Course Number: INFO 5810
Course Title: Data Analysis and Knowledge Discovery
Course Section(s): 001, 005

Huerta, Jose
SYLLABUS
Spring 2021

Table of Contents

COURSE INFORMATION
- Course Name: INFO 5810
- Course Title: Data Analysis and Knowledge Discovery
- Course Section(s): 001, 005
- Credit Hours: 3 Hours
- Class Meeting: Online
- Course Pre-requisites: None

Instructors Contact Information
- Name: Huerta, Jose
- Ph.D. Candidate
- UNT Email: Josehuerta@my.unt.edu
- Online Office Hours:
    - January 11 thru January 22nd: Tuesday and Thursday 5:30 PM to 6:30 PM
    - After January 22nd, Thursdays 5:30pm to 6:30pm or by appointment arranged email
    - Please note that office hours will close after the first 15 minutes if there are no shows. Office hours are meant to clarify questions you may have about assignments, quizzes, or projects and not intended for lecture.

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
This course will introduce the student to data analysis, data mining, text mining, and knowledge discovery principles, concepts, theories, and practices. It is designed for the aspiring or practicing information professionals and covers the basics of working with data from a hands-on and practical perspective. Classes will incorporate lecture, discussion, the practice of learned concepts, and readings. The student will learn how to approach data and data mining tasks and techniques using Microsoft Excel and Rapid Miner tool through practice exercises and instructor-led training. They will learn the basic principles and theories of data mining, and text mining techniques as well as the business application of data mining and knowledge discovery tools.

- *Course Goals:* The learner will be an active and engaged participant in discussion forums within the learning community by analyzing, constructing/creating, and evaluating information presented within the textbook, external readings/resources, student research, and class activities.
- The learner will be able to demonstrate an understanding of the fundamental principles, concepts, theories, and practices of data analysis, data mining, and knowledge discovery and discuss the interplay between them.
- The learner will apply gained knowledge to solve real problems with datasets provided using skills developed in the course. The learner will be able to construct an appropriate bibliography in APA6 from scholarly sources of material for study and research.

Materials
Required Textbooks

1. Winston, W. (2016). Microsoft Excel 2016 Data Analysis and Business Modeling, 5th Edition. Microsoft Press, ISBN 9781509304219.

2. North, M. (2018). Data Mining for The Masses. Third Edition: With Implementations in RapidMiner and R. ISBN-10: 0615684378, ISBN-13: 978-1727102475

Textbook materials: https://sites.google.com/site/dataminingforthemasses3e/
http://docs.rapidminer.com/downloads/DataMiningForTheMasses.pdf

Recommended for Excel beginners
    1. Etheridge, D. (2010). Excel Data Analysis: Your visual blueprint for creating and analyzing data, charts and Pivot tables. (3rd ed.). Indianapolis: Wiley Publishing, Inc.
        • Available at: http://iii.library.unt.edu/record=b4558404~S12

Additional study guide: LinkedIn Learning or Lynda.com
Lynda.com is an online course website where you can view expert lead tutorials on various topics. UNT has subscribed for this resource and it is available for currently enrolled students. We will be using these videos as additional help for some hands-on exercises. Please refer to this link to know more: https://it.unt.edu/lynda

TECHNICAL SUPPORT AND ASSISTANCE

The UNT University Information Technology provides student IT services and technical support, including Canvas.
    • UIT Homepage: https://it.unt.edu/uit
    • Email: helpdesk@unt.edu
    • Phone: 940.565.2324
    • In Person: Sage Hall, Room 330
Canvas technical requirements: https://clear.unt.edu/supported-technologies/canvas/requirements

Devices and Tools required:
-A personal computer, either Windows or Mac OS, capable of accessing Canvas. -**Microsoft Excel 2013 or above**.
-Check your computer's compatibility to RapidMiner software:
https://my.rapidminer.com/nexus/account/index.html#downloads

If you have Mac machine, you might run into issues with the software. You are encouraged to use the software by remote access to Lab B205 at discovery park.

Link to remote labs: https://itservices.cas.unt.edu/covid-19/labs

Minimum Technical Skills Needed
Students should be able to use the learning management system – Canvas to access course related materials and resource, keep up with emails regularly, create, modify or submit files according to instructors direction, such as proper file format, be able to download and install software when needed, and utilize the basics of the Microsoft Suite (Word, Excel, Power Point).

