ENCE 3503: Data & Information Engineering

Spring 2025

Table of Contents

Table of Contents
Contacts
Course Description
Grading
Attendance
Participation
Homework
Team Project
Deadline Policy
Examination
Schedule
Appendix I: Communication Rules
Appendix II: Policies and Resources
Revisions

Contacts

Instructor(s)[1]		
Name	**Email**	**Office**
Mr. Ismayil Shahaliyev	ishahaliyev@ada.edu.az	B320

Academic Advisors[2]		
Name	**Email**	**Program**
Ms. Firuza Mammadzada	fmammadzada@ada.edu.az	BSCS
Ms. Gunay Aghamaliyeva	gaghamaliyeva@ada.edu.az	BSCE & BSEEE
Ms. Kamala Hashimova	khashimova@ada.edu.az	BSIT & BSMATH

Course Description

This course offers an introduction to data science, focusing on fundamental concepts and practical skills in Python programming, data analysis, and machine learning. The course emphasizes practical data manipulation and visualization techniques using Python, and introduces machine learning concepts including data preprocessing, supervised and unsupervised learning, and model evaluation. Students will also gain experience in SQL for data management and AutoML tools like Dataiku.

Learning Objectives

Students are expected to understand, know and/or be capable of doing/applying/coding the following after this course:

- Fundamental data science/machine learning techniques and concepts widely used in the industry.
- Essential statistical methods for data analysis.
- Visualization of different types of data samples using fundamental Python libraries.
- Basic SQL for data management and AutoML tools like Dataiku.

Prerequisite(s)

- Desire to learn

Suggested Reading

- *Python for Data Analysis. Data Wrangling with Pandas, NumPy, and IPython* (O'Reilly) by *William McKinney*, Online Open Access HTML edition (the latest edition).
- Other readings/resources may be shared throughout the semester.

[1] Read: Communication Rules
[2] Office: B building 1st floor

Technological Requirements

You should arrive at each class with your **laptop** and always be ready to solve a given task. Python (3.9 or later) and git version control system should be installed on your computer. You should have a Google (for Google Colab), Github and CodeAcademy accounts. **Your CodeAcademy profile should be public**. You should have a prior understanding of git and Github (please see any youtube tutorial). In addition to Google Colab, if you want to run code locally on your device, you will need to install Jupyter.

Grading

Item	**100%**
Attendance	5
Participation	10
Homework	10
Team Project	25
Midterm Exam	25
Final Exam	25

You shouldn't disrupt learning and teaching by asking for a grade increase. Any attempt by a student on **grade bargaining** may be considered a **violation of the Student Code of Conduct** and reported by faculty. If you have legitimate reasons, you may appeal your grade. For that, you should consult with your Academic Advisor for the cases eligible for **grade appeal**.

Attendance

Attendance will be checked before the start of each class. There is no late attendance mark. In case of **illness or other emergency**, students are expected to submit **documented proof** to their **advisors.** Advisors then will forward the document to the students' relevant instructors, who will decide whether to accept the absence as excused or not.

Participation

Students are expected to read the weekly material and **be prepared for practical sessions**. Students should bring their **laptops** to each class.

During class time, random questions will be asked, and a task covering that week's **learning outcomes** will be assigned to a randomly chosen student. If a student demonstrates certain knowledge when solving the problem, they will receive 1 (otherwise 0) for participation. If a student is called and is absent, the grade is 0. For improper behavior in the class, a student may get -1.

When solving practical problems, students should share their screens via blackboard. Students are expected to have a **backup plan**, solve laptop related issues themselves, and always **be ready to code** (e.g., they should not spend time opening an IDE when called).

Students will have limited answer attempts. Their in-class participation grade will be calculated as the ratio of $\frac{points\_received}{total\_attempts}$. For example, in case a student is called 5 times during the semester, then grading *1, 1, 0, A (absent), -1* means that the student gets 20% of the total in-class participation grade.

$$\frac{1+1+0+0-1}{5} \cdot 100\% = 20\%$$

In case of difficulties in communication and personal issues, you are encouraged to consult your academic advisor and seek psychological support. For that, please refer to Contacts and Policies and Resources sections.

Homework

CodeAcademy courses will be supplementary to your education and will help you to understand relevant topics step by step. You are required to pass only **free** lessons, modules, etc. If you are retaking the course, you are expected to reset your progress and redo the tasks. You may be assigned non-CodeAcademy based homework if instructors find it necessary.

