STOR 520-001
Introduction to Data Science
Fall 2023

Course Description

This course provides hands-on experience working with data sets provided in class and downloaded from certain public websites. Lectures cover basic topics such as R programming, visualization, data wrangling and cleaning, exploratory data analysis, web scraping, data merging, predictive modeling, and elements of machine learning. Programming analyses in more advanced areas of data science.

Plan to come to every class with your computer and ready to work with others. Using resources around you is a key component of successful data analysis. This includes the internet and people.

Students may not receive credit for both STOR 320 and STOR 520.

Course Goals and Learning Objectives

This course will enable you to:

- Establish proficiency in the statistical programming language R making the student immediately competitive in the data science job market.
- Acquire both structured and unstructured data for the purpose of gathering insight on well-crafted research hypotheses.
- Clean, transform, and merge datasets.
- Summarize data using professionally developed tables and various visualizations.
- Model relationships between variables using a variety of techniques, including linear regression, nonlinear regression, logistic regression, and various machine learning techniques.
- Evaluate and compare predictive accuracy from competing modeling methods using cross-validation.
- Honestly interpret results from predictive analytics using creative visualizations and tables.
- Effectively communicate insights both verbally and in writing to a non-technical audience.
- Share data science ideas with a worldwide audience using HTML web pages and dynamic web applications using RMarkdown and RShiny.

Course Details

Number	STOR 520-001
Title	Introduction to Data Science
Credit hours	4
Course Format	The course format will include 3 lectures per week. Lectures will be supplemented with in-class programming and practical discussion. Students will also be required to register for a weekly lab with required attendance in person.
Prerequisites	STOR 435 or MATH 535 (MS/PhD Students Welcome)
Target Audience	All MS or PhD students interested in advancing data science skills. Given the similarity to STOR 320, undergraduate students are recommended to take STOR 320 and not STOR 520.
Instructor	Dr. Mario Giacomazzo Office: Hanes 134 Phone: 480-489-1398 (Cell) Email: mgiacoma@email.unc.edu Office Hours: W, 9AM – 11 AM / 12PM – 2PM
Lab Instructors	Dan Meskill (404, 405) Email: dmeskill@unc.edu Office Hours: TTh, 11AM – 12PM, Hanes B-04 Yuhao Zhou (406, 407) Email: yuhaoza@live.unc.edu Office Hours: TTh, 1:15PM – 2:15PM, Remote Zoom: Link
Course Website	https://supermariogiacomazzo.github.io/STOR320_WEBSITE/
Class Days, Times, Location	TTh, 5:00PM – 6:15PM, Hanes 120
Lab Days, Times, Location	404: W, 9:05AM - 9:55AM, Phillips 222 405: F, 9:05AM - 9:55AM, Phillips 247 406: W, 4:40PM - 5:30PM, Hanes 107 407: F, 4:40PM - 5:30PM, Hanes 107
Course Texts	*R for Data Science*, Hadley Wickham. Legally free online, but can be purchased for less than $40 on Amazon.

Course Assessments

Assignments	Percentages
Attendance	2%
Labs	15%
Homework	0%
Mini Projects	25%
Midterms	20%
Final Project	28%
Group Involvement	10%

Grading Scale

Your final grade is based on a weighted average according to the previously addressed breakdown. Curving on individual/group assessments should not be expected. A curve may be applied to the final grades depending upon the class average. Conversion to a letter grade will be based on the table below:

H	90 to 100	F	0 to 69.99
P	80 to 89.99		
L	70 to 79.99		

A	93 to 100	B	83 to 86.99	C	73 to 76.99	D	60 to 66.99
A-	90 to 92.99	B-	80 to 82.99	C-	70 to 72.99	F	0 to 59.99
B+	87 to 89.99	C+	77 to 79.99	D+	67 to 69.99		

Assignment Descriptions

Attendance:
Attendance will be taken every class using the UNC Check-in App. You will need to install the UNC Check-in app to your mobile device and bring it to every class. Starting at the beginning of class, you will have 15 minutes to check-in using the mobile app. Instructions for installing and using the UNC Check-in App are available at https://unccheckin.unc.edu/. Go to this website to learn more. You need to attend at least 70% of the classes to get credit for attendance, otherwise you will receive a 0 for your attendance grade. If you need to miss class for a reason permitted by the university and you don't want to be penalized, you will need to get a university approved absence at https://uaao.unc.edu/. If you cannot get a university approved absence and don't want to be penalized, you must notify your instructor of the reason and provide documentation of the reason in email. The reason should line up with UNC'S definition of a university approved absence. For example, a job interview would not be approved by the university or me.

Labs:
Attendance to all labs is mandatory. Every week, your lab instructor will take attendance. If you are there for the entire class, you will receive 5 points. During this period, students are required to work on a lab assignment. Each lab assignment will be based on the topics discussed in lecture or related to your final project. Students are responsible to turn in their own labs, but are encouraged to work in teams.

