Double Standards

The impact of disparate state approaches to standardized testing

Why Standardized Tests

Why standardized tests matter in admissions:

-Efficiency: Standardized test scores provide an empirical method to narrow down an unmanageably large applicant pool and to quickly decide between otherwise similarly promising candidates

-Consistency: Standardized test scores provide a standard metric to students whose schools use wildly different grading metrics

-Equity: Standardized tests can be a way for students with limited access to internships, AP classes and applications coaching to distinguish themselves

-Accountability: Statistically rigorous admissions criteria make it possible to audit admissions decisions for signs of bias or corruption

What’s wrong with standardized tests?

-Inefficiency: Students often spend months preparing to take these tests, building skills that may have little applicability to other endeavors. This is particularly a problem when the tests are administered to students who are not planning to apply to a selective university

-Inconsistency: Normalization makes it possible to rank students across schools on their performance in a standardized testing environment. They cannot be normalized based on the abstract qualities they are meant to test. Different approaches to standardized testing at the household, institution, district and state level may effect the the degree to which test performance correlates to college readiness

What’s wrong with standardized tests?

-Inequity: The standardized test prep industry in the US is valued at 30 billion dollars[1] including textbooks, workshops, classes and private tutoring. This suggests that, at the very least, applicants **believe** that their test scores come down to factors other than “college readiness” such as having time and money to prepare.

-Unaccountability: There’s no incentive to demonstrate or correct for any effect that differences in preparation have on test scores.

1. https://www.forbes.com/sites/forbestechcouncil/2020/04/29/why-the-test-preparation-industry-may-finally-get-out-of-the-classroom/

For standardized testing to live up to it’s potential, schools need empirical tools to gauge the effects of preparation

Can we develop tools to quantify and correct for bias introduced by test prep?

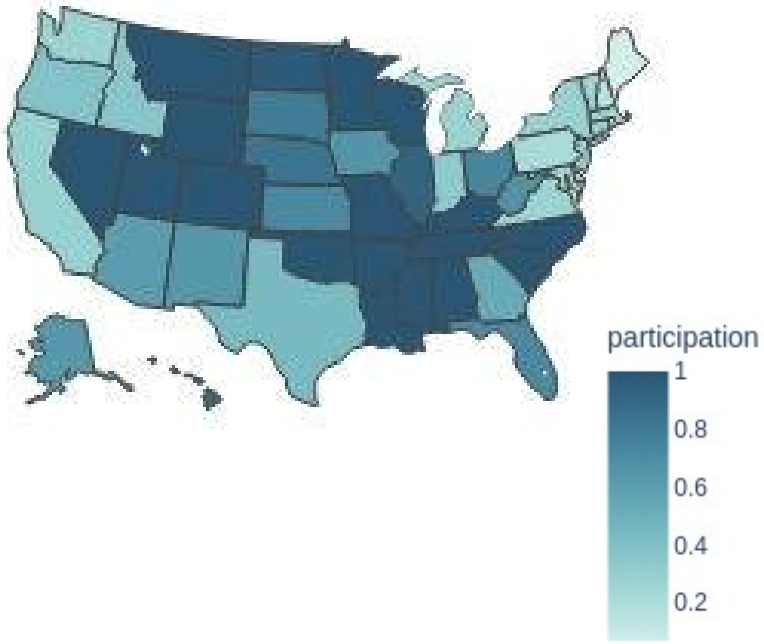
Methods:

- This project Is a preliminary exploratory analysis of the effect of state test participation rate on scores.
- Test participation was used as a rough estimate of preparation: it was assumed that higher participation would be correlated with more institutional support for the test and consequently more participation
- This analysis considered three years of data on the two major undergraduate admissions tests: ACT data from 2016, 2017, and 2018 and SAT data from 2017, 2018 and 2019

