INFO 3010 / Introduction to Data Science

Instructor Contact

Name: Lingzi Hong Pronouns: she/her

Office Location: DP E292E Phone Number: (940)369-8545

Office Hours: Thursdays 11:00am-12:00pm

Email: lingzi.hong@unt.edu

Communication Expectations: While I want to make myself as available as possible to each of you, I do have to place some limitations on when I can be contacted. I would prefer that most general questions go through the Q & A forum in the Discussion Board area. If you have a general question about the course or assignments, please post it there. Either I will answer it, or, one of your classmates will. This way we can all benefit from questions asked, and they can be answered in a venue that the whole class can see. You may also want to find someone in class to be a "buddy" with. This will give you at least one other person who you can email with questions.

If you have a private question, please contact me via email and I will respond within 24 hours on weekdays (usually sooner). Please do not expect a response over the weekend. Please use my phone number as a last resort - but, also, please use it if you need to!

Normally, I will return feedback on all written assignments within 1 week of the due date. However, if I see that I will be unable to return your feedback that quickly I will post an Announcement to let everyone know when it can be expected. You can expect to see me participate in the discussion board after all student original posts have been posted - usually on the Friday of the first week of the module.

Welcome to UNT!

As members of the UNT community, we have all made a commitment to be part of an institution that respects and values the identities of the students and employees with whom we interact. UNT does not tolerate identity-based discrimination, harassment, and retaliation. UNT's full Non-Discrimination Policy can be found in the UNT Policies section of the syllabus.

Course Description

The course is designed to introduce students the concepts, theories, and techniques that are involved in different stages of data lifecycle: data access, collection and management, data description and summarization, data modeling and evaluation. The course aims to present a broad view of data science in industry and research, which prepares students a solid base to develop specializations in data science. Students will acquire an integrated skill set that cover web and database technologies, programming, statistics, machine learning, while gain hands-on experience using R for applied data analysis and datadriven research.

Course Structure

The is an online course. In this semester, we have 16 weeks of study. Each week, we meet on Thursday from 2:00pm to 5:00pm on Zoom.

Topic: INFO 3010 Meeting Room

Join Zoom Meeting

https://unt.zoom.us/j/2504112817

Meeting ID: 250 411 2817

One tap mobile

+13462487799,,2504112817# US (Houston)

+16699006833,,2504112817# US (San Jose)

In the first half of the meeting, we discuss concepts, methods, and case studies in data science. The second half is lab session. Students are required to work on in-class practice. The instructors guide students and help them complete the in-class tasks. This course covers 12 topics in introductory data science. Each week focuses on one topic. In the last 4 weeks, students are teamed up to complete a course project.

Course Prerequisites or Other Restrictions

MATH 1650, MATH 1680, CSCE 1030, CSCE 1040; or permission of instructor.

Course Objectives

After successfully completing this course you will be able to:

- 1. Have a thorough understanding of stages in data lifecycle.
- 2. Build a mindset of managing and integrating different sources of information for decisionmaking through data.
- 3. Understand the generation of different types of data and how to obtain data from webpages and application databases.
- 4. Handle basic algorithms in machine learning for data modeling, including supervised methods and unsupervised clustering.

- 5. Foster critical thinking of ethical issues in data science projects.
- 6. Master the use of R for data scrapping, data manipulation, exploratory data analysis and data modeling.
- 7. Learn how to document and report data science projects.

Materials

Textbooks:

- 1. Irizarry, Rafael A. (2019). Introduction to Data Science: Data Analysis and Prediction Algorithms with R. Available at: https://rafalab.github.io/dsbook/
- 2. Cielen, D., Meysman, A., & Ali, M. (2016). Introducing data science: big data, machine learning, and more, using Python tools. Manning Publications Co.. Available at: https://livebook.manning.com/book/introducing-data-science/about-this-book/

Required Technology:

Laptop - We will do live programming exercises during most classes, so bring your laptop and be prepared to write code. Any reasonably current operating system can be used. If you don't have access to a laptop, contact me before the first class.

