

SAT & ACT Analysis

Geoffrey Liu





The background

Company

College Board is the organization that administers the SAT (Scholastic Assessment Test)

Changes

SAT had a new format released in 2016

Context

College Board would like to utilize data analysis on SAT participation rate across the US

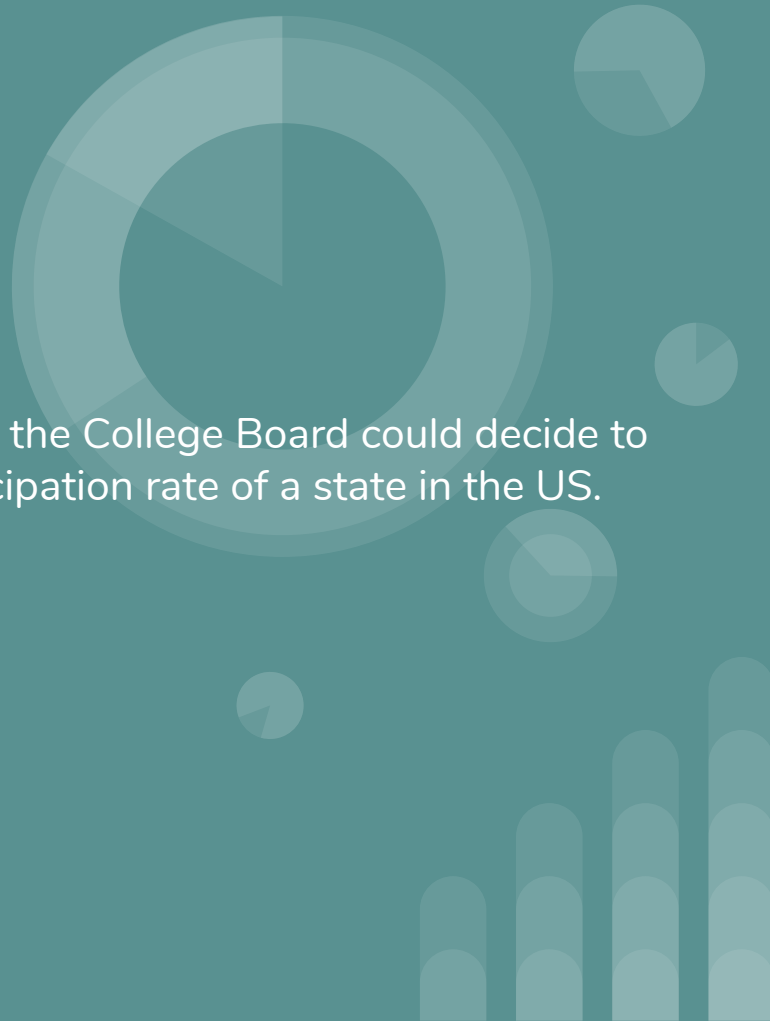
Contents

- Problem Statement
- Exploratory Data Analysis
- Data Visualization
- Descriptive and Inferential Statistics
- Outside Research
- Conclusions and Recommendations



Problem Statement

The project is to make recommendations about how the College Board could decide to allocate funding to work and increase the SAT participation rate of a state in the US.



Data we are working with

The data that will be analyzed upon for the years 2017 and 2018:

- ACT Participation Rates
- ACT English Scores
- ACT Maths Scores
- ACT Reading
- ACT Science
- ACT Composite Scores

- SAT Participation Rates
- SAT Evidence-based Reading & Writing Scores
- SAT Maths Scores
- ACT Total Scores

Exploratory Data Analysis

	state	act2017_participation_rate
0	Alabama	1.00
17	Kentucky	1.00
49	Wisconsin	1.00
44	Utah	1.00
42	Tennessee	1.00
40	South Carolina	1.00
36	Oklahoma	1.00
33	North Carolina	1.00
28	Nevada	1.00
26	Montana	1.00
24	Mississippi	1.00
23	Minnesota	1.00
18	Louisiana	1.00
25	Missouri	1.00
50	Wyoming	1.00
5	Colorado	1.00
3	Arkansas	1.00

-100% ACT participation rate occurs in 17 out of 50 states

Compared to 100% SAT participation rates for just about 4 out of 50 states

-We can see ACT participation rates tend to fare better than SAT

-We will further touch on this in data visualization

Continued

-If median is less than the mean the data is right skewed and if the median is greater than the mean the data is left skewed.

-ACT PR for 2017 and 2018 have their means and medians only slightly apart from each other, therefore skewness is faint.

