# Prompting the Professoriate: A Qualitative Study of Instructor Perspectives on LLMs in Data Science Education

Ana Elisa Lopez-Miranda [1], Tiffany Timbers [2], Rohan Alexander [1]

[1]University of Toronto [2]University of British Columbia

## Abstract

Large Language Models (LLMs) have shifted in just a few years from novelty to ubiquity, raising fundamental questions for data science education. Tasks once used to teach coding, writing, and problem-solving can now be completed by LLMs, forcing educators to reconsider both pedagogy and assessment. To understand how instructors are adapting, we conducted semi-structured interviews with 42 instructors from 33 institutions in 10 countries in June and July 2025. Our qualitative analysis reveals a pragmatic mix of optimism and concern. Many respondents view LLMs as inevitable classroom tools—comparable to calculators or Wikipedia— while others worry about de-skilling, misplaced confidence, and uneven integration across institutions. Around 58 per cent have already introduced demonstrations, guided activities, or make extensive use of LLMs in their courses, though most expect change to remain slow and uneven. That said, 31 per cent have not used LLMs to teach students and do not plan to. We highlight some instructional innovations, including AI-aware assessments, reflective use of LLMs as tutors, and course-specific chatbots. By sharing these perspectives, we aim to help data science educators adapt collectively to ensure curricula keep pace with technological change.

## Objective

How should statistics and data science pedagogy change, given that many substantive coding and writing tasks can now be done by AI?

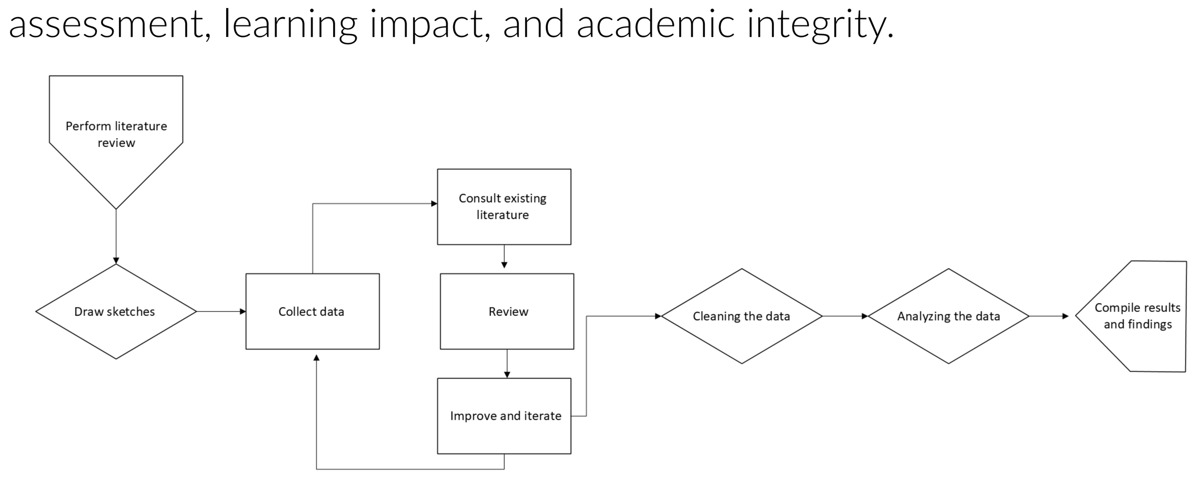
## Background

Teaching has always required adaption to new tools, whether calculators, the internet, or Wikipedia. Each innovation reshaped what students needed to know and how they came to learn it. Large Language Models (LLMs) — a type of artificial intelligence trained on vast amounts of text to understand and generate human-like language, with ChatGPT being the most well-known example — likely represent the most disruptive change yet. For instructors, the challenge is twofold. On the one hand, LLMs promise efficiency, personalization, and expanded access to expertise. On the other, they raise serious concerns about ethics, bias, academic integrity, and the erosion of core skills. We imagine that many data science educators have made some promising adjustments to their teaching in response to LLMs, such as encouraging or restricting LLM use in assessments, adding new content to our curriculum (e.g., transformers and LLMs), and changing approaches to pedagogy. However, in a single instructor's teaching practices, these changes can feel small and incremental. We suspect we are not alone in this feeling. This paper is grounded in the hypothesis that many educators are navigating similar uncertainties. By sharing what instructors across institutions are thinking, feeling, and doing, we hope to foster a sense of community, spark and cross-pollinate new ideas across institutions, and support one another in reimagining our teaching during this pivotal moment.

