The Role of LLM-Based Tools in Shaping Data Science Student Behavior

Felipe de S. Siqueira[1], **Claudio E. C. Campelo**[1]

[1]Systems and Computing Department – Federal University of Campina Grande (UFCG)
58429-900 – Campina Grande – PB – Brazil

felipe.siqueira@ccc.ufcg, campelo@dsc.ufcg.edu.br

Abstract. *This study evaluates the impact of tools based on Large Language Models (LLMs) in data science education, focusing on how reliance on these tools affects students' analytical and technical skills. The research examines two main aspects: students' perceptions of the tools' impact on their analytical abilities and their behavior during data science activities. The study involved 27 students who participated in a practical exercise using a chatbot and completed questionnaires to assess their perceptions of such tools. The results suggest that students with higher levels of reliance tend to have a more positive perception of the tools, although this perception does not necessarily correlate with better performance in analytical and visualization tasks. Moreover, students who actively explore alternative approaches and modify the generated code perform better in data clustering tasks. These findings suggest that LLM-based tools can support data science education, provided that students are encouraged to develop critical thinking and go beyond the solutions automatically generated by the models.*

1. Introduction

The growing use of Large Language Models (LLM) is increasingly prominent in various domains, including education. In data science courses, LLM-based tools such as ChatGPT[1] are commonly employed to support tasks related to data analysis and visualization. However, a critical issue in this context involves understanding how reliance on such tools influences students' development of technical skills, particularly in areas such as selecting clustering methods and designing effective visualizations. Investigating this impact is fundamental to optimizing the integration of emerging technologies into educational practices and ensuring their effective use in fostering student learning.

The Cognitive Load Theory [Sigolo and Casarin 2024] highlights that learning, cognitive flexibility, and decision-making directly depend on the limited processing capacity of working memory and on how cognitive resources are allocated during task performance. When cognitive load is distributed inadequately—whether due to intrinsic, extrinsic, or learner-related factors—the consolidation of information into long-term memory is compromised, reducing both learning quality and individual autonomy. In this sense, the inappropriate use of LLMs may hinder the development of essential skills such as critical thinking, mental flexibility, and problem-solving—fundamental abilities for data scientists—since the systematic outsourcing of cognitive effort to the tool diminishes the student's need to actively engage their own cognitive resources. The result is an

[1]ChatGPT - `https://chatgpt.com/`

DOI: 10.5753/sbie.2025.12715

artificial balance of cognitive load: while the extraneous load of the task decreases, the exercise of the intrinsic load necessary for meaningful learning is also limited. Thus, instead of fostering the expansion of cognitive structures, excessive dependence may crystallize a pattern of surface learning, hindering the transfer and flexible application of knowledge in new contexts [Sigolo and Casarin 2024].

Moreover, the convenience offered by these technologies may lead to excessive reliance, discouraging student engagement with traditional and interactive learning methods [Budhiraja et al. 2024, Liu and Gonzalez 2024] and reducing meaningful interaction between students and instructors [Budhiraja et al. 2024].

Recent research on active learning methodologies indicates that fostering students' autonomy and innovation depends on pedagogical practices that position them as protagonists of their formative process [Lopes et al. 2024]. In this learner-centered model, students are able to exercise their cognitive capacities more fully, promoting the meaningful internalization of concepts and the flexibility required for creative problem-solving.

Another concern is the impact of AI tools like ChatGPT on academic integrity. When misused, these technologies can facilitate plagiarism and cheating in assessments, ultimately lowering educational quality [Budhiraja et al. 2024, Baker and Smith 2023]. This is particularly detrimental for early-stage students, who may miss the opportunity to engage with simpler questions that are crucial for building a solid theoretical foundation. As LLMs continue to evolve and reshape the field of data science, educators and policymakers must anticipate future implications and adapt their teaching strategies accordingly [Malinka et al. 2023].

