Syllabus for STOR 320-001
Introduction to Data Science
Summer 2020

Instructor: Dr. Mario Giacomazzo   **Office hours**: M-F 11:00AM - 11:30AM
E-mail: mgiacoma@email.unc.edu   M-F 1:00PM - 1:30PM
Phone: 480-489-1398
Office: Wherever you want it to be

Assistant: Peter Lin
E-mail: pwlin@live.unc.edu
Office Hours: M-F from 2:30PM to 3:30PM

Pavlos Zoubouloglou
E-mail: pavlos@live.unc.edu
Office Hours: M-F from 8:00AM to 9:00AM

Lectures: M-F from 11:30AM – 1:00PM

Course URL: Website: https://supermariogiacomazzo.github.io/STOR320_WEBSITE/
Assignment Submission: https://sakai.unc.edu/ and login with your Onyen

Zoom Links: Due to the pandemic, lectures and office hours will be hosted live online via Zoom.
Lectures will also be recorded and linked to course website.

Lectures: https://unc.zoom.us/j/98063448751
Instructor Office Hours: https://unc.zoom.us/j/94479071959
Peter's Office Hours: https://unc.zoom.us/j/631940383
Pavlos's Office Hours: https://unc.zoom.us/j/2903979720

Description: This course is an application-driven introduction to data science. Statistical and
computational tools are valued throughout the modern workplace from Silicon Valley
startups, to marine biology labs, to Wall Street firms. These tools require technical
skills such as programming and statistics. They also require professional skills such as
communication, teamwork, problem solving, and critical thinking.

You will learn these tools and hone these skills through hands-on experience working
with datasets provided in class and downloaded from certain public websites. During
the first part of the semester, we will focus on R programming skills and data
visualization. Later topics will include: exploratory data analysis, web scraping, data
wrangling, modeling, and effective communication of results.

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.

Textbook: *R for Data Science*, Hadley Wickham. Legally free online, but can be purchased for less than $40 on Amazon. Additional suggested texts are provided on the website. All texts used in this course are free and downloadable from course website.

Software: A personal laptop with a working copy of **R** and **R Studio** is required. You are required to bring your laptop to every class. Directions for free downloads of these materials will be provided.

Prerequisites: STOR 155 (Introductory Statistics) or equivalent. If you haven't had a programming course, then you really need this course or something similar and you should expect this course to be extremely time demanding for the first 6 weeks.

Final Grade: Labs (30%)
Analyses (40%)
Final Project (30%)

Labs: Labs are constructed using problems from the course textbook, *R for Data Science*. Each lab will be worth 20 points. These labs are to be completed using RMarkdown and submitted as an HTML file on Sakai. Expect labs to be due at **5PM**. If you submit your lab late, expect a **25% deduction** for less than 1 day late, **50% deduction** between 1 and 2 days late, **75% deduction** between 2 and 3 days late, and **100% deduction** more than 3 days late. To submit a lab late, email it directly to mgiacoma@email.unc.edu.

Analyses: Analyses are constructed using customized problems from real life data sets. These analyses allow you to practice the techniques learned from lab assignments. Each analysis will be worth 40 points. These analyses are to be completed using RMarkdown and submitted as an HTML file on Sakai. Expect analyses to be due at **5PM**. If you submit your analyses late, expect a **25% deduction** for less than 1 day late, **50% deduction** between 1 and 2 days late. Homework will not be accepted beyond 2 days. To submit a lab late, email it directly to mgiacoma@email.unc.edu.

Final Project: The final project is done in groups of at least 4 and 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 middle of the semester after groups have been designated. The second part, the **Exploratory Data Analysis**, is worth **20 points** and will be due sometime towards the end 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 Sakai by **5:00PM** on **Friday, July 23**. 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 **11:30AM** on **Monday, July 27.** Slides must be submitted by **11:30AM** on **Monday, July 27**.

