## Instructor Contact

**Name: Yixun Xing, PhD**
**Office Location: GAB 110D**
**Office Hours: Tu 5-5:50pm (FRLD 238);**
   **Th 5-5:50pm (GAB 110D);**
   **Fri 2-4pm (GAB 110D);**
   **you may reserve a time slot to meet with me online on https://calendly.com/yxing/office-hour ;**
   **other time by appointment via email (also available on most weekends).**
**Email: yixun.xing@unt.edu**
**Communication Expectations:** The best way to contact me is via UNT email (not the Canvas email tool). Emails will be answered in timely manner, usually within 24 hours. Many part-time students work on assignments during weekends, so I also answer emails on weekends. Please include your course and section number in the email as I teach several courses/sections. Please send emails via your UNT account because external emails may be routed to the junk folder and emails are expected to follow professional etiquette standards as these are formal communications.
**About the Professor:** Dr. Xing earned a Ph.D. degree in Statistical Science from SMU and had been a post-doctoral researcher at UTSW. She also had been adjunct faculty and a full-time assistant professor teaching data science and statistics for years. Besides academia, she has variety of working experience in pharmaceutical, financial services, and asset management companies.

## Course Description

This course focuses on using advanced analytics in practical case studies to help students develop the skills needed to address complex challenges in industry and business. Students will select projects based on provided datasets, formulate insightful questions, and devise effective solutions that align with project goals. They will choose appropriate methods from various alternatives and implement relevant techniques and technologies in real-world scenarios. Emphasis is placed on mastering the deployment of analytics and improving professional communication skills.

## Course Structure

This is an 8-week course. We will meet on Zoom every Mon 7:30-9:00pm CT. Participation in the weekly meeting is optional, but you are required to watch the recordings if you couldn't attend. Students are also expected to participate in various online activities such as reading papers, watching videos, and asynchronous discussions.

## Course Prerequisites or Other Restrictions

This course requires that the student has successfully completed ADTA 5130 Data Analytics 1, ADTA 5230 Data Analytics 2, ADTA 5250 Large Data Visualization, ADTA 5240 Harvesting, Storing and Retrieving Data, ADTA5340 Discovery and Learning with Big Data. Please check the department website for the latest policy changes regarding the course prerequisite policies.

## Course Objectives

By the end of the course, students should be able to:
1. Evaluate various research and project methodologies, identifying their suitability for addressing complex problems in industry and business contexts.

2. Formulate clear, focused, and actionable questions that guide data-driven projects in professional/academic settings.
3. Conduct targeted literature and market analysis using academic and industry databases to identify relevant trends, findings, and best practices.
4. Design a practical and methodologically sound approach for analyzing real-world data in a business or academic context.
5. Draft a data collection plan for a research or applied project.
6. Apply appropriate data analysis techniques to the collected data (e.g., statistical analysis, coding for qualitative data).
7. Interpret data results and create visual representations (graphs, charts) to highlight key findings.
8. Structure a research paper by organizing the introduction, methods, results, and discussion sections. Present research findings clearly and effectively using PowerPoint or another presentation tool. Submit a final research project report that meets academic standards for clarity, structure, and content.

## Materials

**Recommended textbook:**

**James, G., Witten, D., Hastie, T., Tibshirani, R. , & Taylor, J. (2023).** *An Introduction to Statistical Learning: with Applications in Python (Springer Texts in Statistics)*. **Springer.**
The *free* eBook can be downloaded from https://www.statlearning.com/

**Required software: Python**

**Optional readings:**

UNT Willis Library provides free PDF and EPUB of some of below materials.

- L Kuhn, Max, and Kjell Johnson (2018). *Applied predictive modeling*. Vol. 26. New York: Springer. (https://link-springer-com.libproxy.library.unt.edu/book/10.1007/978-1-4614-6849-3 )
- Hastie, T., Tibshirani, R., and Friedman, L. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. (https://hastie.su.domains/ElemStatLearn/)
- Wickham,H. & Grolemund, G. (2019). *R for Data Science*. (https://r4ds.had.co.nz/)
- Chang, W. (2020). *R Graphics Cookbook*, *2nd Ed*. (https://r-graphics.org/)
- Molnar, C. (2021). *Interpretable machine learning. A Guide for Making Black Box Models Explainable*. (https://christophm.github.io/interpretable-ml-book/)
- Lantz, Brett. (2019). *Machine Learning with R, 3rd* Ed.
- van Buuren, S. (2012). *Flexible Imputation of Missing Data*. Boca Raton: CRC Press. (https://stefvanbuuren.name/fimd/ )
- Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). *Learning from imbalanced data sets.* Berlin: Springer. (https://link-springer-com.libproxy.library.unt.edu/book/10.1007/978-3-319-98074-4)

Some Modules in Canvas may contain some supplemental materials.

