Syllabus – STOR 320.001

Spring 2021 (January 19 – May 5)
TuTh 2:00pm – 3:15pm
remote instruction

Instructor: Jan Hannig	**Phone:** (919) 962-7511
Office: 330 Hanes	**E-mail:** jan.hannig@unc.edu
Office Hours: MW 2:00 – 3:00PM and by appointment.	**Course home page on** https://sakai.unc.edu/portal/site/stor320.001.sp21
Zoom link for Office hours: https://unc.zoom.us/j/91787829144?pwd=bFQwUW9HUG85d2VaZkZvRVBFV1gwdz09 Meeting ID: 917 8782 9144 Passcode: HanesHall	**Zoom link for lectures** https://unc.zoom.us/j/94535557682?pwd=YWEwYzVWN0VSNHB1QnZHUTR1T2pkZz09 Meeting ID: 945 3555 7682 Passcode: HanesHall

Lab instructors

Section	Lab:	Office Hour
320.400: **Taylor Petty**	W 5:45-6:35pm	M 2:30-3:30pm W 1:30-2:30pm
tmpetty@live.unc.edu	https://unc.zoom.us/j/91440986085?pwd=WUpKeW5hVG0rNGRXaXlZZStJL2xuUT09	
320.401 Pavlos Zouboulouglou	W 4:40 – 5:30pm	M 10-11am F 1-2pm
pavlos@live.unc.edu	https://unc.zoom.us/j/93735742250	https://unc.zoom.us/j/2903979720
320.402 **Sam Booth**	F 4:40 – 5:30pm	TuTh 4-5pm
slbooth@live.unc.edu	https://unc.zoom.us/j/5891075192	

Target Audience:

Undergraduate students interested in data science who have successfully completed STOR155 or equivalent.

Course Objective:

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. During the first part of the semester, we will focus on R programming skills and data visualization. Later topics will include: exploratory data analysis, data wrangling, modeling, and effective communication of results.

Course Format:
This class is in lecture/lab format. I will cover these topics and provide you opportunity for some hands-on experience working with datasets provided in class and downloaded from certain public websites. More hands-on experience will be available in labs.
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.

Required Text:

- R for Data Science, by Hadley Wickham. available free online https://r4ds.had.co.nz/

Assessment:
Your grade will be based on a Lab Attendance (10%), Lab Assignments (15%), Homework (45%) and Final Project (30%).

Homework:
There will be two types of homework assignments: Regular homework will be based on problems from the course textbook, R for Data Science. Each regular homework assignment will be worth 20 points. Data analysis homework are constructed using customized problems from real life data sets. Each analysis will be worth 40 points. These analyses allow you to practice the techniques learned from the course.

- You may discuss homework with classmates and teaching staff. But you must submit your own work.
- You may and often should search online for solutions to coding problems. This is perfectly fine and encouraged.
- However, copying responses from students who have taken the course, including from sources online, is unacceptable and could be treated as an honor code violation.
- Homework must be submitted as the PDF output from an R Markdown file on Gradescope. In other words, your homework submission must be a .pdf file with

all code and writing, as produced in R Markdown. Submissions that do not 'knit' will not be accepted. Such cases most often result from errors in the code, which students must correct before submission.

- Late homework submitted less than 24 hours from when it was due will have its score reduced 50%.
- Homework later than 24 hours or a failure to adhere to the rules above will result in a score of zero for that assignment.

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 10 points. At the end of the semester, the lab attendance grades will be curved by 10 points allowing you to miss a single lab and receive a 100% on your lab attendance grade. If you show up to every lab, you will get above 100% on your lab attendance grade.

During the lab session, students are required to complete a lab assignment that will be due **30 minutes after the lab ends**.

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 and help each other. 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. Lab assignment must be submitted as the PDF output from an R Markdown file on Gradescope.

