

Embracing the AI Revolution: ChatGPT’s Role in Advancing Data Science Consultation Services

ALP TEZBASARAN^{1,2,3,*} AND SHANNON RICCI^{2,*}

¹*Department of Nuclear Engineering, North Carolina State University, Raleigh, NC, USA*

²*NC State University Libraries Data Science Services, North Carolina State University, Raleigh, NC, USA*

³*Data Science and AI Academy, North Carolina State University, Raleigh, NC, USA*

Abstract

The Data Science Consulting Program at North Carolina State University Libraries, in partnership with the Data Science and AI Academy, provides comprehensive support for a wide range of tools and software, including R, Python, MATLAB, ArcGIS, and more, to assist students, faculty, and staff with their data-related needs. This paper explores the integration of generative AI, specifically ChatGPT, into our consultation services, demonstrating how it enhances the efficiency and effectiveness of addressing numerous and diverse requests. ChatGPT has been instrumental in tasks such as data visualization, statistical analysis, and code generation, allowing consultants to quickly resolve complex queries. The paper also discusses the program’s structured approach to consultations, highlighting the iterative process from initial request to resolution. We address challenges like prompt engineering and response variability, offering best practices to maximize the tool’s potential. As AI technology continues to evolve, its role in our data science consultations is expected to expand, improving service quality and the consultant’s ability to handle increasingly complex tasks. The study concludes that ChatGPT is a valuable asset in academic data science, significantly streamlining workflows and broadening the scope of support provided by our program.

Keywords *ChatGPT; data science consultations; generative AI; productivity tool*

Disclaimer

NC State University Libraries Data Science Services and NC State University Data Science and AI Academy have **no affiliation** with the developers of ChatGPT or OpenAI; the opinions and experiences shared in this study reflect solely authors’ view. Additionally, our use of ChatGPT in our data science consultation service is strictly **non-commercial** and for productivity enhancement purposes only.

1 Introduction

The application of data science approaches is becoming widespread across many disciplines, from engineering to life sciences to humanities. In academic settings, particularly at research institutions, students, faculty, and staff may need assistance using emerging technologies or analysis approaches to explore, visualize, and synthesize their ever-growing datasets (Kirk, 1991).

*Corresponding author. Email: alptezbasaran@ncsu.edu or swbrown@ncsu.edu.

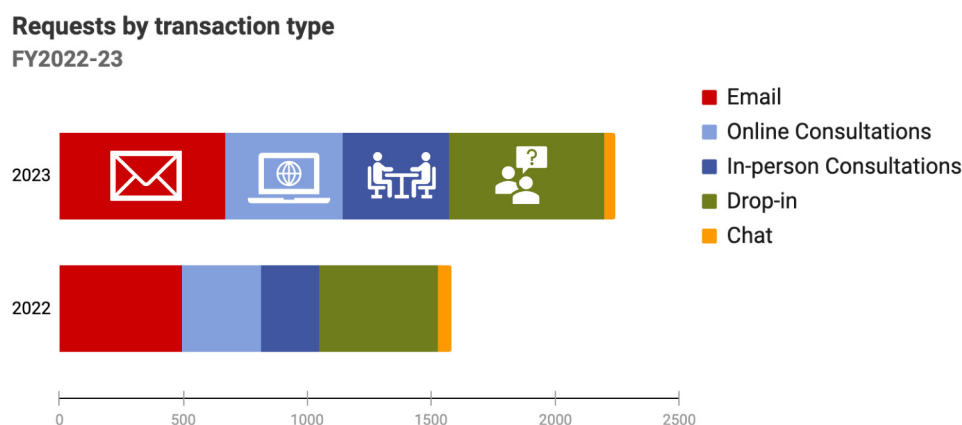


Figure 1: Total number of transactions our team had in fiscal years 2022 and 2023 (Marchesoni et al., 2024).

At North Carolina State University (NC State), the Libraries' Department of Data Science Services and the Data Science and AI Academy offer free data science consultations to any NC State affiliate through their data science consulting service. This service receives data help requests through a dedicated email inbox, scheduled consultations, and drop-in consultations within the Libraries. We have graduate students as consultants working part-time under the supervision of full-time non-tenure track faculty specializing in data science. A more complete description of NC State's data science consulting program and staffing model can be found in Marchesoni et al. (2024).