## 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**:
Website: https://aits.unt.edu/support/
Below hours are subject to the Current Hours on their website.
Walk-In (Sage Hall 330): 8am-5pm
Call/Chat/Email: • Monday-Thursday: 8am-9pm • Friday: 8am-5pm • Saturday-Sunday: 11am-3pm
Laptop Checkout: 8am-7pm

## Course Requirements and Grading

Your final grade will be determined based on the assignments noted in the table below.

| Assignments | Percentage of Final Grade |
|---|---|
| **Homework** | 40% |
| **Quizzes**<br>5 quizzes (lowest dropped) | 10% |
| **Data Camp**<br>7 Data Camp courses (lowest dropped) | 20% |
| **Final Project**<br>1 final group project report | 15% |
| **Recorded Presentation** | 15% |
| **Total** | 100% |

| Final Grade | A | B | C | D | F |
|---|---|---|---|---|---|
| **Final Percentage** | 89.5-100% | 79.5-89.4% | 69.5-79.4% | 59.5-69.4% | Below 59.5% |

**Quizzes**
Quizzes will not be timed, and you can save answers and revisit any time until the due date. Quiz may be taken twice (unlimited attempts for the syllabus quiz) with the highest score counted toward the grade. The lowest will be dropped.

**Homework**
This component includes a literature review report, a project proposal report, and several labs.
The *literature review report* aims to enhance your skills in constructing a literature review using AI tools, and critically evaluating the accuracy and relevance of the sources cited by the AI. You will interact with two AI tools such as ChatGPT and Gemini to collect information for a draft of a literature review on a specified topic and then conduct a thorough fact-checking exercise to assess the validity and reliability of the references cited.
In *project proposal*, we aim to write a project proposal, which can be completed in groups of 2-3 students. The proposal should outline a clear research question, describe the dataset to be used, and specify the potential analytical methods. This assignment serves as the foundation for your final project.
You will also gain hands-on experience in the *labs*, which may include questions to be answered about a specific concept, analysis using provided data sets, interpretation of the results of the analysis, or questions related to the course material. Written responses are expected to be free of grammatical errors. You must turn in required files via Canvas.

**Data Camp**
Complete the required courses on Data Camp to earn full marks of the Data Camp assignment. If you only complete a fraction of the course (say 50%) by the deadline, your grade will be that fraction of the full score.

Your score on the DataCamp activities will not be factored in your grade. In Data Camp, you may get the answers to the exercises, but make as many attempts as possible so you get sufficient practice. If you request the answer, please review the code to understand the solution. Note that you must register on Data Camp using your UNT e-mail to obtain the free 6-months subscription.

**Final Project**

- Project assignment will give you real experience of application of advanced data analytics. Groups of 2-3 students will be randomly formed in Canvas to complete a project. Team members are jointly responsible for the team's work, including academic integrity violations.
- **Peer review** Each student will provide a peer review on team members' contribution to the final product, and your grade will be determined by the quality of the product and your contribution. If everyone contributes equally, all members will receive the same credit. Team members who contribute insufficiently will be deducted points. The professor reserves the right to adjust the team's recommended contribution.

