

Extended Syllabus

Course Title	인공지능을 위한 데이터분석 (Data Analysis for AI)	Semester	Fall 2026
Credit	3	Course Number	ENG3510-01
Class Time	Wed/Fri 13:30 ~ 14:45	Enrollment Eligibility	DISCOURAGED for students from engineering with coding experience

Instructor's Photo	Name: Yoonseok Heo (허윤석)	Homepage: https://haicolab.sogang.ac.kr
	E-mail: nlpysheo@sogang.ac.kr	Telephone: N/A
	Office: J825 Office Hours: Mon-Fri 10:00~18:00 (*By appointment via email in advance)	

I. Course Overview

1. Description
NOTE: This course is designed primarily for students in the English Division and the Humanities. Students from other fields – particularly Computer Science, Artificial Intelligence, or Electrical Engineering, as well as those with significant programming experience – are strongly advised to pursue more advanced programming courses instead. Should such students wish to enroll, they must first arrange a separate meeting with the instructor before enrollment can be approved. NOTE 2: Students are strongly encouraged to subscribe to a paid AI tool, such as ChatGPT Pro or Claude Pro, to ensure stable access to the advanced features needed for course projects. Subscription is not mandatory; however, all projects will be evaluated according to the same criteria, and limitations associated with free-tier access will not constitute grounds for grade adjustments. Students are responsible for securing an appropriate tool environment, whether through a paid subscription or suitable alternatives, to complete the assigned work. ----- This course is designed primarily for students in the Department of English Literature and the Humanities, with the aim of introducing fundamental experiences in data analysis. At the outset, students will learn how to handle diverse forms of data, particularly textual data, through Python programming, and will lower the entry barrier to coding by engaging in vibe coding as a learning experience. Building upon this foundation, students will conduct preprocessing and exploratory analysis of large-scale text datasets, developing the ability to critically interpret underlying patterns and meanings. In the latter part of the course, the focus will shift to the application of generative AI, where students will study and practice prompt engineering as a means of cultivating effective language-based communication with AI systems. Ultimately, this course seeks to integrate literary perspectives with data-driven methodologies, enabling students to gain insights grounded in AI-assisted data analysis and to explore its academic and practical applications. * The finalized course syllabus will be distributed on the first day of class.

2. Prerequisites

There are no formal prerequisites for this course. However, having a basic understanding of python programming and AI will be helpful for following the lectures and hands-on exercises.

Recommended prerequisite: Introduction to Linguistics (ENG2003), Linguistics with AI (ENG2112)

3. Course Format (%)

Lecture	Discussion	Experiment /Practicum	Field study	Presentations	Other
45%	%	45%	%	10%	%

4. Evaluation (%)

Mid-term Exam	Final exam	Quizzes	Presentations	Projects	Assignments	Participation	Other
25%	%	%	10%	50%	10%	5%	0%

II. Course Objectives

Knowledge:

- Understand the fundamentals of data analysis with a focus on linguistic data.
- Gain introductory knowledge of Python programming concepts
- Recognize the role of generative AI and prompt engineering in analyzing and interpreting language data.

Skill:

- Apply basic Python programming to preprocess and analyze large-scale text datasets.
- Utilize libraries and tools for exploratory data analysis.
- Design and test effective prompts to interact with generative AI models, and implement simple agent-based workflows.

Attitude:

- Develop confidence in approaching programming and AI without prior technical background.
- Cultivate curiosity and openness to integrating literary/linguistic perspectives with computational methods.
- Foster a critical and responsible mindset when using AI tools for academic or creative purposes.
- Students are STRONGLY ENCOURAGED to make active use of office hours (OH) for questions, feedback, and additional support throughout the semester.

