Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
ASEE
Paper ID #47277

Harnessing the Power of GenAI: A New Era for Data Science Education for Civil and Environmental Engineering

Matthew Yukio Takara, Carnegie Mellon University

Matthew Yukio Takara is a Ph.D. student in the civil and environmental engineering department at Carnegie Mellon University. He holds a B.S. in civil engineering with a minor in data science from the University of California, Berkeley and a M.S. in civil engineering from Carnegie Mellon University. In addition to his interest in engineering education research, his thesis research focuses on the sustainable and ethical use of AI and sensing technologies in smart infrastructure.

Dr. Fethiye Ozis P.E., Carnegie Mellon University

Dr. Fethiye Ozis is an Associate Teaching Professor in the civil and environmental engineering department at Carnegie Mellon University. Dr. Ozis holds a B.S. in environmental engineering from the Middle East Technical University, Ankara, Turkey and M.S. and Ph.D. degrees from the University of Southern California. Dr. Ozis is a licensed Professional Engineer, Environmental, in Arizona. She is also an ENVISION Sustainability Professional. Before joining CMU, Dr. Ozis was a faculty member at Northern Arizona University.

Allison E. Connell Pensky, Carnegie Mellon University

Dr. Allison Connell Pensky is a Cognitive Psychologist and Neuroscientist by training. She is currently a Data Science Research Associate at the Eberly Center for Teaching Excellence & Educational Innovation at Carnegie Mellon University. She uses her training and years of experience as a faculty member to support graduate students, faculty, and programs on educational research projects and making data-driven improvements to courses.

Harnessing the Power of Generative AI: A Case Study on Teaching Data Science to Environmental Engineering Students

1 Introduction

Data science has played a transformative role across research, industry, and education in the 21st century. While computational fields such as computer science and electrical engineering have long embraced data-driven approaches, interdisciplinary domains like civil and environmental engineering (CEE) are increasingly integrating data science into their education and practice. In addition, while the programming skills used in computational fields often lend themselves well to data science practice, there is more often a gap in skills for practitioners in other interdisciplinary domains. For instance, the traditional CEE curriculum could benefit from a greater emphasis on popular open source programming languages such as Python. This shift reflects a growing need for future CEE practitioners to have the skill sets and tools to analyze and understand large datasets generated by the built environment, including air quality, structural health, and energy consumption data. The field has been further disrupted by the emergence of powerful generative AI (GenAI) tools in the last few years, such as OpenAI's GPT-4 which powers the popular ChatGPT chatbot. These large language models can perform sophisticated tasks like interpreting large datasets, writing code for wrangling and analyzing data, brainstorming ideas, and explaining complex statistical and mathematical concepts in ways that closely mimic natural human language. In CEE education, GenAI holds the potential to make data analysis and programming more accessible to students who may lack a strong background or interest in these areas. However, this raises critical questions about how these tools impact student learning and problem-solving approaches in domain-specific contexts. This paper examines the role of GenAI in a junior-level undergraduate CEE course, where students cleaned, visualized, and analyzed air quality data collected using air quality sensors they deployed themselves around the university's campus. We compare two course offerings: one conducted before GenAI tools were widely available (Spring 2023) and another where students had the option to use GenAI tools (Spring 2024). Through analysis of student assignments, reflections, and outcomes, we address two key research questions: (1) How do GenAI tools affect students' ability to process and interpret large datasets in CEE education? (2) What are students' attitudes, perceptions, and experiences with using GenAI for these tasks? By giving students full agency in choosing whether or not to use GenAI, we aim to balance inclusive teaching practices with understanding how students utilize these tools in an unstructured manner, which could be representative of how students might use these tools without proper instruction or even instructor consent in other courses. By exploring these questions, we aim to provide insights into the opportunities and challenges of integrating GenAI into CEE education, ultimately contributing to the broader understanding of how emerging technologies can transform STEM curricula.

2 Literature review

This literature review aims to provide a comprehensive overview and analysis of data literacy education and how generative AI technologies have shaped its trajectory in recent years. We begin by reviewing previous approaches to teaching data literacy and the emergence of data science as an educational discipline (Section 2.1). Next, we explore the emergence of GenAI tools in data analysis and visualization, including case studies identifying strategies and implementations of generative AI tools in educational settings (Section 2.2). Finally, we gain insight into student attitudes and perceptions towards GenAI tools, and how their use impacts learning experiences and outcomes overall (Section 2.3).

2.1 Rise of data science and current approaches to teaching data literacy

Data literacy can be defined as the ability to read, work with, analyze, and argue with data [1]. Reading data refers to understanding what aspects of a given system or phenomenon is represented. Working with and analyzing data refers to the technical skills needed to acquire, process, and perform other operations to create insights from data. Finally, arguing means the use of data to support a narrative and to communicate this narrative and insights with a target audience [1]. Data science is a growing field that includes an increased emphasis on advanced computational tools and technologies to apply aforementioned data literacy skills in various domains in academia and industry. Popular, established approaches in industry often use spreadsheets, such as Microsoft Excel, for handling and analyzing data. These, however, are limited in terms of scalability when working with large datasets and lack flexibility and customizability. Open-source programming languages like Python and R, provide this flexibility and customizability [2], but they generally have steeper learning curves and lack the user-friendly interfaces that come with spreadsheet software.

The demand for data scientists and data science skills in industry has been rapidly growing over the past decade. The US Bureau of Labor Statistics [3] projects a 36% growth in data scientist employment between 2023 and 2033, a rate that is faster than the average of all occupations. This also does not account for the number of jobs not explicitly labeled as "data scientists," "data analyst", "data engineer", etc. that will require some degree of data literacy skills in the coming years. [4] performed an extensive search of job advertisements across engineering industries (including civil and environmental engineer) that found a significant demand for programming skills such as Python and are associated with higher salaries than jobs without. The growing job market for data-related jobs has been met by a similar increase in the number of data science programs and courses at the university level. [5] conducted a comprehensive qualitative and quantitative analysis of data science programs and found over 300 programs worldwide in 2016. In contrast, just 6 years later, in 2022, [6] found nearly 800 data science programs with over 9000 individual courses in US universities alone. Many of these programs have an interdisciplinary approach in which students can supplement their fundamental data science coursework with domain-application-specific courses. For instance, a recent course in Florida Gulf Coast University's civil engineering department, has an explicit focus on teaching R programming for data wrangling, visualization, modeling and other data literacy skills with applications in civil infrastructure systems [7].