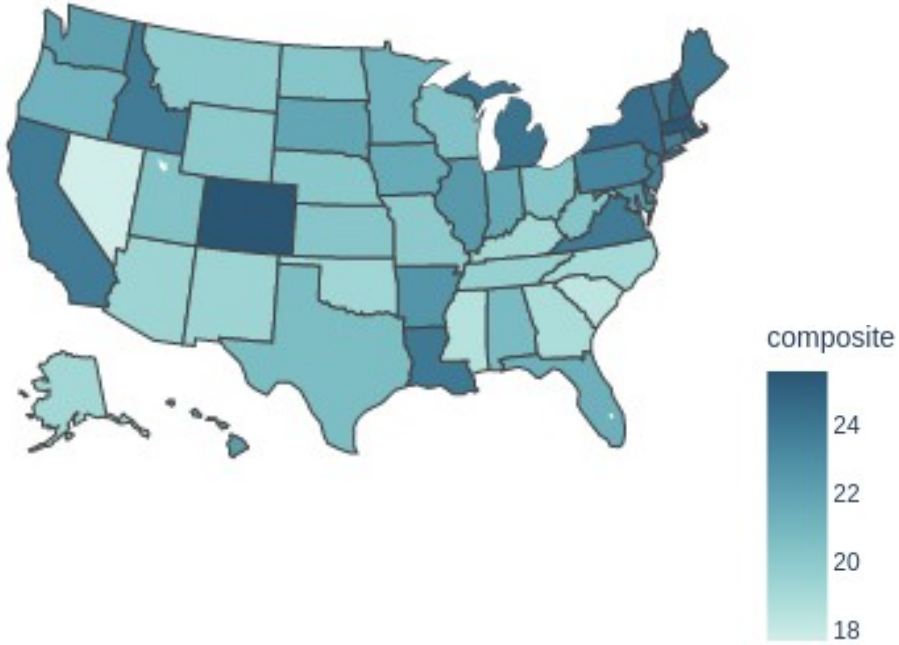
Key Findings

Data From ACT 2017

ACT Participation Rate



ACT Composite Score

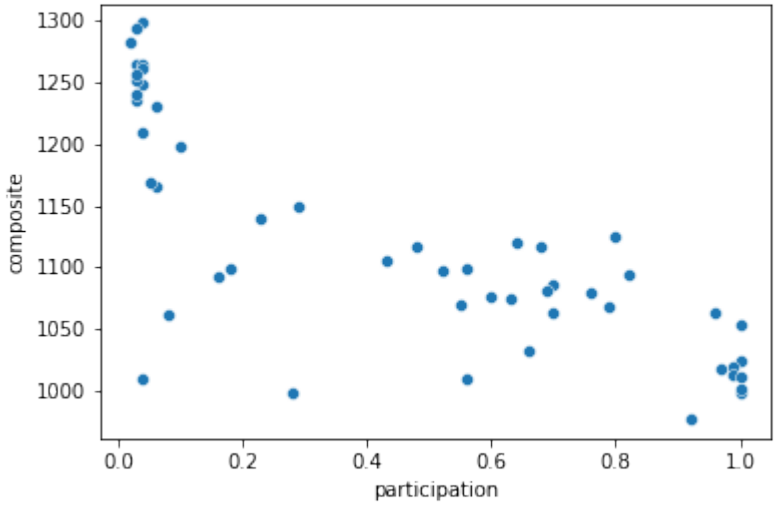


Inverse Correlation

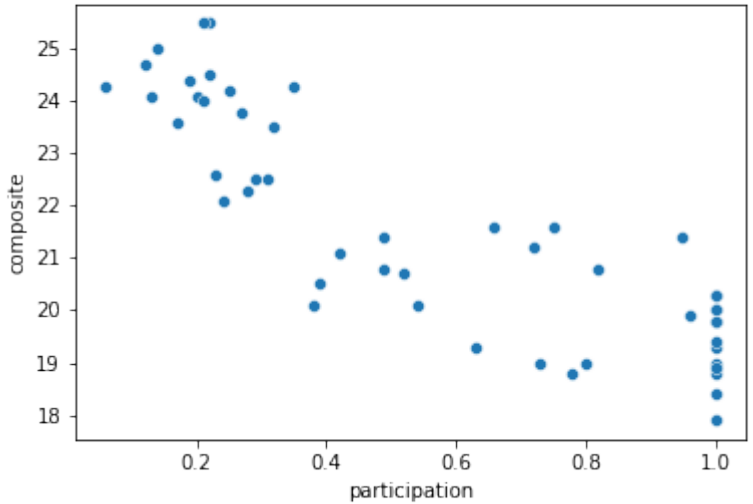
- The map on the left, which shows the participation rate of each state, looks like the negative of the one on the right, the composite score in each state
- This shows a strong inverse correlation between participation rate and score
- This is unsurprising since low participation rate states select for students most likely to attend college