These assignments are to be completed using RMarkdown and submitted as an HTML file on Canvas by 9:05AM on the following Wednesday. A lab instructor will be provided to help students in the completion of the lab and to facilitate group work. Every lab is worth 10 points and no late lab assignments will be accepted. You will need to get a university approved absence and notify your lab instructor to prevent a loss of points in these weekly labs if you miss class.

Homework:
Homework will be based on problems from the course textbook, *R for Data Science*. These assignments are recommended, but not required.

Mini Projects:
Mini Projects are larger assignments designed to test your data science skills while also helping you build a professional portfolio of data science work for when you leave. Some of the material needed for these assignments will be taught in lecture or found in the textbook. Other material needed will be provided by your instructor. You will be expected to read all the material provided for each mini project. The amount of time you get for these mini projects and the point value will vary based off the difficulty of the assignment. For each mini project, you will use RMarkdown and submit an HTML file on Canvas. If you submit your mini project late, expect a 25% deduction for less than 1 day late, 50% deduction between 1 and 2 days late, and 100% deduction more than 2 days late. All late mini projects must be submitted to Canvas as soon as they are completed.

In most cases, these mini projects need to be completed individually without any help from other students, AI tools (e.g. ChatGPT), or uploaded work from other students (e.g. CourseHero). If you cheat, you will receive a 0 and the alleged violation will be reported to UNC.

Midterms:
There will be 2 midterm exams during the semester. The date and material covered of each midterm will be announced at least 1 week before the midterm occurs. Midterms will be completed during class and submitted to Canvas as a knitted HTML by the end of class.

You will need to have your computer in class for these midterms. These midterms will be "open book" meaning you will be able to use all provided course material, the internet, and R to help you. You cannot get help from other students. You will be penalized for not submitting on-time.

Make-up midterms are only possible for University approved absences or consent from the instructor due to extenuating circumstances. Unless you have accommodation as determined by the university or university approved absences, you must take the midterms at the dates and times provided. If you have accommodation, please provide the formal notification before the second week of classes.

Final Project:
The final project is done in groups of 4-5 students and is worth a total of 100 points. There will be 4 parts of varying point values submitted throughout the semester. The first part, the Project Proposal, is worth 10 points and will be due sometime in the beginning of the semester after groups have been designated. The second part, the Exploratory Data Analysis, is worth 20 points and will be due after the middle of the semester after the Project Proposal has been completed. The third part, the Final Paper, is worth 40 points and must be submitted on Canvas by 11:59PM on Wednesday, December 6. The fourth part, the Final Presentation, is worth 30 points and will take place during our designated final exam time

according to the university calendar. For our class, this is 4PM to 7PM on Friday, December 15. Slides must be submitted by 4PM on December 15 to Canvas.

Group Involvement:
Since the final project is a group project that is worth a tremendous amount of points, it is very important that each group member fulfills their obligation to their group. Four times during the semester, there will be a survey sent out to the class, where you will score your group members on a scale from 0 (Terrible) to 5 (Excellent). In extreme circumstances, Dr. Mario has the right to remove individuals from their group and force them to submit their own project.

Course Policies and Resources

Syllabus Changes
Dr. Mario reserves the right to make changes to the syllabus, including all due dates. These changes will be announced as early as possible so that students can adjust their schedules.

Attendance Policy
University Policy: As stated in the University's Class Attendance Policy, no right or privilege exists that permits a student to be absent from any class meetings, except for these University Approved Absences:

1. Authorized University activities: University Approved Absence Office (UAAO) website provides information and FAQs for students and FAQs for faculty related to University Approved Absences
2. Disability/religious observance/pregnancy, as required by law and approved by Accessibility Resources and Service and/or the Equal Opportunity and Compliance Office (EOC)
3. Significant health condition and/or personal/family emergency as approved by the Office of the Dean of Students, Gender Violence Service Coordinators, and/or the Equal Opportunity and Compliance Office (EOC).

Instructors may work with students to meet attendance needs that do not fall within University approved absences. For situations when an absence is not University approved (e.g., a job interview, illness/ flu or club activity), instructors are encouraged to work directly with students to determine the best approach to missed classes and make-up assessment and assignments. Dr. Mario will require digital documentation in situations that are not University approved, but instructor approved. Without documentation, approval to miss class will definitely not be given, and Dr. Mario has the right to inspect the validity of the documentation. Dishonest documentation is a violation of UNC's honor code.

Technology Use
Students are required to bring their cell phone and computer to every class. Students must have R and RStudio installed on their computer. Directions for free downloads of this software will be provided. The professor or teaching assistant will occasionally request computers to be closed for dynamic discussion and guest speakers.

Time Limit for Grade Disputes
You only have 1 week after the grading of an assessment is complete to dispute your grade and possibly receive points back for that assessment. The only grade that should be disputed and modified after the

last day of class is the final exam. Dr. Mario makes mistakes and desires that all students receive a fair grade that is correct. However, it is the student's responsibility to review their grades on assignments in a timely manner so that Dr. Mario can make fair decisions and modify grades quickly.

Accessibility Resources

[Accessibility Resources and Service](#) (ARS – [ars@unc.edu](#)) receives requests for accommodations, and through the Student and Applicant Accommodations Policy determines eligibility and identifies reasonable accommodations for students with disabilities and/or chronic medical conditions to mitigate or remove the barriers experienced in accessing University courses, programs and activities.