With the increased focus on computational methods, modern data science courses often spend a considerable amount of time teaching programming skills in addition to fundamental concepts in statistics and mathematics [7–11]. Despite the growing emphasis on programming skills in data science education, students often enter these courses and programs with varying levels of prior experience and proficiency, and it is often the skill that students struggle the most with, especially in applied settings. In [7], surveys of students found that the majority of students struggle most with basic programming skills and applying those skills in new problem settings. Broad studies of instructors uphold these findings, where it is reported that the biggest challenge in teaching data science is technology adoption, which includes tools like GitHub and languages like R and Python [11]. Recent literature has found that the best learning outcomes are achieved when data literacy skills are taught in the context of some domain application (e.g., air quality monitoring). [12] found that 5th grade students given an interdisciplinary curriculum teaching environmental science and data literacy skills showed an increased understanding of and engagement with both the environmental science and data literacy concepts. Similarly, in higher education, research has found that students who learn programming or data literacy skills in conjunction with a domain subject such as physics [13] or biology [14] perform better in both domain-specific tasks and computational-skill-related tasks and have higher self efficacy compared to students who only learn one. This demonstrates that from K-12 to higher education, interdisciplinary learning can help students develop data literacy skills while improving their understanding and appreciation for their domain subject. As the demand for data literacy and programming skills continues to grow, it becomes increasingly important to teach these skills efficiently and effectively. Interdisciplinary approaches have shown promise in addressing this challenge by contextualizing programming within specific domains. Additionally, emerging tools like AI offer new opportunities to support students' learning in a streamlined fashion, which will be discussed in the next section.

2.2 The emergence of AI-powered tools in data science

AI in education (AIEd) research has existed since the 1980s as primitive virtual learning companions, which users could interact and compete with to aid their learning [15]. AI tools for educational purposes have been steadily developed in the decades since, with the majority of AIEd research focusing on instructor-side AI tools used for administrative tasks such as automated grading, feedback, and content creation [16–20]. There are some cases of AI tools developed for learner use, but they are more application-specific and are very different from the modern, more robust generative AI tools of the present study [21, 22]. Recent years have seen the emergence of powerful generative AI tools, such as OpenAI's ChatGPT and Google's Gemini (formerly Bard) released in late 2022 and early 2023, respectively. These tools are examples of powerful large language models (LLMs) capable of interpreting human language inputs and generating outputs resembling natural human language. The emergence of these highly sophisticated LLM-based chatbots has generated significant interest in their potential applications in educational settings. In the context of data science and data literacy education, they are capable of performing complex analyses, creating visualizations, and writing code for wrangling and cleaning data. They can also simplify explanations of difficult technical concepts in statistics and probability [23].

However, despite the purported benefits of GenAI usage in the classroom, there are important practical limitations and ethical considerations that come with its usage. GenAI models "learn" by finding statistical patterns and relationships in their training data, such as text or images. This means that they do not have a true, complex understanding of concepts and are therefore still capable of producing biased, inaccurate, vague, or nonsensical responses, or hallucinations [24]. Students and educators should be aware of the limitations and risks associated with utilizing GenAI in their learning and instruction. Additionally, there are several ethical considerations that must be made when employing GenAI tools in the classroom. Users must be aware that models may be trained on biased data and may use personal data. There are also risks related to privacy and data security, such as students' work being collected and added to the model's training data [25]. There are also concerns over job security for instructors and over reliance, cheating, or plagiarism by students [26]. However, [27] argues that instead of replacing jobs and skills, generative AI has transformed the role of data scientists and educators. The future role of a data scientist, or in this case a CEE professional applying data science skills, should be to focus more on understanding, interpreting, verifying, and managing the outputs of generative AI tools rather than performing all of the preprocessing, computation, analyses, etc themselves [27]. It can then be inferred that a new role for instructors is to help students develop GenAI literacy skills such as how to appropriately utilize generative AI tools, by understanding how to create effective prompts and understanding the ethical and security risks associated with using these tools [23, 28]. Because the technology is still being integrated into curricula, there is relatively limited literature on real-world case studies in which students are permitted to use generative AI tools.

2.3 Student attitudes and perceptions towards GenAI tools

With the potential to integrate ChatGPT and other generative AI tools in the classroom, there is growing interest in understanding the current attitudes and perceptions that students and educators have towards these tools. Similar themes appear across multiple studies assessing student perspectives and experiences on GenAI usage in education. In general, students hold many positive attitudes about GenAI usage. Among many benefits, students generally believe that it can improve their learning experience by summarizing complex ideas, providing immediate personalized feedback, improving confidence in technical skills like programming, and streamlining literature reviews and online research [25, 29]. In addition to integrating GenAI tools into their education, students have shown an interest in integrating them into their future careers as well, indicating their understanding or belief that GenAI tools will become an important part of industry practice in the near future [30]. It is also found that students who have a greater prior knowledge of or familiarity with using AI tools are more likely to have a positive attitude and acceptance of these tools in their learning [31]. Students have also demonstrated an understanding of the limitations of GenAI citing specific concerns and/or weaknesses of the tool. One study found that around 40% of students chose not to use GenAI for their assignments, citing multiple reasons such as distrust in the accuracy of GenAI tools' outputs, lack of alignment with assignment needs, and lack of guidance on how to use the tools [32]. Generally, students tend to share common concerns about GenAI usage in the classroom. For instance, they have an awareness of the possibility of receiving inaccurate or biased responses from AI chatbots [29]. Students generally have a decent understanding of the risks and ethical issues regarding AI usage. They also often express concerns about overreliance on GenAI and GenAI usage leading to a lack of originality and innovation, reduction in critical thinking skills, and intentional and unintentional plagiarism [33]. While students demonstrate an awareness and understanding of the limitations of GenAI, the extent to which this awareness translates into effective usage remains underexplored.

2.4 Synthesis and identification of gaps in literature

Prior literature on teaching practices demonstrate the efficacy of interdisciplinary approaches to teaching data literacy. However, with the advent of GenAI tools, there is still limited research on how these tools fit into this paradigm. In more open-ended unstructured assignments, it is not clear how students should use GenAI to help them solve complex, multi-step, interdisciplinary problems, or whether they should use GenAI in the first place. For example, in environmental engineering laboratory courses where students collect their own data, how do they use GenAI to help them decide what to use the data for, what approaches to use to process the data, and then execute their plan (e.g., Excel workflow or Python code)? The present study aims to address these gaps by examining what motivates students to choose to use GenAI tools or not and how students choose to integrate GenAI tools into their project workflow working with unstructured data and programming tasks. For those who do choose to use GenAI, we explore how it affects their learning, problem solving approach, and overall performance on these assignments? Finally, by focusing on CEE education, this study aims to learn insights into GenAI technology adoption in an area of engineering that is more closey tied with government regulation and therefore traditionally slower to adopt new technologies compared to other fields such as information technology and finance [34]. The present study aims to understand how students in civil and environmental engineering view the usage of GenAI in their data analytics related coursework.