Requests from our patrons (i.e., NC State University affiliates), span the entirety of the data science lifecycle, and require knowledge of a wide range of tools and software. We currently support tools/software including but not limited to R, Python, MATLAB, ArcGIS Pro/ArcGIS Online, Google Sheets/Excel, SAS, SPSS, and Stata and concepts such as high performance computing, version control, statistical analysis, and geospatial analysis. We answer questions related to study design, finding data, data preparation, analysis, data visualization, and software and or environment help. In the past two fiscal years of 2022 and 2023, the consulting program at NC State University has answered more than 3500 requests for support from across the university (Marchesoni et al., 2024). Figure 1 shows our transactions for 2022 and 2023 and their intake methods.

While total numbers are essential when evaluating capacity, they do not provide a complete picture of the time and resources involved in this work (Blake et al., 2024). Consulting is an iterative process, and involves both technical and communication skills throughout (Kirk, 1991; Kenett and Thyregod, 2006). A request may start as an intake email, which requires some initial research by the consultant to attempt to answer the request, find the correct person to answer the request, or find resources to support the patron in answering the request. After this investigation phase, the patron may meet with a consultant and discuss their problem in detail. This often reveals additional components to the initial request, requiring more research for the consultant to resolve the question. This iterative approach may require returning to multiple stages (email, research, meetings) before a final resolution.

In the rapidly evolving field of artificial intelligence (AI), generative AI has emerged as a groundbreaking technology, reshaping the way we interact with and harness data. At its core, generative AI refers to sophisticated algorithms capable of creating new, diverse content –

ranging from textual outputs to complex images – by learning from vast datasets. This technology has revolutionized numerous sectors, fostering innovation and efficiency (Meyer et al., 2023; Azaria et al., 2024).

ChatGPT, a state-of-the-art language model developed by OpenAI is a prime example of generative AI’s prowess. ChatGPT, built on the GPT-4 architecture, stands as a pinnacle of conversational AI. It is designed to understand and generate human-like text, enabling seamless and intuitive interactions. Trained on a diverse range of internet text, ChatGPT can engage in discussions, answer questions, and even mimic stylistic writing nuances, making it an invaluable tool in various applications.

In the context of data science consultations, the capabilities of ChatGPT take on a transformative role. Its ability to process and analyze large datasets, generate insightful observations, and assist in model development positions it as an invaluable asset in data science. This introduction sets the stage for an in-depth exploration of how ChatGPT can be leveraged in data science consultations, highlighting its potential to streamline analysis, foster innovation, and drive informed decision-making in this field.

In Section 2, we briefly describe the data science consultation process we adopted over the years, Section 3 shows the example use cases and discusses how ChatGPT helped; Section 4 summarizes our learnings, and the best practices that apply to consultation style hands-on work, and finally in Section 5 we conclude the study with a discussion on how ChatGPT and generative AI can advance data science, and data science communication. For convenience:

Patron requests we received are highlighted via verbatim.

ChatGPT prompts are highlighted in lstlisting snippets.

2 Data Science Consultation Approach

While it may seem counterintuitive, technical expertise is not the only skill required for a successful data science consultation. Data science consulting shares a methodological approach to statistical collaboration in that a series of steps should be followed throughout the interaction (Kenett and Thyregod, 2006; Kirk, 1991; Vance et al., 2022; Barman et al., 2022). Effective communication and problem definition are critical components of this process, surrounding the obvious technical steps involving working with and analyzing data. Our general consultation steps are as follows:

1. Initial request
2. Preparation
3. Meeting with the patron
4. Iteration on the problem
5. Resolution

The first step, Initial Request, involves the client reaching out with a specific problem or need, often accompanied by their dataset or a broad description of the issue. This step sets the stage for the entire process and requires clear communication to understand the patron’s objectives and constraints. Next, during the Preparation phase, the data science consultant gathers relevant background information, reviews the provided data, identifies necessary tools and techniques, and often conducts preliminary analyses. This step is crucial for identifying potential challenges and formulating initial solution steps. The Meeting with the Patron follows, where