Despite these concerns, LLM adoption is rapidly growing among undergraduate students, while evidence of their specific impact on education remains limited. Most existing research focuses on areas such as programming and mathematics [Zheng 2023], leaving a gap in the literature regarding their role in data science education. To address this gap, we conducted an in-classroom experiment to empirically assess the impact of LLM use on the learning process in data science. This study offers insights into how these technologies can both support and hinder students when used as educational tools in this discipline.

Our experiment seeks to answer the following research questions:

RQ1 – How do students with varying levels of reliance on chatbots perceive their impact on analytical skills?

RQ2 – How do students with varying levels of reliance on chatbots behave when performing data visualization and clustering tasks?

2. Related Work

AI-based tools — especially LLMs — are increasingly used across fields and in education. In data-science courses, such tools commonly support analysis and visualization. A central question, however, is how reliance on these tools shapes students' skill development, particularly in technical choices such as selecting clustering methods and designing effective visualizations. Clarifying this impact is essential to optimize pedagogy and ensure students derive the most benefit from emerging technologies.

In this context, Zheng [Zheng 2023] investigated the use of ChatGPT specifically in data science education, highlighting the opportunities and challenges that arise when incorporating this tool into a data science course. The researchers evaluated ChatGPT's effectiveness in various scenarios, including code generation and supporting the understanding of data science concepts. Their findings indicate that the tool can be a valuable for learning technical tasks such as building machine learning models and explaining parameters. However, the study also pointed out limitations, such as excessive reliance on the tool, which can hinder the development of critical skills — an issue also addressed in the present work, which investigates the impact of ChatGPT on students analytical skills.

Another study [Shen et al. 2024] explored the impact of ChatGPT on data science activities, evaluating its performance in generating code for complex tasks across different levels of data science courses. The study highlighted the importance of prompt engineering to improve the quality of the generated responses, using techniques such as breaking problems into smaller steps and adding context about datasets. The authors noted that without such interventions, ChatGPT struggled with more advanced tasks like data analysis and machine learning, reinforcing the need for prompt adjustments and active engagement with generated code to achieve better educational outcomes.

The integration of LLM in software engineering education was explored in [Kirova et al. 2024]. The study emphasized that chat bots can be effective in routine tasks such as debugging and generating basic solutions, but also warned of the need to focus more on developing students' critical skills. According to the authors, excessive reliance on language models can impair deep conceptual learning, which is essential for solving complex problems and making informed decisions. The article also highlights the importance of training students to identify limitations in the generated responses, such as inaccuracies and biases, fostering a more reflective and critical approach to using these tools in educational settings.

A complementary study [Tu et al. 2023] explores the impact of LLMs on data science education, highlighting their potential to automate tasks like data cleaning, model development, and report generation. This automation shifts the focus of data scientists toward project management. The authors advocate for curricular changes emphasizing critical thinking, AI-assisted programming, and interdisciplinary skills, while warning that early reliance on LLMs may hinder the development of foundational conceptual understanding.

This study extends existing research by investigating how reliance on a LLM-based tool (i.e., ChatGPT) shapes student behavior, a dimension that has received little to no attention in prior work.

3. Methodology

This is a quasi-experimental study [Creswell and Creswell 2017], with a post-test design, analytical in nature and employing a quantitative approach, conducted with a convenience sample of 27 students from the Descriptive Data Science course in the undergraduate Computer Science program at the Federal University of Campina Grande. Although rigorous control is not possible in quasi-experimental studies, it is possible to observe what happens, when it happens, and to whom it happens, thereby facilitating cause-effect analyses [Gil 2008]. No sociodemographic data were collected to characterize the sample.

3.1. Experiment Design

The activity was designed to simulate real-world data analysis scenarios, incorporating challenges that require both technical and analytical skills. The setup was crafted to assess how students interact with LLMs when facing data science problems, particularly regarding analytical thinking, data visualization, and clustering techniques.

3.2. Preparatory Class

A 30-minute preparatory session (n=27) preceded the experiment, delivered as a slide-supported lecture to ensure a clear understanding of ChatGPT and its effective use in data-science tasks. It covered handling tabular data (manipulation, interpretation, analysis), interpreting model outputs, and crafting context-rich, precise prompts to improve relevance and accuracy, and concluded with an open Q&A.