Success in the Online Course
*While the online classroom shares many similarities with the face-to-face classroom, success in online education requires certain skills and expectations that students should be aware of. Read more about "*How to Succeed as an Online Student.*"*
Student Academic Support Services
    • Code of Student Conduct: provides Code of Student Conduct along with other useful links

- **Office of Disability Access:** exists to prevent discrimination based on disability and to help students reach a higher level of independence
  - **Counseling and Testing Services:** provides counseling services to the UNT community, as well as testing services; such as admissions testing, computer-based testing, career testing, and other tests
- **UNT Libraries**
- **UNT Learning Center:** provides a variety of services, including tutoring, to enhance the student academic experience
- **UNT Writing Center:** offers free writing tutoring to all UNT students, undergraduate and graduate, including online tutoring
- **Succeed at UNT:** information regarding how to be a successful student at UNT

COURSE MODULES, TOPICS AND LECTURES

INFO 5810 is organized into 7 Learning Modules:
- Module 1: Data Analysis Introduction
- Module 2: Data Modeling Basics
- Module 3: Operating on Data
- Module 4: Knowledge Discovery Introduction
- Module 5: Data Mining Fundamentals
- Module 6: Visual Display of Data
- Module 7: Advanced Data Mining Techniques

ASSESSMENT & GRADING

Assignments and Assessments

Quizzes

For each learning module there will be a quiz with a combination of multiple-choice questions, one word answer and short answer questions.

Assignments for Learning Modules

The course is comprised of seven learning modules. Each learning module has an overview, assigned readings, and supplementary readings, and there are hands-on exercises in 5 of the 7 modules.

Midterm Assessment

The midterm assessment is sectioned into two parts. Part A is a take home quiz from modules 1 to 4. Part B is a video presentation not more than 10 minutes summarizing your data goals and analysis from assignments 1 to 3. Detailed instructions will be posted on Canvas under the title 'Midterm'.

The midterm assessment is a take home quiz from modules 1 to 4.

Data Analysis Project

The final project for the course will be a text mining assignment using the data mining software, Rapid Miner and data summarization using concepts learned in MS Excel or preferred data visualization tool.

The format and detailed instructions for the project will be provided prior to the distribution of the project.

Project idea and execution (100)
Project presentation (50)
Project report (50)

Grading

Grading Table

	Points Possible
Quizzes	30x6 + 20x1 =200
Assignments	100x5 = 500
Midterm Exam	100x1 =100
Data Analysis Project	200x1 =200
TOTAL	**1000**

Total Points Possible for Semester/Grading Scale = 1000

1000-900 = A
899-800 = B
799-700 = C
699-600 = D
599-500 = F

CALENDAR

Table 1. Study Schedule

Module	Topics	Readings
1	**Data Analysis Introduction** Principles, Concepts, and Practices of Data Analysis **Hands On**: Excel introduction, Internet Resources, Library Resources, Entering data in a spreadsheet and creating charts and graphs.	Readings for Module 1 are located in Course Content, Module 1. They cover a variety of introductory concepts including formatting data for analysis, data types and display types, visual considerations etc. Book Chapters: Text 1: 1 – Basic Spreadsheet Modeling 2 – Range Names 3 – Lookup Functions 4 – Index Function 5 - Match Function 14 – The Paste Special Command, 25 – Sorting in Excel 26 - Tables
2	**Data Modeling Basics** Database and Data warehouse concepts **Hands-on**: Getting data into Excel and working with Excel. Installing and exploring RapidMiner.	Book Chapters: Text 1: 43 – Using Pivot Tables and slicers to describe data 44 – The Data Model 49 – Filtering data and removing duplicates 51 – Creating subtotals 88 – Array Formulas and Functions Reading Topics: Conceptual modeling of data, sets, tables, spreadsheets, and cubes
3	**Operating on Data** Functions	Book Chapters: Text 1: 6 – Text Functions, 7 – Date and Time Functions, 13 – Time & time functions 41 – Summarizing Data by Using Histograms and Pareto Charts, Text 2: Chapter 3: Data Preparations Reading Topics: Functions and formulas

4	**Knowledge Discovery Introduction** Topic: Principles, concepts, theories and practices of knowledge discovery	Book Chapters: Text 1: 38 – Importing Data from a Text File or Document, 40 – Validating Data Reading Excel Macros – Materials on Canvas
	Hands-on Topics: No topics.	Text 2: Chapter 4: Correlation Methods Chapter 5: Association Rules
5	**Data Mining Fundamentals** and basics of working with multi- dimensional data Topic: Semantic distance, Similarity, Classification, Clustering, Association, Regression, Online analytical processing - OLAP, ROLAP, MOLAP, and Cubes	Book Chapters: Text 1: None Excel Macros – Materials on Canvas Text 2: Chapter 6: K-Means Clustering Chapter 7: Discriminant Analysis, k-nearest neighbors and Naïve Bayes
6	**Visual Display of Data** Topic: Illustration, graphics, quantitative displays Hands-on Topics: Textual Analysis with RapidMiner.	Book Chapters: Text 1: 24 – Conditional Formatting, 47 – Sparklines, 52 -Charting Tricks Text 2: Chapter 12: Text Mining
7	**Advanced data and data mining techniques** Topic: Rule induction (if then else), Neural networks (trainable nets for decision making/prediction), Decision trees (classification), Genetic algorithms Hands-on Topics: No topic.	Book Chapters: Text 1: 28- The Analytics Revolution, Introduction to Excel Dashboards - Materials on Canvas Text 2: Chapter 10: Decision Trees Chapter 11: Neural Networks Reading Topics: Types and examples of advanced data analysis