the consultant engages directly with the client to discuss their findings, refine the problem statement, and ensure alignment on goals and expectations. This interaction helps clarify ambiguities and allows for incorporating the client's insights and feedback. Moving forward, the Iteration on the Problem involves a cyclical process of developing models, analyzing results, and continuously refining approaches based on iterative feedback and emerging data patterns. This step is highly dynamic and often the most time-consuming. Finally, the Resolution stage brings the consultation to a close, with the consultant presenting the finalized solution, delivering reports, and, if necessary, providing guidance for implementation. This step ensures that the patron has a clear understanding of the outcomes and actionable insights derived from the consultation, leading to informed decision-making and successful application of the results. Each of these steps is critical in ensuring that the data science consultation is thorough, collaborative, and ultimately effective in addressing the patrons's needs.

Our two biggest challenges are **time constraints** and both **in-breadth** and **in-depth knowledge requirement** about the concepts and tools we support. Like most academic research institutions, a wide variety of research is being done, and we may receive questions from researchers across various research disciplines (Kirk, 1991). To elaborate, by in-breadth, we mean that our consulting program must maintain expertise in many cutting-edge concepts and tools, for they are the most relevant to the data science lifecycle. Furthermore, the in-depth knowledge requirement signifies the need for deep understanding and experience in particular topics. How to convert one's simulation code to utilize parallel computing can be an excellent example of an in-depth knowledge requirement in scientific computation.

3 ChatGPT Use and Example Cases

3.1 ChatGPT Use

ChatGPT is designed to be a chatbot, and for the majority of our use-cases we utilize it as such. OpenAI has an available API too for more coding heavy requests, but this study focuses solely on the chat service which is available for free.

Before submitting a prompt to ChatGPT, the consultant must assess if this tool can assist. ChatGPT serves as a valuable tool for addressing a variety of tasks, particularly those that involve general explanations, code generation, and methodological advice. It is well-suited for responding to requests such as clarifying statistical concepts, offering basic interpretations of data trends, generating or correcting code snippets, and recommending appropriate tools or resources. Additionally, ChatGPT can provide high-level guidance on workflows, such as structuring data pipelines or selecting suitable methodologies for handling missing data. However, its utility diminishes when faced with more complex or domain-specific tasks, such as in-depth data analysis, debugging intricate technical issues, or interpreting proprietary frameworks. Moreover, ChatGPT is not appropriate for handling sensitive or confidential data due to privacy concerns, nor is it advisable to rely on it for real-time decision-making or critical operations that require nuanced understanding and accountability. Once decided as suitable, we use a prompt that is *descriptive enough* and upload any related files (only applies to the paid version of chatbot, not a necessity), keeping the privacy of our patrons. If the response can be used with little to no modification, we immediately utilize it. However, some cases require more in-depth discussion and a few more steps toward the desired goal. In that case, we modify our follow-up prompts by considering the previous parts of the discussion and what needs to be modified. We identified the following guidelines for a quick and concise issue resolution.

Table 1: Custom Chat-GPT instructions used for the threads used in this study.

What traits should ChatGPT have?	If applicable, you must also provide an example with Python with synthetic data. The provided code must be working correctly, complete, and ready to execute. The code must have the proper comments and debug print statements, with an option to disable printing.
Anything else ChatGPT should know about you?	I am a data scientist. You must evaluate every query from a data scientist’s perspective and use critical thinking to provide the best possible answer. You must avoid wrong practices and always briefly promote the best practices. If there is an uncommon term, you must elaborate.

3.2 Example Use Cases

This section introduces the example cases in which ChatGPT was employed for various reasons. Each example shows the tool’s vast capabilities and versatility. These cases are not synthetically generated but picked from actual requests our program has previously resolved. We slightly modified the outputs to keep our patrons’ privacy without forsaking reproducibility.

While a comparison between answering these requests using AI and versus without would be interesting, the examples described here were completed without a goal in mind to publish. Consultants in our program are working on numerous requests, so it is unlikely we would assign the same request to two different consultants – one using AI tools and another without. So while we cannot give a comparison in time to completion for each scenario, we have added a description of a general process of how you could approach the requests below without the use of generative AI. Also, we mention below about custom instructions. And we would like to list the ones used to generate responses in this paper. You can enable Custom Instructions by going to Settings → Personalization → Custom Instructions → Toggle on Enable for new chats. The relevant fields are shown in Table 1.

