# Course Number, Section and Title

##### MIS 71605/71610, DATA SCIENCE PROGRAMMING

###### Summer 2023

Class Meets: MWTTH 10-12 and 2-4 in GSB 3.104

**Instructor:** Deepayan Chakrabarti (Deepayan)

Office: CBA 6.462

Pronouns: he/him Office hours: by appointment

Email: deepay@utexas.edu

**TA:** Rahul Nandakumar

Email: rahul.nandakumar@utexas.edu   
Office hours: TBA

# Course Description

### *Data-driven analysis has wrought a quiet revolution in business. As disk storage and computing power have become cheaper, companies have started maintaining detailed logs of inventories, sales, and customer activity, among others. Yet, this is only half the job; the real need is for insights, and this course teaches you the tools for that. This course uses Python & Pandas.*

### What will I learn?

##### Introductory Python, where we learn the basic language syntax, and gain familiarity with general-purpose tools such as string manipulation,

##### Pandas, which is a powerful data analysis toolkit (similar to R) that makes it easy to explore and visualize data,

##### Classification, where we develop an understanding of how to make predictions,

##### Clustering, where we learn how to discover the major groups or components of a given dataset, and

##### Other Topics, including regression and hypothesis testing.

### How will I learn?

The class will be face-to-face.

### Pre-requisites for the course:

None, but prior knowledge of programming is strongly recommended.

### How to succeed in this course:

This is a coding course, and the only way is to success is through practice. Try to work out all the problems in the assignments, even if they are group assignments. This will help during the exams.

### MCCOMBS STATEMENT ON ACADEMIC RIGOR:

At McCombs, *you can expect to feel intellectually challenged in your classes*. This may be experienced as ambition, boldness, curiosity, doubt, excitement, fear, gratitude, or something else entirely. We encourage you to embrace these feelings as indicators of authentic learning and growth. Our mission is to nurture the intellectual exploration and academic tenacity that will help you build the knowledge, skills, and values as foundations to thrive in your forthcoming professional pursuits.

To achieve this mission, we implement a rigorous curriculum designed to guide you in mastering complex concepts across diverse disciplines and contexts. We emphasize high standards and empower you to develop the in-demand skills needed for success in today’s workforce as well as the durable skills required to adapt and lead tomorrow’s workforce.

Although it is normal to experience difficult coursework in your McCombs classes, we are dedicated to supporting your development and academic success. Our goal is to help each of you live up to your potential and leave here knowing what it means to produce high-quality, credible work. Our priority is that each of you who starts here at McCombs grows to become a curious, life-long learner equipped to change the world.

# Course Requirements

### Required Materials

### There are no required books for this course. The following books are optional:

### [*Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*](https://wesmckinney.com/book/), by Wes McKinney (for analytics programming)

### *Applied Predictive Modeling*, by Max Kuhn and Kjell Johnson (for machine learning)

### [*Think Python*](http://www.greenteapress.com/thinkpython/thinkpython.pdf), by Downey (for introductory python)

### Required Devices

### A laptop on which you can install software. We will discuss what needs to be installed during the lectures.

### Classroom expectations

**Class attendance.** Attendance is optional except for project presentations.

**Behavior expectations.** Be nice, y’all.

**McCombs Classroom Professionalism Policy.** [delete or modify if provided online] The highest professional standards are expected of all members of the McCombs community. The collective class reputation and the value of the Texas MS experience hinges on this.You should treat the Texas MS classroom as you would a corporate boardroom. Faculty are expected to be professional and prepared to deliver value for each and every class session. Students are expected to be professional in all respects. The Texas MS classroom experience is enhanced when:

• Students arrive on time. On time arrival ensures that classes are able to start and finish at the scheduled time. On time arrival shows respect for both fellow students and faculty and it enhances learning by reducing avoidable distractions.

• Students display their name cards. This permits fellow students and faculty to learn names, enhancing opportunities for community building and evaluation of in-class contributions.

• Students are fully prepared for each class. Much of the learning in the Texas MS programs take place during classroom discussions. When students are not prepared, they cannot contribute to the overall learning process. This affects not only the individual, but their peers who count on them, as well.

• Students respect the views and opinions of their colleagues. Disagreement and debate are encouraged. Intolerance for the views of others is unacceptable.

• Students do not confuse the classroom for the cafeteria. The classroom (boardroom) is not the place to eat meals (i.e.-breakfast tacos, salad, sandwich). Recognizing that back-to-back classes sometimes take place over the lunch hour, energy bars and similar snacks are permitted. Please be respectful of your fellow students and faculty and plan to eat around your class schedule.