Final Project:
The final project is done in groups of 5 and worth a total of 100 points. There will be 4 parts of varying point values submitted throughout the semester. See website for details on submission instruction and due dates.

1. Part I: Project Proposal, is worth 10 points.
2. Part II: Exploratory Data Analysis, is worth 20 points.
3. Part III: Final Paper, is worth 40 points.
4. Part IV: Final Presentation, is worth 30 points.

Course Outline:

1. Core programming and data science skills
    a. R Markdown
    b. data frame creation and manipulation
    c. summary statistics
    d. visualization
    e. exploratory data analysis

    f. ‘tidy’ and relational data
    g. functions and functional programming
    h. string manipulation and regular expressions
2. Modeling
    a. cross-validation
    b. linear and generalized linear models
    c. classification techniques
    d. clustering
3. Advanced topics
    a. Shiny
    b. more advanced modeling with support vector machines and tree-based methods
    c. web scraping

Exact dates are available on the website.

COVID-19 Each of us has a responsibility to know and act on these standards and policies in a way that maximizes a safe and healthy environment for us to teach, work, learn and live. To this end, we are developing a set of community standards and policies for our students, faculty, staff and visitors. We are all in this together, and we believe that together, we can face the challenges presented by COVID-19 with resilience, determination and great support for our community.

See https://carolinatogether.unc.edu/community-standards/ lines that we all need to follow to reduce the spread of COVID-19:

- Understand How COVID-19 Spreads
- Wash Your Hands Often
- Practice Physical Distancing
- Wear a Face Mask
- Maintain Clean, Safe Spaces
- Engage in Smaller Group and Virtual Settings
- Follow Immunization Recommendations
- Protect the community
- Provide Medical Return Clearance
- Regularly participate in asymptotic testing.

See the Carolina Together Roadmap at https://carolinatogether.unc.edu/ for more information on the University’s plans regarding COVID-19.

Delivery Method This course is classified as synchronous remote instruction. More details on this delivery method can be found at https://carolinatogether.unc.edu/ .

Etiquette for Zoom These are my five expectations of you on Zoom:

- Have your camera turned on.
- Mute your microphone unless answering or asking a question.
- Communicate by unmuting yourself or using the chat feature.
- Be mindful of background noise when not muted.
- Limit your distractions and avoid multi-tasking.

UNC Attendance Policy No right or privilege exists that permits a student to be absent from any class meetings, even online classes, except for these university approved absences: authorized university activities, and disability/religious observance/pregnancy, as required by law and approved by Accessibility Resources and Service and/or the Equal Opportunity and Compliance Office (EOC). 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).

Accessibility Resources UNC-Chapel Hill facilitates the implementation of reasonable accommodations for students with learning disabilities, physical disabilities, mental health struggles, chronic medical conditions, temporary disability, or pregnancy complications, all of which can impair student success. See the ARS website for contact and registration information: https://ars.unc.edu/about-ars/contact-us

Counseling and Psychological Services CAPS is strongly committed to addressing the mental health needs of a diverse student body through timely access to consultation and connection to clinically appropriate services, whether for short or long-term needs. Go to their website: https://caps.unc.edu/ or visit their facilities on the third floor of the Campus Health Services building for a walk-in evaluation to learn more.

Title IX 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. Please contact the Director of Title IX Compliance (Adrienne Allison - Adrienne.allison@unc.edu), Report and Response Coordinators in the Equal Opportunity and Compliance Office (reportandresponse@unc.edu), Counseling and Psychological Services (confidential), or the Gender Violence Services Coordinators (gvsc@unc.edu; confidential) to discuss your specific needs. Additional resources are available at https://safe.unc.edu

Honor Code Statement Students are bound by UNC's honor code (https://studentconduct.unc.edu/information/policies) in taking exams and in written work, and the submission of said 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.

Legal The instructor reserves the right to make any changes he considers academically advisable. It is your responsibility to attend classes and keep track of the proceedings.