COURSE TITLE/SECTION: Data Science I (COSC 3337)   *November 21, 2023*

TIME: TT 11:30a-1p in S105

FACULTY: Christoph F. Eick   **OFFICE HOURS**: TU 4:10-5p TH 8:50-10a
E-mail: ceick@uh.edu   **Phone: 33345 (use e-mail!!)**   **FAX: 33335**

I. Course *Data Science I (COSC 3337)*
Credit Hours: 3.0
Lecture Contact Hours: 3   *Lab Contact Hours:* 0

Formerly COSC 4335.
Prerequisite: A grade of C- or better in COSC 2306 or COSC 2436, and MATH 3339 and declared COSC major, COSC minor, or Data Science minor.
Description
Data science concepts including exploratory data analysis, data visualization, statistical inference and modeling, machine learning, clustering, post-processing and interpreting results.

II.   Course Objectives

Upon completion of this course, students
  a. will know what the goals and objectives of data science are and how to conduct a data science project
  b. will have a sound knowledge of basic statistics and basic machine learning concepts
  c. will have some sound knowledge about exploratory data analysis and data visualization techniques used in exploratory data analysis
  d. will have knowledge of popular classification techniques, such as decision trees, support vector machines, and neural networks
  e. will have some sound knowledge about how to construct distance functions
  f. will have detailed knowledge of popular clustering algorithms, such as K-means, DBSCAN, EM and hierarchical clustering and cluster evaluation.
  g. will get some exposure to deep learning centering on autoencoders
  h. will get hands-on exposure in the course problem sets and group project how to apply data analysis techniques to real world data sets. They will obtain valuable experience in learning how to interpret data analysis results, how to select parameters of data analysis tools, and data storytelling
  i. will learn on how to use the data analysis and visualization environments, such as **R** and its popular libraries and how to develop software on the top of R.
  j. will have some basic knowledge concerning outlier

detection and association analysis.

Course Webpage: [COSC 3337---Data Science I (uh.edu)](uh.edu)
MS Team: 3337-Class

III.   Course Content

1. Introduction to Data Analysis
2. Exploratory Data Analysis—how to Visualize and Compute Basic Statistics for Datasets and How to Interpret the Obtained Findings
3. How to Interpret the Findings Introduction to R/Python for Data Science
4. Density Estimation
5. Introduction to Supervised Learning
6. Introduction to Deep Learning centering on Autoencoder
7. Similarity Assessment—how to Obtain Distance Functions
8. Introduction to Clustering
9. Anomaly and Outlier Detection.
10. Data Storytelling
11. Data Preprocessing
12. Association Analysis

There will be 3 Problem Sets each consisting of individual tasks, centering on:
Problem Set1: Exploratory Data Analysis and Classification
Problem Set2: Clustering and Outlier Detection
Problem Set3: Neural Networks and Autoencoders

There will be a 6-week long group project Sept. 28-November 11, 2023 centering on analyzing, mapping, change analysis and hotspot discovery of solar flares.

IV.   Course Structure and Evaluation

22 lectures
2-3 labs
3 exams (October 3, November 14, December 7 2023)
3 problem sets
1 group project
1 group homework credit task[1]

1-2 presentation (maybe group project and as part of group homework credit)

The tentative weights of the different course elements are as follows:

---

1 Groups of students present solutions to homework style tasks or demo a tool during the lecture.

Problem Sets (3), Group Project and Group Homework Credit and Attendance: 51%;
Exams (3): 49% (14%, 15%, 20%)
Tentative weights of non-exam tasks: Problem Sets: 30%, Group Project: 15%, Group Homework Credit: 3% ; Attendance: 3%

MS Team 3337-Class will be used to support teaching the course; it also serves as a discussion forum. Course material can be downloaded from the course webpage,

V.   Textbooks

Recommended Text:
P.-N. Tang, M. Steinback, and V. Kumar *Introduction to Data Mining*, Addison Wesley, 2018.