3 Methods

3.1 Course structure and study context

Our institution, Carnegie Mellon University, offers a series of vertically-scaffolded, sequential courses designed to produce Civil and Environmental Engineering (CEE) graduates proficient in designing sensing systems for various applications and environments. These courses incorporate interpreting large datasets and using data to understand and control infrastructure systems, and enhance infrastructure management strategies by implementing smart technologies. The details about the structure of these courses, alignment, and contents can be found in [35]. Our study took place across two offerings of a junior-level undergraduate civil and environmental engineering sensing lab course project offered in Spring 2023 (S23) and Spring 2024 (S24). The title of the course is Experimental & Sensing Systems Design and Computation for Infrastructure Systems (12-333). The class met once a week, over a 14-week semester. This sensing lab course, where field-based CEE problems are posed, requires students to choose and deploy appropriate off-the-shelf sensing systems to collect data to use for further analysis of the problem.

3.2 Student projects and main assignment

3.2.1 Course project description

The course project aims to establish a sensor network to monitor air quality and answer questions posed by students or the client, including identifying outdoor exercise locations, optimal times for activities, and comparing air quality on campus and in the region. Monitoring and improving indoor air quality is also crucial for our client, the university facilities director, and our community's well-being. Therefore, the relationship between indoor and ambient air quality was also included in the project's scope.

In this project, students calibrated and installed Purple Air ® air quality sensors around our urban campus located in Pittsburgh, Pennsylvania, a city notorious for its air quality issues. A total of 18 sensors were deployed with each sensor taking measurements about every minute across a four-week period. These sensor readings formed the data set student teams were expected to work with. Students were then given an open-ended task of preprocessing, cleaning, analyzing, and visualizing their air quality data using any methods they choose, though we recommended using either Python or spreadsheet software, as students are most familiar with these tools.

3.2.2 Project execution and timeline

The project was launched in the third week of the semester for both iterations, completed in 5 weeks in S23, and in 6 weeks in S24. In both semesters, teams submitted their final paper during finals week. The project execution timeline can be seen in Table 1. In both S23 and S24 semesters, students were assigned pre-class readings before the Purple Air ® - Air Quality lab was launched. During the first class meeting, students discussed pre-class readings, identified their team's research question, and developed ideas about their experimental methodology. In both S23 and S24, students mounted the sensors in the classroom at the same height from the ground level to start the two-day initial sensor calibration period in the first project meeting.

After that, both cohorts' students were expected to review two day's worth of air quality data from their team's sensors and compare it with the rest of the class's sensors (total 18) to confirm initial sensor performance and identify any inconsistencies for calibration purposes. During the second meeting in S23 (the third meeting in S24), teams presented on their pre-deployment plan, proposed experimental methodology, and proposed data analysis methodology. Any potential issues about access to the teams' proposed locations on campus, and equipment needed for

Table 1: Timeline comparison between two semesters of the project

Project Meeting	S23	Project Meeting	S24
1	Air Quality - PM2.5 Sensors, Review Pre-class readings	1	Air Quality - PM2.5 Sensors, Review Pre-class readings, Mount Sensors to Calibrate
2	Determining Methodology for a Study & Mount Sensors to Calibrate	2	Determining Methodology for a Study
3	Initial Presentations (Pre-deployment plan and methodology)	3	Determining Methodology for Data Analysis, Walk Through with Facilities Management
4	Deploy Sensor System @ Scheduled Times & Start Collecting Data	4	Initial Presentations (Pre-deployment plan and methodology)
5	Air Quality Sensing Lab - Final Presentations	5	Deploy Sensor System @ Scheduled Times & Start Collecting Data
Finals Week	Air Quality Sensing Lab - Briefing Paper, CATME Peer Evals	6	Air Quality Sensing Lab - Final Presentations
		Finals Week	Air Quality Sensing Lab - Final Paper, Individual Reflection, CATME Peer Evals

deployment (e.g., electric extension cables, screws, drills) were noted by the client.

Sensor deployments took place during assigned class time (week 4 and 5 in S23 and S24, respectively), teams' sensors were deployed with help from the university facilities team. Sensors started collecting data, and remained in their locations, for approximately four weeks. Students were asked to perform a weekly sensor checkin to make sure of continuous data collection. The project culminated in a final report and a presentation in which student teams present a description of their data analysis and assessment, conclusions, and recommendations. Teams also provided feedback on the learning value of the activity and suggested improvements. All assigned tasks related to these activities can be found in section 3.5.

3.2.3 Similarities between semesters

The majority of the project related activities stayed exactly the same in both semesters (S23 and S24). Students in both semesters had access to CSV files containing all of the data for each sensor, and both cohorts were asked to perform data cleaning and exploratory analysis in addition to creating data visualizations for their final papers. In S24, three additional tasks were created based on the experience and student feedback from S23. These differences were also due to the addition of permitted GenAI use in S24. Between sensor deployment and final paper, an additional task (Task 4 in S24, see table 2) was created for individual data cleaning, analysis, and visualization based on the first two weeks of data collection. In S23 there was no individual data clean up task like Task 4 in S24, and the team aspect of data cleaning and analysis was part of the S23 Task 4 Briefing Paper. Some tasks were slightly modified to collect feedback on student use of GenAI. For instance, students in S23 were not asked to reflect on GenAI-aided vs manual data analysis, because GenAI usage was permitted for those students. Table 2 below presents the project tasks, assignments, and comparison between two semesters.