3.3. Dataset Used

The practical activity used a diabetes dataset containing medical characteristics and test results, selected for its applicability to real-world health scenarios. One of its analytical challenges was the presence of biologically implausible zero values in variables like "bloodPressure," deliberately included to evaluate students' ability to detect and address inadequate or incomplete data — an essential data science skill. Additionally, two categorical columns with textual values were added to test whether students could identify notable features and understand the role of categorical data in analysis. These elements aimed to assess students' competence in recognizing and handling variables that may require special treatment or transformation prior to deeper investigation.

[2]

3.4. Practical Activity

The activity comprised three sequential tasks targeting analytical and technical skills: (1) data evaluation — students inspected a dataset with anomalies (e.g., biologically implausible zeros in `bloodPressure`), judged readiness, and decided on cleaning; (2) visualization —selected appropriate charts for categorical and binary variables to communicate results; and (3) clustering—normalized features with Min–Max, chose an algorithm, visualized clusters, assessed quality/interpretability, and examined each cluster's relation to diabetes to derive actionable insights. Evaluation was binary (0/1) using the following criteria: Q1.1—dataset not ready; Q1.2—removed `BloodPressure`=0; Q2.1—identified suitable categorical variable and plot; Q2.2—adequate plots for remaining variables; Q3.1—cluster visualization with ≥ 2 well-separated groups; Q3.2—explicit interpretation of each cluster.

3.5. Questionnaire

A post-activity questionnaire assessed the perceived impact of ChatGPT on data-science coursework. Items covered students' trust in model outputs and the tool's perceived influence on analytical problem-solving; self-reported reliance on LLM answers and effects on understanding of core concepts; the frequency of incorrect or inadequate responses; and ChatGPT's perceived contribution to solving the specific tasks in the activity.

[2]Supplementary materials are available at `osf.io/a8cw6`. They include: slide presentation, English version of the questionnaire, original Portuguese version, participant chat transcripts, collected CSV data, and the Python notebook used for analysis.

3.6. Results Evaluation

Questionnaire responses and ChatGPT interaction data were stored in CSV files, and students were grouped by self-reported reliance: high ("Totally/Highly") vs low ("Slightly/Very slightly"). For RQ1, we cross-analyzed reliance against perceived impact on analytical skills and computed correlations among reliance, perceived analytical impact, and performance on analytical tasks (answer justification, natural-language formulation). For RQ2, using the same groups, we related membership to confidence in ChatGPT responses, task correctness, and visualization/cluster quality, and explored interaction patterns to see how reliance shaped engagement and problem-solving.

4. Results and Discussion

This section presents and analyzes the results of the experiments conducted to answer the two guiding research questions outlined in Section 1.

4.1. Perception of Impact on Analytical Skills in Data Science Activities

To address RQ1 (How do students with different levels of reliance on ChatGPT or similar tools perceive its impact on their analytical skills?), data were collected regarding student performance during the activity, including the precision of their answers, the adequacy of their visualizations, their trust in the tool, correlations between trust and performance, whether they experimented with different code approaches and if they tested different methods or numbers of clusters. The analysis focused specifically on questions 1 and 3.2 of the practical activity. The correctness criteria in these questions were defined as follows:

- **Question 1**: Identifying that the data was not ready for analysis and justifying this conclusion based on the context provided in the activity.
- **Question 3.2**: Interpreting the results of the clustering process and being able to describe the composition of each group.

To explore how ChatGPT reliance influenced students' perception of its impact on their analytical skills, participants were divided into two groups:

- **High reliance Group**: Students who reported being "Totally" or "Highly" dependent on the tool.
- **Low reliance Group**: Students who considered themselves "Slightly" or "Very slightly" dependent.

Table 1 presents the perceptions of the students about ChatGPT's impact, classified as positive, indifferent, or negative for both groups, with percentages calculated based on the total of 27 participants. The goal of this analysis was to identify differences between groups and investigate whether reliance significantly affects students' perceived development of analytical abilities.