**Presentations**

- Groups of students will also provide a presentation on the project during the semester.
- **Peer review** Each student will provide a peer review on team members' contribution to the final product, and your grade will be determined by the quality of the product and your contribution. If everyone contributes equally, all members will receive the same credit. Team members who contribute insufficiently will be deducted points. The professor reserves the right to adjust the team's recommended contribution.

| Tentative Schedule: **All due time is 11:59pm CT of the designated date unless otherwise noted (**It is your responsibility to check for changes in the schedule and important information in the posted Announcement in Canvas.**)** | | | |
|---|---|---|---|
| **Module** | **LECT Date** | **Topic / Learning Activities** | **Assignments (Due date)** |
| Week 1 | 01/12 | Review machine learning, statistical learning, and Python Basics/ <br> Optional Reading: ISLP – C1, 2 | • Data Camp register and Data Camp 1 (01/18) <br> • Syllabus quiz (01/18) |
| Week 2 | 01/19 | Understanding the Context and Data <br> Developing a Research Proposal | • Quiz Basics (01/25) <br> • Data Camps 2 and 3 (01/25) <br> • HW-COVID data (01/25) |
| Week 3 | 01/26 | Literature Review / <br> Required Reading on Canvas | • HW-Project proposal (02/01) <br> • HW-Literature review report (02/01) |
| Week 4 | 02/02 | Professional & Academic Communication in Data Analytics 1/ | |
| Week 5 | 02/09 | Professional & Academic Communication in Data Analytics 2/ <br> Optional Reading: ISLP – C5 | • Presentation of EDA Recording (02/15) <br> • Quiz Resampling (02/15) |
| Week 6 | 02/16 | Review Data and Select Methodology 1 <br> Optional Reading: ISLP – C3, 6 | • Data Camps 4 and 5 (02/22) <br> • Quiz Model selection and regularization (02/22) |
| Week 7 | 02/23 | Review Data and Select Methodology 2 <br> Optional Reading: <br> •   ISLP – C4 <br> •   APM - C16 Remedies for Severe Class Imbalance <br> •   ISLP – C8 | • Data Camps 6 and 7 (03/01) <br> • Quiz Classifier (03/01) <br> • HW Classifier (03/01) |
| Wee 8 | 03/02 | Final project and Course Review | • Final project report (03/05) |

## Course Policies

### Late Work

The students are expected try best to submit assignments on time and will be responsible for the consequence if they start the assignments late at a time approaching the deadline. Unless pre-approved by the instructor, any assignment submitted after that time will receive a highest possible score of 80% through one (1) days past the deadline and no points will be awarded after that. Additional points may be deducted based on the quality of the submission. However, late final project will not be accepted. Please do not lose valuable points due to late work.

### AI Tools Policy

AI tools may be used in a limited capacity to support learning in this course. All submissions must reflect the student's own understanding and effort. Prohibited uses include, but not limited to, the followings:

- Submitting AI-generated work as your own without proper attribution.
- Using AI-generated content without critical review or confirmation, such as including false or fabricated references.
- Using AI during exams or generating full assignments without approval.

Guidelines: If AI significantly contributes to your work, disclose its use (e.g., "This section was assisted by [Tool Name]"). Review and refine AI outputs to ensure accuracy and originality.

Misuse of AI may violate academic integrity policies and result in disciplinary action. For questions, consult the instructor.

### Attendance Policy

The student is responsible for regular and punctual attendance and is expected to participate in all courses in which the student is enrolled. It is important that you communicate with the professor and the instructional team prior to being absent, so you, the professor, and the instructional team can discuss and mitigate the impact of the absence on your attainment of course learning goals. Please inform the professor and instructional team if you are unable to attend class meetings because you are ill, in mindfulness of the health and safety of everyone in our community.

### Class Participation

Students must login regularly to the Canvas learning management system. The instructor can use the tracking feature in Canvas to monitor student activity. Students are also expected to participate in all learning activities such as discussion board and projects.

### Group Project Policy

Group projects are an essential part of this course. These projects help you develop collaboration skills that are essential in the workplace. Group projects also contribute to learning and retention of class content. Other benefits include practice with time management and communication skills, giving and receiving constructive feedback, sharing perspectives in a respectful manner, and developing conflict management skills.

Ground rules: • Start the project early. Begin by discussing and defining project goals, group leadership, time schedule, and preferred means of communication • Respect each group member, attend group meetings prepared and on time • Each group member should complete a fair share of the work • Contact your instructor (sooner rather than later) If there is an issue or group dynamic that cannot be resolved in a timely manner.