ARS also offers its Testing Center resources to students and instructors to facilitate the implementation of testing accommodations.

Faculty and instructors with any concerns or questions about accommodations and/or their implementation, are invited to [reach out to ARS](#) to discuss.

Undergraduate Testing Center

The College of Arts and Sciences provides a secure, proctored environment in which exams can be taken. The center works with instructors to proctor exams for their undergraduate students who are not registered with ARS and who do not need testing accommodations as provided by ARS. In other words, the Center provides a proctored testing environment for students who are unable to take an exam at the normally scheduled time (with pre-arrangement by your instructor). For more information, visit http://testingcenter.web.unc.edu/.

Honor Code

Remember that as a student of UNC-Chapel Hill, you are bound by the University's Honor Code, which states that "It shall be the responsibility of every student at The University of North Carolina at Chapel Hill to obey and support the enforcement of the Honor Code, which prohibits lying, cheating, or stealing when these actions involve academic processes or University students or academic personnel acting in an official capacity." An especially serious Honor Code violation is plagiarism. You may wish to take a tutorial on plagiarism that was developed by librarians at UNC, Duke, NCSU, and NCCU. If you have questions, please consult your instructor.

You are not permitted to download or upload any content from this course to the web in any form, including but not limited to Chegg, Course Hero, Coursera, Google Drive, etc. If you post my course content, you may be violating my intellectual property rights. If you post your own work from this course, you are allowing sites to profit from your intellectual property. In utilizing web sources to upload or download course content, you risk violating the University's Honor Code.

IT Acceptable Use Policy

By attending the University of North Carolina at Chapel Hill, you agree to abide by the University of North Carolina at Chapel Hill policies related to the acceptable use of IT systems and services. The Acceptable Use Policy (AUP) sets the expectation that you will use the University's technology resources responsibly, consistent with the University's mission. In the context of a class, it's quite likely you will participate in online activities that could include personal information about you or your peers, and the AUP addresses your obligations to protect the privacy of class participants. In addition, the AUP addresses matters of others' intellectual property, including copyright. These are only a couple of typical

examples, so you should consult the full Information Technology Acceptable Use Policy, which covers topics related to using digital resources, such as privacy, confidentiality, and intellectual property.

Additionally, consult the Safe Computing at UNC website for information about data security policies, updates, and tips on keeping your identity, information, and devices safe.

Counseling and Psychological Services

UNC-Chapel Hill is strongly committed to addressing the mental health needs of a diverse student body. The Heels Care Network website is a place to access the many mental health resources at Carolina. CAPS is the primary mental health provider for students, offering timely access to consultation and connection to clinically appropriate services. Go to their website https://caps.unc.edu/ or visit their facilities on the third floor of the Campus Health building for an initial evaluation to learn more. Students can also call CAPS 24/7 at 919-966-3658 for immediate assistance.

Title IX and Related Resources

Any student who is impacted by discrimination, harassment, interpersonal (relationship) violence, sexual violence, sexual exploitation, or stalking is encouraged to seek resources on campus or in the community. Reports can be made online to the EOC at https://eoc.unc.edu/report-an-incident/ or by contacting the University's Title IX Coordinator (Elizabeth Hall, titleixcoordinator@unc.edu) or the Report and Response Coordinators in the Equal Opportunity and Compliance Office (reportandresponse@unc.edu). Confidential resources include Counseling and Psychological Services and the Gender Violence Services Coordinators (gvsc@unc.edu). Additional resources are available at safe.unc.edu.

Policy on Non-Discrimination

The University is committed to providing an inclusive and welcoming environment for all members of our community and to ensuring that educational and employment decisions are based on individuals' abilities and qualifications. Consistent with this principle and applicable laws, the University's Policy Statement on Non-Discrimination offers access to its educational programs and activities as well as employment terms and conditions without respect to race, color, gender, national origin, age, religion, genetic information, disability, veteran's status, sexual orientation, gender identity or gender expression. Such a policy ensures that only relevant factors are considered, and that equitable and consistent standards of conduct and performance are applied.

If you are experiencing harassment or discrimination, you can seek assistance and file a report through the Report and Response Coordinators (email reportandresponse@unc.edu or see additional contact info at safe.unc.edu) or the Equal Opportunity and Compliance Office at https://eoc.unc.edu/report-an-incident/.

Learning Center

Want to get the most out of this course or others this semester? Visit UNC's Learning Center at http://learningcenter.unc.edu to make an appointment or register for an event. Their free, popular programs will help you optimize your academic performance. Try academic coaching, peer tutoring, STEM support, ADHD/LD services, workshops and study camps, or review tips and tools available on the website.

Writing Center

For free feedback on any course writing projects, check out UNC's Writing Center. Writing Center coaches can assist with any writing project, including multimedia projects and application essays, at any stage of the writing process. You don't even need a draft to come visit. To schedule a 45-minute appointment, review quick tips, or request written feedback online, visit http://writingcenter.unc.edu.