: Wednesday ~ March 23, 2022**	
Week 11 3/30	9	Chapter 7	**Nonlinear Modeling; R Lab**	Course Project
Week 12 4/6	10	Chapter 8	**Tree-Based Methods: Bagging, Random Forests, Boosting**	Homework 5
Week 13 4/13	11	Chapter 9	**Support Vector Machines**	
Week 14 4/20	12	Chapter 10	**Unsupervised Learning**	Homework 6
Week 15 4/27			**Students Project Presentation**	Deadline of the project report

Week 16 5/4			**Reading Day 1**	
Week 17 5/11			**FINAL EXAM:** **Wednesday ~ May 11, 2022**	