You are expected to follow UNT's Code of Student Conduct which is intended to "foster a safe environment conducive to learning and development. Students and student groups are expected to conduct themselves in a manner that demonstrates respect for the rights and property of others and upholds the integrity and values of

the University community." The Code of Student Conduct can be found at https://policy.unt.edu/policy/07-012 . You are also expected to follow UNT's Student Academic Integrity Policy.

### Examination Policy

Exams must be completed independently. Students that engage in academic dishonesty will suffer the consequences per department guidelines. If you lose your internet connection during an exam, contact the help desk immediately and notify me as well. If necessary, I can reset your exam.

### Assignment Policy

Assignment due time (all in Central Time) and dates are in the syllabus and on Canvas. Any changes to due dates will be updated on Canvas and communicated in an announcement. For assignments that require you to upload your work, the submission should be in one of the following formats: .doc, .docx, .csv, .xlsx, .sas or .pptx. Do NOT submit .pages files. TurnitIn will be utilized to address plagiarism issues in all formal scholarly writing. All works submitted for credit must be original works created by the scholar uniquely for the class. It is considered inappropriate and unethical, particularly at an advanced undergraduate/graduate level, to make duplicate submissions of a single work for credit in multiple classes, unless specifically requested by the instructor. It is also considered inappropriate and unethical to work together on individual assignments or share work that is to be created on an individual level. Work submitted at the senior/graduate level is expected to demonstrate higher-order thinking skills and be of significantly higher quality than work produced at the lower undergraduate levels.

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.

### Turnitin Notice

Turnitin is used as a tool to assist students in their scholarly writing to address plagiarism issues. It is recommended that students use TurnitIn to ensure their work is free of copyright issues prior to final submission of their projects.

### Syllabus Change Policy

While the plan is to follow this syllabus as written, adjustments may be made when necessary or due to events outside of control. Any change will be announced.

### UNT Policies

### Academic Integrity Policy

The University of North Texas promotes the integrity of learning and embraces the core values of trust and honesty. Academic integrity is based on educational principles and procedures that protect the rights of all participants in the educational process and validate the legitimacy of degrees awarded by the University. In the investigation and resolution of allegations of student academic dishonesty, the University's actions are intended to be corrective, educationally sound, fundamentally fair, and based on reliable evidence. 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 (including violating course policies on AI usage), and sabotage. A finding of academic dishonesty will result in a range of

academic penalties or sanctions ranging from admonition to expulsion from the University. ADTA students must read and adhere to the university, department, and course Academic Integrity expectations. The consequences of any academic misconduct, including the first-time violation, are outlined below.

| | Penalty | Other |
|---|---|---|
| 1st Academic Integrity Offense | The minimum penalty is a 0 for the assignment AND a deduction of one letter grade from the final grade for the course. Other penalties may be assessed by the course instructor up to course failure, depending on the severity of the offense. | All Academic Integrity offenses will be reported to the UNT Academic Integrity Office following the UNT policy process. |
| 2nd Academic Integrity Offense | Suspension from the ADTA program. | A second offense is defined as a separately reported offense either in the same class as the 1st offense or in a different course. Students suspended for a second Academic Integrity violation will not be allowed to enroll in ADTA courses for 1 calendar year. For students who had a single Academic Integrity violation prior to this current semester, a second violation will result in suspension from the ADTA program. |
| 3rd Academic Integrity Offense | Dismissal from the ADTA program. | Students committing a 3rd Academic Integrity offense will be dismissed from the program. For students who had multiple Academic Integrity violations prior to this current semester, any additional violation will result in dismissal from the ADTA program. |

## 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 reasonable accommodation for every semester and must meet with each faculty member prior to implementation in each class. Students are strongly encouraged to deliver letters of reasonable accommodation during faculty office hours or by appointment. Faculty members have the authority to ask students to discuss such letters during their designated office hours to protect the privacy of the student.  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.

## Copyright Notice

Materials used in connection with this course may be subject to copyright protection. Materials may include, but are not limited to: documents, slides, images, audio, and video. Materials in this course Web site are only for the use of students enrolled in this course, for purposes associated with this course, and may not be retained for longer than the class term. Unauthorized retention, duplication, distribution, or modification of copyrighted materials is strictly prohibited by law. For more information, visit http://policy.unt.edu/policy/08-001