Teaching Philosophy

The instructor will explain concepts supplemented with case studies to guide students' understanding of stages in data lifecycle. The instructor takes heuristic methods to guide students thinking about questions and strategies in working with data. The instructor will work together with students on inclass coding practices to understand how to use R for data manipulation, data scrapping, data visualization and data modeling. The instructor will monitor the progress of students and is open to suggestions from students. Students are expected to participate actively in class, work together and help each other to achieve success in class. The students should submit their assignments on time to achieve satisfactory class performance. Interaction between the student and the instructor/TA is guaranteed and strongly encouraged. Students who don't have knowledge and experience in statistics or programming are expected to spend extra hours on this course.

Notes

Our time together in class is precious. To use it effectively, you must come to class on time and be prepared. Being prepared for class means that you have:

- 1. Completed the required readings.
- 2. Attempted the installation of required packages and read the documentations for hands-on practice in class. Either successfully completed them or submitted your questions the night before class, so I have time to prepare and answer in class.
- 3. Arrived 5 minutes before class starts, with R or RStudio ready. You have downloaded any notes, materials or datasets for the day from Canvas. Any paper assignments are ready to hand in. You are ready to take notes.

Computers are used when needed to accomplish a class objective. No phones or tablet devices are permitted during our class meetings. Researchers have found that such devices present an irresistible distraction and detract from the cooperative learning environment.

We will use Canvas Discussions as an online forum that you can use to ask/answer questions, get clarifications, point out my mistakes, etc. Be sure to check it regularly.

Here is my suggested general strategy for working on assignments:

- 1. Start early don't wait. That will give you time to work through the problems and get help as needed.
- 2. When you run into a problem, first trying to solve the problem by yourself. You can also search Q&A sites (e.g. StackOverflow) for solutions to technical questions, however search answers to assignment questions or sharing assignments online to get help is cheating.
- 3. You are encouraged to post the problem and how you learn to solve it on discussion board. You will get bonus points for actively participate in sharing information.
- 4. You are welcome to post your question in discussion board. I will be monitoring and will respond as soon as I am able, usually within a day (longer during weekends, travel, etc.).

Course Technology & Skills

Minimum Technology Requirements

- Computer
- Reliable internet access
- Software: Microsoft Office Suite, R, RStudio
- <u>Canvas Technical Requirements</u> (https://clear.unt.edu/supportedtechnologies/canvas/requirements)

Computer Skills & Digital Literacy

- Using Canvas
- Using email with attachments
- Downloading and installing software
- Using spreadsheet programs
- Using presentation and graphics programs

Technical Assistance

Part of working in the online environment involves dealing with the inconveniences and frustration that can arise when technology breaks down or does not perform as expected. Here at UNT we have a Student Help Desk that you can contact for help with Canvas or other technology issues.

UIT Help Desk: <u>UIT Student Help Desk site</u> (http://www.unt.edu/helpdesk/index.htm)

Email: helpdesk@unt.edu
Phone: 940-565-2324

In Person: Sage Hall, Room 130 Walk-In Availability: 8am-9pm

Telephone Availability:

• Sunday: noon-midnight

• Monday-Thursday: 8am-midnight

Friday: 8am-8pmSaturday: 9am-5pmLaptop Checkout: 8am-7pm

For additional support, visit <u>Canvas Technical Help</u> (https://community.canvaslms.com/docs/DOC-10554-4212710328)

Rules of Engagement

Rules of engagement refer to the way students are expected to interact with each other and with their instructors. Here are some general guidelines:

- While the freedom to express yourself is a fundamental human right, any communication that
 utilizes cruel and derogatory language on the basis of race, color, national origin, religion, sex,
 sexual orientation, gender identity, gender expression, age, disability, genetic information,
 veteran status, or any other characteristic protected under applicable federal or state law will
 not be tolerated.
- Treat your instructor and classmates with respect in any communication online or face-to-face, even when their opinion differs from your own.

- Ask for and use the correct name and pronouns for your instructor and classmates.
- Speak from personal experiences. Use "I" statements to share thoughts and feelings. Try not to speak on behalf of groups or other individual's experiences.
- Use your critical thinking skills to challenge other people's ideas, instead of attacking individuals.
- Avoid using all caps while communicating digitally. This may be interpreted as "YELLING!"
- Be cautious when using humor or sarcasm in emails or discussion posts as tone can be difficult to interpret digitally.
- Avoid using "text-talk" unless explicitly permitted by your instructor.
- Proofread and fact-check your sources.
- Keep in mind that online posts can be permanent, so think first before you type.