• Students minimize unscheduled personal breaks. The learning environment improves when disruptions are limited.

• Students attend the class section to which they are registered. Learning is enhanced when class sizes are optimized. Limits are set to ensure a quality experience and safety.

• Technology is used to enhance the class experience. When students are surfing the web, responding to e-mail, instant messaging each other, and otherwise not devoting their full attention to the topic at hand they are doing themselves and their peers a major disservice. Those around them face additional distraction. Fellow students cannot benefit from the insights of the students who are not engaged. Faculty office hours are spent going over class material with students who chose not to pay attention, rather than truly adding value by helping students who want a better understanding of the material or want to explore the issues in more depth. Students with real needs may not be able to obtain adequate help if faculty time is spent repeating what was said in class. There are often cases where learning is enhanced by the use of technology in class. Faculty will let you know when it is appropriate.

• Phones and wireless devices are turned off. Not only is it not professional, it cuts off the flow of discussion when the search for the offender begins. When a true need to communicate with someone outside of class exists (e.g., for some medical need) please inform the professor prior to class.

### Assignments

There will be three homework assignments. All assignments must be handed in electronically before the beginning of the class (submit online on Canvas).

Use of resources: You can use any online resources, including Google searches, Stackoverflow, and ChatGPT. BUT MAKE SURE YOU UNDERSTAND THE CODE before you copy-paste from these resources. You are also welcome to discuss your answers with others. But please do not look at others’ code before doing it yourself.

### Project

You must develop a group project on any relevant topic that interests you. You must submit a project report. Also, each group must give a presentation on their work (order decided by lottery).

### Final Exam

There will be a final exam. You will need a laptop. The exam is open-book, open-notes, open-slides, and open-web.

### Grading for this Course

The following table represents how you will demonstrate your learning and how we will assess the degree to which you have done so.

|  |  |  |
| --- | --- | --- |
| Assignments | Point Possible | Percent of Total Grade |
| 3 Group Assignments | 100 points each | 3\*15% = 45% |
| 1 Group Project | 100 points | 25% |
| 1 Final Exam | 100 points | 30% |

The groups for the assignments and the project will be randomly assigned.

### Course Outline

All instructions, assignments, readings, rubrics and essential information will be on the Canvas website at

[utexas.instructure.com](https://utexas.instructure.com/). Check this site regularly and use it to ask questions about the course schedule.

**Changes** to the schedule may be made at my discretion and if circumstances require. It is your responsibility to note these changes when announced (although I will do my best to ensure that you receive the changes with as much advanced notice as possible).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Week | Date | Day | Class Topic | Out of Class Activities | Assignments Due |
| 1. | 07/10 | M | Introduction and Python I | Values and variables, control flow, and functions  (Think Python chapters 2, 3, and 5) |  |
| 1 | 07/11 | T | Python II and III | Data Structures (Think Python chapters 10, 11, and 12); Data structures detailed example |  |
| 1 | 07/12 | W | Python III (contd.) and IV | Strings and regular expressions; Files  (Think Python chapters 8 and 14) | First assignment released |
| 1 | 07/13 | Th | Series and DataFrames I | McKinney chapters 5 and 6 |  |
| 2 | 07/17 | M | Series and DataFrames II | Examples using the NYC Complaints dataset |  |
| 2 | 07/18 | T | Series and DataFrames III |  |  |
| 2 | 07/19 | W | Data Wrangling I | Merging, concatenation, reshaping, pivoting (McKinney chapter 7) | First assignment due;  Second assignment released |
| 2 | 07/20 | Th | Data Wrangling II |  |  |
| 3 | 07/24 | M | Visualization | Plotting; Histograms  (McKinney chapter 8) |  |
| 3 | 07/25 | T | Grouping data | McKinney chapter 9 |  |
| 3 | 07/26 | W | Grouping (contd.) and Time Series | McKinney chapter 10 | Second assignment due;  Third assignment released |
| 3 | 07/27 | Th | Regression and Intro to classification |  |  |
| 4 | 07/31 | M | Nearest Neighbors classification |  |  |
| 4 | 08/01 | T | Naïve Bayes and Logistic Regression |  |  |
| 4 | 08/02 | W | Decision Trees and Ensemble methods |  | Third assignment due |
| 4 | 08/03 | Th | K-Means clustering |  |  |
| 5 | 08/07 | M | Project Presentations I |  | Project report due |
| 5  5 | 08/08  08/09 | T  W | Project Presentations II  Review session |  |  |
| 5 | 08/10 | Th | (free for office hours) |  |  |
|  |  |  | **FINAL EXAM TBA** |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# Policies

## Classroom Policies

##### Statement on Learning Success

Your success in this class is important to me. We will all need accommodations because we all learn differently. If there are aspects of this course that prevent you from learning or exclude you, please let me know as soon as possible. Together we’ll develop strategies to meet both your needs and the requirements of the course. I also encourage you to reach out to the student resources available through UT. Many are listed on this syllabus, but I am happy to connect you with a person or Center if you would like.