SAT and 2017 Scatter plots composite vs participation

SAT



ACT

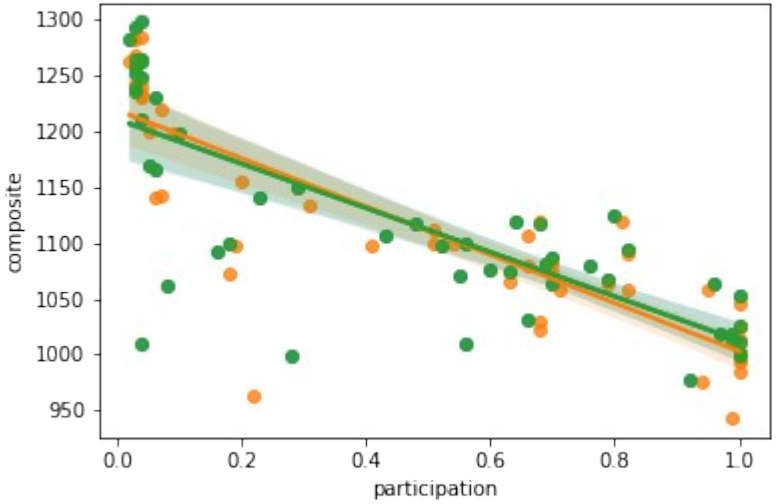


Scatter Plots

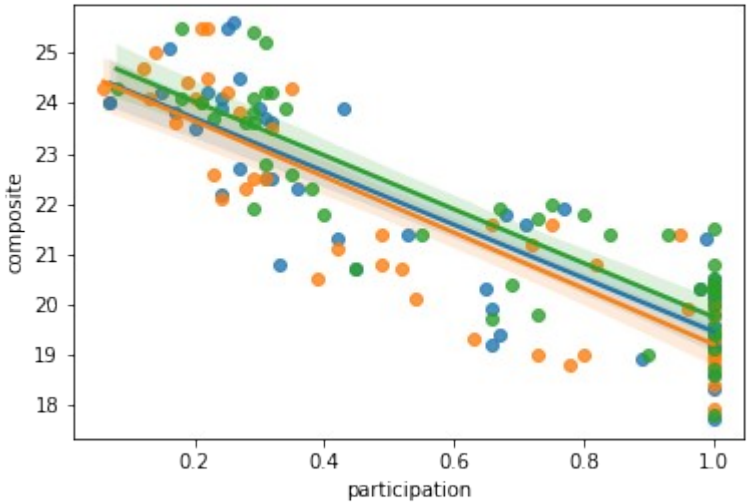
- These scatter plots reinforce the idea that there is a strong inverse correlation
- The distributions follow a different pattern though: There seems to be much more spread in the ACT distribution and less in the SAT.
- The relationship does not appear to be linear in either case, but levels off as the participation rate increases
- This would seem to indicate that, past a certain point, having more students take the test causes less of a drop off, which could be the result of more resources being devoted to test preparation

Regression plot: All years

SAT Regression



ACT Regression



Regression plots

- These scatter plots combine all cases of each test and have simple regression lines
- Briefly, regression lines are the lines that most closely go through all data, and help us visualize the general trend
- Despite the differences in how the data is distributed, the regression lines are very similar
- This model is limited because, on visual inspection, the data does not look to be linearly distributed
- Future research could try to fit the data to a more complicated regression model, such as polynomial regression

The Participation-Score-Index

- A fast way to approximate the effectiveness of a given state’s test preparation is to divide the mean score by the participation rate. I call this measurement the participation score index
- The details of this analysis is included in the EDA notebook included in this packet
- The short version is that the same the highest PSI scores were consistent across years but not across tests

The Participation-Score-Index

- On the ACT the top states by PSI were: Maine, New Hampshire, Rhode Island and Delaware
- On the SAT they were: North Dakota, Wyoming, Nebraska and Iowa

Takeaways

Takeaways

- There does seem to be an interesting relationship between test participation rate and composite score
- This might indicate broader trend in how preparation impacts performance
- Measures like the PSI as well as more complex statistical tools like polynomial regression could be useful for colleges attempting to correct for the effects of preparation on test performance
- More research is needed