The specific details about the homework, as well as the deadline for completion, will be shared with you throughout the semester. Your total homework grade will be calculated based on your average performance. For example, if there were five homeworks during the semester (there can be more or less), and you completed four of them fully and 30% of the last one, then you will get 86% of the total homework grade.

$$\frac{100+100+100+100+30}{5}\% = 86\%$$

Team Project

The goal of the team project is to test the practical skills you have acquired during the semester. You will form a group of 3-4 students. You will make your submissions to a Github Classroom repository which will be shared with you. Further details will be communicated to you during the semester.

Deadline Policy

Deadlines will never be extended. When submitting, a student must consider possible internet connection issues, uploaded file size, etc. Late submissions of one minute until thirty minutes will receive a 10% penalty, late submissions of thirty minutes until one day

will receive a 25% penalty with no claim for bonuses (if there are any). For each next day a further 25% penalty will be applied.

Examination

You will have two exams: midterm and final. Both exams will be computer-based (unless noted otherwise) with questions of various types (e.g. multiple-choice, fill-in-the-blank, code/essay). More specific rules will be communicated to you before the exam.

You are expected to be aware of the Exam Rules and Regulations. If you have special needs or health issues, you are strongly recommended to contact the University's Student Academic Support Services well ahead of the examination date. It is your responsibility to manage conflicts in your schedule and notify your instructors about it **at least two weeks before** the examination date. Your instructors may accept or disregard your request.

Schedule

Disclaimer: The schedule is subject to change. Please consistently check the ADA Academic Calendar for modifications as well. The last day of the **add/drop** period, holidays, etc. are noted in the calendar.

Lecture	Theme	Learning Outcomes	Assessments
1	Introduction to Data and Information Engineering Course	What we will learn during the semester. Syllabus.	Fill pre-semester survey and Github/CodeAcademy accounts forms (if announced). Follow Technological Requirements. Read Syllabus.
2	Introduction to Statistical Ideas and Methods	Concepts: data, information, knowledge. Summarizing Data: One Variable, Five Number Summary, Spread of Data, Shape of the Data	
3	Introduction to Statistical Ideas and Methods	Summarizing Data: Relationships Between Variables, Sampling, Biases, Data collection: Anecdotes, Observational studies, experiments	
4	Python Basics	Hello world, data types, conditional statements	
5	Python Basics	Loops, functions, classes	

6	Python Data Structures	List, tuple, dictionary, set	
7	Python Data Structures	List, tuple, dictionary, set	
8	Array Operations, Matrices	Python arrays, creating arrays, accessing elements, operations, indexing and slicing	
9	Array Operations, Matrices	Numpy arrays: Why numpy?, numpy ndarray, array creation, numpy data types, operations, indexing and slicing	
10	Data Manipulation	pandas: Creating dataframe, importing data, basics, modify/add columns, deleting columns, renaming columns, groupby	
11	Data Manipulation	pandas: sorting, slicing: loc/iloc, missing values, treating missing values, add/drop rows, merge, concatenate	
12	Visualizing Data	matplotlib: Basic graph, Bar chart, Line graph	
13	Data summaries	Data summaries / Five number summary / Mean median variance / Correlation between columns / Kurtosis	
14	Data summaries	Histograms, left skew, right skew / Outliers and outlier removal techniques	
MIDTERM EXAM			TBA (early March)
15	Guest Lecture I	TBA	
16	Introduction to Web Scraping	Best Practices / User Agent / API tokens / robots.txt / Website Structure / Python libraries for Web Scraping	
17	Introduction to Web Scraping	Best Practices / User Agent / API tokens / robots.txt / Website Structure / Python libraries for Web Scraping	
18	Introduction to Machine Learning	Data Preprocessing/ Introduction to Machine Learning/ Applications of Machine Learning/ Supervised Learning/Unsupervised Learning	

19	Introduction to Machine Learning	Model Evaluation and Validation: Metrics, Bias Variance tradeoff, overfitting vs. underfitting	
20	Introduction to Machine Learning	Practice	
21	SQL Constraints, SQL Operators, SQL Clauses	Overview of databases, Database Management Systems, DBMS components, Data storage in DBMS, Database design, Relational Data Model, DDL, DML, DCL	
22	SQL Constraints, SQL Operators, SQL Clauses	Primary keys: Surrogate keys vs. natural keys, foreign keys, Create tables, Insert values, Retrieve data, Where clause, in, between, distinct, like (wildcards), order by, aliases	
23	SQL Grouping Data / Joins	Min, Max, Average, count, sum, group by, having, join, update, delete	
24	SQL Grouping Data / Joins	With Clause, Subqueries, Normalization: 1NF, 2NF, 3NF	
25	AutoML-Dataiku (https://www.dataiku.com/product/get-started/)		
26	Guest Lecture II	TBA	
27	Project Presentations		
28	Project Presentations		
FINAL EXAM			TBA (May 13-22)

CRN Schedule

Day	**Time**	**20576**
Mon	10:00	A202
Fri	8:30	A101

Note: you will receive a blackboard announcement in case we change classrooms to B012 Lab.