##### Grading Policies

The grade will be based on the standard cutoffs (shown below) after applying a curve on your overall score at the end of the course.

|  |  |
| --- | --- |
| Grade | Cutoff |
| A | 94% |
| A- | 90% |
| B+ | 87% |
| B | 84% |
| B- | 80% |
| C+ | 77% |
| C | 74% |
| C- | 70% |
| D | 65% |
| F | <65% |

### Late work

Late work will be handled on a case-by-case basis.

### religious holy days

By UT Austin policy, you must notify me of your pending absence at least fourteen days prior to the date of observance of a religious holy day. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, you will be given an opportunity to complete the missed work within a reasonable time after the absence.

### Classroom safety

To help preserve our in-person learning environment, the university recommends the following.

* [Vaccinations (and boosters) are widely available](https://www.healthyhorns.utexas.edu/coronavirus_vaccination.html), free and not billed to health insurance. The vaccine helps protect against viral transmission to others and reduces serious symptoms in those who are vaccinated.
* What if you’re sick?
  + For *any* illness, students should stay home if they are sick or contagious, not only to stop the spread but also to promote their personal wellness. Also, please use the exposure action chart below to help you know what to do in the event you are sick:
    - <https://www.healthyhorns.utexas.edu/coronavirus_exposure_action_chart.html>
    - UHS maintains up-to-date resources on COVID, which can be found at: <https://www.healthyhorns.utexas.edu/coronavirus.html>

Sharing of Course Materials is Prohibited

Sharing of Course Materials is Prohibited: No materials used in this class, including, but not limited to, lecture hand-outs, videos, assessments (quizzes, exams, papers, projects, homework assignments), in-class materials, review sheets, and additional problem sets, may be shared online or with anyone outside of the class unless you have my explicit, written permission. Unauthorized sharing of materials promotes cheating. It is a violation of the University’s Student Honor Code and an act of academic dishonesty. I am well aware of the sites used for sharing materials, and any materials found online that are associated with you, or any suspected unauthorized sharing of materials, will be reported to Student Conduct and Academic Integrity in the Office of the Dean of Students. These reports can result in sanctions, including failure in the course.

FERPA and Class Recordings

Class recordings are reserved only for students in this class for educational purposes and are protected under FERPA. The recordings should not be shared outside the class in any form. Violation of this restriction by a student could lead to Student Misconduct proceedings.

### **Student Rights & Responsibilities**

* You have a right to a learning environment that supports mental and physical wellness.
* You have a right to respect.
* You have a right to be assessed and graded fairly.
* You have a right to freedom of opinion and expression.
* You have a right to privacy and confidentiality.
* You have a right to meaningful and equal participation, to self-organize groups to improve your learning environment.
* You have a right to learn in an environment that is welcoming to all people. No student shall be isolated, excluded or diminished in any way.

With these rights come responsibilities:

* You are responsible for taking care of yourself, managing your time, and communicating with the teaching team and with others if things start to feel out of control or overwhelming.
* You are responsible for acting in a way that is worthy of respect and always respectful of others. • Your experience with this course is directly related to the quality of the energy that you bring to it, and your energy shapes the quality of your peers’ experiences.
* You are responsible for creating an inclusive environment and for speaking up when someone is excluded.
* You are responsible for holding yourself accountable to these standards, holding each other to these standards, and holding the teaching team accountable as well.

### Diversity and Inclusion

It is my intent that students from all diverse backgrounds and perspectives be well served by this course, that students’ learning needs be addressed and that the diversity that students bring to this class can be comfortably expressed and be viewed as a resource, strength and benefit to all students. Please come to me at any time with any concerns.

### Personal Pronoun Preference

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender, gender variance, and nationalities. Class rosters are provided to the instructor with the student’s legal name. I will gladly honor your request to address you by a name different than what appears on the roster, and by the gender pronouns you use. Please advise me of this preference early in the semester so that I may make appropriate changes to my records.

