Journal of Data Science 24 (1), 269–271
January 2026
DOI: 10.6339/25-JDS1208F
Education in Data Science

Discussion of "Addressing the Challenges of AI-Generated Assignment Submissions in Education: Insights and Strategies"☆

Xiaoming Huo[1] and Xuelei Sherry Ni[2,*]
[1]*H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA*
[2]*School of Data Science and Analytics, Kennesaw State University, Kennesaw, GA, USA*

Keywords *AI auditing; assessment redesign; GenAI/ChatGPT; model diagnostics*

Wang et al. (2025) argue that the surge of AI-generated assignments is fundamentally a pedagogical problem rather than merely an integrity or detection problem, and that durable solutions must focus on assessment and curriculum design. Recent work in higher education reinforces this view: scoping surveys and systematic reviews show that tools like ChatGPT are now widely used by students (Freeman, 2025), with mixed but substantial impacts on how they approach problem solving and assessment. In other words, AI is not an occasional shortcut but an embedded part of students' learning ecosystem, and our curriculum and assessments must be designed with that reality in mind.

In regression analysis and other modeling courses, this reality is especially stark. Generative AI can now produce syntactically correct R or Python code, fit models, generate polished tables, and even supply plausible narrative interpretations. If homework primarily asks students to "run this regression and interpret the output," it becomes easy to submit work that appears competent without internalizing the underlying model, its assumptions, or the properties of the estimators. Furthermore, simply hiding data, randomizing parameters, or relying on detection tools does not fundamentally change the situation, because students can still feed prompts and data into AI systems. For us as instructors of graduate modeling courses, the central competency at risk is **model literacy**: the ability to reason from assumptions, to understand concepts such as bias, efficiency, and consistency, and to exercise judgment about model choice and use.

A further complication is that AI capability is a moving target: what is difficult for AI today (e.g., subtle causal interpretation) may be routine tomorrow. Any pedagogical strategy that relies on current AI weaknesses is therefore inherently short-lived. The only sustainable response is to emphasize skills that are uniquely human and hard to automate: engaging with the axiomatic foundations of models, articulating and critiquing assumptions (e.g., Gauss–Markov conditions), and applying metacognitive judgment about the ethical and contextual appropriateness of modeling decisions. Building on Wang et al. (2025), we therefore seek assessment strategies in graduate regression courses in which high-stakes credit depends primarily on students' own reasoning rather than on AI-generated text or code. This aligns with Hu's (2025) call to "grade the thinking, not just the text," as well as with El Khoury's "waves of assessment and GenAI" framing, which urges a shift from purely product-based assessment toward approaches that foreground process and judgment. In the following, we outline concrete homework-level strategies designed to enforce student engagement and remain robust as AI tools continue to evolve.

☆Main article: https://doi.org/10.6339/25-JDS1208.

*Corresponding author. Email: sni@kennesaw.edu.

1. Model Defense Briefs
   One concrete strategy is to shift grading weight from coding to argument through "Model Defense Briefs." Students may use AI to help write R or Python code for Ordinary Least Squares (OLS), Instrumental Variables (IV), or Generalized Least Squares (GLS), but the assessment focuses on their written justification of the model in statistical terms. A brief might require:
   - **Assumption mapping:** Identify which Gauss-Markov assumption(s) are plausibly violated in a baseline OLS specification (e.g., heteroscedasticity, omitted variables, correlated errors), with reference to residual plots or domain considerations.
   - **Method justification:** Explain, using core concepts such as bias, variance, and efficiency, why a particular remedy (e.g., GLS, robust standard errors, transformation, or IV) is appropriate and what problem it solves.

