

Written Assignment

The analytic responses didn't surprise me at all because I don't put much trust in anonymously submitted data. Working with it is fast (no formal validation), but it's messy and biased. The sample is self-selected and may be affected by bragging or volunteer bias, meaning people are more likely to report strong scores, and rejections often go unreported. Because it's anonymous, verifying anything is difficult. I believe it is almost impossible. There can be duplicates, incomplete fields, or inflated/inaccurate values. Formats aren't consistent either; for example, GRE Analytical Writing had a few out of range values.

The average GRE Quantitative score is about 157 from legitimate sources, while the scraped data shows around 165. That supports my view. People tend to post after a better retake, some round up, and many ignore rejections or low scores. I tried to reduce noise with range checks and other cleaning, but self-selection bias remains. I read these results as comparative, not absolute – useful for patterns (program popularity, timing, etc.), but not a replacement for official numbers. Overall, this anonymously submitted dataset is great for directional trends, not for official statistics.