ISSN: 2473-4901
v11 n6312

# Bridging Data Analytics Education With Generative AI: A Study of ChatGPT-Assisted Learning in Data Analytics Using Python

Mandy Dang
Mandy.Dang@nau.edu

Yulei Gavin Zhang
Gavin.Zhang@nau.edu

Yiyan Stella Li
Yiyan.Li@nau.edu

Howard Qi
Hao.Qi@nau.edu

The W. A. Franke College of Business
Northern Arizona University
Flagstaff, AZ 86011, USA

Xihui "Paul" Zhang
xzhang6@una.edu
Sanders College of Business and Technology
University of North Alabama
Florence, AL 35632, USA

## Abstract

The rapid advancement of generative AI tools, particularly large language models like ChatGPT, presents new opportunities for enhancing data analytics education. This study explores the integration of ChatGPT as a supplementary learning aid in a business data analytics course focused on data mining and machine learning algorithms. To support students with diverse programming backgrounds, three ChatGPT-assisted Python labs were introduced alongside traditional tools. A post-course survey completed by 260 students revealed overall positive perceptions across five key learning constructs. Comparisons between undergraduate and graduate students showed limited differences, with graduates reporting higher task-technology fit and undergraduates reporting higher perceived learning performance. Among undergraduates, female students reported significantly higher ratings across all dimensions, while gender differences among graduates were limited. These findings suggest ChatGPT was well-received across student groups, though perceived benefits varied by academic level and gender. The study offers empirical insights into integrating generative AI tools into analytics education alongside traditional instruction.

**Keywords:** Generative AI, ChatGPT, Data analytics education, Data mining, Python programming

 ISSN: 2473-4901
v11 n6312

# Bridging Data Analytics Education With Generative AI: A Study of ChatGPT-Assisted Learning in Data Analytics Using Python

*Mandy Dang, Yulei Gavin Zhang, Yiyan Stella Li, Howard Qi, and Xihui "Paul" Zhang*

## 1. INTRODUCTION

In recent years, generative artificial intelligence (GenAI) (Corchado et al., 2023) has experienced unprecedented growth, emerging as a transformative force across multiple sectors. The rapid advancements in large language models (LLMs) (Corchado et al., 2023; Myers et al., 2024) and multimodal systems have enabled GenAI to be widely deployed in various domains, such as healthcare, finance, law, and education (Yang et al., 2024). For instance, in the financial domain, GenAI has been developed and adopted in enhancing tasks such as financial modeling, risk assessment, fraud detection, and customer service (Remolina, 2024). In addition, GenAI has also been utilized in the law sector to streamline legal research, automate document analysis, and improve case management, leading to increased efficiency and productivity (Wilkins, 2024). In education, GenAI has facilitated personalized learning experiences and adaptive teaching methods, contributing to improved student engagement and learning outcomes (Denny et al., 2024).

Alongside advancements in GenAI, the field of data analytics has also experienced significant growth in recent years, becoming an essential component of modern business operations. Organizations across various industries increasingly rely on data-driven decision-making to enhance efficiency, competitiveness, and innovation. This surge in demand has led to a considerable expansion of data analytics programs in higher education. Many business schools across the US have responded by developing interdisciplinary curricula that combine critical technical skills with essential business knowledge, preparing graduates to meet the evolving needs of the industry (Olsen et al., 2022).

Building on both perspectives, this study explores the synergy between GenAI and data analytics education. While a substantial body of empirical research has recently emerged analyzing the impact of GenAI tools, primarily in the domains of foreign language education and programming, relatively fewer studies have focused on data analytics. Therefore, this study specifically aims to examine how GenAI tools, such as ChatGPT (Corchado et al., 2023; OpenAI, 2022), can be leveraged to assist students in overcoming technical challenges associated with data analytics tasks, particularly within the domain of data mining. By enabling natural language-driven programming support, GenAI has the potential to lower barriers for students with limited coding experience, thereby enhancing their engagement with complex analytical methods. Accordingly, this study presents and discusses the integration of ChatGPT in a data analytics course to support student learning of data mining algorithms using Python.

The paper is organized as follows. In Section 2, we review the related literature on the use of GenAI in education. Section 3 describes the design of the learning tasks. Section 4 presents the data analysis based on the survey results. Finally, Section 5 discusses the research contributions and implications, and the study's limitations and suggestions for future research.