Students with a high reliance on ChatGPT were found to have a predominantly positive perception of the impact of the tool, with 80% indicating a positive influence. In contrast, among students with low reliance, only 62.5% reported a positive perception. Notably, the low-reliance group also shows a higher proportion of indifferent (25%) and negative (12.5%) responses, suggesting a more balanced or critical view of ChatGPT's impact on their analytical abilities. Despite these percentage differences, the Fisher–Freeman–Halton exact test for the 2×3 table (High/Low × Negative/Neutral/Positive) did not indicate a significant association (p = 0.809; Cramér's V = 0.19; n = 23).

Table 1. Perception of ChatGPT's influence on students' analytical skills, categorized by level of dependency on the tool, expressed as percentages based on a total of 27 students.

Level of dependency	Positive (%)	Indifferent (%)	Negative (%)
High	80.0	10.0	10.0
Low	62.5	25.0	12.5

During the analysis, it was observed that a more positive perception of ChatGPT's impact does not necessarily translate into higher performance on analytical tasks. The patterns by reliance level shown in Table 2 should be regarded as preliminary, as the small and the multiple-testing adjustment reduce statistical sensitivity and the self-reported measures may introduce noise; consistent with this, Fisher's exact tests detected no statistically significant differences between high- and low-reliance groups (Q1.1: p = 0.193; Q3.2: p = 1.000; FDR-adjusted q-values 0.386 and 1.000).

Table 2. Accuracy rates by dependency group.

Dependency level	Right Answer (Question 1) (%)	Right Answer (Question 3.2) (%)
High dependency	30.8	53.8
Low dependency	80.0	60.0

These proportions suggest that, for Question 3.2, participants with low reliance showed slightly higher accuracy (60.0%) than those with high reliance (53.8%). For Question 1, the difference was more pronounced (80.0% vs. 30.8%).

Table 3. Evaluation of ChatGPT's influence on students analytical skills and performance in two questions, categorized by level of dependency on the tool (Q1: $X^2(3)$=1,46; $p \approx 0,692$; V=0,232, Q3.2: $X^2(3)$=3,37; $p \approx 0,338$; V=0,353).

Dependency Level	Inf. Analytical Skills	Correct (Q1)	Correct (Q3.2)
High dependency	Very positively	50.0%	75.0%
High dependency	Positively	25.0%	37.5%
High dependency	Indifferent	0.0%	100.0%
Low dependency	Very positively	100.0%	0.0%
Low dependency	Positively	80.0%	60.0%
Low dependency	Indifferent	100.0%	100.0%
Low dependency	Negatively	50.0%	50.0%

Table 3 reveals a trend indicating that participants with low dependency on ChatGPT who feel indifferent to its influence tend to have high accuracy rates, suggesting they may be more autonomous or less influenced by ChatGPT's responses. Those with high dependency show greater variation in their accuracy rates, possibly indicating a stronger influence of ChatGPT on their decision-making and analytical processes. In order to better understand the relationships between ChatGPT dependency and students' performance on analytical tasks, a correlation matrix was calculated among four key variables: the level of dependency on ChatGPT, the perceived impact of the tool on analytical skills, accuracy in the justification of the first question, and accuracy in the natural language response to the fourth question. Figure 1 presents these correlations.

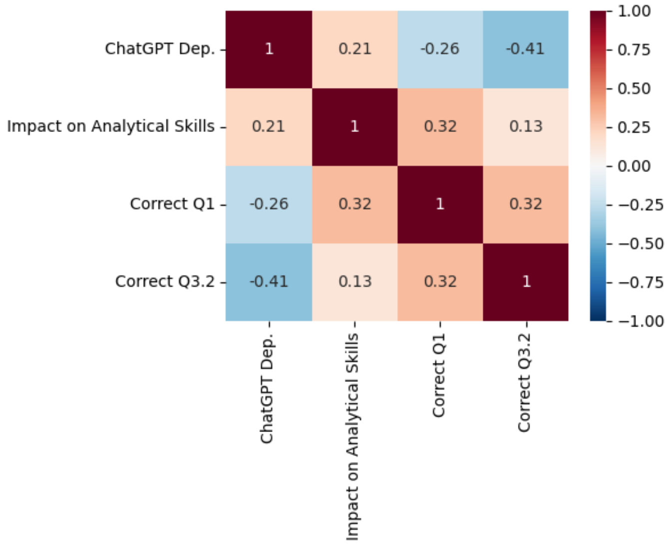


Figure 1. Correlation between dependency, perceived impact, and performance on the questions.