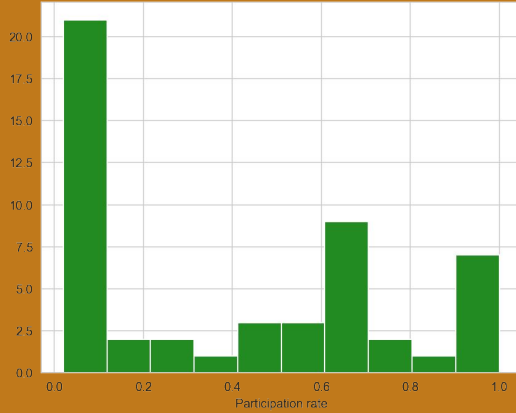
-its noticeable that while SAT 2017 PR mean of 0.39 is only slightly higher than its median of 0.38, SAT 2018 PR saw some left skew with its median of 0.52 and mean of 0.45.

-For the various subtests of ACT and SAT, their means and medians do not differ much from each other

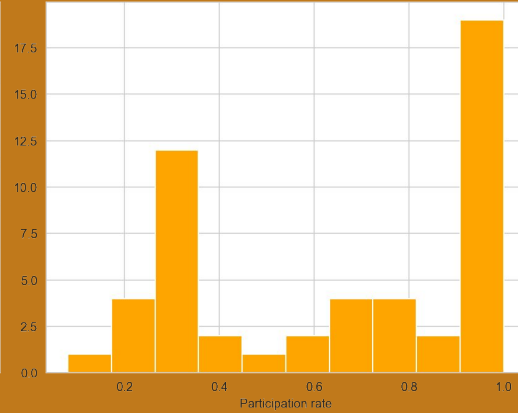
	count	mean	std	min	25%	50%	75%	max
act2017_participation_rate	51.0	0.652549	0.321408	0.08	0.310	0.69	1.000	1.0
act2017_avg_english_score	51.0	20.931373	2.353677	16.30	19.000	20.70	23.300	25.5
act2017_avg_math_score	51.0	21.182353	1.981989	18.00	19.400	20.90	23.100	25.3
act2017_avg_reading_score	51.0	22.013725	2.067271	18.10	20.450	21.80	24.150	26.0
act2017_avg_science_score	51.0	21.450980	1.739353	18.20	19.950	21.30	23.200	24.9
act2017_composite_score	51.0	21.519608	2.020695	17.80	19.800	21.40	23.600	25.5
sat2017_participation_rate	51.0	0.398039	0.352766	0.02	0.040	0.38	0.660	1.0
sat2017_avg_evidence_based_reading_and_writing_score	51.0	569.117647	45.666901	482.00	533.500	559.00	613.000	644.0
sat2017_avg_math_score	51.0	556.882353	47.121395	468.00	523.500	548.00	599.000	651.0
sat2017_avg_total_score	51.0	1126.098039	92.494812	950.00	1055.500	1107.00	1212.000	1295.0
act2018_participation_rate	51.0	0.616471	0.340810	0.07	0.285	0.66	1.000	1.0
act2018_avg_english_score	51.0	20.988235	2.446356	16.60	19.100	20.20	23.700	26.0
act2018_avg_math_score	51.0	21.125490	2.035765	17.80	19.400	20.70	23.150	25.2
act2018_avg_reading_score	51.0	22.015686	2.167245	18.00	20.450	21.60	24.100	26.1
act2018_avg_science_score	51.0	21.345098	1.870114	17.90	19.850	21.10	23.050	24.9
act2018_composite_score	51.0	21.486275	2.106278	17.70	19.950	21.30	23.550	25.6
sat2018_participation_rate	51.0	0.457451	0.373143	0.02	0.045	0.52	0.775	1.0
sat2018_avg_evidence_based_reading_and_writing_score	51.0	563.686275	47.502627	480.00	534.500	552.00	610.500	643.0
sat2018_avg_math_score	51.0	556.235294	47.772623	480.00	522.500	544.00	593.500	655.0
sat2018_avg_total_score	51.0	1120.019608	94.155083	977.00	1057.500	1098.00	1204.000	1298.0