## 2. RELATED LITERATURE

In recent years, a growing body of research has emerged examining the influence of GenAI tools on higher education (Xia et al., 2024). As the adoption of large language models and related technologies accelerates, scholars have increasingly investigated both the positive and negative impacts of these tools on teaching, learning, and academic integrity (Denny et al., 2024; Kakhki et al., 2024). Some studies focus on broader discussions of how GenAI alters educational practices and student engagement, highlighting opportunities for personalized learning and concerns about misuse and overreliance (Gill et al., 2024; Kakhki et al., 2024). Others take a more focused approach by exploring the practical application of specific GenAI technologies, such as ChatGPT and other language models, in supporting teaching tasks, improving student writing, and assisting with coding and data analysis (Kasneci et al., 2023).

For instance, as a comprehensive guidance study, Kakhki et al. (2024) investigated the affordances of ChatGPT in higher education and examined how AI technologies could reshape learning

ISSN: 2473-4901
v11 n6312

functions within academic institutions. Using a grounded theory approach, the study analyzed academic panel discussions to examine perceptions of ChatGPT and related AI tools in higher education. The study presents its findings through a framework consisting of four categories of affordances: (1) mitigating challenges in traditional learning environments, (2) enhancing effective educational practices, (3) transforming traditional learning approaches, and (4) negatively impacting current effective educational practices.

As another guidance paper, Gill et al. (2024) discussed the wide range of benefits that ChatGPT could bring to the education sector, as well as its potential risks. The authors categorized the transformative effects of ChatGPT on modern education into six dimensions: (1) educating with ChatGPT, (2) important ethics, (3) transforming online education, (4) higher education risks, (5) treatment required immediately, and (6) potential challenges.

In addition to examining the overall impact of ChatGPT in higher education, a recent body of research has emerged investigating how ChatGPT can assist student learning across various domains, such as foreign language acquisition (Koraishi, 2023; Warschauer et al., 2023), computer programming (Kosar et al., 2024; Yilmaz & Yilmaz, 2023), mathematics (Sánchez-Ruiz et al., 2023), business writing (Kétyi et al., 2025), data science (Zheng, 2023), and data analytics (Zhong & Kim, 2024).

For instance, Warschauer et al. (2023) investigated the use of AI-generated text, including ChatGPT, in supporting second language (L2) English writing. Through surveys and interviews with students and instructors, the study identified key benefits such as reducing writing anxiety, providing immediate feedback, and offering model texts to enhance L2 learning. However, it also noted risks, including overreliance on AI outputs, inaccuracies, and diminished critical thinking and independent writing skills. The authors emphasize the need for structured pedagogical strategies to ensure that AI tools are used as a complement, not a replacement, to traditional L2 writing instruction.

Similarly, Koraishi (2023) investigated the application of ChatGPT in English as a Foreign Language (EFL) education, with an emphasis on material development and assessment. The study demonstrated how ChatGPT can aid teachers by generating customized reading texts, adapting materials across different proficiency levels, integrating targeted vocabulary, and creating assessments and lesson plans. These capabilities have been shown to reduce teacher workload and provide more personalized learning experiences. However, the author also emphasizes the need for teacher supervision, as the tool occasionally produces inaccuracies or content requiring refinement.

In another study examining ChatGPT's role in computer science education, Kosar et al. (2024) investigated the impact of ChatGPT on novice programmers in an undergraduate object-oriented programming course. The study involved 182 students, divided into a control group and an experimental group using ChatGPT for programming assignments. Results showed no significant performance differences between the two groups, suggesting ChatGPT, when properly guided, does not compromise or improve learning performance. The authors note that ChatGPT can reduce frustration and provide immediate assistance, but emphasize that structured instructional strategies are essential to prevent overreliance and ensure the development of critical coding skills.

In another study, Yilmaz and Yilmaz (2023) explored student perceptions of ChatGPT as a learning tool in an Object-Oriented Programming II course involving 41 undergraduates over eight weeks. Students highlighted benefits such as rapid responses, assistance with debugging, improved problem-solving skills, and increased confidence. However, some noted risks of overreliance, occasional inaccuracies, and reduced independent thinking. The authors conclude that, with proper instructional design and supervision, ChatGPT can be a valuable supplement to programming education.

In mathematics education, Sánchez-Ruiz et al. (2023) evaluated the impact of ChatGPT in a blended learning mathematics course for engineering students. Students reported that ChatGPT supported their understanding of concepts and problem-solving, though trust was higher for theoretical explanations than numerical accuracy. In business writing education, Kétyi et al. (2025) explored ChatGPT's use for drafting business letters and dialogues, with students reporting improved writing efficiency and confidence alongside concerns about overreliance and content accuracy. As to data science education, Zheng (2023) examined ChatGPT's use in a graduate-level course and found that it aided learning in areas such as coding explanations and tool recommendations, while limitations remained in fostering critical thinking

ISSN: 2473-4901
v11 n6312

and problem-solving without human guidance.