The correlation analysis reveals that dependency on ChatGPT has a negative relationship with students' performance on more analytical questions. The moderate negative correlation (-0.408) between ChatGPT dependency and accuracy on Question 3.2 (natural language responses) suggests that the more students rely on the tool, the worse their performance on tasks that require more detailed and complex explanations. Similarly, there is a weak negative correlation (-0.258) with accuracy on Question 1, which involves the justification of the first question, indicating a slight trend of lower performance as dependency increases.

On the other hand, the perception of a positive impact of ChatGPT on analytical skills is moderately associated with accuracy on Question 1 (0.324), suggesting that students who believe the tool enhances their analytical abilities tend to justify their answers better. However, this perception has a very weak correlation (0.128) with performance on Question 3.2, indicating that the perceived positive impact does not necessarily influence the ability to formulate complex responses.

4.2. Behavior in Data Visualization and Clustering Questions

To answer QP2, the study analyzed performance data related to Questions 2.1, 2.2, and 3.1, focusing on response accuracy, visualization quality, confidence in ChatGPT, and whether students experimented with code or clustering methods. Question 2 assessed students' ability to select appropriate visualizations for a categorical and a binary variable.

Question 3.1 evaluated their ability to apply clustering and generate interpretable visualizations. These tasks were designed to show how reliance on ChatGPT relates to students' choices in visualization and clustering. Using Pearson correlations on 1–5 recoded scales, dependency aligns with higher perceived usefulness for visualization (Q2; $r \approx .24$), while the association reverses for clustering (Q3; $r \approx -.30$). This pattern suggests that students who rely more on the tool recognize its value when selecting visuals, yet judge it more critically when tasks demand interpretive decisions.

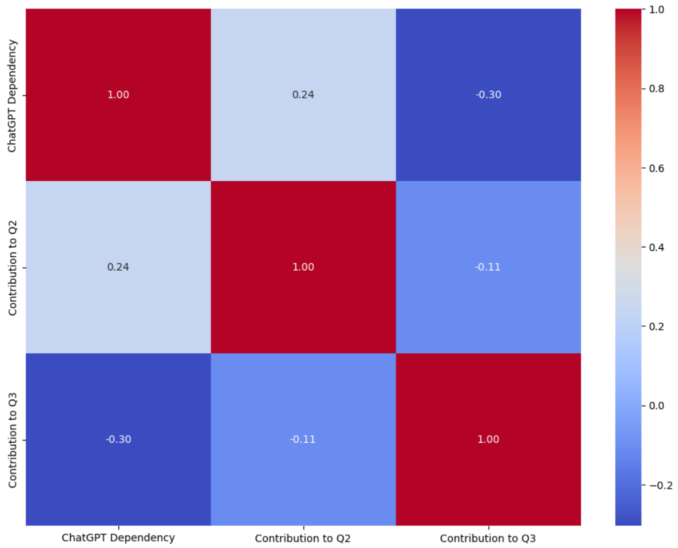


Figure 2. Correlation Matrix between LLM Dependency and Evaluations of ChatGPT. pQ2=0,179 (p=0,358); pQ3=0,091 (p=0,646)

Perceived usefulness appears to be task-specific rather than generic: ratings for Q2 and Q3 are essentially uncorrelated (Q2 vs. Q3; $r \approx -.11$), indicating that students distinguish where the tool adds value. Consistent with this profile, Table 4 shows that higher-dependency students more often supplied richer context (e.g., variable descriptions or concise summaries such as `describe()`), a behavior typically associated with stronger downstream responses.

