CS 105, "Data Analysis Methods"
Fall'25

Syllabus

Course Description: An introduction to fundamental concepts and methods in data analysis and visualization essential to a variety of data science tasks. Designed to provide preparation for the data science major and for advanced courses in data analysis and applications of data science.

Prerequisites: CS009B CS010B (or equivalent). The prerequisites are strictly enforced.

Schedule:

Lecture:

Tuesday, Thursday 5:00 PM - 6:20 PM

Instructor: Elena Strzheletska, office: WCH 135, **elenas@cs.ucr.edu (mailto:elenas@cs.ucr.edu)** .

Office hours:

- Tuesday: 12:30 PM - 1:30 PM
- Thursday: 12:30 PM - 1:30 PM
- or by appointment

Lab:

Wednesday 05:00PM - 06:50PM, (sec. 21), WCH 143.

Wednesday 07:00PM - 08:50PM, (sec. 22), WCH 143.

Teaching Assistant: Nghia Nguyen, office: WCH 461, **nnguy498@ucr.edu. (mailto:nnguy498@ucr.edu) (mailto:nnguy498@ucr.edu)**

Office hours:

- Monday: 4 PM - 5 PM
- Friday: 4 PM - 5 PM

Readers:

- , Calendly:
- , Calendly:

Recommended textbooks:

- *Data Science from Scratch, J, Grus*
- *Introduction to Machine Learning with Python*, A. C. Müller, S. Guido

- *An Introduction to Statistical Learning: with Applications in Python*, G. James, D. Witten

- *Data Mining:* Concepts and Techniques, J. Han, M. Kamber.

- *Applied Text Analysis with Python*, Benjamin Bengfort

Assignments and tests:

Labs: Lab attendance is required; six - eight lab assignments; can be completed individually or in groups of two (strongly recommended). Each group submits one assignment. Both students will receive the same credit (unless requested otherwise). Each lab must be demoed. If a student (or a group) fails to submit or demo the assignment, he/she receives a "0".

Quizzes: Five-seven quizzes (some of the quizzes will be take-home).

Reflective assignments: Four -five take-home assignments.

Midterms: Two (in-person) midterm exams:

        Midterm 1:   (CS Test Center)

        Midterm 2: (CS Test Center)

Final project and presentation: see Project Description

Attendance:

Regular attendance at lectures is strongly advised. Some of the presented material may not be covered in the posted lecture notes. Students are also strongly encouraged to take advantage of the office hours.

Grading:

Labs: 20% (10% + 10%),

Quizzes: 15%,

Midterms: 40%,

Reflective assignments: 5%,

Final project/presentation: 20%.

Extra credit: 1%.

Course grades are expected to be determined as follows: A (A-, A, A+) = 90-100%, B = 80-89%, C = 70-79%, D = 60-69%. Minor adjustments of this scale can be made at the end of the quarter.

Extra credit: Small extra credit assignments will be given during lectures and labs. You can complete all of these assignments and apply up to 85% of the maximum available number of points towards your final grade (1%).

Academic Integrity: Zero-tolerance policy on plagiarism is enforced. Cheating on any assignments or tests will result in an F grade for the course and a disciplinary action, independently of the extent of plagiarism.

Copyright: See UC Copyright Policies.

Additional Class Resources:

1. **Slack** (for communication): **https://join.slack.com/t/cs105dataanal-ryl3619/shared_invite/zt-3e6teeuhf-fGUPlCGytR2OhiyEZ5e5hQ** ➡️ **(https://join.slack.com/t/cs105dataanal-ryl3619/shared_invite/zt-3e6teeuhf-fGUPlCGytR2OhiyEZ5e5hQ)**

2. **Gradescope** (for assignment submission): If you don't see CS 105 among your courses in Gradescope, please sign in with your *First name, Last name* - exactly as in Canvas. Code (if needed): VWNYEZ (Gradescope.com).

3. **Google Drive folder** (for surveys and projects): **https://drive.google.com/drive/folders/1RwJVwAmA8b1Oh8wcGzVF0XymzaJ7dUXP?usp=sharing** ➡️ **(https://drive.google.com/drive/folders/1RwJVwAmA8b1Oh8wcGzVF0XymzaJ7dUXP?usp=sharing)**

Helpful UCR resources:

https://registrar.ucr.edu/ (https://sdrc.ucr.edu/ucr-resources-students)

https://sdrc.ucr.edu/ucr-resources-students (https://sdrc.ucr.edu/ucr-resources-students)

https://counseling.ucr.edu/ (https://counseling.ucr.edu/)

Tentative Schedule

Week 0	Thursday, September 25	

Week 1	Tuesday, September 30 Thursday, October 2	Introduction Statistics for DS
Week 2	Quiz 1 (Statistics) - Monday Tuesday, October 7 Thursday, October 9 Reflection 1 is due Thursday	Probability for DS Data Collection and Preprocessing
Week 3	Quiz 2 (Probability) - Monday Tuesday, October 14 Thursday, October 16	Mini-Project introduction Data Visualization
Week 4	**Midterm 1** Tuesday, October 21 Thursday, October 23 Reflection 2 is due Thursday	Brief review for Midterm 1 Mathematical Models in DS Loss Functions
Week 5	Tuesday, October 28 Thursday, October 30	Gradient Descent Regression Analysis Classification (KNN) KNN Regression
Week 6	Tuesday, November 4 Thursday, November 6 Reflection 3 is due Thursday	KNN Unsupervised Learning
Week 7	Tuesday, November 11 Thursday, November 13	Unsupervised Learning
Week 8	Tuesday, November 18 Thursday, November 20	Text mining
Week 9	Tuesday, November 25 Thursday, November 27 **Thanksgiving Day - no class**	Recommender Systems Review
Week 10	**Midterm 2** Tuesday, December 2 Thursday, December 4	Midterm 2

Other Books

- *Learning Python,* 5th edition, Mark Lutz
- *Data Mining and Analysis:* Fundamental Concepts and Algorithms Mohammed J. Zaki and Wagner Meira Jr.
- *The Wall Street Journal Guide to Information Graphics: The Dos and Dont's of Presenting Data, Facts, and Figures,* Dona M. Wong
- *Storytelling with data : a data visualization guide for business professionals,* Cole Nussbaumer Knaflic
- *Practical statistics for data scientists : 50 essential concepts,* Peter C. Bruce, Andrew Bruce, Peter Gedeck
- *Hands-on exploratory data analysis with python: perform EDA techniques to understand, summarize, and investigate your data,* S.K. Mukhiya
- *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking,* Foster Provost, Tom Fawcet
- *Text Mining: Predictive Methods for Analyzing Unstructured Information,* S.M. Weiss, N. Indurkhya, T. Zhang, F. Damerau.
- *Numsense! Data Science for the Layman: No Math Added,* Annalyn Ng, Kenneth Soo
- *Doing Data Science: Straight Talk from the Frontline,* Cathy O'Neil, Rachel Schutt
- *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results,* Bernard Marr