Table 2: Comparison of subtasks in the Purple Air assignment between S23 and S24

S23	**S24**	**Included in this study (Y/N)**
Task 1. One-page Memo	Task 1. Calibration Memo	N
Task 2. Presentation: Plan for Deployment and Method	Task 2. Presentation: Plan for Deployment and Method	N
Task 3. Deployment	Task 3. Deployment	N
N/A	Task 4. Individual Data Cleanup	Y
N/A	Task 5. Team Data Cleanup	Y
Task 4. Briefing Paper	Task 7. Final Paper	Y
Task 5. Pecha Kucha Final Presentation	Task 6. Pecha Kucha Presentation	N
Optional last item on Task 4. Briefing Paper	Task 8. Individual Reflection	Y

The primary difference between the semesters was the addition of permitted GenAI tool use. Students in the S23 semester were not permitted to use GenAI tools to complete any of their assignments. In the S24 semester, students were given the option to use GenAI tools to help complete their data analysis tasks. Because it was optional, we did not want to overemphasize the use of GenAI. Thus, students were provided with only the following brief instructions in their assignment handout:

> "You can also use GenAI to do this. We have tried a few different ones and identified that PerplexityAI seems to be the best fit. It is free and available, unlimited & conversational. You are allowed to use GenAI in any way you choose, you can use it as a consultant to deepen your learning, you can use it as an assistant. GenAI can provide you with methods and ideas to clean the data. You can use it as a coach to help you get additional practice and training. It can also help you clean the data but remember sometimes it hallucinates (we recommend you upload and test with sample data to validate what it is doing). It won't plot the data, but it can provide you with a code to create the plot in Python (and other ways if you ask it) so that you can check that out. We don't expect you to substitute your own data processing with AI data processing but support your data processing task. Please note that it can't handle everything. Document the prompts and ways you used GenAI if you chose to do so."

We primarily recommended using GenAI to aid in data analysis—both directly and by writing code that can process the data in software, such as Python or Excel. We also suggested other uses of GenAI, such as brainstorming and for explaining concepts, but did not require the students to use the tool in any particular way. In addition, we warned of hallucinations and emphasized the need to verify results.

3.3 Study design and participant details

Participants in this study were registered third year undergraduate civil engineering students at our university. Students were informed about the study at the beginning of the semester and were free to opt out at any time. A total of 51 (17 in S23 and 34 in S24) students participated in the study. While student demographic data was not a factor in our analysis, we will provide the student population demographics for the final version of the paper. This sample size and demographic breakdown was generally representative of the university as a whole. This project was conducted under the exempt IRB protocol STUDY2016_00000148 approved by the Carnegie Mellon University Institutional Review Board.

S24 students (N = 34) served as the intervention group (permitted use of GenAI). They were compared to S23 students (N = 17) who served as the comparison group (no permitted use of GenAI).

3.4 Assignment structure and data collection

Firstly, student grades from equivalent data analysis assignments between S23 and S24 were collected. In addition, the contents of each assignment were collected in order to understand how students approached the tasks using or not using GenAI. While the final analyses and reports in both semesters were done in teams, S24 students were required to attempt the data handling tasks individually before working with their team. They answered reflection questions regarding their approach and experience using or not using GenAI.

In addition to using the teams' grades for their final analysis report, we blinded both semesters' assignments and rescored them using a more detailed research rubric to better capture aspects of their performance and problem solving approaches. While the submissions were blinded, in some cases, the scorer was able to form "hypotheses" as to which semester and therefore condition the submission was from. Because the final analysis report was a team assignment, the sample size became N = 5 (S23) and N = 7 (S24) for this comparison. This rescoring was performed after both semesters had ended and did not impact the students' grades.

3.5 Performance and attitudes assessment

To answer RQ 1 (How do GenAI tools affect students' ability to process and interpret large datasets in CEE education?), the grades of students' submissions in S23 and both grades and the contents of students' submissions in S24 for each assignment were collected. Overall grades between S23 and S24 were compared to look at whether the addition of permitted GenAI use impacted student grades, while the contents of student assignments in S24 were examined in order to gain a more detailed understanding of how students approached the tasks using or not using GenAI. Additionally, a more in depth qualitative evaluation was conducted to better capture the intricacies of students' work compared to numeric grades. This evaluation was conducted after the course was completed and did not impact students' grades.

S24 students provided additional data sources through assignment reflections in order to answer the RQ 2: What are students' attitudes, perceptions, and experiences with using GenAI for these tasks? Students answered questions related to their overall course project experience such as challenges performing data analysis tasks, lessons and values they took away from the assignments, and their reasons for choosing to use or not use GenAI.

Students performed basic analyses and visualizations on the data they had collected, first individually and then in teams. In both scenarios, students and teams were required to submit a report of their analysis along with reflections and documentation of their data analysis approach and whether or not they chose to use GenAI. The assignment for S24 students was designed to be similar to S23 students' assignments for direct comparison.

4 Results

Both quantitative and qualitative analyses were conducted to assess the impact of GenAI tools on students' ability to process and interpret large datasets (RQ1) and learn students' attitudes, perceptions, and experiences using GenAI for these tasks (RQ2). We accomplish this by using the S23 (control, no GenAI) and S24 (treatment, optional GenAI) data.

4.1 How do GenAI tools affect students' ability to process and interpret large datasets in CEE education? (RQ1)

A quantitative analysis of student grades was conducted to address the first research question. Students' data processing and interpretation skills will be referred to as student performance. To evaluate the impact of generative AI (GenAI) on student performance, comparisons were made between Task 4 (S23, control group with no GenAI) and Task 7 (S24, treatment group with optional GenAI) as well as overall course grades. Levene's test indicated unequal variances for both the group assignment scores (F(1,50) = 8.33, p = .006) and overall semester grades (F(1,50) = 6.63, p = .01). Thus, Welch's t-test was employed to compare the two groups. For group assignment scores, no significant difference was found between S23 and S24 (t(35.49) = 1.16, p = 0.26). Similarly, overall semester grades showed no significant difference (t(46.81) = -1.39, p = 0.17) between the S23 control students and the S24 treatment students.

To better understand how student performance may have differed by semester, we analyzed the rescored team assignments (see section 3.5? for more details). We analyzed three rubric criteria. The criterion of word count refers to the total number of words in the team's submission including titles, results, methods, figures, and figure captions, but excluding the references section. The criterion of number of visualizations is the simple count of the number of data visualizations included in the team's report. The criterion of quality of analysis refers to how well the team analyzed their data as scored on a 4-point scale with larger numbers indicating better quality.