Grade 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:

A	94 to 100	B	83 to 86.99	C	73 to 76.99	D	60 to 66.99
A-	90 to 93.99	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		

These are hard break lines and no rounding will be applied to push an individual student up to a more desirable letter grade.

More Help: Programming can be incredibly frustrating and take some time to get used to. Before you email your instructor or IA please spend some time trying to solve/Google the problem on your own. Then, proceed to find someone in the class to ask**.** It is highly possible someone else has experienced the same problem. Finally, if you have not solved the problem, email your instructor or IA.

Ethics: It is my belief that effective programming can involve the heavy use of the work of others. That said, in this course and in your professional work, you must give attribution when using the work of others. This is true if you are borrowing a part of a friend's code, and it is true if you are dragging code off a website. Excessive use of the work of others (my judgment) will impact your grade but cannot have honor code implications if you give attribution.

For avoiding excessive use of others' code, one good rule of thumb is that it is generally ok for another person to look at your code and make suggestions. If you are copying someone's code from this class (or who took STOR 320 last spring), this is excessive use.

You will learn best by suffering through your own programming with the help of others.

Note: This section is long because this can be a touchy topic in programming courses. My real concern is that you are transparent about what you do for graded assignments. I will also try to be transparent about what is expected.

Honor Code: All students are expected to follow the guidelines of the UNC honor code. In particular, students are expected to refrain from "lying, cheating, or stealing" in the academic context. If you are unsure about which actions violate that honor code, please see me or consult honor.unc.edu.

Students are bound by the Honor Code in taking exams and in written work. The Honor Code of the University is in effect at all times, and the submission of work signifies understanding and acceptance of those requirements. Plagiarism will not be tolerated. Please consult with me if you have any questions about the Honor Code.

The University of North Carolina at Chapel Hill has had a student-administered honor system and judicial system for over 100 years. The system is the responsibility of students and is regulated and governed by them, but faculty share the responsibility. If you have questions about your responsibility under the honor code, please bring them to your instructor or consult with the office of the Dean of Students or the Instrument of Student Judicial Governance. This document, adopted by the Chancellor, the Faculty Council, and the Student Congress, contains all policies and procedures pertaining to the student honor system. Your full participation and observance of the honor code is expected (honor.unc.edu).

Attendance: 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
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).

Please communicate with me early about potential absences. Please be aware that you are bound by the Honor Code when making a request for a University approved absence.

Accessibility: The University of North Carolina at Chapel Hill facilitates the implementation of reasonable accommodations, including resources and services, for students with disabilities, chronic medical conditions, a temporary disability or pregnancy complications resulting in difficulties with accessing learning opportunities.

All accommodations are coordinated through the Accessibility Resources and Service Office. See the ARS Website for contact information: https://ars.unc.edu or email ars@unc.edu.

Title IX: Acts of discrimination, harassment, interpersonal (relationship) violence, sexual violence, sexual exploitation, stalking, and related retaliation are prohibited at UNC-Chapel Hill. If you have experienced these types of conduct, you are encouraged to report the incident and seek resources on campus or in the community.

Please contact the Director of Title IX Compliance/Title IX Coordinator (Adrienne Allison, adrienne.allison@unc.edu), Report and Response Coordinators (Ew Quimbaya-Winship, eqw@unc.edu; Rebecca Gibson, rmgibson@unc.edu; Kathryn Winn, kmwinn@unc.edu), Counseling and Psychological Services (CAPs) (confidential) in Campus Health Services at (919) 966-3658, or the Gender Violence Services Coordinators (confidential) (Cassidy Johnson, cassidyjohnson@unc.edu; Holly Lovern, holly.lovern@unc.edu) to discuss your specific needs.

Additional resources are available at safe.unc.edu.

Legal: Dr. Mario reserves the right to make changes to the syllabus, including all lab and analyses due dates, when unforeseen circumstances occur. These changes will be announced as early as possible so that students can adjust their schedules.