However, to the best of our knowledge, relatively few studies have specifically investigated the role of GenAI tools, such as ChatGPT, in supporting student learning in data analytics, particularly in relation to data mining algorithms.

Through an extensive review of the literature, we identified one relevant study conducted by Zhong and Kim (2024). Their study examined the use of ChatGPT to support business students in learning logistic regression in R. By guiding students through data preparation, model building, and evaluation using AI-generated code within the CRISP-DM framework, the study aimed to help lower technical barriers and enhance students' data analytics and problem-solving skills.

Our study differs from theirs in three key ways: (1) we focus on multiple data mining algorithms rather than solely on logistic regression; (2) our course uses Python instead of R, both of which are widely adopted programming languages in business analytics education; and (3) while their study presents a teaching case with detailed assignment information, our study also provides empirical results to demonstrate the effectiveness of the proposed instructional design.

## 3. STUDY DESIGN

### The Data Analytics Course

Our data analytics course is a technical course focused on teaching students data mining and machine learning techniques and their applications in solving business problems. The major data mining algorithms covered include linear regression, logistic regression, association analysis, k-nearest neighbors (k-NN), decision trees, artificial neural networks, and clustering. The course is offered to both senior undergraduate and graduate students. Specifically, it is a required course for all business analytics, information systems, and marketing undergraduate majors, and an elective for other majors within the College of Business. At the master's level, it is a required course for business analytics students and an elective available to students from other colleges, such as engineering, including those majoring in computer science, information technology, electrical engineering, and others. Both undergraduate and graduate students use the same learning materials (textbooks and lectures) and are assessed using the same tools (weekly quizzes and lab projects). However, master's students are also required to complete a more comprehensive, term-long project, which is not required for undergraduates.

Given the diverse academic backgrounds of students enrolled in the course, not all students possess prior programming experience. To accommodate this variability and ensure equitable learning outcomes, the course has traditionally incorporated RapidMiner, a user-friendly data analysis software platform (https://academy.rapidminer.com/), for regular lab projects. RapidMiner's intuitive drag-and-drop interface allows students to focus on understanding data mining concepts and workflows without the steep learning curve associated with coding. However, we also recognize the value of providing students with exposure to data analytics using programming languages, such as Python, to broaden their skill set and prepare them for industry expectations. While software tools like RapidMiner offer accessibility and ease of use, they can limit students' flexibility in handling complex data tasks or customizing algorithms. Conversely, programming approaches offer greater control, scalability, and opportunities for problem-solving, but often associated with a higher initial learning barrier. Therefore, in our course, we aim to combine both perspectives to leverage the strengths of RapidMiner in enabling students to quickly grasp the fundamental concepts of data mining algorithms, while also providing them with experience in performing data analytics using Python.

To implement this idea, we decided to leverage the capabilities of a GenAI tool, specifically ChatGPT, to support student learning and engagement with programming-based data analytics. We designed a series of three targeted lab sessions, distributed throughout the semester, to provide students with progressive exposure to data analytics using Python. These labs were strategically placed at the beginning, middle, and end of the course to gradually introduce students to Python-based data analytics while reinforcing key data mining concepts. The first ChatGPT-assisted Python data analytics lab was conducted early in the semester and focused on multiple linear regression. The second lab took place in the middle of the semester and addressed association analysis, while the last lab, conducted towards the end of the semester, focused on artificial neural networks. We selected these three algorithms because they were introduced at different points in the course and, more importantly, they represent distinct categories of data mining algorithms.

 ISSN: 2473-4901
v11 n6312

To design the three ChatGPT-assisted Python labs, we utilized Google Colab, a cloud-based programming environment that allows users to write and execute Python code. Following the approach of Zhong and Kim (2024), we provided students with structured ChatGPT prompts to guide them in generating Python code for data analytics tasks step by step — from data understanding and exploration to model development and results interpretation. Students were instructed to use two web browsers during each lab session: one to access ChatGPT and input the prompts to obtain relevant code, and the other to open Google Colab, where they would paste, run, and explore the code and interpret the results. In addition to running the code, students were required to read and understand the detailed explanations of the logic and functions that ChatGPT provided alongside the generated Python code. They were also encouraged to make necessary adjustments and minor modifications to the prompts to ensure the successful completion of the lab tasks.

Due to space limitations, we present only the final ChatGPT-assisted Python data analytics lab below, with the hope that it offers inspirational insights for educators in data analytics.