Maybe:
Cathy O'Neil and Rachel Schutt. Doing Data Science, Straight Talk from the Frontline. O' Reilly. 2014

VII.   Evaluation and Grading

Translation number to letter grades:
A:100-91 A-:91-87 B+:87-84 B:84-80 B-:80-76 C+:76-71
C: 71-66 C-:66-62 D+:62-58 D:58-54 D-:54-50 F: 50-0

Students may discuss course material and assignments but must take special care to discern the difference between **collaborating** in order to increase understanding of course materials and collaborating on the homework / course project itself. We encourage students to help each other understand course material to clarify the meaning of homework problems or to discuss problem-solving strategies, but it is **not** permissible for one student to help or be helped by another student in working through homework problems and in the course project. If, in discussing course materials and problems, students believe that their like-mindedness from such discussions could be construed as collaboration on their assignments, students must cite each other, briefly explaining the extent of their collaboration. Any assistance that is not given proper citation may be considered a violation of the Honor Code, and might result in obtaining a grade of F in the course, and in further prosecution.

COSC 3337 exams will be open book/notes paper exams; however, the use of cellphones and computers during exams is not allowed.

---

2 *Online Credit*: Different Homework-style problems will be assigned to groups of about 6 students (each group gets a different task) and each group presents their solutions in a presentation during the lecture and shares them in the respective channel of this course's MS Teams page.

VIII. Course Exams

Course exams will be paper exams which you take in a UH classroom, likely our assigned classroom. If you are experiencing any COVID-19 symptoms that are not clearly related to a pre-existing medical condition, do not come to the exam. Please see Student Protocols for what to do if you experience symptoms and Potential Exposure to Coronavirus for what to do if you have potentially been exposed to COVID-19. Consult the (Graduate Excused Absence Policy) for information regarding excused absences due to medical reasons.

Moreover, if you miss a course exam for reasons that are not covered by the Undergraduate Excused Absence Policy you will get a grade of 'F' for the missed exam.

IX. Mental Health and Wellness Resources

The University of Houston has a number of resources to support students' mental health and overall wellness, including CoogsCARE and the UH Go App. UH Counseling and Psychological Services (CAPS) offers 24/7 mental health support for all students, addressing various concerns like stress, college adjustment and sadness. CAPS provides individual and couples counseling, group therapy, workshops and connections to other support services on and off-campus. For assistance visit uh.edu/caps, call 713-743-5454, or visit a Let's Talk location in-person or virtually. Let's Talk are daily, informal confidential consultations with CAPS therapists where no appointment or paperwork is needed. The Student Health Center offers a Psychiatry Clinic for enrolled UH students. Call 713-743-5149 during clinic hours, Monday through Friday 8 a.m. - 4:30 p.m. to schedule an appointment. The A.D. Bruce Religion Center offers spiritual support and a variety of programs centered on well-being.

Need Support Now?

If you or someone you know is struggling or in crisis, help is available. Call CAPS crisis support 24/7 at 713-743-5454, or the National Suicide and Crisis Lifeline: call or text 988, or chat 988lifeline.org.

X. Title IX/Sexual Misconduct

Per the UHS Sexual Misconduct Policy, your instructor is a "responsible employee" for reporting purposes under Title IX regulations and state law and must report incidents of sexual misconduct (sexual harassment, non-consensual sexual contact, sexual assault, sexual exploitation, sexual intimidation, intimate partner violence, or stalking) about which they become aware to the Title IX office. Please know there are places on campus where you can make a report in confidence. You can find more information about resources on the Title IX website at https://uh.edu/equal-opportunity/title-ix-sexual-misconduct/resources/.

XI. Reasonable Academic Adjustments/Auxiliary Aids
The University of Houston complies with Section 504 of the Rehabilitation Act of 1973 and the Americans with Disabilities Act of 1990, pertaining to the provision of reasonable academic adjustments/auxiliary aids for disabled students. In accordance with Section 504 and ADA guidelines, UH strives to provide reasonable academic adjustments/auxiliary aids to students who request

and require them. If you believe that you have a disability requiring an academic adjustments/auxiliary aid, please contact the Justin Dart Jr. Student Accessibility Center (formerly the Justin Dart, Jr. Center for Students with Disabilities).