An independent-samples t-test found no impact of semester (control S23 vs treatment S24) on word count (t(10) = 0.25, p = 0.81). The S23 students included marginally more visualizations than the S24 students (t(10) = 1.78, p = 0.106). The S23 students provided significantly higher quality analysis than the S24 students (t(10) = 2.28, p = 0.045)

While S23 control students performed better in quality of analysis and number of visualizations, this may reflect differences in task execution (i.e., GenAI use) in the S24 treatment students. We therefore split the S24 teams based on whether they opted to use GenAI or not. Follow-up one-way, between-subjects ANOVA analyses explored the effect of these three conditions (required manual (S23), optional manual (S24), and optional GenAI analysis (S24)) on the three rubric criteria. Due to the small sample size of the number of teams, none of these analyses were statistically significant (see Table 3 and Figures 1, 2, and 3), however, they all had a large effect size.

Table 3: Rubric Criterion and Statistical Results

Rubric Criterion	**Statistical Result**
Word Count	$F(2,9) = 0.92, p = 0.43, \eta_p^2 = 0.17$
Number of Visualizations	$F(2,9) = 1.65, p = 0.25, \eta_p^2 = 0.27$
Quality of Analysis	$F(2,9) = 2.7, p = 0.12, \eta_p^2 = 0.38$

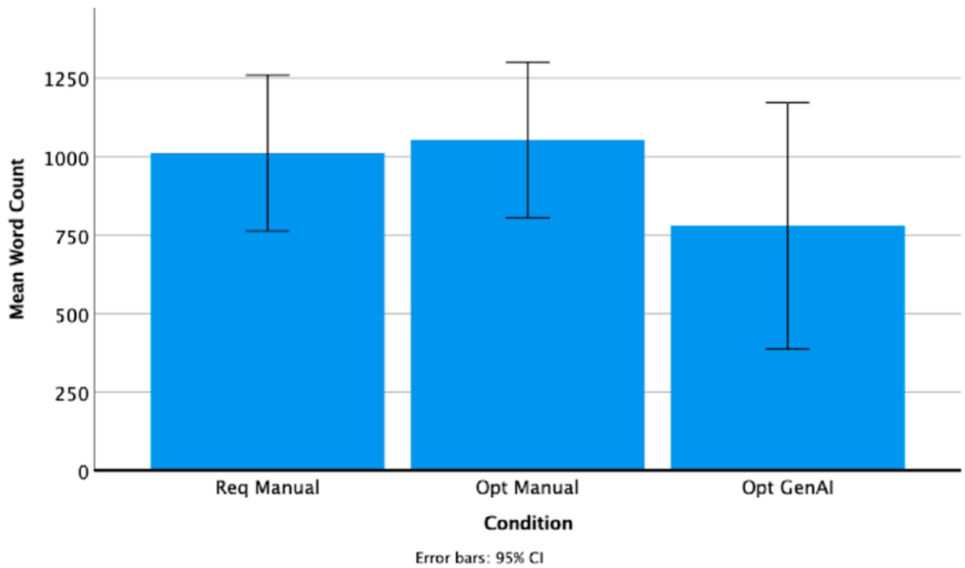

Figure 1: Statistics for word count between the three groups: required manual (S23), optional manual (S24), and optional GenAI (S24). Although both the required manual and optional manual groups wrote slightly more than the GenAI users, we found no statistically significant differences between the groups, but they all had large effect sizes.

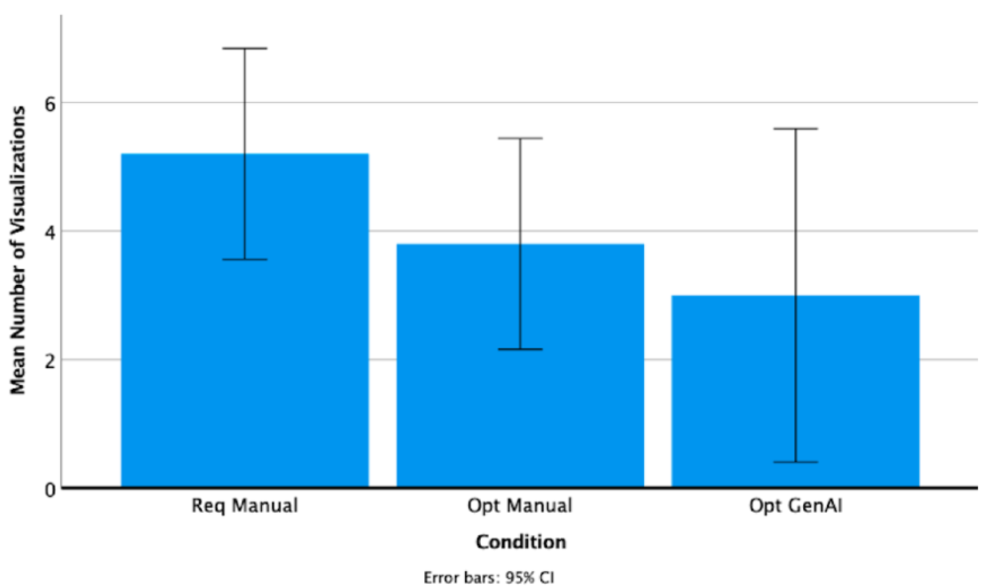

Figure 2: Statistics for number of visualizations between the three groups: required manual (S23), optional manual (S24), and optional GenAI (S24). The required manual students (S23) created marginally more visualizations than the students in S24. However, we found no statistically significant differences between the groups, but they all had large effect sizes.

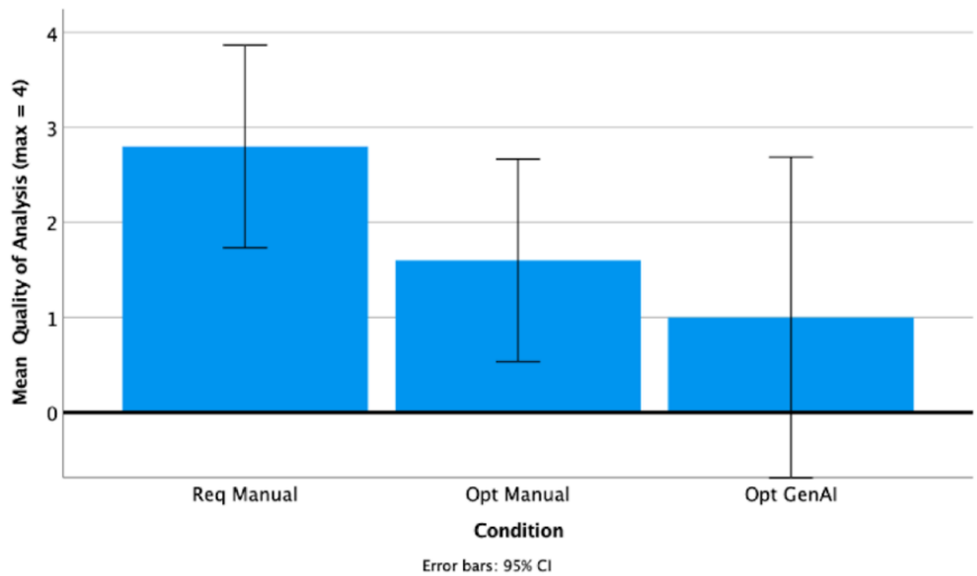

Figure 3: Statistics for the quality of data analysis between the three groups: required manual (S23), optional manual (S24), and optional GenAI (S24). We found that in general, the required manual students (S23) performed higher quality analyses than the students in S24, and within S24, the optional manual students performed better than the optional GenAI students. However, we found no statistically significant differences between the groups, but they all had large effect sizes.