III. Course Format

(* In detail)

This course will combine short introductions to key concepts with extensive hands-on practice in a computer lab or online coding environment (e.g., Google Colab). Early sessions will emphasize guided exercises through vibe coding to lower the entry barrier to programming. As the semester progresses, students will engage in data preprocessing and analysis of large-scale text datasets, followed by interactive workshops on prompt engineering and simple AI agent design. Active participation, regular practice, and collaborative discussion will be essential components of the learning process.

Because in-class practice may occur, students are strongly encouraged to bring their own laptops to each session.

In addition, recitation sessions may be scheduled during the semester if necessary. These sessions are not mandatory, but will be designed to provide hands-on practice, extended Q&A, technical setup (e.g., programming environments), and topics that cannot be fully covered during lecture hours. The specific times and details will be announced later.

IV. Course Requirements and Grading Criteria

Course Requirements

- Students are expected to bring a laptop to each class for hands-on practice (Google Colab will be used to minimize setup). If bringing a laptop is difficult, students must notify the instructor in advance. A computer lab option is currently under consideration.
- Active participation in coding sessions, discussions, and workshops is essential.
- Assignments and projects must be completed on time to ensure steady progress.

Grading Criteria

- Mid-term Exam: 25%
- Assignments (Coding & Data Analysis Exercises): 30%
- Final Projects: 30%
- Presentation: 10%; After assignments, presenters will be randomly selected to share their work in class. Volunteer presenters may receive extra credit within the presentation score.
- Attendance: 5%

Grading policy: The relative grading scheme will follow the university's official regulations, but the detailed weight of each component may be adjusted slightly as the course progresses.

Examples of curriculum-based writing homeworks/activity: all writing exercises requested by classes, including reports, discussion and presentations

V. Course Policies

- The FA policy and all course regulations will be applied based on the university's academic policies. Additional course-related regulations will follow the guidelines announced on the cyber campus.
- Any form of academic dishonesty will result in an F grade.
- Except for officially recognized reasons for absence (e.g., military reserve duty, family emergencies), NO make-up quizzes or exams will be given, and no extensions for assignment submissions will be allowed.
- AI Tools Policy: Collaboration with generative AI tools such as Co-Pilot and ChatGPT is allowed, treating them as collaborators in the problem-solving process. However, the direct solicitation of answers or copying solutions, whether from peers or external sources, is strictly prohibited. Note that indiscriminate reliance on AI tools can seriously harm learning efficiency, and this may be revealed as a significant gap between high assignment performance and low exam scores. Therefore, students are strongly encouraged to use AI tools as collaborators rather than as substitutes for their own work.

VI. Materials and References

There is no official textbook for this course. The primary course materials will be lecture slides, which are developed based on the most up-to-date research and resources available.

VII. Course Schedule

(* Subject to change)

Week 1 9/2, 9/4	Learning Objectives	Introduction to AI and Vibe coding 1
	Topics	Vibe coding, Python basics (variables, strings, loops)
	Class Work (Methods)	Lecture and Hands-on exercise
	Materials (Required Readings)	Lecture Slides
	Assignments	To be announced(TBA)
Week 2 9/9, 9/11	Learning Objectives	Introduction to AI and Vibe coding 2
	Topics	Vibe coding, Python basics
	Class Work (Methods)	Lecture and Hands-on Exercise

	Materials (Required Readings)	Lecture slides
	Assignments	TBA
	NOTE	9/9 – No class(Backup: Online recoding on Cybercampus)
Week 3 9/16 9/18	Learning Objectives	Data Processing 1
	Topics	File I/O, string normalization, regex, tokenization with sample datasets
	Class Work (Methods)	Lecture and Hands-on Exercise
	Materials (Required Readings)	Lecture Slides
	Assignments	TBA
Week 4 Sep 23, Sep 25	Learning Objectives	Data Processing 2
	Topics	File I/O, string normalization, regex, tokenization with sample datasets
	Class Work (Methods)	Lecture and Hands-on Exercise
	Materials (Required Readings)	Lecture slides
	Assignments	TBA
	NOTE	9/25 – No class(Chuseok, National Holiday)
Week 5 9/30 10/1	Learning Objectives	Data analysis with Python 1
	Topics	Pandas basics, word frequency, simple visualizations
	Class Work (Methods)	Lecture and Hands-on Exercise
	Materials (Required Readings)	Lecture Slides
	Assignments	TBA
Week 6 10/7 10/9	Learning Objectives	No class (conference attendance, 한글날)
	Topics	