**ChatGPT-Assisted Python Data Analytics Lab on Artificial Neural Networks**

**Dataset:** The dataset used in this lab is the Diagnostic Wisconsin Breast Cancer Database, which contains 30 features and 569 instances. The features are derived from a digitized image of a fine needle aspirate (FNA) of a breast mass and describe various characteristics of the cell nuclei present in the image.

**Part 1: Data understanding and exploration**

1) Import the data file onto Google Colab.
2) Import the necessary Python libraries for data analysis and visualization. Prompt: ***Import the libraries for data analysis and visualization.***
3) Read data from the data file. Prompt: ***Read data from the data file.***
4) After loading the data, describe the dataset step by step as follows:
    a. Prompt: ***Display the number of records and the number of attributes in the dataset.***
    b. Prompt: ***Display the names of the attributes in the dataset.***
    c. Prompt: ***Calculate the minimum, maximum, average, and standard deviation of each attribute in the dataset.***
5) Check for missing values in the dataset. Prompt: ***Check if there is any missing value in each column of the dataset.***
6) Explore the distribution of the dependent variable on a histogram. Prompt: ***Visualize the dependent variable on a histogram.***
7) Visualize the data using a parallel plot based on all features, excluding the dependent variable. Use color to differentiate between the two categories, based on the dependent variable. Prompt: ***Visualize the data using a parallel coordinates plot for all features, excluding the dependent variable. Use color to differentiate between the two categories based on the dependent variable.***
8) Calculate the correlation values between each feature and the dependent variable, and sort the results in descending order. Prompt: ***Calculate correlation values between each feature and the dependent variable, and sort the results in descending order.***
9) To further explore the relationships across variables, visualize the correction matrix on a heat map. Prompt: ***Create a heat map based on the correlation matrix.***

**Part 2: Model development and evaluation**

Part 2 focuses on creating the artificial neural network model. To do this, students first need to split the entire dataset into a training dataset (containing 75% of the data points) and a testing dataset (containing the remaining 25%). They will then use the training dataset to build the model and the testing dataset to evaluate the model's performance.

1) Split the entire dataset into a training dataset (containing 75% of the data points) and a testing dataset (containing the remaining 25%). Prompt: ***Split the entire dataset into a training dataset (containing 75% of the data points) and a testing dataset (containing the remaining 25%).***
2) Use the training dataset to build an artificial neural network model with a total of 3 layers: an input layer, a hidden layer, and an output layer. The input layer should have 30 nodes, and the output layer should have 1 node. The hidden layer is set to have 16 nodes in this example. Prompt: ***Use the training dataset to build an artificial neural network model with a total of 3 layers: an input layer, a hidden layer, and an output layer. The input layer should have 30 nodes, and the output layer should have 1 node. Make the***

 

 ISSN: 2473-4901
v11 n6312

***hidden layer with 16 nodes.***

3) Use the testing dataset to evaluate the model's performance. Calculate the confusion matrix and accuracy. Prompt: ***Use the testing dataset to evaluate the model's performance. Calculate the confusion matrix and accuracy.***

**Part 3: Further exploration with model comparisons**

To gain an in-depth understanding through further exploration, students are asked to use the same training and testing datasets to create another classifier using the decision tree algorithm and evaluate its performance. They are then asked to compare which algorithm, the artificial neural network or the decision tree, yields a higher accuracy. After that, they are asked to further explore a third model — a 4-layer artificial neural network model — to once again compare the performance across the three models.

1) Create a decision tree model using the same training dataset, and then evaluate the model's performance using the same testing dataset. Prompt: ***Create a decision tree model using the same training dataset, and then evaluate the model's performance using the same testing dataset. Calculate the confusion matrix and accuracy.***
2) **Discussion question:** Based on the results from the two models, which one has a higher accuracy on this dataset: the artificial neural network or the decision tree model?
3) Using the same training dataset, build another artificial neural network model with a total of 4 layers: an input layer, two hidden layers, and an output layer. The input layer should have 30 nodes, and the output layer should have 1 node. In this example, use 20 nodes for the first hidden layer and 16 nodes for the second hidden layer. Then, use the same testing dataset to evaluate its performance. Prompt: ***Use the same training dataset to build another artificial neural network model with a total of 4 layers: an input layer, two hidden layers, and an output layer. The input layer should have 30 nodes, and the output layer should have 1 node. Make the first hidden layer with 20 nodes and the second one with 16 nodes. Then, evaluate the model's performance using the same testing dataset. Calculate the confusion matrix and accuracy.***
4) **Discussion question**: So far, we have created 3 models: a 3-layer artificial neural network, a decision tree, and a 4-layer artificial neural network. Which one has the highest accuracy on this dataset, and what is that accuracy value?
5) **Discussion question**: What insights have you gained from this lab and the comparisons among the performances of models?