Table 4. Frequency of combinations between ChatGPT dependency and adequate context. Fisher's exact: p = .414.

	Yes (%)	No (%)
High Dependency	25.9	14.8
Low Dependency	7.4	29.6

The subsequent analysis aimed to investigate students' behavior regarding the exploration of different clustering and data visualization methods based on suggestions provided by ChatGPT. A crucial aspect of this analysis was evaluating how often students modified the provided code—testing different numbers of clusters or exploring alternative clustering algorithms.

The correlation matrix shown in Figure 3 reveals the relationships between ChatGPT dependency and these actions, including the use of methods to determine clusters, testing different numbers of clusters, and exploring alternative algorithms. Analyzing these correlations helps to understand whether ChatGPT dependency influences students' initiative to go beyond the suggested solutions and how these choices may impact outcomes in data visualization and clustering tasks.

- *Dependency on ChatGPT-generated responses* was abbreviated to **Dep. on ChatGPT**.
- *Applied any method to choose clusters* was shortened to **Meth. for clusters**.
- *Tested a different number of clusters than suggested by the chat?* was renamed to **Tested diff. nr**.
- *Explored other clustering algorithms?* was renamed to **Explored others**.

First, a weak negative correlation is observed between Dep. on ChatGPT and Meth. for clusters (-0.070), suggesting that students who are more dependent on ChatGPT tend to apply fewer self-directed methods for choosing clusters. The correlation between Dep. on ChatGPT and Tested diff. nr is almost nonexistent (-0.013), indicating that dependency on the tool is not associated with testing different numbers of clusters beyond those suggested. In contrast, there is a slight positive correlation (0.088) between Dep. on ChatGPT and Explored others, suggesting that students who rely more on the tool are slightly more inclined to explore other clustering algorithms.

Another relevant finding is the moderate negative correlation between Meth. for clusters and Explored others (-0.346), indicating that students who apply methods for choosing clusters tend to explore fewer alternative clustering algorithms. On the other hand, there is a weak positive correlation (0.229) between Tested diff. nr and Explored others, suggesting that students who test different numbers of clusters are also more likely to experiment with other algorithms.

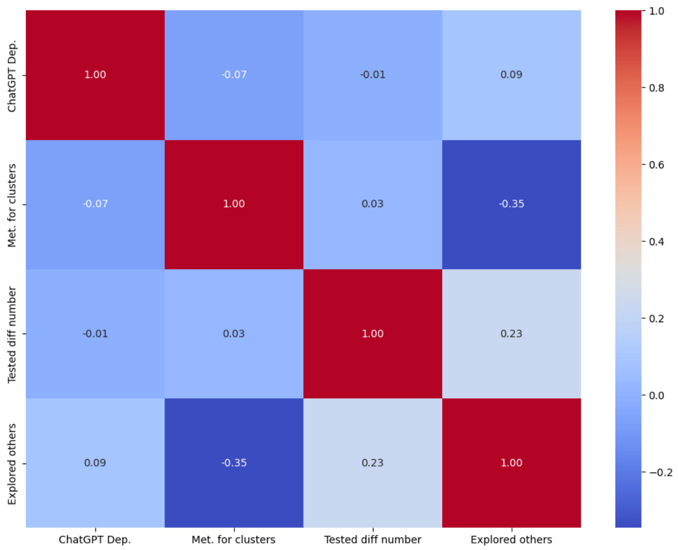
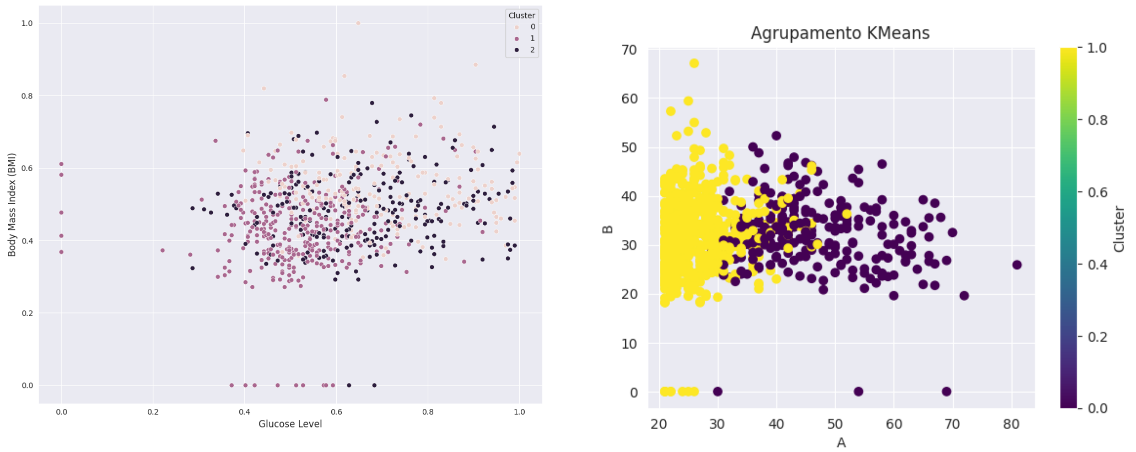