3.2.1 Ashby Plot

Our patron initiated her request through email. She wrote:

I need an Ashby plot for the tenacity and specific modulus of selected materials. Your expertise in this matter would be highly appreciated.
Thank you.

Our team had never heard of Ashby plots before. Named after its developer, Michael Ashby, Ashby plots are scatter plots that show at least two or more properties of materials or material groups. The points are stretched into an elliptical shape to show the range of the properties for each material. These can include material properties such as modulus vs. density, fracture toughness vs. strength, and property vs. cost charts (Ashby, 2011). From a data science consultant perspective, there are material properties; some can include uncertainty, hence ranges. Utilizing the shape, marker, color, transparency, and other customization options provides more

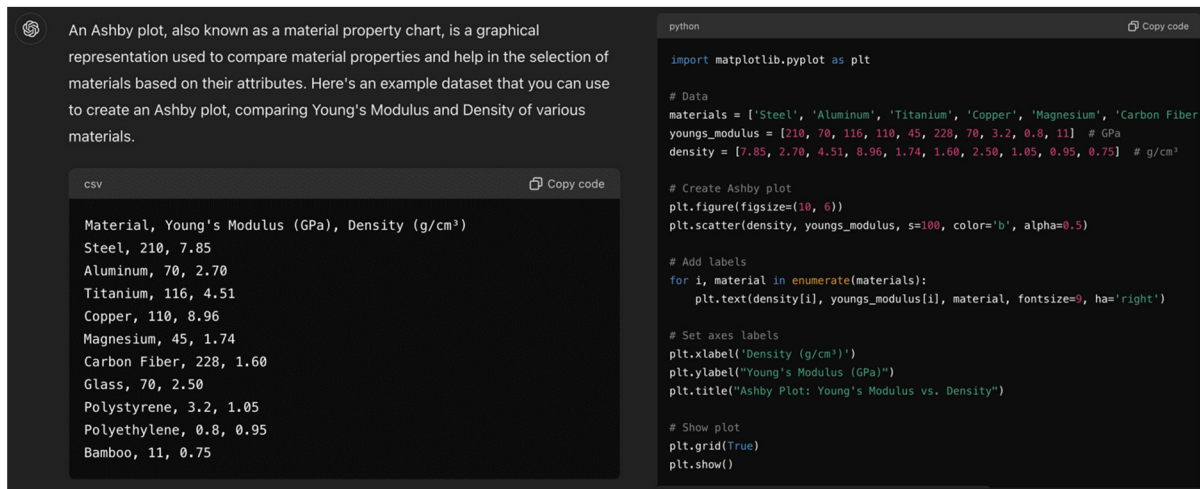


Figure 2: Initial response to Ashby plot inquiry by ChatGPT.

information and dimensions on the plot. Once we scrutinized what an Ashby plot constitutes with the help of ChatGPT (<https://chatgpt.com/share/df0fac0a-133e-4d02-9803-947d0eb44d76>, ChatGPT Version 4o), we started the iterative consultation process. After initial searches and checking the source material, we went to ChatGPT and prompted the following simple question:

example data for ashby plot

ChatGPT then provided a brief description, example data, and a code to generate the plot in Python, shown in Figure 2. Beware that we gave ChatGPT custom instructions to output a code snippet in Python when applicable, which is not the default behavior. We could prepare a script to generate the desired plot with the proper customizations before our patron prepared her final data with ChatGPT's help. Once she sent us her data, generating and delivering the plot and steps to reproduce such figures took only minutes. Figure 3 shows the reproduced modified final plot. We iterated with the chatbot a few more times with the following prompts for the final product:

i have ranges for some of the materials

instead of error bars, can we change the size of the marker?

what about different ranges in different properties, the markers should be ellipses

this is a good start. can we add color and a legend instead of annotating. it is overlapping

To approach this request without using AI, a consultant would use Google (or another search tool), to look up information on Ashby plots. They would then Google how to generate this plot using python packages and would have to review multiple tutorials or links.

