

MEJO 571.001: Social Media Analysis Using Python
Fall 2020

Time: Mondays and Wednesdays, 9:45am – 11:00am (Eastern Standard Time)

Location: Zoom (<https://unc.zoom.us/j/94105383290>)

Instructor information:

Jacob Rohde

jarohde@unc.edu

Office Hours: Zoom only; by appointment

COURSE DESCRIPTION

Social media analysis is in high demand across the communications field, both professionally and academically. Identifying insights and patterns in social media data can be critical to the success of many businesses and organizations. Specifically, analyzing social media data can help communication researchers understand the ideas and actions of relevant users, measure audience interactions, and characterize user diversity. Such information can be used to create and disseminate more meaningful content.

This course will teach students the fundamentals of the Python programming language and its application in communications research, particularly in regard to social media analysis. During this course, students will develop the skills to design, execute, and critique various computational research methods, such as sentiment and term-based dictionary analysis. As a final deliverable, students will apply their skills in a comprehensive research project on a social media data set of their choosing.

COURSE GOALS

Upon completion of this course, students will be able to:

- Read, write, organize, debug, and critique Python code
- Understand the knowledge of basic programming concepts and skills, and their uses in various communications fields
- Recognize the limitations in various computation methodologies
- Design, execute, and evaluate a big data social media research project using Python
- Discuss ethical considerations of social media inquiry

COURSE GUIDELINES

Recommended text: Downey, A. B. (2015). *Think Python: How to think like a computer scientist* (2 Ed). Needham, MA: Green Tea Press.

You can download a free PDF version of the recommended text here:

<http://greenteapress.com/thinkpython2/thinkpython2.pdf>

I may also assign additional readings to supplement certain class lectures. These readings will be accessible in Sakai at least one week prior to their assigned day.

Attendance: Students are expected to attend all virtual classes. I recognize this may not be possible, particularly given the circumstances surrounding the novel coronavirus pandemic. Those who are not able to log into our Zoom meeting room synchronously (i.e., at our scheduled meeting time) due to any reason are given the ability to view class lectures asynchronously at their own pace. This option will **not**

affect students' academic standing in my class. All class lectures will be uploaded to the Sakai site immediately after each class period. To stay up to date with lectures, I recommend students view them the same day they are uploaded.

Students will be allowed to miss up to **four** classes without being penalized. I do not need an excuse (e.g., sick, mental health day, etc.) for these absences. Absences past four will reduce your overall grade by 10%. Please be mindful when using an absence. Excused absences will be recognized for NCAA or similar events. If you believe your absence should be excused for a relevant event or religious holiday, please send me an email detailing the event beforehand.

Participation: Students are asked to participate regularly during class by expressing their opinions and raising questions about the course material. There will also be regularly updated Sakai forums for students to pose questions/seek advice. If you do not feel comfortable speaking during class, posting questions on Sakai will count toward student participation. Please be respectful and open to others' opinions and questions. In all, participation will account for 10% of students' overall course grade.

Technology: Laptops are required for class participation. We will be learning and executing all code using PyCharm, which is a third-party Python editor. While I recommend using laptops, students who *only* have tablets available can use Google Colab, which is a notebook Python editor. Google Colab is part of the Google suite of applications and can run on tablet devices through the web (though not well unless you have an iPad with a keyboard). Using any device for recreation (e.g., social media) during class time is prohibited. Cellphones are also not permitted during class except for emergencies. Lastly, please mute zoom when you're not talking.

More information about PyCharm can be found here:

<https://www.jetbrains.com/pycharm/>

More information about Google Colab can be found here:

<https://colab.research.google.com/notebooks/intro.ipynb>

Email and Communication: I check email regularly during the week but I cannot guarantee an immediate response. I will try my best to get back to emails within a reasonable time (24-48 hours). Please treat emails to me as professional correspondence (i.e., no typos).

