

CS 210: Introduction to Data Science

Syllabus, Fall 2020

Data science topics span a large variety of disciplines and require a collection of skills. This course is intended to cover data science's fundamental principles and techniques, emphasizing data-centric quantitative thinking. We will tour the basic data science techniques from manipulation and summarizing the essential characteristics of a data set, basic statistical modeling, visualization, and prediction.

Prerequisites: IF 100 and MATH 203

Contact Information

Instructor: Oznur Tastan otastan@sabanciuniv.edu
Instructor Office Hours: By appointment.

TAs: Hasan Alp Boz bozhasan@sabanciuniv.edu
Mert Gurkan mertgurkan@sabanciuniv.edu
Busra Oz boz@sabanciuniv.edu
Rana Bengisu Kalkan ranakalkan@sabanciuniv.edu

TA Office Hours: Will be announced.

Course Webpage

We will be using SuCourse+. Please regularly check the SuCourse+ of the course for lecture notes, homework assignments, project information, discussions and announcements.

Textbook: No dedicated textbooks. There will be required readings, videos put on SuCourse+.

Learning outcomes

- Learning the fundamentals of data science pipeline
- Learning how to explore and experiment with data
- Learn basic statistics (sampling techniques, mean, variance, outliers, distributions) and machine learning techniques (clustering) that are necessary to analyze data: big and small
- Perform statistical analysis on sample socio-economic data
- Building an understanding of data analytics techniques (data collection, cleaning, exploratory techniques, modeling, and presentation) Develop competency in the Python programming language within the course projects and assignments
- Design and run experimental tests to evaluate hypotheses about data

Delivery format We will have synchronous lectures and recitations on Zoom. Your active participation is expected.

Disclaimer

- Students who are registered to the course with time-conflict override accept the responsibility of any inconvenience that might occur due to missed content and/or quizzes. No make-up will be available for missed quizzes/content.
- This syllabus and course details might need to be updated throughout the semester because of the uncertainties the pandemic brings. Any modification will be announced at SUCourse and also during the class. Students are responsible from following the announcements.

Grading

The final grades will be based on the following, and is subject to change if necessary due to COVID-19 related regulations or problems arise due:

- In-class quizzes, a mixture of pop up and non-pop up (25%).
- One final exam (20%).
- Homework assignments (30%).
- A term project (25%).
- There may be oral exams to ensure a fair assessment in homework, quizzes and the exams. The student will be called for an exam in short notice. The instructor might ask questions from a broader range of topics to evaluate student's understanding of the course material.

IMPORTANT: One of the following conditions will result in an automatic F regardless of other grades:

1. Average of the homework assignment is below 30.
2. Not submitting a project report.
3. Being absent in a project demo without a medical report.
4. Missing the final exam without a medical report.

*** Not falling in one of the conditions does not guarantee passing the course. If your overall performance is poor, you might fail the course.

Homework submissions: Both written and programming. You are expected to program in Python. We will provide submission details.

Homework late day policy (IMPORTANT: Each student will have a total of 4 free late (calendar) days to use for homework assignments. You do not need to explain why you are submitting late and no need to notify us. ≤ 24 hours late counts as 1 day late, etc. Once these total of four late days are exhausted, any assignments turned in late will be penalized and will incur a reduction of 33% in the final score for each day (or **part thereof**) it is late. For example, if an assignment is up to < 24 hours late, it incurs a penalty of 33%. Else if it is up to more than 24 hours and less than 48 hours late, it incurs a penalty of 66%. And if it is 72 or more hours late, it will receive no credit.

Regrade policy: You may object a grade within 14 days after the grades are announced. If you feel that an error was made in grading, please get an appointment to discuss it. Please note your grade may go up or down based on the evaluation.

Honor code: This course follows the Sabanci University Disciplinary Rules and Regulation. Violations of the rules will not be tolerated.

Project

The purpose of the project is to increase your knowledge about data science and get hands-on practical experience. The grade for the project will include a peer grade. We will provide more details on the project content, deliverables and the key dates in the upcoming week.