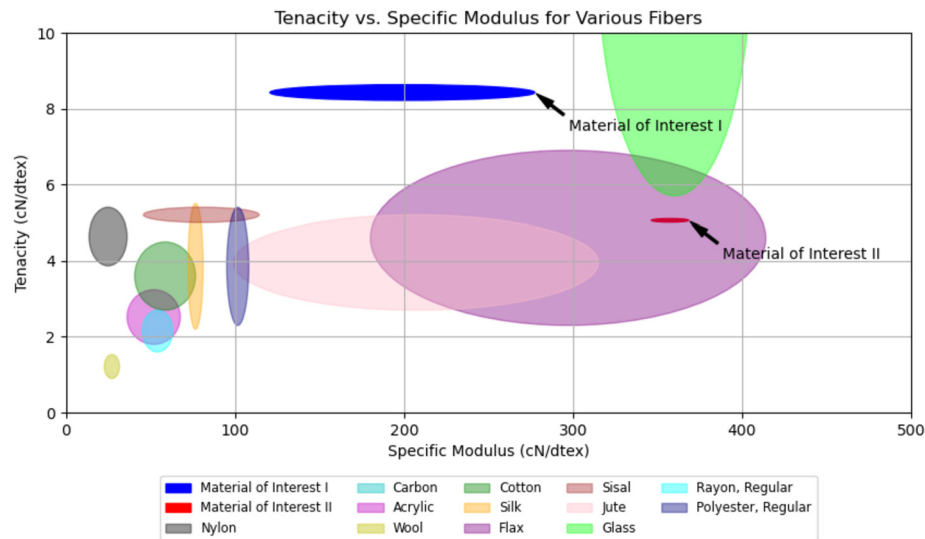


Figure 3: Final version of Ashby plot (modified from original).

Once an approach is decided on, the consultant would read the documentation for the suggested package, and likely would have to search for information on annotation, how to change colors, and how to make the desired edits requested by the patron. This approach would result in numerous tabs or browser windows and would require reading materials from multiple sources.

3.2.2 Optical Character Recognition (OCR)

This walk-in request was from a graduate student who wanted to extract measured temperature values from a section in a video for each frame. He wanted to use the time-dependent data for further analysis. He needed to perform this operation error-free and multiple times for different experiments. We asked our patron to send us the relevant information and an example video. Videos were at least one minute long and formatted in a specific way shown in Figure 4. Our patron specifically wanted the numerical values shown on the right-hand side, above the legend for each frame, highlighted in the rectangle. After our initial interactions, we defined this short project's goal, scope, and deliverables. He was interested in performing extraction using Python, and the extracted data should be in the CSV file format, including the frame's timestamp and the numerical temperature value. As a follow-up, the patron emailed the following:

I am extracting temperature data from my thermal camera. The data is output onto the screen, on a fixed area continuously throughout the recording. I am trying to create a python code with the tesseract and open CV libraries that extracts each frame from the video and graph the data in a temperature vs. time graph.

The video is 320x240 pixels with the location of the text I am extracting (using OCR) at pixels 287, 41 pixels in a box of width 33x28, where the top text is located.

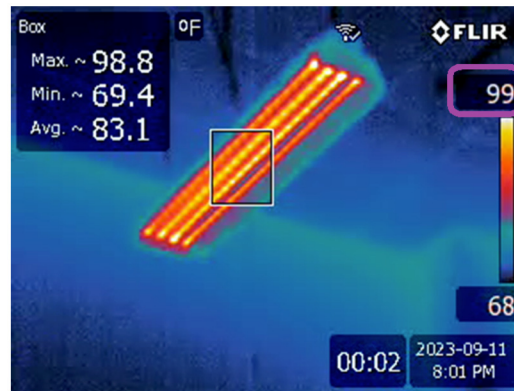


Figure 4: Frame of a video recorded by a temperature measurement camera.