## 4. EMPIRICAL ASSESSMENT

To assess students' perceptions of the ChatGPT-assisted Python data analytics labs and their effectiveness, we conducted a survey. After obtaining IRB approval, a survey invitation was sent to all students enrolled in different sections of the course approximately one week before the end of the semester, after the completion of the final lab. Participation was voluntary, and a small amount of extra credit (approximately 1.5% of the total possible course points) was offered as an incentive for completing the survey. In total, the survey invitation was sent to 410 students; 264 students participated, and 260 completed the survey, yielding a response rate of 64.39% and a completion rate of 63.41%. The average age of participants who completed the survey was 23.45 years. Among them, 134 identified as male, 121 as female, 3 as non-binary, and 2 preferred not to specify their gender. Additionally, 117 were undergraduate students and 143 were graduate students.

In this study, we evaluated the effectiveness of our instructional design, which incorporated ChatGPT to assist students in learning data analytics using Python, based on five theoretical constructs. These constructs have been widely adopted in the information systems education literature to assess student learning. They include learning effort expectancy, task-technology fit, difficulty management, learning satisfaction, and perceived learning performance.

For four of these constructs, we adopted measurement items from existing literature with modifications to fit the context of our study. Specifically, the measure for learning effort expectancy was adapted from Sabeh (2024); the measure for task-technology fit was adapted from Chen et al. (2023). To assess difficulty management, we drew on the conceptual framework of Wall and Knapp (2014) and developed our own items. The measure for learning satisfaction was adopted from Almulla (2024), and the measure for perceived learning performance was adapted from Islam (2013). All questionnaire items were rated on a 7-point Likert

ISSN: 2473-4901
v11 n6312

scale, ranging from 1 for "strongly disagree" to 7 for "strongly agree." Appendix A lists all the measurement items.

Table 1 summarizes the descriptive statistics of student responses across the five measured constructs. Overall, the results indicate a positive evaluation of the ChatGPT-assisted Python data analytics labs. The mean scores for all assessment dimensions exceeded the midpoint value of 4 on the 7-point Likert scale, suggesting generally favorable student perceptions. Specifically, perceived learning performance (M = 5.578, SD = 1.267) and learning effort expectancy (M = 5.507, SD = 1.251) received the highest ratings, indicating that students felt ChatGPT meaningfully supported their learning process and reduced the effort needed to conduct data analytics tasks using Python. The other dimensions also demonstrated strong positive responses, with means ranging from 5.341 to 5.501 across task-technology fit, difficulty management, and learning satisfaction. These findings collectively indicate the overall effectiveness of integrating ChatGPT as a supplementary instructional tool in data analytics education.

| Assessment | Mean/Std dev |
|---|---|
| Learning Effort Expectancy (EE) | 5.507/1.251 |
| Task-Technology Fit (TTF) | 5.341/1.342 |
| Difficulty Management (DM) | 5.501/1.261 |
| Learning Satisfaction (SAT) | 5.479/1.323 |
| Perceived Learning Performance (PLP) | 5.578/1.267 |

**Table 1: Overall Statistics (N=260)**

To further assess the effectiveness of our instructional design, we also compared the results between undergraduate and graduate students. Within each group, we then conducted additional comparisons between male and female students. Table 2 presents the results of the comparison between undergraduate and graduate students across the five measured dimensions.

| | Under (N=117) M/SD | Master (N=143) M/SD | p-value |
|---|---|---|---|
| EE | 5.498/1.092 | 5.514/1.368 | 0.418 |
| TTF | 5.242/1.338 | 5.422/1.342 | 0.031* |
| DM | 5.507/1.185 | 5.497/1.321 | 0.453 |
| SAT | 5.513/1.276 | 5.451/1.360 | 0.454 |
| PLP | 5.664/1.134 | 5.508/1.363 | 0.044* |

Note: * indicates statistically significant at 0.05.