Table 2. Study Schedule and Due Dates

Academic Week	Dates	Study Focus	Assignment/Project/Quiz	Due Dates
1	Jan 11 – Jan 17	Syllabus Module 1	Self-Introduction (not graded) Quiz 1	Jan 17 23:59
2	Jan 18 – Jan 24	Module 1	Assignment 1	Jan 24 23:59
3	Jan 25 – Jan 31	Module 2	Quiz 2	Jan 31 23:59
4	Feb 1 – Feb 7	Module 2	Assignment 2	Feb 7 23:59
5	Feb 8 – Feb 14	Module 3	Quiz 3	Feb 14 23:59
6	Feb 15 – Feb 21	Module 3	Assignment 3	Feb 21 23:59
7	Feb 22 – Feb 28	Module 4	Quiz 4	Feb 28 23:59
8	Mar 1 – Mar 7	Mid Review	Midterm Assessment/ Midterm Quiz	Mar 7 23:59

9	Mar 8 – Mar 14	Module 5	Quiz 5	Mar 14 23:59
10	Mar 15 – Mar 21	Module 5	Assignment 4	Mar 21 23:59
11	Mar 22 – Mar 28	Module 6	Quiz 6	Mar 28 23:59
12	Mar 29 – Apr 4	Module 6	Assignment 5	Apr 4 23:59
13	Apr 5 – Apr 11	Module 7	Quiz 7	Apr 11 23:59
14	Apr 12 – Apr 18	Term Project	Work on Term Project	
15	Apr 19 – Apr 25	Term Project	Work on Term Project	-
16	Apr 26 – May 1	Term Project	Term Project Due !!! ->->->->->	Apr 27 23:59

Final Examination:
There is no final exam in this course.

COURSE EVALUATION

Student Perceptions of Teaching (SPOT) is the student evaluation system for UNT, implemented in summer 2015, to comply with the State of Texas House Bill 2504. Each semester student will receive an email from UNT to their student UNT email to response to the anonymous survey. For more information, visit Office of the Provost, SPOT page: https://vpaa.unt.edu/spot

COURSE POLICIES

Assignment Policy

This course uses Turnitin for assignment submissions. 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. The instructor and the UNT Student Help Desk will work with the student to resolve any issues at the earliest possible time.

Examination Policy
The quizzes and assignments are open-book.

Instructor Responsibilities and Feedback
The instructor will post feedback on quizzes and assignments after 48 hour of the deadlines of the assignment.

Late Work and Missed Work
Students are expected to submit assignments and projects on time. **The due dates for quizzes and assignments are Sundays 11:59 pm unless mentioned otherwise.** 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 as soon as possible before the due date. **Late work without the permission of the instructor will receive a grade with a 10% penalty for graduates (or 10 points out of 100) per day after the due date. However, the late submission policy may apply on the prior circumstance.** A student who is having a trouble with the assignments is strongly encouraged to contact the instructor as early as possible for personal advising.

Course Incomplete Grade
The UNT Graduate Catalog (https://registrar.unt.edu/grades/incompletes) describes and explains grading policies. A grade of Incomplete (I) will be given only for a justifiable reason and only if the student is passing the course. The student is responsible for meeting with the instructor to request an incomplete and discuss requirements for completing the course. If an incomplete is not removed within the time frame agreed to by instructor and student, the instructor may assign a grade of F.

Attendance Policy
You are required to attend any online meetings when necessary.

Class Participation
Students are expected to actively participate in the online discussions and check the course content on canvas.

Students' Responsibility for Their Learning
To achieve the above learning objectives, students are expected to study 9 - 12 hours per week for this course.

Syllabus Change Policy
Any changes to the syllabus will be notified to the students via announcements on Canvas and/or emails from the instructor.

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 at disability.unt.edu.

Course Safety Procedures (for Laboratory Courses)
Students enrolled in [insert class name] are required to use proper safety procedures and guidelines as outlined in UNT Policy 06.038 Safety in Instructional Activities. While working in laboratory sessions, students are expected and required to identify and use proper safety guidelines in all activities requiring lifting, climbing, walking on slippery surfaces, using equipment and tools, handling chemical solutions and hot and cold products. Students should be aware that the UNT is not liable for injuries incurred while students are participating in class activities. All students are encouraged to secure adequate insurance coverage in the event of accidental injury. Students who do not have insurance coverage should consider Standard Syllabus Statements Related Policy 06.049 Course Syllabi Requirements obtaining Student Health Insurance. Brochures for student insurance are available in the UNT Student Health and Wellness Center. Students who are injured during class activities may seek medical attention at the Student Health and Wellness Center at rates that are reduced compared to other medical facilities. If students have an insurance plan other than Student Health Insurance at UNT, they should be sure that the plan covers treatment at this facility. If students choose not to go to the UNT Student Health and Wellness Center, they may be transported to an emergency room at a local hospital. Students are responsible for expenses incurred there.

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.

Optional UNT Policies Statements

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 Blackboard 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. 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. 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.

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 at 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.