This time, we recognized some details we could omit and generalized the problem before asking for help. This generalization is not an abstraction, meaning that we still keep the main aspects of the request, such as the data being in video format (frames), there is a need to keep timestamps, needed OCR on a specific region, and the output format. So we prompted the following (<https://chatgpt.com/share/528aaaa4-1799-4bea-8864-f974692b049e>, ChatGPT Version 4o-mini):

I have a video and my goal is to collect data from each frame. Fortunately the video content is structured and I am interested in the number located in a specific area on the frame. can you give me an example code please

ChatGPT responds with the steps required:

Certainly! Let's break the process down step-by-step:

- **Extract Frames from Video:** We'll use *OpenCV* to read the video and extract frames.
- **Crop the Region of Interest (ROI):** Once we have the frames, we can crop to the specific area where the number is located.
- **Extract Text from the Cropped Image:** We can use Optical Character Recognition (OCR) to extract the text from the cropped image. For this, the *pytesseract* library is a popular choice.

- Code snippet to perform the operation.

Notes:

- You'll need to have both *opencv-python* and *pytesseract* libraries installed.
- The *pytesseract* library requires the Tesseract binary to be installed on your system. You can follow the installation guide for your OS from the official Tesseract GitHub repository.
- Adjust the *roi* (Region of Interest) tuple to match the location and size of the area where the number is located in your video frames.
- The *-psm 6* configuration for *pytesseract* assumes that the text in the image is sparse. Depending on your image, you might need to adjust this configuration for better accuracy.

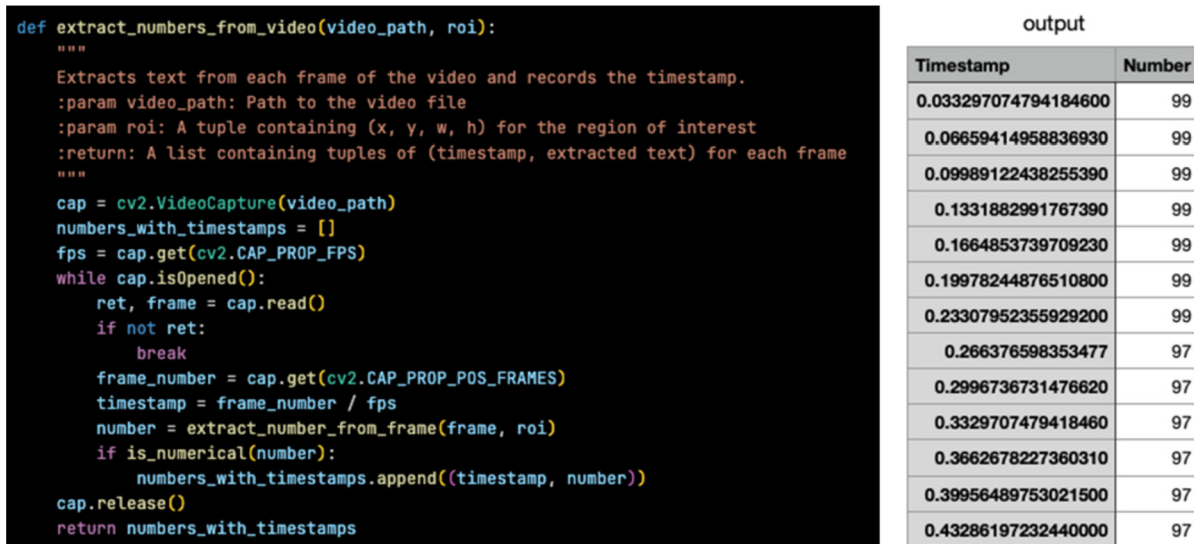


Figure 6: Final Python function used in the extraction script (left), and the example CSV output (right).

3.2.3 Data Analysis

I have data collected from field experiments.
I'm looking at the effects of different water stress levels on the yield and nutrient uptake. I'd like to find out if there are any differences between hybrids in yield and nutrient uptake.

We prompted our patron's request without modifying it to gain insights into the overall framework. Summary of the ChatGPT's response is as follows (<https://chatgpt.com/share/6522449e-3ce6-4e0f-87d9-cd92c82b2ec1>, ChatGPT Version 4o):

To analyze the effects of different water stress levels on yield and nutrient uptake and to determine if there are differences between hybrids, you can follow these steps:

1. **Data Preparation**
2. **Exploratory Data Analysis**
3. **Hypothesis Testing:** Formulate hypotheses for the differences in yield and nutrient uptake between hybrids under different water stress levels.
4. **Statistical Analysis**
 - (a) Analysis of Variance (ANOVA)
 - (b) Post-hoc Tests
 - (c) Regression Analysis
5. **Interaction Effects**
6. **Data Visualization**
7. **Interpretation and Reporting**
8. Example Python script that uses synthetic data

