

Module 3: Limitations

Taylor J. Reininger
trinin2@jh.edu
EN.605.256
2025-06-08

Written Assignment Requirements:

*Having carried out this assignment, please write two paragraphs about the inherent limitations of carrying out analytics over anonymously submitted data items. Did the analytic responses surprise you? How does this differ from standards? For example, the average GRE quantitative reasoning score was 157 for 2023-2023 and was nearly 165 for grad school entries submitted (see sample output). Why do you think that is? What might cause this to occur? Please place your essay into a file called **limitations.pdf***

Answer:

Doing analytics on anonymously submitted data has a few major issues:

1. The data can be in the wrong format or lie outside of expected bounds because users are allowed to freely type in many of the fields.
2. Some fields are empty, meaning the data is incomplete.
3. There are likely to be biases that you can't account for. For example: students with low GPAs may not self-report them, where students with high GPAs could be more likely to include them.

I was a little surprised at how high the GPAs were for the students across the board. I wonder if this is skewed by some of the concerns mentioned above or if the type of people that use Grad Cafe tend to have higher GPAs than others. One obvious reason may be that students are essentially self-selecting. If a student had a GRE score significantly lower than what they see on Grad Cafe, they may be discouraged from posting their score due to their embarrassment or other emotional responses. It could also indicate a different type of selection – students who get lower GRE scores may be less likely to do extra research on graduate schools and therefore may be less likely to visit or post on Grad Cafe at all. This is entirely unsupported speculation and would require further interrogation to validate.