	Class Work (Methods)	
	Materials (Required Readings)	
	Assignments	
Week 7 10/14 10/16	Learning Objectives	Data analysis with Python 2
	Topics	Word embedding, Visualization
	Class Work (Methods)	Lecture and Hands-on Exercise
	Materials (Required Readings)	Lecture Slides
	Assignments	TBA
Week 8 10/21 10/23	Learning Objectives	Apply computational methods to examine linguistic patterns in text.
	Topics	POS tagging, concordance, collocation
	Class Work (Methods)	Lecture and Hands-on Exercise
	Materials (Required Readings)	Lecture Slides
	Assignments	TBA
Week 9 10/28 10/30	NOTES	10/21 – in class (10/7 backup class) 10/23 – No class (Mid-term)
	Learning Objectives	Apply computational methods to examine linguistic patterns in text.
	Topics	POS tagging, concordance, collocation
	Class Work (Methods)	Lecture and Hands-on Exercise
	Materials (Required Readings)	Lecture Slides
	Assignments	TBA
	NOTES	10/28 – No class (International Conference Attendance)

Week 10 Nov 4, Nov 6	Learning Objectives	Understand and practice basic sentiment analysis on text data.
	Topics	Sentiment Analysis
	Class Work (Methods) Materials (Required Readings)	Lecture and Hands-on Exercise Lecture Slides
	Assignments	TBA
Week 11 11/11 11/13	Learning Objectives	Prompt Engineering 1
	Topics	Large Language Model, Prompt Engineering
	Class Work (Methods) Materials (Required Readings)	Lecture and Hands-on Exercise Lecture Slides
	Assignments	TBA
Week 12 11/18 11/20	Learning Objectives	Prompt Engineering 2
	Topics	Chain-of-thought
	Class Work (Methods) Materials (Required Readings)	Lecture and Hands-on Exercise Lecture Slides
	Assignments	TBA
	NOTES	11/20 – No class (수시)
Week 13 11/25 11/27	Learning Objectives	Prompt Engineering 3
	Topics	Retrieval-augmented Generation
	Class Work (Methods) Materials (Required Readings)	Lecture and Hands-on Exercise Lecture slides

	Assignments	TBA
Week 14 12/2 12/4	Learning Objectives	Data-centric AI 1
	Topics	Data collection/cleaning, Synthetic Data Generation
	Class Work (Methods)	Lecture and Hands-on Exercise
	Materials (Required Readings)	Lecture slides
	Assignments	TBA
Week 15 12/9 12/11	Learning Objectives	Final Project Presentation (NOTE: All students MUST attend the class and participate in peer-review process. Otherwise get 0 scores on presentation and projects)
	Topics	
	Class Work (Methods)	
	Materials (Required Readings)	
	Assignments	
Week 16 12/16 12/18	Learning Objectives	Final Exam Period
	Topics	
	Class Work (Methods)	
	Materials (Required Readings)	
	Assignments	

VIII. Special Accommodations

Students are STRONGLY ENCOURAGED to make active use of office hours (OH) for questions, feedback, and additional support throughout the semester.

In addition, recitation sessions may be scheduled during the semester if necessary. These sessions are not mandatory, but will be designed to provide opportunities for extended Q&A, clarification of lecture materials, and preliminary guidance for upcoming assignments. The specific times and details will be announced later.

IX. Aid for the Challenged Students

Everyone will be welcomed, and I will try to make the class environment accessible for anyone with any type of challenge. Please do not hesitate to contact me.