The actual response is slightly more detailed than the list shown above. The extra information includes descriptions for each step and expands sections like ANOVA into One-way ANOVA

and Two-way ANOVA, or Regression Analysis into Linear and Multiple regressions. It also has visualization options, among other information. This request might seem too simple, and the steps towards the solution may be apparent. However, the consultants might have some ideas on approaching the problem; they might be uninitiated in the field and the collected data. This problem translates to a loss of valuable time because consultants will have to do extensive preparation, which in most cases is not feasible. Talking to ChatGPT, in this case, can help with ideas that one might not think about for the analysis, assist with the terminology, and find relevant resources. It can also provide an example analysis with synthetic data, speeding up the overall consultation process since the only change will be using the patron data and interpreting the respective results. Here, we would like to warn our readers. Even though the chatbot will answer the original prompt, questions such as statistical analysis are a cautionary tale. Problems like data cleaning or coding generally have tangible and gaugeable outcomes. However, statistical analysis is different by nature. It involves complexities that are often not as straightforward, with outcomes that may be more interpretative, context-dependent, and influenced by the assumptions or models used. While tasks like coding and data cleaning provide concrete results, statistical analysis demands a deeper understanding of the data context, potential correlations, and the limitations of the chosen methods, making it a more nuanced and interpretative process. Generative AI tools should only be regarded as a reference or suggestion.

4 Lessons Learned and Best Practices

The following insights have been curated from our experiences in academic data science consultations and short-term projects. It is important to note that these lessons and best practices are specifically tailored to the data science lifecycle and are relevant to the time of publication. As ChatGPT and other generative AI tools continue to evolve rapidly, our practices must adapt accordingly.

Version Variability One of the key takeaways from our study is the variability in responses generated by different versions of ChatGPT. For instance, responses from versions 3 and 4 may differ significantly. This variability, rather than being a limitation, can be leveraged to gain broader insights. Utilizing multiple versions can provide a more comprehensive view of potential solutions and ideas, enhancing the overall consultation process.

Prompt Engineering Crafting effective prompts is critical to obtaining relevant and accurate responses from ChatGPT. We have found that clear and descriptive prompts, which include specific constraints such as environmental factors, formatting requirements, and desired methods, lead to more effective problem resolution. For example, when extracting temperature data using Optical Character Recognition (OCR) technology, specifying these constraints upfront proved essential for success.

Iterative Refinement Achieving the desired outcomes often requires patience and iterative refinement of prompts. It is rarely the case that a single prompt will yield a perfect response. Through multiple iterations and careful adjustments, the quality of the responses can be significantly improved, making this approach a cornerstone of effective AI-driven consultations.

Politeness and Clarity An interesting observation during our interactions with ChatGPT is that politeness and clarity in prompts can influence the quality of responses. Although AI does not have emotions, clear and polite language seems to facilitate better processing and interpretation of the query, leading to more coherent and relevant outputs. We should note that

this finding was first anecdotal, then further investigation revealed that independent researchers reached similar conclusions, Yin et al. (2024), Nawar (2024).

To maximize the benefits of ChatGPT in academic consultations, we have developed several best practices:

- **Descriptive Initial Prompts:** Begin with clear and concise prompts that include relevant constraints, such as desired coding languages, methods, and output formats. This approach ensures that the AI's responses are closely aligned with the client's needs.
- **Maintaining Focus:** It is crucial to maintain a focus on the core objective of the consultation. If a conversation diverges too far from the original goal, starting a new session can help realign the discussion and avoid unnecessary confusion.
- **Handling Data Inputs:** When providing data inputs, avoid using confusing separators and clearly explain any unnecessary information, such as columns or variables that are not needed. This enhances the clarity of the AI's outputs and reduces the risk of errors.
- **Continuous Learning and Adaptation:** The integration of ChatGPT into our services has underscored the importance of continuous learning and adaptation. Expanding our knowledge of relevant tools and topics enables us to better guide the AI's responses and refine its outputs to suit the specific needs of our consultations.