## Methodology

185 data science and applied statistics instructors were contacted. We conducted semi-structured interviews with 42 instructors teaching data science, and closely related, courses from 33 unique institutions in 9 countries, between 9 June and 9 July 2025. During those interviews, instructors were asked about their thoughts and approaches to LLMs in education. These interviews were conducted by one of the authors and focused on three aspects:

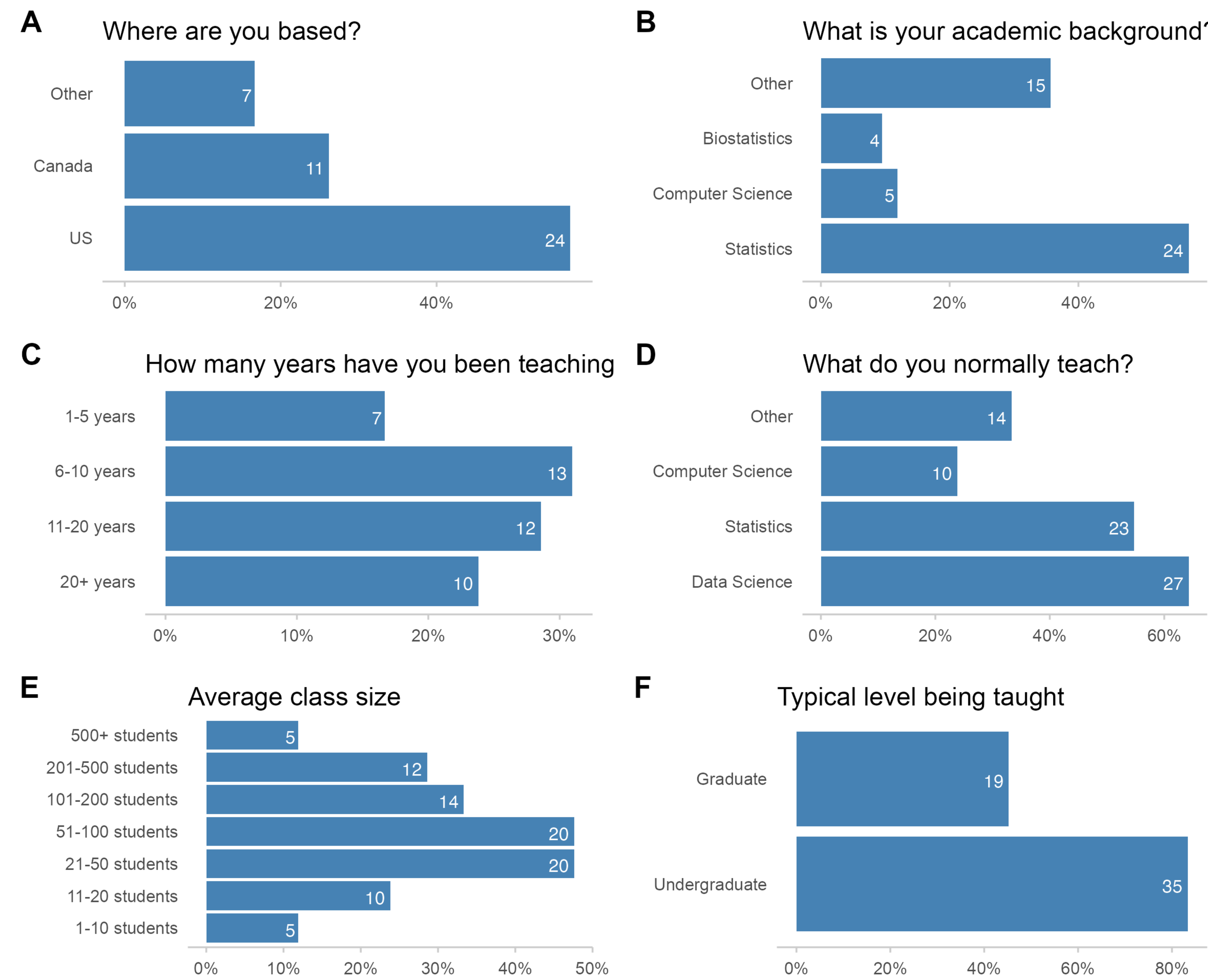
1. instructor and course context;
2. LLM adoption and classroom practice changes;
3. assessment, learning impact, and academic integrity.