## University Policies

### Academic Integrity

Each student in the course is expected to abide by the University of Texas Honor Code: “As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity.” **Plagiarism is taken very seriously at UT**. Therefore, if you use words or ideas that are not your own (or that you have used in previous class), you must cite your sources. Otherwise you will be guilty of plagiarism and subject to academic disciplinary action, including failure of the course. You are responsible for understanding UT’s Academic Honesty and the University Honor Code which can be found at the following web address: [deanofstudents.utexas.edu/conduct](http://deanofstudents.utexas.edu/conduct/)

### University Resources for Students

Your success in this class is important to me. We will all need accommodations because we all learn differently. If there are aspects of this course that prevent you from learning or exclude you, please let me know as soon as possible. Together we’ll develop strategies to meet both your needs and the requirements of the course. There are also a range of resources on campus:

##### Services for Students with Disabilities

This class respects and welcomes students of all backgrounds, identities, and abilities. If there are circumstances that make our learning environment and activities difficult, if you have medical information that you need to share with me, or if you need specific arrangements in case the building needs to be evacuated, please let me know. I am committed to creating an effective learning environment for all students, but I can only do so if you discuss your needs with me as early as possible. I promise to maintain the confidentiality of these discussions. If appropriate, also contact Services for Students with Disabilities, 512-471-6259 (voice) or 1-866-329- 3986 (video phone). [diversity.utexas.edu/disability/about](http://diversity.utexas.edu/disability/about/)

##### Counseling and Mental Health Center

Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. [cmhc.utexas.edu/individualcounseling.html](http://cmhc.utexas.edu/individualcounseling.html)

Student Emergency Services:

If at any time you experience an emergency that necessitates your absence from a class requirement (e.g., attendance, assignment submission, or exam), please report your circumstances and absence via the Student Emergency Services website: [deanofstudents.utexas.edu/emergency](http://deanofstudents.utexas.edu/emergency/)

## Important Safety Information

If you have concerns about the safety or behavior of fellow students, TAs or Professors, call BCAL (the Behavior Concerns Advice Line): 512-232-5050. Your call can be anonymous. If something doesn’t feel right – it probably isn’t. Trust your instincts and share your concerns.

### Title IX Reporting

Title IX is a federal law that protects against sex and gender-based discrimination, sexual harassment, sexual assault, sexual misconduct, dating/domestic violence and stalking at federally funded educational institutions. UT Austin is committed to fostering a learning and working environment free from discrimination in all its forms. When sexual misconduct occurs in our community, the university can:

1. Intervene to prevent harmful behavior from continuing or escalating.
2. Provide support and remedies to students and employees who have experienced harm or have become involved in a Title IX investigation.
3. Investigate and discipline violations of the university’s [relevant policies](https://titleix.utexas.edu/relevant-polices/).

Faculty members and certain staff members are considered “Responsible Employees” or “Mandatory Reporters,” which means that they are required to report violations of Title IX to the Title IX Coordinator. **I am a Responsible Employee and must report any Title IX related incidents** that are disclosed in writing, discussion, or one-on-one. Before talking with me, or with any faculty or staff member about a Title IX related incident, be sure to ask whether they are a responsible employee. If you want to speak with someone for support or remedies without making an official report to the university, email [advocate@austin.utexas.edu](mailto:advocate@austin.utexas.edu) For more information about reporting options and resources, visit [titleix.utexas.edu](http://titleix.utexas.edu/) or contact the Title IX Office at [titleix@austin.utexas.edu](mailto:titleix@austin.utexas.edu).

The following recommendations regarding emergency evacuation from the Office of Campus Safety and Security, 512-471-5767, [operations.utexas.edu/units/csas](http://operations.utexas.edu/units/csas/)

Occupants of buildings on The University of Texas at Austin campus are required to evacuate buildings when a fire alarm is activated. Alarm activation or announcement requires exiting and assembling outside.

* Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building.
* Students requiring assistance in evacuation shall inform their instructor in writing during the first week of class.
* In the event of an evacuation, follow the instruction of faculty or class instructors. Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.
* Link to information regarding emergency evacuation routes and emergency procedures can be found at:

[emergency.utexas.edu](http://emergency.utexas.edu/)

LAND ACKNOWLEDGEMENT

(I) We would like to acknowledge that we are meeting on Indigenous land. Moreover, (I) We would like to acknowledge and pay our respects to the Carrizo & Comecrudo, Coahuiltecan, Caddo, Tonkawa, Comanche, Lipan Apache, Alabama-Coushatta, Kickapoo, Tigua Pueblo, and all the American Indian and Indigenous Peoples and communities who have been or have become a part of these lands and territories in Texas, here on Turtle Island. <https://liberalarts.utexas.edu/nais/land-acknowledgement/what-is-la.php>