4.2 What are students' attitudes, perceptions, and experiences with using GenAI for these tasks? (RQ2)

To answer research question 2 about student attitudes, perceptions, and experiences, students' individual reflections (N= 34) on their process and learning at the conclusion of the Purple Air lab assignment were analyzed from the Sp24 treatment course. These reflections were coded using thematic analysis, which provided insight into students' choices and perceptions of GenAI. Thirteen students chose to use GenAI. The types of use that they reported were primarily for advice and troubleshooting rather than task automation. Manual cleaning remained slightly preferred (N= 18), with students often citing that GenAI was unnecessary for the nature of the tasks. Both groups reported similar perceived learning, though their focus areas varied. GenAI users emphasized technical and collaborative skills, while manual users highlighted task-specific insights such as air quality monitoring. Students generally recommended their chosen approach (manual or GenAI) for future tasks. Manual users emphasized the need for more class time to collaborate, while GenAI users requested additional instructional support for the data analysis tasks.

5 Discussion

The need for engineering students to build their data literacy has been a growing area of focus for many years [4–6]. Research has shown that hands-on, concrete, experiential learning is effective in building students' skills in this area [12–14]. Student performance when given unrestricted and unstructured access to GenAI tools may be indicative of how students in not computationally heavy disciplines such as CEE engineering adopt and utilize new technology (RQ1), as well as students' attitudes and perceptions towards these tools (RQ2). The present study sought to test how introducing the availability of generative AI tools during an open-ended data task in a CEE class would impact students' performance and attitudes. We found that overall, the difference in students' final course grades between treatment groups was not statistically significant. Further investigation of teams' submissions and a more rigorous scoring criteria found no significant difference in three criteria between the three groups (Figure 1, 2, and 3). However, since this was a team-based assignment, the sample sizes in each group were small (5, 5, and 2 teams respectively). Nevertheless, each of rubric metrics had large effect sizes suggesting that with a larger sample size or under more controlled experimental conditions we may find lower quality analyses, fewer word count, and fewer visualizations in submissions of teams who choose to use GenAI. Despite the potentially lower quality analysis from teams who chose to use GenAI, their reflections indicate that they understood the value of and wide range of use cases of GenAI beyond task automation and code writing, such as troubleshooting, idea-generation, and generating advice.

In addition to our quantitative analysis, qualitative analysis of student reflections showed that when given the opportunity to use GenAI, many students still chose to manually complete their analysis tasks and believed that GenAI was unnecessary given the nature of the tasks. We found that without structured guidance, some students opted not to use the tool because they did not find it to be useful or necessary. Regardless of whether students completed their analysis task with or without GenAI, they gave feedback that they would have liked to have had more instruction for data cleaning and using data software. Manual students also requested more time to complete their assignments and were more likely to submit their work late suggesting that despite a gap in data literacy skills, some students were still reluctant to use GenAI tools that could have potentially streamlined their work.

A strength of the present study is its ecological validity, because it was conducted with real students in an actual engineering class. The cost of this ecological validity is reduced internal validity because we were unable to randomly assign students to treatment conditions. Specifically, students were allowed to self-select their use of the GenAI tool. Additionally, students were permitted to decide how they would utilize the tool. Allowing students this choice is considered an inclusive teaching practice [36] and promotes more authentic student use. Because students were given this freedom of choice, our study provides some insight into how students choose to utilize GenAI tools in a real-world classroom involving open-ended projects. In alignment with recent literature, we showed that students display a varied level of acceptance and adoption of GenAI tools - some more readily use it, while others are more skeptical about its reliability or necessity [25, 29, 30, 33].

6 Limitations and Future Directions

This study has a few limitations. Being conducted in a real classroom setting, the sample size was limited to the number of enrolled students (N = 52 across two semesters). This small sample size restricted our statistical power, reducing our ability to detect potential effects. Despite mostly non-significant effects, we observed large effect sizes in all three metrics of the re-scoring rubric (see section 4.1) suggesting a potential Type II error. The observed large effect sizes for word count, number of visualizations, and quality of analysis suggest that with a larger sample size, clearer trends might emerge. For instance, GenAI users demonstrated a trend of more concise writing, a potentially valuable skill in technical contexts. However, this concision appeared to come at the cost of quality in analysis, underscoring the need for more robust pedagogical support for critical use of GenAI tools. Additionally, the pedagogical choice to allow students to self-select whether or not to use the GenAI tools may have introduced selection bias in which the students who chose to use GenAI may have differed in background or prior knowledge compared to those who opted for manual methods. This means that any observed differences between our GenAI and manual students may be attributable to student characteristics rather than the impact of the tool use. However, our choice to allow students to choose their approach was intentional in order to align with inclusive teaching practices, respecting individual preferences and learning styles. Further, the study found no negative impact on final course grades when comparing pre-GenAI (S23) and post-GenAI (S24) cohorts, demonstrating that the integration of such tools does not inherently disadvantage students. A more rigorous scoring rubric revealed valuable insights into areas where GenAI students may need additional pedagogical support. Student reflections confirmed that those who chose to use GenAI explicitly expressed a need for more instructional guidance to effectively use the tools [37–40]. In the next iteration of the course, the instructional team provided structured guidance to foster a better understanding of the tool's capabilities and address disparities in students' readiness, bridging gaps in skill levels and ensuring all students can potentially benefit from these tools. In order to better test the cause-and-effect relationship between GenAI use and student performance or attitudes, an experimental research design would need to be used. Students would need to be randomly assigned to one of the conditions so that their performance and attitudes could be compared in the absence of underlying demographic differences. Alternatively, if an instructor wants to maintain the inclusive teaching practice of student choice, they should consider collecting and analyzing demographic data on their sample. This could help identify if students with different demographic characteristics tend to choose a GenAI or manual approach to data tasks and whether their attitudes differ. In particular, it would be helpful to understand students' prior perceptions of GenAI as well as their background and prior experience with GenAI tools and new technology correlate with tool use.