**Table 2: Undergraduate vs. Graduate Students**

As shown in Table 2, overall, both groups reported similarly positive perceptions of the ChatGPT-assisted labs, with all mean scores exceeding the midpoint of the 7-point Likert scale. Statistically significant differences were found in task-technology fit (TTF) and perceived learning performance (PLP). Graduate students rated TTF significantly higher than undergraduates (M = 5.422 vs. M = 5.242; $p$ = 0.031), indicating a greater perceived alignment between ChatGPT and their learning tasks. Conversely, undergraduates reported significantly higher perceived learning performance (M = 5.664 vs. M = 5.508; $p$ = 0.044). These contrasting results may reflect differences in academic experience and learning needs. Graduate students, often possessing more familiarity with complex tools and independent learning strategies, may have found ChatGPT to align more effectively with their task requirements, resulting in higher task-technology fit scores. In contrast, undergraduate students, who generally have less prior experience with data mining and programming, may have perceived greater learning gains from the structured support provided by ChatGPT, contributing to higher perceived learning performance. No significant differences were observed for learning effort expectancy, difficulty management, or learning satisfaction between the two groups.

We further examined gender differences within both the undergraduate and graduate student groups. Tables 3 and 4 present the results for each group, respectively.

| | Under-male (N=59) M/SD | Under-female (N=55) M/SD | p-value |
|---|---|---|---|
| EE | 5.364/1.190 | 5.623/0.983 | 0.006* |
| TTF | 4.949/1.505 | 5.539/1.068 | <0.001* |
| DM | 5.181/1.328 | 5.842/0.930 | <0.001* |
| SAT | 5.233/1.429 | 5.800/1.032 | <0.001* |
| PLP | 5.463/1.206 | 5.867/1.039 | 0.001* |

Note: The 3 non-binary students were excluded in this table. * indicates statistically significant at 0.05.

**Table 3: Undergraduate Male vs. Female Students**

As illustrated in Table 3, female undergraduate students reported significantly higher scores than male undergraduate students on all dimensions, with all differences reaching statistical significance ($p$ < 0.05). Specifically, female undergraduate students rated learning effort expectancy (EE), task-technology fit (TTF),

ISSN: 2473-4901
v11 n6312

difficulty management (DM), learning satisfaction (SAT), and perceived learning performance (PLP) more positively than male undergraduate students. These findings suggest that female undergraduate students perceived greater benefits and support from the ChatGPT-assisted labs compared to their male peers. One possible explanation is that female undergraduate students, who may have had lower initial confidence or less prior exposure to programming tasks, found the structured and supportive nature of the ChatGPT-assisted approach particularly valuable. The AI-driven guidance may have helped reduce anxiety and increase engagement, contributing to their overall positive evaluation of the learning experience.

| | **Master-male (N=75) M/SD** | **Master-female (N=66) M/SD** | **p-value** |
|---|---|---|---|
| EE | 5.362/1.603 | 5.527/1.328 | 0.480 |
| TTF | 5.265/1.623 | 5.535/1.212 | 0.086 |
| DM | 5.293/1.571 | 5.611/1.190 | 0.041* |
| SAT | 5.194/1.671 | 5.587/1.183 | 0.023* |
| PLP | 5.361/1.579 | 5.586/1.279 | 0.191 |

Note: The 2 students who preferred not to specify their gender were excluded in this table. * indicates statistically significant at 0.05.

**Table 4: Graduate Male vs. Female Students**

As indicated in Table 4, while female graduate students reported higher mean scores on all dimensions, statistically significant differences were observed only for difficulty management (DM; $p$ = 0.041) and learning satisfaction (SAT; $p$ = 0.023). Female graduate students rated DM and SAT more positively, suggesting they experienced greater confidence in managing the technical demands of the labs and were more satisfied with the learning experience. No significant differences were found for learning effort expectancy, task-technology fit, or perceived learning performance. These results may indicate that at the graduate level, both male and female students possess relatively high baseline skills and comfort with learning technologies, minimizing gender disparities. However, female students may still have particularly benefited from the structured support and clear guidance provided by the ChatGPT-assisted labs, which could explain their higher satisfaction and better perceived ability to manage lab challenges.

## 5. CONCLUSIONS

### Research Contributions and Implications

This study makes several key contributions to the literature on data analytics education and the pedagogical application of GenAI tools. First, it advances our understanding of how GenAI tools can be systematically leveraged to support students with diverse academic and technical backgrounds in learning highly technical content such as data mining and machine learning using Python. Data analytics courses often present a significant learning barrier for students with limited programming experience. Our approach demonstrates that AI-assisted coding guidance, integrated within a supportive instructional framework, can reduce these barriers and provide an equitable learning experience. The positive student perceptions across learning effort, satisfaction, and perceived learning performance highlight the potential of GenAI to enhance access to complex analytics education.

