

Limitations of Using Self-Reported Admissions Data

A major limitation of using Grad Cafe and similar platforms as data sources is that the information is entirely self-reported and anonymous. Applicants choose whether and how to report their outcomes, test scores, and academic metrics, which introduces significant selection bias. Users with unusually strong profiles or notable outcomes (acceptances or rejections from elite programs) are more likely to post than average applicants, while many entries omit key fields such as GPA or GRE scores entirely. Additionally, because submissions are anonymous, there is no mechanism to verify the accuracy of the reported data, and duplicate or exaggerated entries may exist. These factors limit the representativeness and reliability of the dataset.

The analytic results produced in this project were not entirely surprising given these limitations. For example, the average GPA and GRE scores observed in the dataset are notably higher than national averages reported by ETS. This discrepancy is likely caused by selection bias, as applicants to highly competitive programs and international students—who often face higher admissions thresholds—are overrepresented on Grad Cafe. Furthermore, applicants may selectively report strong metrics while omitting weaker ones, inflating observed averages. As a result, while Grad Cafe data can be useful for exploratory analysis and trend observation, it should not be interpreted as a statistically representative sample of the overall applicant population.