7 Conclusions

The present study adds on to existing literature by exploring the use of GenAI in domain-specific data science coursework for CEE students. We provide early insights into how CEE students interact with GenAI, adding on to the relatively small number of case studies on real-world GenAI integration into the classroom. The integration of GenAI tools into educational settings, even as an optional feature, is a promising step toward inclusive and innovative teaching practices that may support building students' data literacy. Although optional GenAI use did not significantly improve or detract from student performance overall, it was a good starting point in supporting data literacy skills. The successful integration of GenAI would benefit from more structured pedagogical support teaching students its strengths and limitations. By refining experimental methods and increasing sample sizes, future studies can better evaluate the nuanced impacts of GenAI on student performance and learning experiences.

References

[1] C. D'Ignazio and R. Bhargava, "Approaches to building big data literacy," in *Proceedings of the Bloomberg data for good exchange conference*, vol. 6, 2015. [Online]. Available: https://www.media.mit.edu/publications/approaches-to-building-big-data-literacy/

[2] D. Incerti, H. Thom, G. Baio, and J. P. Jansen, "R You Still Using Excel? The Advantages of Modern Software Tools for Health Technology Assessment," *Value in Health*, vol. 22, no. 5, pp. 575–579, May. 2019. [Online]. Available: https://doi.org/10.1016/j.jval.2019.01.003

[3] Bureau of Labor Statistics, U.S. Department of Labor, "Occupational outlook handbook: Data scientists," 2024, accessed: Jan. 15, 2025. [Online]. Available: https://www.bls.gov/ooh/math/data-scientists.htm

[4] G. C. Fleming, M. Klopfer, A. Katz, and D. Knight, "What engineering employers want: An analysis of technical and professional skills in engineering job advertisements," *Journal of Engineering Education*, vol. 113, no. 2, pp. 251–279, Apr. 2024. [Online]. Available: https://doi.org/10.1002/jee.20581

[5] T. Wiktorski, Y. Demchenko, A. Belloum, and A. Shirazi, "Quantitative and Qualitative Analysis of Current Data Science Programs from Perspective of Data Science Competence Groups and Framework," in *2016 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*, Dec. 2016, pp. 633–638. [Online]. Available: https://doi.org/10.1109/CloudCom.2016.0109

[6] N. Nellore and M. Zimmer, "Towards a Systematic Review of Data Science Programs: Themes, Courses, and Ethics," in *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 2*, ser. SIGCSE 2023. New York, NY, USA: Association for Computing Machinery, Mar. 2023, p. 1413. [Online]. Available: https://doi.org/10.1145/3545947.3576356

[7] A. Badir, S. Tsegaye, and L. D. Nguyen, "Data Science in the Civil Engineering Curriculum," Jun. 2023. [Online]. Available: https://doi.org/10.18260/1-2--42870

[8] D. Asamoah, D. Doran, and S. Schiller, "Teaching the Foundations of Data Science: An Interdisciplinary Approach," Dec. 2015, arXiv:1512.04456 [cs]. [Online]. Available: https://doi.org/10.48550/arXiv.1512.04456

[9] I.-Y. Song and Y. Zhu, "Big data and data science: what should we teach?" *Expert Systems*, vol. 33, no. 4, pp. 364–373, 2016. [Online]. Available: https://doi.org/10.1111/exsy.12130

[10] T. Munasinghe and A. Svirsky, "Engaging Students in Data Literacy: Lessons Learned from Data Intensive Classrooms," in *Companion Publication of the 13th ACM Web Science Conference 2021*, ser. WebSci '21 Companion. New York, NY, USA: Association for Computing Machinery, Jun. 2021, pp. 40–43. [Online]. Available: https://doi.org/10.1145/3462741.3466665

[11] A. Schwab-McCoy, C. M. Baker, and R. E. Gasper, "Data Science in 2020: Computing, Curricula, and Challenges for the Next 10 Years," *Journal of Statistics and Data Science Education*, vol. 29, no. sup1, pp. S40–S50, Jan. 2021. [Online]. Available: https://doi.org/10.1080/10691898.2020.1851159

[12] C. Baek, D. Saito-Stehberger, S. Jacob, A. Nam, and M. Warschauer, "Computer science framework to teach community-based environmental literacy and data literacy to diverse students," Oct. 2023, arXiv:2309.14098 [cs.CY]. [Online]. Available: https://doi.org/10.48550/arXiv.2309.14098

[13] Y.-T. Lin, M.-T. Wang, and C.-C. Wu, "Design and Implementation of Interdisciplinary STEM Instruction: Teaching Programming by Computational Physics," *The Asia-Pacific Education Researcher*, vol. 28, no. 1, pp. 77–91, Feb. 2019. [Online]. Available: https://doi.org/10.1007/s40299-018-0415-0

[14] J. P. Gibson and T. Mourad, "The growing importance of data literacy in life science education," *American Journal of Botany*, vol. 105, no. 12, pp. 1953–1956, Dec. 2018. [Online]. Available: https://doi.org/10.1002/ajb2.1195

[15] R. Luckin and W. Holmes, "Intelligence Unleashed: An argument for AI in Education," UCL Knowledge Lab, London, UK, Report, Feb. 2016. [Online]. Available: https://www.pearson.com/content/dam/corporate/global/pearson-dot-com/files/innovation/Intelligence-Unleashed-Publication.pdf

[16] M. Moundridou, N. Matzakos, and S. Doukakis, "Generative AI tools as educators' assistants: Designing and implementing inquiry-based lesson plans," *Computers and Education: Artificial Intelligence*, vol. 7, p. 100277, Dec. 2024. [Online]. Available: https://doi.org/10.1016/j.caeai.2024.100277

[17] H. Shi and V. Aryadoust, "A systematic review of AI-based automated written feedback research," *ReCALL*, vol. 36, no. 2, pp. 187–209, May. 2024. [Online]. Available: https://doi.org/10.1017/S0958344023000265

[18] A. Farazouli, "Automation and Assessment: Exploring Ethical Issues of Automated Grading Systems from a Relational Ethics Approach," in *Framing Futures in Postdigital Education: Critical Concepts for Data-driven Practices*, A. Buch, Y. Lindberg, and T. Cerratto Pargman, Eds. Cham: Springer Nature Switzerland, 2024, pp. 209–226. [Online]. Available: https://doi.org/10.1007/978-3-031-58622-4_12