Second, the study contributes a novel instructional design that differs from prior research. Building on earlier exploratory efforts, we implemented a staged design featuring three distinct ChatGPT-assisted labs spaced strategically throughout the semester. This design exposed students to progressively complex data mining algorithms (including multiple linear regression, association analysis, and artificial neural networks) rather than focusing on a single technique as in some earlier studies (e.g., Zhong & Kim, 2024). The inclusion of both traditional data mining tools (RapidMiner) and AI-supported Python coding offered students a dual learning pathway that accommodates varying skill levels while promoting deeper understanding of data analytics concepts and methodologies. This may serve as a practical framework to be adapted in other technical and business education contexts.

Third, our empirical findings offer actionable insights for data analytics educators. The survey results provide evidence that most students responded positively to the integrated instructional approach, with overall ratings above the scale midpoint across five learning dimensions. Additionally, subgroup analysis revealed meaningful differences between student groups, such as stronger positive perceptions among female undergraduate students. These insights suggest the importance of considering learner diversity when integrating AI tools into data analytics curricula. For example, instructors may consider providing differentiated levels of scaffolding, offering more structured guidance and example prompts for novice programmers, while encouraging more advanced students to engage in exploratory and independent coding tasks to maximize learning outcomes for all

ISSN: 2473-4901
v11 n6312

groups.

**Limitations and Future Research Directions**

This study has several limitations. First, it was conducted at a single institution, which may limit the generalizability of the findings. Second, the research relied on self-reported survey data, which may be subject to bias. Third, while the study focused on students' perceptions, actual learning performance and long-term skill retention were not measured. Future research should explore longitudinal impacts of GenAI-assisted learning and examine outcomes across diverse institutional and cultural contexts. Experimental studies comparing AI-assisted instruction with traditional methods could also provide deeper insights into the pedagogical value and limitations of GenAI in data analytics education. Additionally, this study compared undergraduate and graduate students; future research may further explore comparisons based on prior coding experience and prior GenAI exposure. Future research may also conduct in-depth interviews to gain deeper insights into students' learning experiences.

## 6. REFERENCES

Almulla, M. A. (2024). Investigating influencing factors of learning satisfaction in AI ChatGPT for research: University students' perspective. *Heliyon, 10*, e32220. https://doi.org/10.1016/j.heliyon.2024.e32220

Chen, J., Zhuo, Z., & Lin, J. (2023). Does ChatGPT play a double-edged sword role in the field of higher education? An in-depth exploration of the factors affecting student performance. *Sustainability, 15*(24), 16928. https://doi.org/10.3390/su152416928

Corchado, J. M., F, S. L., V, J. M. N., S., R. G., & Chamoso, P. (2023). Generative artificial intelligence: Fundamentals. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 12*(1), e31704. https://doi.org/10.14201/adcaij.31704

Denny, P., Gulwani, S., Heffernan, N. T., Käser, T., Moore, S., Rafferty, A. N., & Singla, A. (2024). Generative AI for education (GAIED): Advances, opportunities, and challenges. *arXiv*. https://arxiv.org/abs/2402.01580

Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., . . . Buyya, R. (2024). Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. *Internet of Things and Cyber-Physical Systems, 4*, 19–23. https://doi.org/10.1016/j.iotcps.2023.06.002

Islam, A. K. M. N. (2013). Investigating e-learning system usage outcomes in the university context. *Computers & Education, 69*, 387-399. https://doi.org/10.1016/j.compedu.2013.07.037

Kakhki, M. D., Oguz, A., & Gendron, M. (2024). Exploring the affordances of chatbots in higher education: A framework for understanding and utilizing ChatGPT. *Journal of Information Systems Education, 35*(3), 284-302. https://doi.org/10.62273/UIRX9922

Kasneci, E., Sessler, K., Kühn, S., Kasneci, G., & Bannert, M. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kétyi, A., Géring, Z., & Dén-Nagy, I. (2025). ChatGPT from the students' point of view – Lessons from a pilot study using ChatGPT in business higher education. *Society and Economy, 47*(1), 1–21. https://doi.org/10.1556/204.2024.00007

Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education & Technology, 3*(1), 55-72. https://langedutech.com/letjournal/index.php/let/article/view/48

Kosar, T., Ostojić, D., Liu, Y. D., & Mernik, M. (2024). Computer science education in ChatGPT era: Experiences from an experiment in a programming course for novice programmers. *Mathematics, 12*(5), 629. https://doi.org/10.3390/math12050629

Myers, D., Mohawesh, R., Chellaboina, V. I., Sathvik, A. L., Venkatesh, P., Ho, Y.-H., . . . Jararweh, Y. (2024). Foundation and large language models: Fundamentals, challenges, opportunities, and social impacts. *Cluster Computing, 27*, 1–26. https://doi.org/10.1007/s10586-023-04203-7

Olsen, T., Chudoba, K. M., & Dupin-Bryant, P. A. (2022). Examining trends in business analytics education from 2011 to 2021. *Journal of Information Systems Education, 33*(3), 232–244.