   AI can generate code, but it cannot substitute for a student's ability to articulate a coherent, model-based defense. The brief thus becomes the primary evidence of competency.
2. Process-Focused Documentation
   Secondly, following calls in Wang et al.'s (2025) paper to emphasize process as well as product in AI-era assessment, in regression and modeling courses, this documentation, called decision log, must strictly track the logical progression of analytical decisions. This requires students to articulate the justification for all key modeling choices. For example: "The decision to include a quadratic term was not arbitrary but based on the initial residual plot suggesting non-linearity in the relationship, consistent with the theoretical model of diminishing returns." This ensures that students own the analytical decision-making, providing evidence of the internal statistical reasoning, which is the true measure of their competency.
3. Forensic "AI-Auditing" Assignments
   A third strategy is to reposition students as "forensic statisticians." In line with Wang et al.'s suggestion to use intentionally flawed AI-generated solutions, instructors can provide a regression analysis (possibly produced or partially produced by AI) that contains realistic but consequential problems, for example, ignoring heteroscedasticity, misinterpreting an interaction, or extrapolating far outside the data range. Students are then asked to act as an auditor, using appropriate plots and tests to diagnose the model's inadequacy for inference (e.g., invalid standard errors, misleading effect sizes), and subsequently propose a sound corrective action. This forces critical evaluation and debugging skills, the ability to **critique and repair** a model produced by a machine. Instructors can extend this strategy by incorporating brief peer review, asking students to comment on one another's model defense briefs or decision logs, further sharpening their ability to recognize and explain conceptual errors in regression practice.
4. Additional Homework Extensions
   Finally, we propose to layer in several complementary structures that reinforce the same goals while explicitly leveraging AI as a learning partner. One extension is a linked homework-quiz sequence in which AI-assisted homework is followed by short, low-stakes concept quizzes on new but structurally similar scenarios. Here, students may use AI freely on the homework, but must subsequently demonstrate transfer (identifying violated assumptions, choosing diagnostics, or selecting remedies) on fresh examples while AI usage is restricted (e.g., in-class quiz). Another extension is occasional micro "viva" checks in which students briefly explain, in their own words, key modeling choices or interpretations from their submissions, which follows Wang et al.'s idea of incorporating oral examinations and presentations. These checks

signal that understanding, not merely polished text or code, remains the central object of evaluation. Further, "design-before-prompt" tasks can ask students first to outline their own analysis plan (model specification, transformations, diagnostics) without AI, and only then query an AI tool, compare the two plans, and justify which elements of the AI's proposal they accept or reject. Together, these add-ons extend the framework proposed by Wang et al. (2025) by ensuring that homework remains a site of active, reflective learning: students are not merely generating answers with AI, but designing, explaining, critiquing, and transferring ideas in ways that keep human reasoning at the center of regression education in the AI era.

The challenge presented by Wang et al. (2025) is not simply an integrity problem, but a call for pedagogical elevation. If AI can already automate much of the mechanical work in regression, then competency must be redefined around what AI cannot reliably provide: coherent justification of models, critical assessment of assumptions, and sound statistical judgment with domain knowledge. The strategies we propose aim to make these human competencies the primary object of assessment. Rather than trying to exclude AI, we design tasks in which high-stakes credit depends on students' ability to explain, critique, and refine analyses, even when AI is part of their workflow. In doing so, we align with Wang et al.'s framework while adapting it to regression education, helping students learn not only how to use AI, but how to think beyond it. We believe many of the ideas proposed here can be extended to other data science and analytics courses.

Disclosure of AI Usage

During the preparation of this work, the author(s) used AI-powered tools (such as language models) for drafting assistance, grammar checking, and literature searching. The author(s) reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

El Khoury E (2025, October 6). Two years of generative artificial intelligence in higher education: The seven waves of assessment and GenAI. *The Assessment Review (CUNY)*, 6(1). https://assessatcuny.commons.gc.cuny.edu/2025/10/two-years-of-generative-artificial-intelligence-in-higher-education-the-seven-waves-of-assessment-and-genai/.

Freeman J (2025). Student Generative AI Survey 2025 (HEPI Policy Note 61). Higher Education Policy Institute / Kortext. Retrieved from https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/.

Hu X (2025, October 14). Assessment in the AI era: Grade the thinking, not just the text. University of Arizona, College of Information Science. Retrieved from the InfoSci "Thought Leadership" series. https://infosci.arizona.edu/news/assessment-in-the-ai-era.

Wang S, Xu L, Liu J, Zhai Y (2025). Addressing the challenges of AI-generated assignment submissions in education: Insights and strategies. *Journal of Data Science*, 24(1): 1–7. https://doi.org/10.6339/25-JDS1208