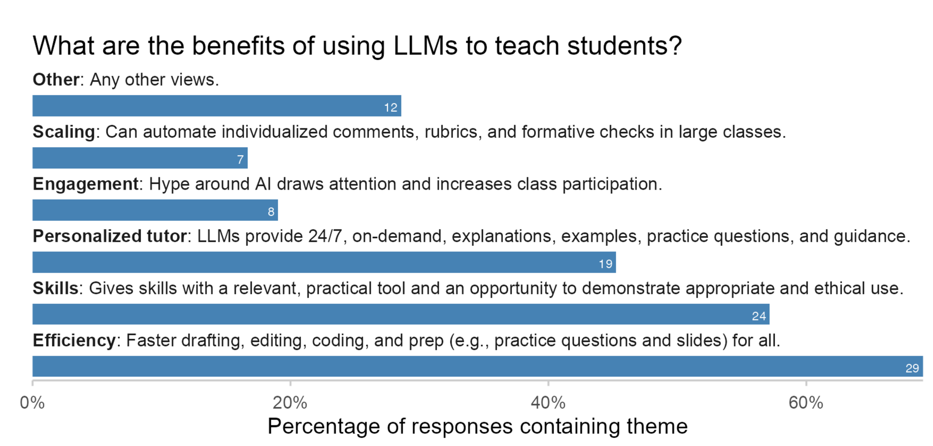


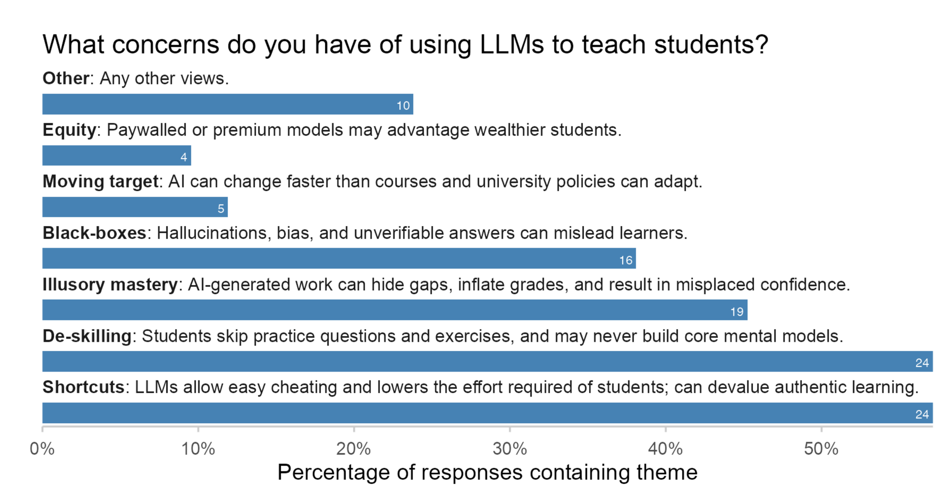

The interviews were then transcribed and anonymized. We created a codebook, coded the qualitative aspects, and then conducted analysis.