By adhering to these best practices, we can ensure that ChatGPT remains a valuable tool in our academic consultations, helping us to achieve more accurate, relevant, and insightful outcomes.

5 Conclusion

The advent of ChatGPT and other generative AI tools has marked a significant transformation in the realm of data science consulting. These technologies have not only enhanced the efficiency and accuracy of consultations but also expanded the scope and depth of services that can be offered. By providing instant access to a vast repository of knowledge and methodologies, ChatGPT allows consultants to quickly generate insights, create reproducible code, and suggest data-driven solutions tailored to the specific needs of each consultation. This capability significantly reduces the time spent on problem-solving and preparation, enabling consultants to focus more on higher-order tasks such as interpretation, strategic decision-making, and client interaction. AI utilization not only accelerates the problem-solving process but also enhances the quality of the solutions provided, making them more robust and aligned with client expectations.

Looking to the future, the integration of generative AI in data science consulting is poised to evolve further. As AI models become more sophisticated, we can anticipate even greater automation of routine tasks, more accurate predictions, and the ability to handle increasingly complex queries. Additionally, with advancements in AI transparency and interpretability, consultants will be able to provide clients with more detailed explanations and justifications for AI-generated outputs, thereby increasing trust and confidence in AI-assisted solutions.

However, this transformation also brings challenges, such as the need for continuous upskilling in AI literacy and the ethical considerations surrounding AI use in sensitive data contexts. Addressing these challenges will be crucial for ensuring that the benefits of generative AI are fully realized in a responsible and sustainable manner.

In conclusion, ChatGPT and similar generative AI tools are revolutionizing data science consultation by drastically improving productivity, enabling more comprehensive analyses, and enhancing the overall quality of service. As we continue to explore the potential of these tech-

nologies, they will undoubtedly play an increasingly pivotal role in shaping the future of data science and its applications across various disciplines.

References

- Ashby MF (2011). *Materials Selection in Mechanical Design*. Butterworth-Heinemann, Burlington, MA, 4th ed edition. OCLC: 708240020.
- Azaria A, Azoulay R, Reches S (2024). ChatGPT is a remarkable tool—for experts. *Data Intelligence*, 6(1): 240–296. https://doi.org/10.1162/dint_a_00235
- Barman A, Chen S, Chang A, Allen G (2022). Experiential learning in data science through a novel client-facing consulting course. In: *2022 IEEE Frontiers in Education Conference (FIE)*, 1–9. ISSN: 2377–634X. <https://doi.org/10.1109/FIE56618.2022.9962532>
- Blake MR, Griffith E, Pierce SJ, Levy R, Parker M, Huebner M (2024). Tell your story: Metrics of success for academic data science collaboration and consulting programs. *Stat*, 13(2): e686. <https://doi.org/10.1002/sta4.686>
- Kenett R, Thyregod P (2006). Aspects of statistical consulting not taught by academia. *Statistica Neerlandica*, 60(3): 396–411. <https://doi.org/10.1111/j.1467-9574.2006.00327.x>
- Kirk RE (1991). Statistical consulting in a university: Dealing with people and other challenges. *The American Statistician*, 45(1): 28–34. <https://doi.org/10.1080/00031305.1991.10475761>
- Marchesoni J, Freeman K, Tezbasaran A, Ricci SW (2024). The student staffing advantage. *Data Science Consulting Service at NC State University Libraries*. *Stat*, 13(2): e702. <https://doi.org/10.1002/sta4.702>
- Meyer JG, Urbanowicz RJ, PCN O'Connor K M, Li R, Peng PC, et al. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, 16(1): 20. <https://doi.org/10.1186/s13040-023-00339-9>
- Nawar E (2024). Do You Say Please or Thank You to ChatGPT? The Subtle. *Influence of Prompt Engineering on Digital Civility. Library Articles and Research*.
- Vance EA, Trumble IM, Alzen JL, Smith HS (2022). Asking great questions. *Stat*, 11(1): e471. <https://doi.org/10.1002/sta4.471>
- Yin Z, Wang H, Horio K, Kawahara D, Sekine S (2024). Should We Respect LLMs? A Cross-Lingual Study on the Influence of Prompt Politeness on LLM Performance. ArXiv preprint: <https://arxiv.org/abs/2402.14531> [cs].