Exploratory Data Analysis
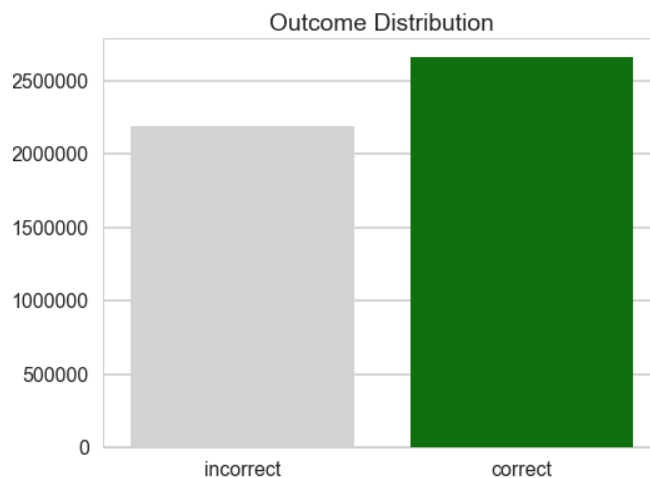## Predicting Mastery in Academic Assessments

All work described in this report can be viewed in the exploratory_data_analysis notebook.

Create data visualizations to explain variables and analyze outliers. Possibly remove or mark outliers. Explore variable dependence and correlations.
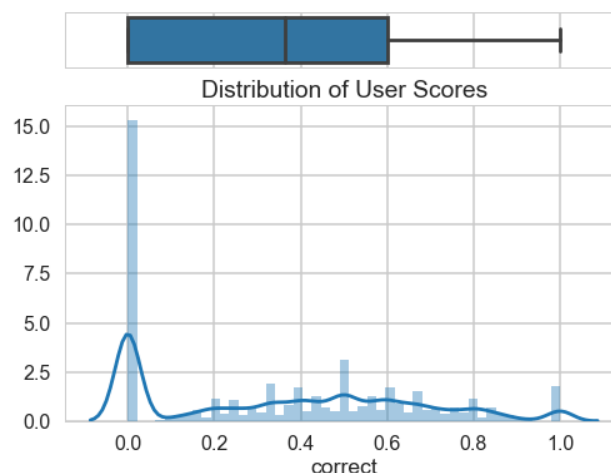
Dataset Statistics

For all records in the clean dataset the mean percentage of correct answers is 54.8% and the mean round duration is 1 minute, 8 seconds.

User Data

The histogram produced by cell 13 in the notebook indicates that a small percent of users attempted a large number of questions and a large percent attempted a small number of questions. This certainly explains the distribution of the ratio of correct answers by user:
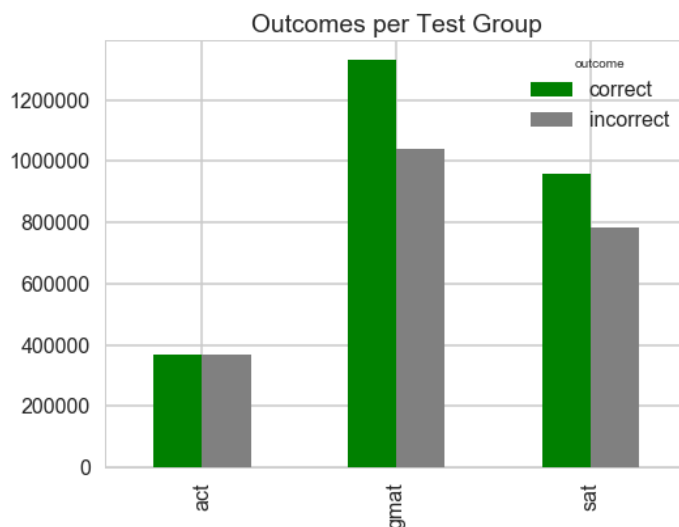
Question Data

The distribution of unique questions showed that a large number are underrepresented in the dataset. This is explained by the distribution of test groups (next section).

A look at prior attempts of unique questions by users revealed that one user attempted one particular question 1268 times. There were at least 12 other questions associated with excessive attempts but the closest count amongst them was for 274 prior attempts.

Group Name Data

There is no overlap in the tracks or subtracks across groups. There are far fewer observations in the ACT group than in the GMAT or SAT groups. The distribution of outcomes within groups is also different for the ACT group:

Given the independence of test groups, the dataframe was divided by group for further exploration. Only a small number of users participated in more than one group.

ACT Data

About 50% of questions in the ACT group were answered correctly, and the average round duration was about 53 seconds.

- Average score for ACT users: 0.347707
- Number of ACT users with no correct answers: 12198
- Average score for dedicated (attempting at least 30 questions) ACT users: 0.507080
- Number of dedicated ACT users with no correct answers: 8
- One user attempted 233 questions and didn't answer any of them correctly. This user was borderline slow in attempting questions, but nothing stood out to explain the poor performance.
- There are obvious outliers above the 2000 question mark with scores below 80% correct. Of those, the user attempting the most questions had the lowest score, but nothing indicated the user should be dropped.

- There is one user who spent a lot of time on the test application with no correct answers. The user's one question attempt was deactivated after 3 days. The question the user attempted may be problematic. Users answered it correctly for little more than 20% of the attempts and the average round_duration for the question is almost 20 minutes. This user's attempt was one of 238 for that question, so it was dropped.

GMAT Data

About 56% of questions in the GMAT group were answered correctly, and the average round duration was about 1 minute 18 seconds.

- Average score for GMAT users: 0.333503
- Number of GMAT users with no correct answers: 28007
- Average score for dedicated GMAT users: 0.547034
- Number of dedicated GMAT users with no correct answers: 2
- Three users never answered their one attempt, so they were dropped. Nothing stood out to indicate the questions attempted by these users were particularly difficult or unusual.
- One user had a very slow round duration average (1 hour, 7 minutes), but attempted 30 questions. For almost half of those, the question was deactivated eventually. The user never participated in a round alone.

SAT Data

About 55% of questions in the SAT group were answered correctly, and the average round duration was about 1 minute.

- Average score for SAT users: 0.381569
- Number of SAT users with no correct answers: 19281
- Average score for dedicated SAT users: 0.524551
- Number of dedicated SAT users with no correct answers: 3
- The user identified earlier for attempting one question 1268 times did so in the SAT group. Nothing stood out to explain this behavior, but it was the most extreme example of such, so this user was dropped.

Summary

While a small amount of users practice in more than one test group, no overlap in questions indicates the test groups are independent. For each test group, outcomes are better across all subjects for users attempting at least 30 questions. This effect is most dramatic for the GMAT group. Of all scenarios, average scores are best amongst the dedicated GMAT users.