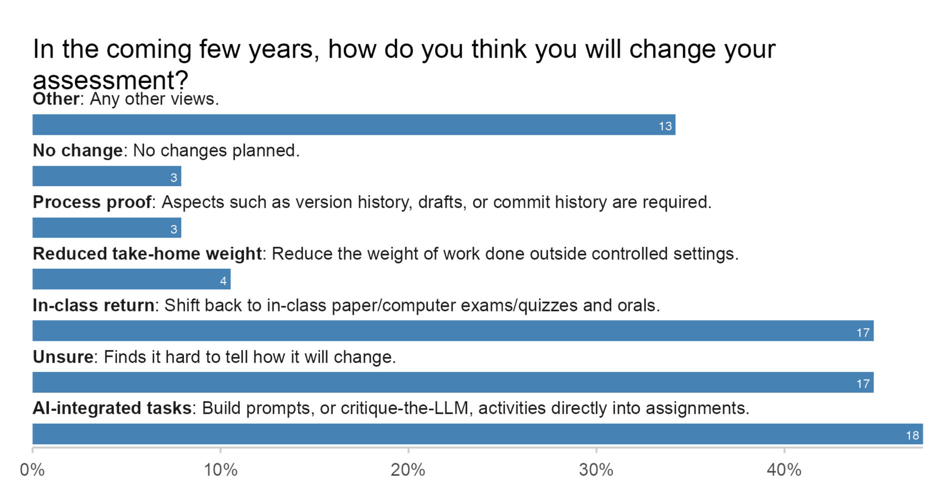
## Findings

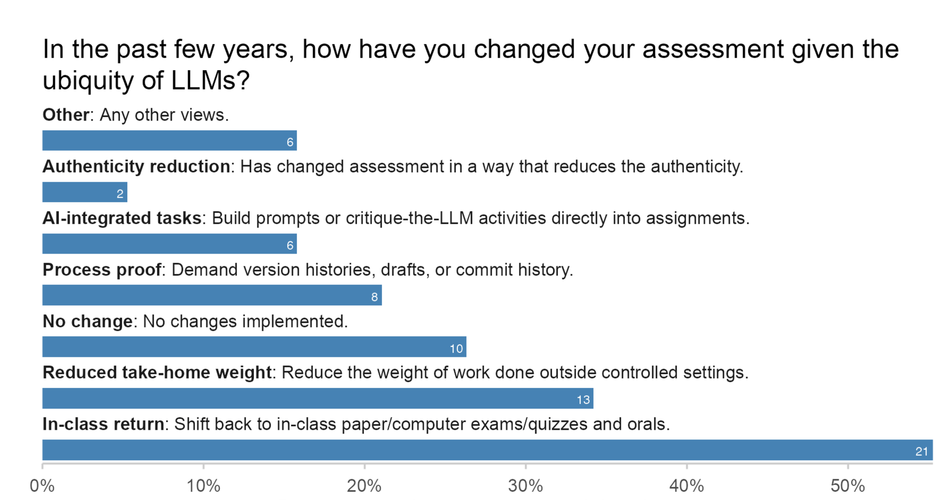
Some responses were coded uniquely, but others could have multiple codes. More importantly, one response could be coded to multiple themes, so our bar graphs do not need to add to 100 per cent. Finally, some questions were not asked of some respondents where their previous responses made it clear it was not relevant.

