Data Limitations

The analysis of anonymously submitted data from GradCafe presents several limitations that affect the reliability and generalizability of our findings. First, there is a significant self-selection bias, as individuals who choose to report their application outcomes on GradCafe are not representative of the entire applicant pool. Successful applicants with higher scores are more likely to share their results, creating an upward bias in metrics like GPA and GRE scores.

Furthermore, the data suffers from incompleteness and inconsistency issues. Many entries lack critical information such as GPA or standardized test scores, limiting the sample size for certain analyses. The free-form nature of submissions also leads to inconsistencies in how information is reported, making it difficult to standardize and compare across entries. These limitations highlight the need for caution when interpreting results from such data sources and underscore the importance of considering these analyses as exploratory rather than definitive.