Figure 3. Correlation Between ChatGPT-Related Variables and Clustering. Spearman (dependency 1–5) and two-sided Fisher's exact test; (—ρ—$\leq$.17; $p_{\text{Fisher}} \geq .67$).

The correlation profile suggests that ChatGPT dependency has little to no influence on students' clustering behaviours (Spearman's ρ with "method for clusters", "tested a different K", and "explored other algorithms": $-.07$, $-.01$, and $.09$, respectively). The clearest pattern is within the behaviours themselves: using a method to select K is associated with less algorithmic exploration ($\rho \approx -.35$), whereas testing different K values relates modestly to exploring other algorithms ($\rho \approx .23$). Overall, strategy choices appear to trade off with one another more than they are driven by dependency, consistent with the idea that students' autonomy in adapting or extending code is what fosters more diverse solutions. Figures 4 and 5 present two examples of clustering that illustrate how exploring and modifying the code can provide valuable resources:



(a) 3 Clusters. Automatic Suggestion by ChatGPT.

(b) 2 Clusters. Student-Initiated.

Figure 4. Comparison between clustering results suggested by ChatGPT and by the student.

The correlation matrix shown in Figure 4 evaluates the relationship between students' dependency on ChatGPT and their visual outcomes when working with clusters and categorical and binary variables.

To facilitate understanding of the data presented in the correlation matrix, some simplifications were made to the variable names:

- *Dependency on ChatGPT-generated responses* was abbreviated to **Dep. on ChatGPT**.
- *Good Visual Outcome for Clusters (Q3.1)* was shortened to **Good Res. for clusters**.
- *Good Visual Outcome for Variables (Q2.2)* was renamed to **Good Res. for variables**.

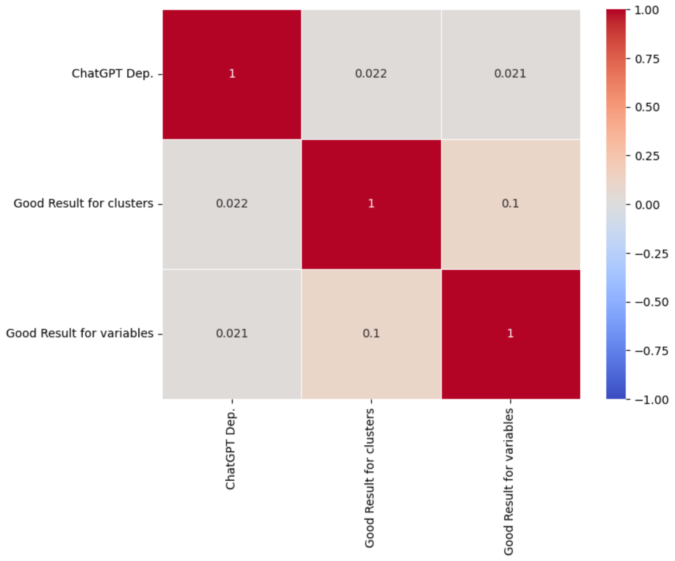


Figure 5. Correlation Matrix: ChatGPT Dependency and Visual Outcomes. $\rho = -0{,}070\ (p = 0{,}748)$

The correlation matrix reveals that dependency on ChatGPT has a weak correlation with achieving good visual outcomes in clustering (0.022) and in categorical and binary variables (0.021), suggesting that simply following the tool's responses without exploring code variations does not lead to significant improvements in these areas.