See these Engagement Guidelines (https://clear.unt.edu/online-communication-tips) for more information.

Course Requirements

List all required assignments and graded activities for the course, along with a short description and the points possible. Best practice is to clearly indicate both points and percentages if you are using both in the course. Here is a table example below:

Assignment	Points Possible	Percentage of Final Grade
Online Discussion	50 points	5%
 5 discussion questions @ 10 points ea. 		
In-class Practice	250 points	25%
 10 submissions @ 25 points ea. 		
Programming Assignments	350 points	35%
 7 submissions @ 50 points ea. 		
Exams	200 points	20%
 2 Exams @ 100 points ea. 		
Final Report	150 points	15%
Total Points Possible	1000 points	100%

Grading

Participation in online discussion (5%)

Students are required to participate at least 5 times in online discussions, either post questions or contribute to ideas to solve questions. Active and constructive participations may lead to bonus points.

In-class Participation (25%)

Students are required to attend class meetings. Prior to the meeting, please preview the readings or download the required R packages. In the meeting, the instructor works together with students to use R for different data processing tasks. After the meeting, students are required to submit code or discussion notes for in-class tasks. Students who cannot attend class meetings due to certain reasons are required to learn by themselves and submit code or response for in-class tasks. Code and discussion notes will be assessed for grades.

Programming Assignments (35%)

The class will have seven assignments. The assignments are designed to help students understand important concepts, methods to work with different types of data and gain hands-on experience of using R for data manipulation, visualization, and modeling. The assignments should be typewritten, and diagrams should be drawn using graphics software packages such as R, PowerPoints and Excel.

Exams (20%)

The exams are designed to assess the students' coding skills and understanding in data processing.

Final Report (20%)

The final report is designed to assess the students' coding skills, understanding in data processing, writing and reporting.

The grading scale (A-F) according to the point totals.

A = 900-1000

B = 800 - 899

C = 700-799

D = 600-699

F = 500-599

Course Evaluation

Student feedback is important and an essential part of participation in this course. The student evaluation of instruction is a requirement for all organized classes at UNT. The survey will be made available during weeks 13, 14 and 15 of the long semesters to provide students with an opportunity to evaluate how this course is taught. Students will receive an email from "UNT SPOT Course Evaluations via IASystem Notification" (no-reply@iasystem.org) with the survey link. Students should look for the email in their UNT email inbox. Simply click on the link and complete the survey. Once students complete the survey they will receive a confirmation email that the survey has been submitted. For additional information, please visit the SPOT website at http://spot.unt.edu/ or email spot@unt.edu.

Course Policies

Assignment Policy

The University is committed to providing a reliable online course system to all users. However, in the event of any unexpected server outage or any unusual technical difficulty which prevents students from completing a time sensitive assessment activity, the instructor will extend the time windows and provide an appropriate accommodation based on the situation. Students should immediately report any problems to the instructor and contact the UNT Student Help Desk: helpdesk@unt.edu or 940.565.2324 and obtain a ticket number. The instructor and the UNT Student Help Desk will work with the student to resolve any issues at the earliest possible time.

Grade Policy

Your grade is determined by your performance on the learning assessments and is assigned individually (not curved). All assessment scores will be posted on the course page. If you would like to review any of your grades, or have questions about how something was scored, please email me to schedule a time for us to meet in my office. Any grade disputes must be submitted in writing and within one week of receiving the grade. The exams are open-book exams. Students are required to take a make-up exam if cannot make the scheduled one.

Instructor Responsibilities and Feedback

- 1. Helping students grow and learn
- 2. Providing clear instructions for projects and assessments
- 3. Answering questions about assignments
- 4. Identifying additional resources as necessary
- 5. Providing grading rubrics
- 6. Reviewing and updating course content
- 7. The instructor and TA will respond to students' emails and questions posted to the discussion boards within two days except for the weekends
- 8. Assignments grades and feedbacks will be returned to the students within one week after the submission deadline.

