Limitations:

Analyzing anonymously submitted data comes with the risk of no checks and balances. The data is self-reported so there is no fact checking. The fields are also open, so inconsistencies in naming convention and spelling errors are present and require sanitization. Users are also not required to fill in every field which means a complete picture of an applicant is not necessarily available. Building on that, users who chose not to fill out a field may enter "filler" data that must be sanitized. I noticed that specifically with the score options, if users did not opt to provide their scores, they sometimes filled out a variation of 0 or 99 which I had to account for when sanitizing.

The analytical results did differ from the average scores. According to ets.org, the average verbal GRE score was 151.21, the average analytical writing score was 3.44 and the average total GRE score was 308.79. The average scores in my data are 161.91 for the verbal GRE, 4.54 for analytical writing and 263.54 for the average GRE. There is a large limitation at play when it comes to the total average GRE score – this technically should be a combination of the verbal and quantitative reasoning scores putting it in the 250+ range; however, my data includes lots of entries in that range as well as entries in the ~150 range which leads me to believe that some people input their analytical reasoning score rather than their total GRE score which is impacting that number. The other scores are higher than average. I also noticed that the average accepted GPA at 3.80 is only 0.02 points higher than the average overall at 3.78. My thought is there is a self-reporting bias for people who were accepted versus rejected and people who were accepted were more likely to report their scores. I did a sanity check on this for the GPAs and found that for the Fall of 2026, 66.56% of applicants who were accepted provided their GPA compared to 58.77% of applicants who were rejected. While this number is not wildly different, there is still an indication that people who are accepted are more likely to report, and I have a hunch that they are more likely to have higher scores which is why the data skews higher than average.