Student Accommodations: If you require special accommodations to attend or participate in this course, please let me know as soon as possible. This includes physical needs as well as less apparent needs such as presentation anxiety. If you need information about accommodations visit the Accessibility Services website at accessibility.unc.edu or the Learning Center website:

<http://learningcenter.unc.edu/ldadhd-services/>

Diversity: UNC is committed to providing an inclusive and welcoming environment for all members of our community and does not discriminate in offering access to its educational programs and activities based on age, gender, race, color, national origin, religion, creed, disability, veteran's status, sexual orientation, gender identity, or gender expression. The Hussman School of Journalism and Media adopted Diversity and Inclusion Mission and Vision statements in spring 2016 with accompanying goals found here: <http://www.mj.unc.edu/diversity-and-inclusion>.

Harassment: The University does not tolerate harassment in any form. We may not always agree with

one another, but we must be mindful and respect everyone's right to this policy. Harassment is a violation of the Honor Code, Title VII of the Civil Rights Act (1964), and Title IX. If you are harassed or feel threatened by another student in the course, please bring it to my attention or contact the Dean of Students (email: dos@unc.edu; phone: (919) 966-4042).

Academic Integrity: All UNC-CH students are expected to adhere to the University's Honor Code, which includes the following on Academic Dishonesty:

It shall be the responsibility of every student enrolled at the University of North Carolina at Chapel Hill to support the principles of academic integrity and to refrain from all forms of academic dishonesty, including but not limited to, the following:

1. Plagiarism in the form of deliberate or reckless representation of another's words, thoughts, or ideas as one's own without attribution in connection with submission of academic work, whether graded or otherwise.
2. Falsification, fabrication, or misrepresentation of data, other information, or citations in connection with an academic assignment, whether graded or otherwise.
3. Unauthorized assistance or unauthorized collaboration in connection with academic work, whether graded or otherwise.

My policy is to turn over any documents that appear to have content from other uncited sources than the author's to the University's Honor Committee. See here for the full text of the Honor Code: <http://instrument.unc.edu/instrument.text.html>

Accreditation: The Hussman School of Journalism and Media's accrediting body outlines a number of values students should be aware of and competencies students should be able to demonstrate by the time you graduate from our program (<http://www2.ku.edu/~acejmc/PROGRAM/PRINCIPLES.SHTML#vals&comps>). No single course can give you all of these values and competencies, but collectively our classes are designed to build abilities in each of these areas. In this course, we will address a number of the values and competencies, with special emphasis on:

- Demonstrate an understanding of professional ethical principles and work ethically in pursuit of truth, accuracy, fairness and diversity
- Think critically, creatively and independently
- Conduct research and evaluate information by methods appropriate to the communications professions in which they work
- Write correctly and clearly in forms and styles appropriate for the communications professions, audiences and purposes they serve
- Critically evaluate their own work and that of others for accuracy and fairness, clarity, appropriate style and grammatical correctness
- Apply basic numerical and statistical concepts
- Apply tools and technologies appropriate for the communications professions in which they work

ASSIGNMENTS & GRADING

Programming assignments: Students will be given **four** programming assignments at various points throughout the semester. These assignments will assess students' ability to write code and apply appropriate computational methods to investigate and explain social media data. Programming

assignments will be due by 11:59pm (Eastern Standard Time) on their assigned date below (also see schedule).

- Assignment 1: 08/26
- Assignment 2: 09/16
- Assignment 3: 10/07
- Assignment 4: 10/26

More information about each of the programming assignments will be available via Sakai at least two weeks prior to when they are due.

Final project: Mid-way through the semester, students will identify a social media data set of their choice. Students will then use the skills developed throughout the course to design and execute a comprehensive research project on their given data set. There are three graded components to this final project: **1)** Python code, **2)** a roughly 4-page written report, and **3)** a virtual class presentation. We will discuss more about the requirements for the final project on the first day of class.

Grade breakdown:	Point total
Assignment 1	10
Assignment 2	10
Assignment 3	15
Assignment 4	15
<i>Final project</i>	
In-class presentation	5
Python code	15
Written research report	20
Class participation/attendance	10
Total = 100	

Grade Scale: Your final grade will be calculated using the University’s grading scale: A, B, C, D, F with the plus or minus option for grades A–D. I follow the University’s grading standards detailed here:

- | | |
|--|-----------------|
| ● A – Mastery of course content at the highest level of attainment that can reasonably be expected of students at a given stage of development. The A grade states clearly that the student has shown such outstanding promise in the aspect of the discipline under study. | Scale |
| | A = 93 - 100 |
| | A- = 90 - 92.99 |
| | B+ = 87 - 89.99 |
| ● B – Strong performance demonstrating a high level of attainment for a student at a given stage of development. The B grade states that the student has shown solid promise in the discipline under study. | B = 83 - 86.99 |
| | B- = 80 - 82.99 |
| | C+ = 77 - 79.99 |
| ● C – An acceptable performance demonstrating an adequate level of attainment for a student at a given stage of development. The C grade states that while not yet showing any unusual promise, the student may continue to study in the discipline with reasonable hope of intellectual development. | C = 73 - 76.99 |
| | C- = 70 - 72.99 |
| | D+ = 67 - 69.99 |
| ● D – A marginal performance in the required exercises demonstrating a minimal passing level of attainment. A student has given no evidence of prospective growth in the discipline; an accumulation of D grades should be taken to mean that the student would be well advised not to continue in the academic field. | D = 60 - 66.99 |
| | F = < 60 |
| ● F – For whatever reasons, an unacceptable performance. The F grade indicates the | |

student's performance has revealed little understanding of the course content. A grade of F should warrant an advisor's questioning whether the student may suitably register for further study in the discipline before remedial work is undertaken.

Grade Questions: Please contact me via email with any grade concerns for all major assignments; though, I recommend waiting at least 24 hours after receiving your grade before emailing me. Please detail your concerns with evidence to support your claim. You have two weeks after a grade has been posted to send a grade concern (this does not apply to the final). General feedback about a grade, however, can be requested at any time during the semester.

Late Work: Late assignments will not be accepted. If an emergency arises that is out of your control, please talk with me and we can work around this policy.

SCHEDULE

Week	What we will be covering	Optional Downey chapters and assignments due
Python Basics		
Aug 10	Introductions, course overview and expectations, set up Python	Readings: None
Aug 12	Python basics day 1: PyCharm editor; executing code; note taking	Readings: None
Aug 17	Python basics day 2: Variables; data types; global functions; user input	Readings: 1.5, 2.1-2.5, 2.7
Aug 19	Strings day 1: Intro to strings; string searching	Readings: 2.6, 8.1-8.3
Aug 24	Strings day 2: Advanced string searching; slicing	Readings: 8.4, 8.9-8.10
Aug 26	Strings day 3: Cleaning string data; regular expressions	Readings: 8.8 Assignment 1 due
Aug 31	Strings day 4: Regular expressions cont.	Readings: none
Sep 2	No class; Labor Day holiday	Readings: none
Sep 7	Conditionals day 1: introduction to lists; if/else statements; boolean data	Readings: 5.2-5.8, 10.1-10.6
Sep 9	Conditionals day 2: For loops; while loops; loading data sets	Readings: 7.1-7.4
Sep 14	Conditionals lab	Readings: none
Sep 16	Dictionaries	Readings: 11.1-11.5 Assignment 2 due
Sep 21	Functions	Readings: 3.1-3.8
Working with Social Media Data		
Sep 23	Pandas 1: File I/O; working with tabular data	Readings: none
Sep 28	Pandas 2: Advanced searching and filtering	Readings: none
Sep 30	Pandas 3: Evaluating expressions; storing data	Readings: none

Oct 5	Pandas lab	Readings: none
Oct 7	Social media day 1: API basics; meta-data	Readings: TBA Assignment 3 due
Oct 12	No class; University day	Readings: none
Oct 14	Social media day 2: Term frequency	Readings: TBA
Oct 19	Social media day 3: Sentiment analysis	Readings: TBA
Oct 21	Social media data lab	Readings: none
Oct 26	Integrating concepts day 1: Automation; visualizing data	Readings: TBA Assignment 4 due
Oct 28	Integrating concepts day 2: Visualizing data cont.; writing results	Readings: TBA
Nov 2	Class workshop	Readings: none
Nov 4	Class workshop	Readings: none
Nov 9	Final presentations (~8 students)	Readings: none
Nov 11	Final presentations (~8 students)	Readings: none
Nov 16	Final Presentations (~9 students)	Readings: none
TBD	No class; final report due	Assignment 5 due (digital copy to be turned in @ TBD)