These findings highlight the importance of an active approach when using ChatGPT. Students who explore and adjust the code provided by the tool may achieve richer visualizations and superior technical performance. This suggests that relying solely on automated responses can be limiting, especially in tasks that require creativity and critical analysis. Encouraging students to modify the generated code and test variations may be an effective strategy to improve both the quality of visualizations and the development of analytical skills.

4.3. Conclusions and Future Work

This study examined the impact of ChatGPT on data science education, focusing on how varying levels of dependency influence students' analytical skills and performance in visualization and clustering tasks. Results indicate that while high-dependency students often report a positive perception of ChatGPT's impact, this does not necessarily lead to better performance. More effective outcomes were observed among students who actively modified and explored the generated code, especially in technical tasks like clustering.

The findings highlight the importance of fostering student autonomy and critical engagement with AI tools. Overreliance may hinder the development of deeper analytical thinking.

Future work will include: (1) deeper analysis of LLM dependency across specific data science domains; (2) development of pedagogical strategies to encourage critical and adaptive use of tools like ChatGPT; (3) longitudinal studies to monitor skill development over time; (4) scale to larger samples to increase precision and enable subgroup analyses and (5) strengthening validity via data triangulation (e.g., surveys, behavioral logs, and artifacts) and richer qualitative analyses of tool–student interactions (e.g., systematic coding of chat transcripts).

5. References

Baker, J. and Smith, E. (2023). The impact of large language models on academic integrity and learning outcomes. *Journal of Educational Technology*, 58:45–60.

Budhiraja, R., Joshi, I., Challa, J. S., Akolekar, H. D., and Kumar, D. (2024). “it’s not like jarvis, but it’s pretty close!”-examining chatgpt’s usage among undergraduate students in computer science. In *Proceedings of the 26th Australasian Computing Education Conference*, pages 124–133.

Creswell, J. and Creswell, J. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.

Gil, A. C. (2008). *Métodos e técnicas de pesquisa social*. 6. ed. Ediitora Atlas SA.

Kirova, V. D., Ku, C. S., Laracy, J. R., and Marlowe, T. J. (2024). Software engineering education must adapt and evolve for an llm environment. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, pages 666–672.

Liu, W. and Gonzalez, M. (2024). Ai chatbots in higher education: Effects on student engagement and performance. In *Proceedings of the International Conference on Educational Data Mining*.

Lopes, C., da Costa, A. B., and Zoppo, B. M. (2024). A influência das metodologias ativas de aprendizagem na promoção da autonomia e inovação em pesquisadores em formação: um estudo transdisciplinar. volume 21, pages e3108–e3108.

Malinka, K., Peresíni, M., Firc, A., Hujnák, O., and Janus, F. (2023). On the educational impact of chatgpt: Is artificial intelligence ready to obtain a university degree? In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, pages 47–53.

Shen, Y., Ai, X., Soosai Raj, A. G., Leo John, R. J., and Syamkumar, M. (2024). Implications of chatgpt for data science education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, pages 1230–1236.

Sigolo, B. d. O. O. and Casarin, H. d. C. S. (2024). Contribuições da teoria da carga cognitiva para compreensão da sobrecarga informacional: uma revisão de literatura. volume 22, page e024027. SciELO Brasil.

Tu, X., Zou, J., Su, W. J., and Zhang, L. (2023). What should data science education do with large language models?

Zheng, Y. (2023). Chatgpt for teaching and learning: an experience from data science education. In *Proceedings of the 24th Annual Conference on Information Technology Education*, pages 66–72.