Late Work

Students are expected to submit assignments and projects on time. The due dates are Mondays 11:59pm of the week specified in Table 2. If an extenuating circumstance such as a medically diagnosed illness or family emergency arises, which prevents you from submitting your assignments, you should contact the instructor and the TA as soon as possible before the due date. Late work without the permission of the instructor will receive a grade with a 10% penalty (or 10 points out of 100) per day after the due date. A student who is having trouble with assignments is strongly encouraged to contact the instructor and the TA as early as possible for personal advising.

Attendance Policy

It is mandatory for students to attend each class meetings. Prior to the meeting, please preview the readings and install required R packages. You may take much longer time to work on in-class submissions or lose points in the submissions if you don't attend class.

Class Participation

The students are required to follow course schedule and finish the class work, assignments, and term projects. Students are expected to study 12-15 hours per week to achieve satisfactory class performance. Students do not have programming experience are required to find extra materials to study.

Syllabus Change Policy

The due dates of assignments may change according to the progress of learning of students.

UNT Policies

Academic Integrity Policy

Academic Integrity Standards and Consequences. According to UNT Policy 06.003, Student Academic Integrity, academic dishonesty occurs when students engage in behaviors including, but not limited to cheating, fabrication, facilitating academic dishonesty, forgery, plagiarism, and sabotage. A finding of academic dishonesty may result in a range of academic penalties or sanctions ranging from admonition to expulsion from the University.

ADA Policy

UNT makes reasonable academic accommodation for students with disabilities. Students seeking accommodation must first register with the Office of Disability Accommodation (ODA) to verify their eligibility. If a disability is verified, the ODA will provide a student with an accommodation letter to be delivered to faculty to begin a private discussion regarding one's specific course needs. Students may request accommodations at any time, however, ODA notices of accommodation should be provided as early as possible in the semester to avoid any delay in implementation. Note that students must obtain a new letter of accommodation for every semester and must meet with each faculty member prior to implementation in each class. For additional information see the ODA website (https://disability.unt.edu/).

Prohibition of Discrimination, Harassment, and Retaliation (Policy 16.004)

The University of North Texas (UNT) prohibits discrimination and harassment because of race, color, national origin, religion, sex, sexual orientation, gender identity, gender expression, age, disability, genetic information, veteran status, or any other characteristic protected under applicable federal or state law in its application and admission processes; educational programs and activities; employment policies, procedures, and processes; and university facilities. The University takes active measures to prevent such conduct and investigates and takes remedial action when appropriate.

Emergency Notification & Procedures

UNT uses a system called Eagle Alert to quickly notify students with critical information in the event of an emergency (i.e., severe weather, campus closing, and health and public safety emergencies like chemical spills, fires, or violence). In the event of a university closure, please refer to Canvas for contingency plans for covering course materials.

Retention of Student Records

Student records pertaining to this course are maintained in a secure location by the instructor of record. All records such as exams, answer sheets (with keys), and written papers submitted during the duration of the course are kept for at least one calendar year after course completion. Course work completed via the Canvas online system, including grading information and comments, is also stored in a safe electronic environment for one year. Students have the right to view their individual record; however, information about student's records will not be divulged to other individuals without proper written consent. Students are encouraged to review the Public Information Policy and the Family Educational Rights and Privacy Act (FERPA) laws and the University's policy. See UNT Policy 10.10, Records Management and Retention for additional information.

Acceptable Student Behavior

Student behavior that interferes with an instructor's ability to conduct a class or other students' opportunity to learn is unacceptable and disruptive and will not be tolerated in any instructional forum at UNT. Students engaging in unacceptable behavior will be directed to leave the classroom and the instructor may refer the student to the Dean of Students to consider whether the student's conduct violated the Code of Student Conduct. The University's expectations for student conduct apply to all instructional forums, including University and electronic classroom, labs, discussion groups, field trips, etc. Visit UNT's Code of Student Conduct (https://deanofstudents.unt.edu/conduct) to learn more.

Access to Information - Eagle Connect

Students' access point for business and academic services at UNT is located at: my.unt.edu. All official communication from the University will be delivered to a student's Eagle Connect account. For more information, please visit the website that explains Eagle Connect and how to forward e-mail Eagle Connect (https://it.unt.edu/eagleconnect).

