

INFO 5502: ANALYTIC TOOLS, TECHNIQUES AND METHODS

Spring 2020

Instructor:	Junhua Ding, Ph.D.	Lecture Time:	Thursday: 2:30pm - 5:20pm
Email:	junhua.ding@unt.edu	Place:	NTDP B190
Office:	Discovery Park: Room E292G	Phone:	940-565-2186

Course Information

- INFO5502 (Sections 002): Analytic Tools, Techniques and Methods, 3 Credit Hours
- Lecture time: Thursday: 2:30pm - 5:20pm.
- Lecture delivering: face to face
- Course management at Canvas: <https://unt.instructure.com/>
- No pre-requisites.

Office Hours

- Wednesday, Thursday: 12:15pm - 2:15pm (in office)
- or by appointment.

Textbooks

1. *Computational and Inferential Thinking: The Foundations of Data Science* by Ani Adhikari and John DeNero, <https://www.inferentialthinking.com/chapters/intro>. Free available online. (**required**)
2. *Principles and Techniques of Data Science*, by Sam Lau, Joey Gonzalez, and Deb Nolan., <https://www.textbook.ds100.org/>. Free available online. (**required**).
3. *Introduction to Computation and Programming Using Python: With Application to Understanding Data*. 2nd ed., by John Guttag. MIT Press, 2016. ISBN: 9780262529624. (**optional**).

Software

The following software will be used in this course.

- Python Jupyter Notebook, JupyterLab, Python Frameworks such as NumPy, SciPy, Pandas; Microsoft Excel; MySQL.

Objectives

The course teaches analytic tools, techniques and methods for data science. It covers the critical concepts and areas in data science. These include languages for transforming, querying and analyzing data; algorithms for machine learning methods including regression, classification and clustering; principles of data visualizations; statistical concepts of measurement error and prediction; and techniques for scalable data processing. We explore key areas of data science including question formulation, data collection and cleaning, visualization, statistical inference, predictive modeling, and decision making. The objectives of this course:

-
1. Students will be able to develop programs using Python tools and libraries for basic data science tasks including data collection, data cleaning, data transformation, data management, data analytics, and data visualization.
 2. Students will be able to select and use analytic tools, techniques, and methods to address real-world data science problems.

Topics

The following is the tentative schedule with the covered topics. Actual schedule may be adjusted according to progress:

1. Introduction
2. Python Programming
3. Data Design
4. Data Cleaning and Exploratory Data Analysis (EDA)
5. Data Visualization
6. Randomness and Monte Carlo
7. Sampling and Empirical Distributions
8. Bootstrapping
9. Testing Hypotheses
10. Comparing Two Samples
11. Modeling and Estimation
12. Prediction
13. Regression
14. Classification
15. Working with Text
16. Databases and SQL
17. Gradient Descent and Numerical Optimization
18. Probability and Generalization
19. Feature Engineering
20. The Bias-Variance Tradeoff
21. Regularization
22. Statistical Inference

Teaching Philosophy

This course will be taught in traditional lectures, and on-class discussions are strongly encouraged. Although software tools such as Python Notebook will be needed to complete the assignments, no lecture will be given on how to use the tools.

Technical Assistance

UIT Help Desk: <http://www.unt.edu/helpdesk/index.htm>. The University of North Texas provides student technical support in the use of Zoom and Canvas and supported resources. The student help desk may be reached at:

- Email: helpdesk@unt.edu
- Phone: 940.565-2324
- In Person: Sage Hall, Room 130

Hours

- Monday-Thursday 8am-midnight
- Friday 8am-8pm
- Saturday 9am-5p
- Sunday 8am-midnight

Canvas technical requirements: <https://clear.unt.edu/supported-technologies/canvas/requirements>

Minimal Technical Skills and Resources Needed

Each student should be able to access computers that have environment for running and developing Python programs with databases Microsoft Excel and mySQL. Google Colab (<https://colab.research.google.com/>) is a recommended environment for developing course projects and running sample codes from the textbook.

Communication

Students can email their questions to the instructor and the teacher assistant (TA). They are also encouraged to talk to the instructor and TA during the office hours. Emails are normally respond within 24 hours, and all assignments, quizzes, projects, and papers should be graded within 10 days after the submission deadline.

Grading Policy Students are required to attend the class on time, complete all assignments, quizzes, exams, and readings on time. Grading will be based on assignments to be assigned as the course proceeds.

Grades will be computed as follows:

- 8 Project Assignments (Python programming for different data science tasks): 70%
- 2 Exams: 20%
- 4 Quizzes: 10%
- **Grading Scale:** A: 90-100; B: 80-89; C: 70-79; D: 60-69; F: 59 or below.

All assignments, projects, quizzes and term paper requirements are post in Canvas, and students are required to uploaded their work into Canvas. Late submission of quizzes won't be graded, but late submission of other work will be graded with reduced points. The final grade is calculated based on grade points of assignments, project, quizzes, and exam.

Incompletes

A grade of incomplete (I) will be given only for a justifiable reason (such as a serious illness or military service) and only if you are passing the course. It is your responsibility to contact the instructor to request an incomplete and discuss requirements for completing the course. If you do not remove the incomplete within the timeframe agreed upon with the instructor or within one calendar year, you will receive a grade of an F. Please refer to <http://essc.unt.edu/registrar/academic-record-incomplete.html> for more information.

Withdrawal

A grade of withdraw (W) or withdraw-failing (WF) will be given depending on your participation and grades to date. If you simply disappear and do not file a formal UNT withdrawal form, you may receive a grade of an F.

ADA Statement

The University of North Texas makes reasonable academic accommodation for students with disabilities. Students seeking reasonable accommodation must first register with the Office of Disability Accommodation (ODA) to verify their eligibility. If a disability is verified, the ODA will provide you with a reasonable accommodation letter to be delivered to faculty to begin a private discussion regarding your specific needs in a course. You may request reasonable accommodations at any time, however, ODA notices of reasonable 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 Office of Disability Accommodation website at <http://www.unt.edu/oda>. You may also contact them by phone at 940.565.4323.

Class Policy

- **Attendance Policy:** You are expected to attend class via Zoom. You are responsible for announcements and assignments given in class. If you miss a class, it is up to you to obtain notes and any other information that was provided in the class. Those who do not attend class or review the recorded lectures in a timely manner can count on doing poorly in this course.
- **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. [Insert specific sanction or academic penalty for specific academic integrity violation.
- **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 at disability.unt.edu.

- **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 Blackboard for contingency plans for covering course materials.
- **APA Style:** When doing the assignment, it is important to provide details of all the sources of information that you have used to prepare your work. All written assignments should follow APA (American Psychological Association) style to ensure that all sources are cited completely, correctly, and with consistency. The purpose of APA style is to (a). give credit to the author whose ideas or research you have used, (b). provide the exact location for sources of information used in the text of your paper, and (c). enable your reader to verify information you have provided or to explore your topic in greater depth. Consult the Publication Manual of the American Psychological Association, 6th edition.
- **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 Center for Student Rights and Responsibilities 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. The Code of Student Conduct can be found at deanofstudents.unt.edu/conduct.
- **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: eagleconnect.unt.edu/
- **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 at <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.

• **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.