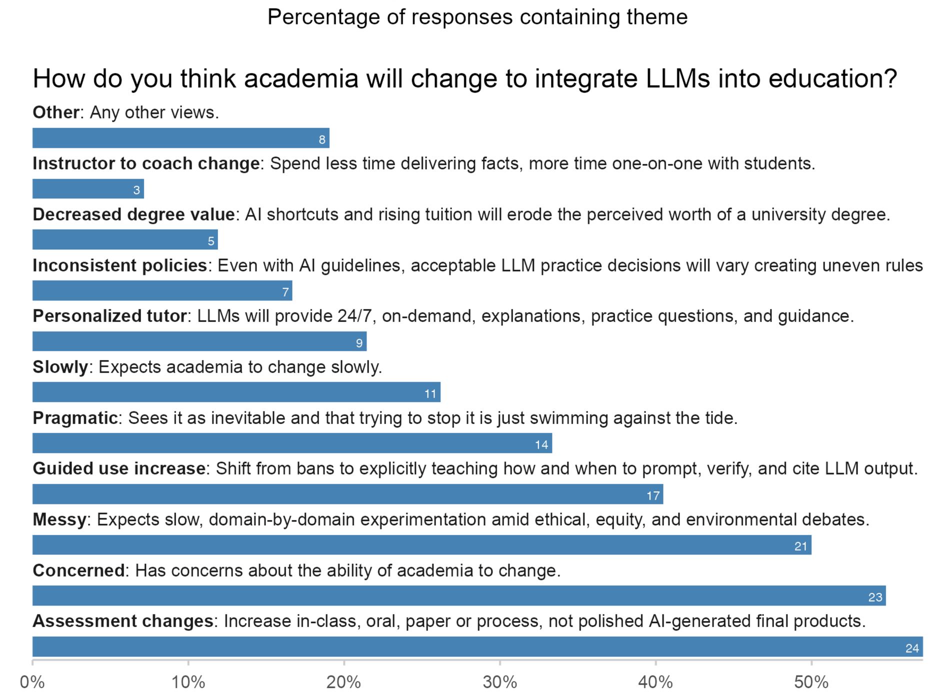


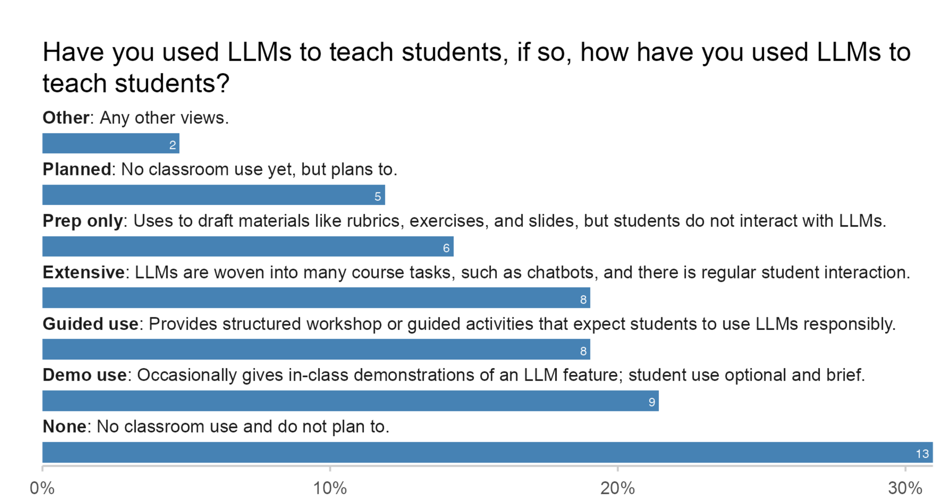




## Responses to LLMs

There are a variety of ways that respondents reported having changed aspects of their teaching because of LLMs.

One common approach was to have a different perspective on LLM-use in different parts of the teaching period. One respondent divided the year into four quarters. In the first quarter, the students were not allowed to use LLMs for anything. In the second quarter, they were encouraged to use LLMs as a tutor, but to try to use it for learning itself. And then in the last half of the year, they were allowed to use LLMs as their own assistant.

Another common approach was to have different categories of assessment and have different LLM-use for each type. One respondent split assessment into three categories: preparation, application, and secure. Assessment in the preparation category included online reading quizzes with unlimited attempts. These could potentially be completed with LLM assistance. Assessment in the application category included assignments and projects that were done individually and were not meant to be done with LLM-assistance, but there was no way of enforcing this. Finally, assessment in the secure category were in the form of an in-person midterm and final, and LLM-assistance was both not possible and enforced. Every question was a variation from the reading quiz or programming assignments. The overall grade in the class was the minimum of the three categories.

A different respondent gave students the option of declaring that they had solved their assignment with AI but by doing so the student would forfeit the option of getting feedback on their work.

Some respondents described how their university has established computer-based assessment in a proctored classroom with a computer on lockdown. Similarly, in programming courses where LLMs can do much of the coding, some respondents had added more exercises requiring students explain and debug code, rather than generate it. Another respondent reported making an assignment where students would have to critically assess LLM outputs. A couple of respondents mentioned having a greater emphasis on student teaching and student presentations so they have to take active agency over the material.

One respondent described adding white-colored writing to their assignment PDF that the LLM would read and interpret, but that a student would not notice. If the student did not engage with the problem at all, they would not notice that the response generated by the LLM had been affected by the hidden text.

One respondent suggested that instructors should instead divide their class into two parts: applied and tools-based learning. In the applied part, students need to process and understand information. The assessments are secure and the questions focus on conceptual understanding. Once the student has a firm foundation, the instructor can introduce tools-based learning. In the tools-based learning part, students can use any tool.

## Conclusion and Future Studies

This study reveals that LLMs are already reshaping data science education. While uncertainty about long-term impacts remains widespread, most respondents view LLMs as an inevitable and as a potentially transformative force requiring thoughtful integration rather than resistance.

Progress can be achieved with Instructors sharing ideas that are successful – through open communication across institutions, students can have a unified view of LLMs in education.

Future work look at what current students think of LLMs, how they are using them, and what concerns they may have. It would also be useful to document how data scientists actually use LLMs in their work.

## Acknowledgements

I thank Professor Rohan Alexander and Professor Tiffany Timbers for their supervision, guidance, and for making this project possible. This work was completed as part of an internship at the Investigative Journalism Foundation. This project is supported by the Data Sciences Institute, University of Toronto.


a.lopez.miranda@utoronto.ca, tiffany.timbers@stat.ubc.ca, rohan.alexander@utoronto.ca