XII. Excused Absence Policy

Regular class attendance, participation, and engagement in coursework are important contributors to student success. Absences may be excused as provided in the University of Houston Undergraduate Excused Absence Policy and Graduate Excused Absence Policy for reasons including: medical illness of student or close relative, death of a close family member, legal or government proceeding that a student is obligated to attend, recognized professional and educational activities where the student is presenting, and University-sponsored activity or athletic competition. Under these policies, students with excused absences will be provided with an opportunity to make up any quiz, exam or other work that contributes to the course grade or a satisfactory alternative. Please read the full policy for details regarding reasons for excused absences, the approval process, and extended absences. Additional policies address absences related to military service, religious holy days, pregnancy and related conditions, and disability.

XIII. Recording of Class

Students may not record all or parts of class, livestream all or part of class, or make/distribute screen captures, without advanced written consent of the instructor. If you have or think you may have a disability such that you need to record class-related activities, please contact the Justin Dart, Jr. Student Accessibility Center. If you have an accommodation to record class-related activities, those recordings may not be shared with any other student, whether in this course or not, or with any other person or on any other platform. Classes may be recorded by the instructor. Students may use instructor's recordings for their own studying and notetaking. Instructor's recordings are not authorized to be shared with *anyone* without the prior written approval of the instructor. Failure to comply with requirements regarding recordings will result in a disciplinary referral to the Dean of Students Office and may result in disciplinary action.

XIV. Syllabus Changes

Please note that the instructor may need to make modifications to the course syllabus and may do so at any time. Notice of such changes will be announced as quickly as possible through (*specify how students will be notified of changes*).

XV. Resources for Online Learning

The University of Houston is committed to student success, and provides information to optimize the online learning experience through our Power-On website. Please visit this website for a comprehensive set of resources, tools, and tips including: obtaining access to the internet, AccessUH, and Blackboard; requesting a laptop through the Laptop Loaner Program; using your smartphone as a webcam; and x If you are experiencing any COVID-19 symptoms that are not clearly related to a pre-existing medical condition, do not come to class. Please see Student Protocols for what to do if you experience symptoms and Potential Exposure to Coronavirus for what to do if you have potentially been exposed to COVID-19. Consult the (select: Undergraduate Excused Absence Policy or Graduate Excused Absence Policy) for information regarding excused absences due to medical reasons.

XVI. UH Email

Please check and use your Cougarnet email for communications related to this course. You will also need your Cougarnet account to access course MS Teams page 3337-Class and for listening to course lecture. To access this email, login to your Microsoft 365 account with your Cougarnet credentials. Your Cougarnet Account also allows you to download Microsoft Office 365 at no cost. For questions or assistance contact UHOnline@uh.edu.

XVII. Academic Honesty Policy

High ethical standards are critical to the integrity of any institution, and bear directly on the ultimate value of conferred degrees. All UH community members are expected to contribute to an atmosphere of the highest possible ethical standards. Maintaining such an atmosphere requires that any instances of academic dishonesty be recognized and addressed. The UH Academic Honesty Policy is designed to handle those instances with fairness to all parties involved: the students, the instructors, and the University itself. All students and faculty of the University of Houston are responsible for being familiar with this policy.

XVIII. Security Escorts and Cougar Ride

UHPD continually works with the University community to make the campus a safe place to learn, work, and live. Our Security escort service is designed for the community members who have safety concerns and would like to have a Security Officer walk with them, for their safety, as they make their way across campus. Based on availability either a UHPD Security Officer or Police Officer will escort students, faculty, and staff to locations beginning and ending on campus. If you feel that you need a Security Officer to walk with you for your safety please call 713-743-3333. Arrangements may be made for special needs.

Parking and Transportation Services also offers a late-night, on-demand shuttle service called Cougar Ride that provides rides to and from all on-campus shuttle stops, as well as the MD Anderson Library, Cougar Village/Moody Towers and the UH Technology Bridge. Rides can be requested through the UH Go app. Days and hours of operation can be found at https://uh.edu/af-university-services/parking/cougar-ride/.

XIX. Use of AI related and Other Tools

The use of AI-related tools, such ChatGPT, is permitted in this course. However, student reports have to cite all AI-related tools and other tools and documents that were used to conduct the task at hand and describe for what they were used. Not citing used tools and documents represents a serious academic honesty violation,

and violators will be prosecuted.

XX Other Helpful Links

Coogs Care: https://uh.edu/dsa/coogscare/

Student Health Center: https://www.uh.edu/healthcenter/