Student Evaluation Administration Dates

Student feedback is important and an essential part of participation in this course. The student evaluation of instruction is a requirement for all organized classes at UNT. The survey will be made available during weeks 13, 14 and 15 [insert administration dates] of the long semesters to provide students with an opportunity to evaluate how this course is taught. Students will receive an email from "UNT SPOT Course Evaluations via IASystem Notification" (no-reply@iasystem.org) with the survey link. Students should look for the email in their UNT email inbox. Simply click on the link and complete the survey. Once students complete the survey they will receive a confirmation email that the survey has been submitted. For additional information, please visit the SPOT website (http://spot.unt.edu/) or email spot@unt.edu.

Sexual Assault Prevention

UNT is committed to providing a safe learning environment free of all forms of sexual misconduct, including sexual harassment sexual assault, domestic violence, dating violence, and stalking. Federal laws (Title IX and the Violence Against Women Act) and UNT policies prohibit discrimination on the basis of sex, and therefore prohibit sexual misconduct. If you or someone you know is experiencing sexual harassment, relationship violence, stalking, and/or sexual assault, there are campus resources available to provide support and assistance. UNT's Survivor Advocates can assist a student who has been impacted by violence by filing protective orders, completing crime victim's compensation applications, contacting professors for absences related to an assault, working with housing to facilitate a room change where appropriate, and connecting students to other resources available both on and off campus. The Survivor Advocates can be reached at SurvivorAdvocate@unt.edu or by calling the Dean of Students Office at 940-565- 2648. Additionally, alleged sexual misconduct can be non-confidentially reported to the Title IX Coordinator at oeo@unt.edu or at (940) 565 2759.

Important Notice for F-1 Students taking Distance Education Courses

Federal Regulation

To read detailed Immigration and Customs Enforcement regulations for F-1 students taking online courses, please go to the Electronic Code of Federal Regulations website (http://www.ecfr.gov/). The specific portion concerning distance education courses is located at Title 8 CFR 214.2 Paragraph (f)(6)(i)(G).

The paragraph reads:

(G) For F-1 students enrolled in classes for credit or classroom hours, no more than the equivalent of one class or three credits per session, term, semester, trimester, or quarter may be counted toward the full course of study requirement if the class is taken on-line or through distance education and does not require the student's physical attendance for classes, examination or other purposes integral to completion of the class. An on-line or distance education course is a course that is offered principally through the use of television, audio, or computer transmission including open broadcast, closed circuit, cable, microwave, or satellite, audio conferencing, or computer conferencing. If the F-1 student's course of study is in a language study program, no on-line or distance education classes may be considered to count toward a student's full course of study requirement.

University of North Texas Compliance

To comply with immigration regulations, an F-1 visa holder within the United States may need to engage in an on-campus experiential component for this course. This component (which must be approved in advance by the instructor) can include activities such as taking an on-campus exam, participating in an on-campus lecture or lab activity, or other on-campus experience integral to the completion of this course.

If such an on-campus activity is required, it is the student's responsibility to do the following:

- (1) Submit a written request to the instructor for an on-campus experiential component within one week of the start of the course.
- (2) Ensure that the activity on campus takes place and the instructor documents it in writing with a notice sent to the International Student and Scholar Services Office. ISSS has a form available that you may use for this purpose.

Because the decision may have serious immigration consequences, if an F-1 student is unsure about his or her need to participate in an on-campus experiential component for this course, s/he should contact the UNT International Student and Scholar Services Office (telephone 940-565-2195 or email internationaladvising@unt.edu) to get clarification before the one-week deadline.

Student Verification

UNT takes measures to protect the integrity of educational credentials awarded to students enrolled in distance education courses by verifying student identity, protecting student privacy, and notifying students of any special meeting times/locations or additional charges associated with student identity verification in distance education courses.

See UNT Policy 07-002 Student Identity Verification, Privacy, and Notification and Distance Education Courses (https://policy.unt.edu/policy/07-002).