 ISSN: 2473-4901
v11 n6312

https://jise.org/Volume33/n3/JISE2022v33n3pp232-244.pdf

OpenAI. (2022). ChatGPT: Optimizing language models for dialogue. *OpenAI Blog*. https://openai.com/blog/chatgpt

Remolina, N. (2024). Generative AI in finance: Risks and potential solutions. *Law Ethics & Technology Review, 2*(1), 1-25. https://doi.org/10.55092/let20240002

Sabeh, H. N. (2024). What drives IT students toward ChatGPT? Analyzing the factors influencing students' intention to use ChatGPT for educational purposes. Paper presented at *the 21st International Multi-Conference on Systems, Signals & Devices* (SSD) (pp. 533–539). IEEE. https://doi.org/10.1109/SSD61670.2024.10548826

Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, a. E. (2023). ChatGPT challenges blended learning methodologies in engineering education: A case study in mathematics. *Applied Sciences, 13*(10), 6039. https://doi.org/10.3390/app13106039

Wall, J. D., & Knapp, J. (2014). Learning computing topics in undergraduate information systems courses: Managing perceived difficulty. *Journal of Information Systems Education, 25*(3), 245-259. https://aisel.aisnet.org/jise/vol225/iss243/248.

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for second language writers of English as a second or foreign language. *Journal of Second Language Writing, 62*, 101071. https://doi.org/10.1016/j.jslw.2023.101071

Wilkins, D. (2024). The legal profession in 2024: AI. *Harvard Law School*. https://hls.harvard.edu/today/harvard-law-expert-explains-how-ai-may-transform-the-legal-profession-in-2024/

Xia, Q., Weng, X., Ouyang, F., & Lin, T. J. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education, 21*(1), 1-20. https://doi.org/10.1186/s41239-024-00468-z

Yang, H., Xu, J., Zhao, L., & Wang, X. (2024). The rapid evolution and deployment of generative AI across industries: Challenges and opportunities. *arXiv*. https://arxiv.org/abs/2405.11029

Yilmaz, R., & Yilmaz, F. G. K. (2023). Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning. *Computers in Human Behavior: Artificial Humans, 1*, 100005. https://doi.org/10.1016/j.chbah.2023.100005

Zheng, Y. (2023). ChatGPT for teaching and learning: An experience from data science education. Paper presented at *the 24th Annual Conference on Information Technology Education* (SIGITE '23). https://doi.org/10.1145/3585059.3611431

Zhong, C., & Kim, J. B. (2024). Teaching case: Teaching business students logistic regression in R with the aid of ChatGPT. *Journal of Information Systems Education, 35*(2), 138–143. https://doi.org/10.62273/DYLI2468

ISSN: 2473-4901
v11 n6312

## APPENDIX A

### Measurement Items

Learning Effort Expectancy

1. Using ChatGPT to learn and understand data analytics in Python is easy for me.
2. I would describe my interaction with the ChatGPT interface for learning/understanding data analytics in Python as being clear and understandable.
3. I find ChatGPT easy to use for data analytics in Python.
4. It has been easy for me to become skilled at using ChatGPT for data analytics in Python.

Task-Technology Fit

1. ChatGPT is well-suited to help me with my learning tasks on data analytics in Python.
2. ChatGPT is a necessary support for my learning and understanding of data analytics in Python.
3. ChatGPT is an appropriate fit for my learning needs.

Difficulty Management

1. ChatGPT has helped make the difficulty of learning data analytics in Python manageable for me.
2. ChatGPT helps me manage the difficulty of learning data analytics in Python.
3. ChatGPT helps me effectively cope with the difficulty of learning data analytics in Python.

Learning Satisfaction

1. Overall, I am satisfied with my ChatGPT-facilitated learning experience.
2. Using ChatGPT enhances my level of satisfaction with learning data analytics in Python.
3. ChatGPT contributes significantly to my overall satisfaction with learning data analytics in Python.
4. I am satisfied with my learning outcomes and experiences using ChatGPT.

Perceived Learning Performance

1. With the support of ChatGPT, I can effectively accomplish my learning tasks on data analytics in Python.
2. With the support of ChatGPT, I can efficiently accomplish my learning tasks on data analytics in Python.
3. Overall, I am satisfied with my performance on my data analytics in Python lab projects.