Appendix I: Communication Rules

Students are expected to strictly follow the communication rules throughout the semester. There are two options to communicate with instructors: 1) via **virtual office hours** on the course blackboard page, 2) via **email**.

Virtual Office Hours

If it is a **general-purpose question**, an answer to which will be **useful to other students**, then a student MUST post it via virtual office hours on the **blackboard discussions page**. Virtual Office Hours is **not** a place for asking for specific help/guidance about your assignments (e.g. how to fix this or that error). Before posting, a student should check if:

1. a similar question **has already been asked** by another student or not.
2. the question has a **meaningful title**.
3. the **answer can easily be found** with the help of google or other resources.

Posted questions not following these rules will be **deleted** without being answered.

Emails

If it is a **personal (private) question**, then a student should write an email. Emails should follow the provided template (see below) and have all the relevant emails in carbon copy (CC). For example, in case of writing an individual question to your instructor, you should **put ALL co-instructor emails in CC.** In case of writing an email regarding your team project, you should put **all co-instructor AND all teammate emails in CC**.

Your emails should have a **meaningful title (subject)**, **introduction**, **body**, and **signature**. The following is one possible template that you can use in your daily email conversations.

Dear **[Mr./Ms./Dr. Your instructor's surname],**

Hope you are doing well. **[your email body].** Thanks in advance.

Best Regards,

[Your Name]
Student, [Your Major]
[Your Course Title & CRN]

Further consider the following:

- The last three lines are your **signature**. You can have different signatures for different courses or different emails. You can also use your signatures as your email templates. You can set up and use your signatures by clicking on the pen icon above

your email (after clicking *New Email*), so that you won't have to copy-paste it every time.

- If your instructor has a **Ph.D. degree**, then refer to your instructor as **Dr.** or **Professor**. Otherwise, use **Mr./Ms.** (unless the person you are writing an email to has a title, such as **M.D.**). Use only **surname** immediately after.
- In case of requesting something, use the phrase *"Thanks in advance"* at the end of your email in order to not write single-line and unnecessary *"Thank you"* emails. **Avoid excess emails** and do not use Outlook as a chat box.
- **Do not** click **REPLY ALL** when your reply concerns only the sender. Use **REPLY** instead. Learn to make proper use of **forwarding an email** when necessary. See the official Microsoft instructions on the matter. You can also **flag** or **pin** important emails if you wish.

Appendix II: Policies and Resources

ADA University has her university-wide policy documents. You are expected to have carefully studied at least the following short documents: Honor Code and Student Code of Conduct. You should also be aware of the Student Assessment Regulations. If you have strong reasons to disagree with the grade you got, you should follow the procedures written in Student Academic Grievance Policy after consulting with your advisor.

Students with special needs or with chronic health issues are strongly recommended to contact the University's Student Academic Support Services. ADA University provides upon request appropriate **academic accommodations** for qualified students with documented disabilities. Any student who feels that they may need accommodation based on the impact of a disability should notify the Office of Disability Services and Inclusive Education about their needs before the start of the academic term. Please contact Office of Student Services at phone: (+994 12) 437 3235 ext. 226 or by email: studentservices@ada.edu.az

Reasonable accommodation is possible for students' religious beliefs, observations, and practices or for foreseeable conflicts because of athletic competition.

Adjusting to student life, pursuing academic and personal goals can be emotionally stressful and challenging. Students are encouraged to make individual appointments with a Counselor to receive professional **psychological support**. Some of the areas to focus on include time management and organizational skills, coping with depression, stress, and anxiety, building healthy relationships and communication skills, developing resilience and emotional regulation. Please contact Ms. Kamala Aghayeva, Student Mental Health Counselor, Student Academic Support Services at phone: (+994 12) 437 3235 ext. 138 or by email: kaghayeva@ada,edu,az | studentsuccess@ada.edu.az

Students are encouraged to consult with the **Writing Center** for checking their papers and assignments. Please visit the center or contact them by email: writingcenter@ada.edu.az

Students with recurring or permanent internet and/or hardware problems are strongly recommended to contact the University's IT Service Desk and/or the Blackboard Administrator. Please contact via emails: itservicedesk@ada.edu.az and blackboard@ada.edu.az

Revisions

v0

- no revisions yet