[19] E. Hegarty-Kelly and A. Mooney, “Analysis of an automatic grading system within first year Computer Science programming modules,” in *Proceedings of the 5th Conference on Computing Education Practice*. Association for Computing Machinery, Jan. 2021, pp. 17–20. [Online]. Available: https://doi.org/10.1145/3437914.3437973

[20] O. Onesi-Ozigagun, Y. J. Ololade, N. L. Eyo-Udo, and D. O. Ogundipe, “Revolutionizing education through AI: A comprehensive review of enhancing learning experiences,” *International Journal of Applied Research in Social Sciences*, vol. 6, no. 4, pp. 589–607, Apr. 2024. [Online]. Available: https://doi.org/10.51594/ijarss.v6i4.1011

[21] M.-L. How and W. L. D. Hung, “Educing AI-Thinking in Science, Technology, Engineering, Arts, and Mathematics (STEAM) Education,” *Education Sciences*, vol. 9, no. 3, p. 184, Sep. 2019. [Online]. Available: https://doi.org/10.3390/educsci9030184

[22] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, “Systematic review of research on artificial intelligence applications in higher education – where are the educators?” *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, p. 39, Oct. 2019. [Online]. Available: https://doi.org/10.1186/s41239-019-0171-0

[23] A. R. Ellis and E. Slade, “A New Era of Learning: Considerations for ChatGPT as a Tool to Enhance Statistics and Data Science Education,” *Journal of Statistics and Data Science Education*, vol. 31, no. 2, pp. 128–133, May. 2023. [Online]. Available: https://doi.org/10.1080/26939169.2023.2223609

[24] K. C. Siontis, Z. I. Attia, S. J. Asirvatham, and P. A. Friedman, “ChatGPT hallucinating: can it get any more humanlike?” *European Heart Journal*, vol. 45, no. 5, pp. 321–323, Feb. 2024. [Online]. Available: https://doi.org/10.1093/eurheartj/ehad766

[25] P. Limna, T. Kraiwanit, K. Jangjarat, P. Klayklung, and P. Chocksathaporn, “The use of ChatGPT in the digital era: Perspectives on chatbot implementation,” *Journal of Applied Learning and Teaching*, vol. 6, no. 1, pp. 64–74, May. 2023. [Online]. Available: https://doi.org/10.37074/jalt.2023.6.1.32

[26] D. Baidoo-Anu and L. O. Ansah, “Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning,” vol. 7, no. 1, pp. 52–62, Dec. 2023. [Online]. Available: https://doi.org/10.61969/jai.1337500

[27] X. Tu, J. Zou, W. Su, and L. Zhang, “What Should Data Science Education Do With Large Language Models?” *Harvard Data Science Review*, vol. 6, no. 1, Jan. 2024. [Online]. Available: https://doi.org/10.1162/99608f92.bff007ab

[28] J. H. Woerner, A. P. Turtova, and A. S. Lang, “Transformative Potentials and Ethical Considerations of AI Tools in Higher Education: Case Studies and Reflections,” in *SoutheastCon 2024*, Mar. 2024, pp. 510–515. [Online]. Available: https://doi.org/10.1109/SoutheastCon52093.2024.10500042

[29] R. Yilmaz and F. G. Karaoglan Yilmaz, “Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning,” *Computers in Human Behavior: Artificial Humans*, vol. 1, no. 2, p. 100005, Aug. 2023. [Online]. Available: https://doi.org/10.1016/j.chbah.2023.100005

[30] C. K. Y. Chan and W. Hu, “Students’ voices on generative AI: perceptions, benefits, and challenges in higher education,” *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 43, Jul. 2023. [Online]. Available: https://doi.org/10.1186/s41239-023-00411-8

[31] C. K. Y. Chan and K. K. W. Lee, “The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers?” *Smart Learning Environments*, vol. 10, no. 1, p. 60, Nov. 2023. [Online]. Available: https://doi.org/10.1186/s40561-023-00269-3

[32] Y. Dai, “Why students use or not use generative AI: Student conceptions, concerns, and implications for engineering education,” *Digital Engineering*, p. 100019, Nov. 2024. [Online]. Available: https://doi.org/10.1016/j.dte.2024.100019

[33] A. R. Malik, Y. Pratiwi, K. Andajani, I. W. Numertayasa, S. Suharti, A. Darwis, and Marzuki, “Exploring Artificial Intelligence in Academic Essay: Higher Education Student’s Perspective,” *International Journal of Educational Research Open*, vol. 5, p. 100296, Dec. 2023. [Online]. Available: https://doi.org/10.1016/j.ijedro.2023.100296

[34] P. Gandhi, S. Khanna, and S. Ramaswamy, “Which Industries Are the Most Digital (and Why)?” *Harvard Business Review*, Apr. 2016. [Online]. Available: https://hbr.org/2016/04/a-chart-that-shows-which-industries-are-the-most-digital-and-why

[35] S. J. Christian, F. Ozis, K. A. Flanigan, J. D. Moore, and G. J. Wang, “Board 36: Case study: Sequential development of sensing skills in a civil and environmental engineering curriculum,” in *2023 ASEE Annual Conference & Exposition*, Jun. 2023. [Online]. Available: https://doi.org/10.18260/1-2--43007

[36] Carnegie Mellon University, “Strategies for creating a positive classroom climate,” https://www.cmu.edu/teaching/designteach/teach/classroomclimate/strategies/index.html, n.d., accessed: 2025-04-23.

[37] W. Cain, “Prompting change: exploring prompt engineering in large language model ai and its potential to transform education,” *TechTrends*, vol. 68, no. 1, pp. 47–57, 2024.

[38] N. Knoth, A. Tolzin, A. Janson, and J. M. Leimeister, “Ai literacy and its implications for prompt engineering strategies,” *Computers and Education: Artificial Intelligence*, vol. 6, p. 100225, 2024.

[39] A. Garg and R. Rajendran, “The impact of structured prompt-driven generative ai on learning data analysis in engineering students.” in *CSEDU (2)*, 2024, pp. 270–277.

[40] Y. Shen, X. Ai, A. G. Soosai Raj, R. J. Leo John, and M. Syamkumar, “Implications of chatgpt for data science education,” in *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, 2024, pp. 1230–1236.