

There are many inherent limitations to using just the data scrapped from a website like gradcafe. The first and most obvious is that the data is self-reported, nothing is don't to confirm or check that any data input is accurate. There may be small controls to make sure the data input is the right type. But there were reports of GREs and GPAs well outside standard bounds which skew data. Self-reporting can also lead to outright lies; there is the potential of someone claiming to be admitted when they were rejected. Although I do hope this is uncommon.

Another inherent limitation of grad café is the sample size. The only people who are posting on gradcafe are those who are aware of gradcafe and desire to post on it. This means that we are pulling data from a much smaller sample size than desired and, if the hopeful gets bad news or bad results, they will be less inclined to post their outcomes on gradcafe. This will strongly skew the results in the positive direction. This can be seen in most effectively in the average GRE score. The average score nationally reported was 157 where the one on average from gradcafe was nearly 165! This is a massive jump indicative of selective reporting. Overall although gradcafe is a good resource that can show some general trends on grad admissions, it should not be used as a highly accurate pool of statistics.