Use of Student Work

A student owns the copyright for all work (e.g. software, photographs, reports, presentations, and email postings) he or she creates within a class and the University is not entitled to use any student work without the student's permission unless all of the following criteria are met:

- The work is used only once.
- The work is not used in its entirety.
- Use of the work does not affect any potential profits from the work.
- The student is not identified.
- The work is identified as student work.

If the use of the work does not meet all of the above criteria, then the University office or department using the work must obtain the student's written permission.

Download the UNT System Permission, Waiver and Release Form

Transmission and Recording of Student Images in Electronically-Delivered Courses

- 1. No permission is needed from a student for his or her image or voice to be transmitted live via videoconference or streaming media, but all students should be informed when courses are to be conducted using either method of delivery.
- 2. In the event an instructor records student presentations, he or she must obtain permission from the student using a signed release in order to use the recording for future classes in accordance with the Use of Student-Created Work guidelines above.
- 3. Instructors who video-record their class lectures with the intention of re-using some or all of recordings for future class offerings must notify students on the course syllabus if students' images may appear on video. Instructors are also advised to provide accommodation for students who do not wish to appear in class recordings.

Example: This course employs lecture capture technology to record class sessions. Students may occasionally appear on video. The lecture recordings will be available to you for study purposes and may also be reused in future course offerings.

No notification is needed if only audio and slide capture is used or if the video only records the instructor's image. However, the instructor is encouraged to let students know the recordings will be available to them for study purposes.

Academic Support & Student Services

Student Support Services

Mental Health

UNT provides mental health resources to students to help ensure there are numerous outlets to turn to that wholeheartedly care for and are there for students in need, regardless of the nature of an issue or its severity. Listed below are several resources on campus that can support your academic success and mental well-being:

- Student Health and Wellness Center (https://studentaffairs.unt.edu/student-health-andwellness-center)
- <u>Counseling and Testing Services</u> (https://studentaffairs.unt.edu/counseling-and-testing-services)
- UNT Care Team (https://studentaffairs.unt.edu/care)
- UNT Psychiatric Services (https://studentaffairs.unt.edu/student-health-and-wellnesscenter/services/psychiatry)
- Individual Counseling (https://studentaffairs.unt.edu/counseling-and-testingservices/services/individual-counseling)

Chosen Names

A chosen name is a name that a person goes by that may or may not match their legal name. If you have a chosen name that is different from your legal name and would like that to be used in class, please let the instructor know. Below is a list of resources for updating your chosen name at UNT.

- UNT Records
- UNT ID Card
- UNT Email Address
- Legal Name

*UNT euIDs cannot be changed at this time. The collaborating offices are working on a process to make this option accessible to UNT community members.

Pronouns

Pronouns (she/her, they/them, he/him, etc.) are a public way for people to address you, much like your name, and can be shared with a name when making an introduction, both virtually and in-person. Just as we ask and don't assume someone's name, we should also ask and not assume someone's pronouns.

You can add your pronouns to your Canvas account so that they follow your name when posting to discussion boards, submitting assignments, etc.

Below is a list of additional resources regarding pronouns and their usage:

- o What are pronouns and why are they important?
- o How do I use pronouns?
- o How do I share my pronouns?
- o How do I ask for another person's pronouns?
- o How do I correct myself or others when the wrong pronoun is used?

Additional Student Support Services

- Registrar (https://registrar.unt.edu/registration)
- Financial Aid (https://financialaid.unt.edu/)
- Student Legal Services (https://studentaffairs.unt.edu/student-legal-services)
- <u>Career Center</u> (https://studentaffairs.unt.edu/career-center)
- Multicultural Center (https://edo.unt.edu/multicultural-center)
- <u>Counseling and Testing Services</u> (https://studentaffairs.unt.edu/counseling-and-testing-services)
- Pride Alliance (https://edo.unt.edu/pridealliance)
- <u>UNT Food Pantry</u> (https://deanofstudents.unt.edu/resources/food-pantry)

Academic Support Services

- <u>Academic Resource Center</u> (https://clear.unt.edu/canvas/student-resources)
- <u>Academic Success Center</u> (https://success.unt.edu/asc)
- <u>UNT Libraries</u> (https://library.unt.edu/)
- Writing Lab (http://writingcenter.unt.edu/)