Data visualization

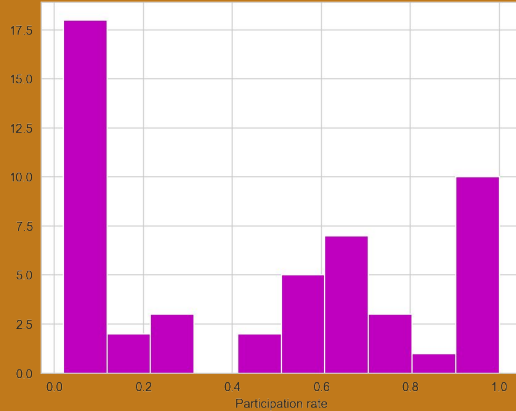
SAT2017 participation rate



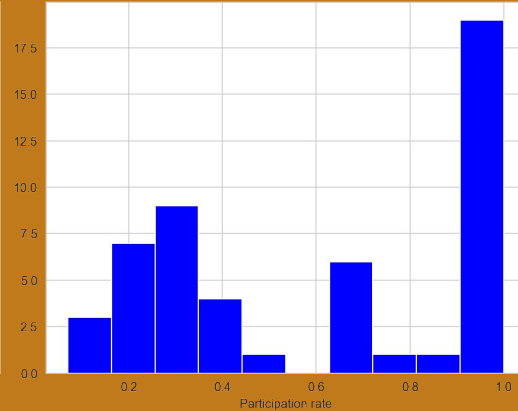
ACT2017 participation rate



SA12018 participation rate



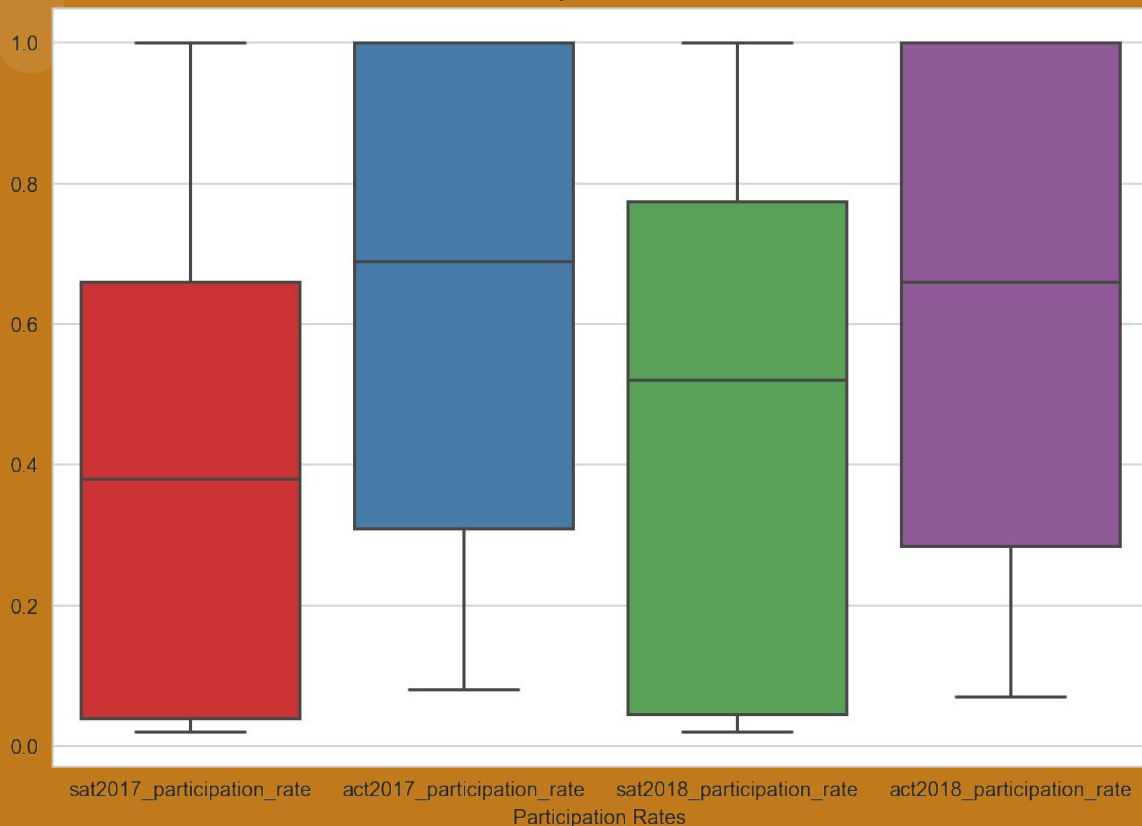
ACT2018 participation rate



- Notice the SAT participation rate histograms having left sided peaks
- Indicates low participation close to 0% is most frequently occurring
- Notice the ACT participation rate histograms having right sided peaks
- Indicates high participation close to 100% is most frequently occurring
- Could possibly suggest data might be biased

Continued

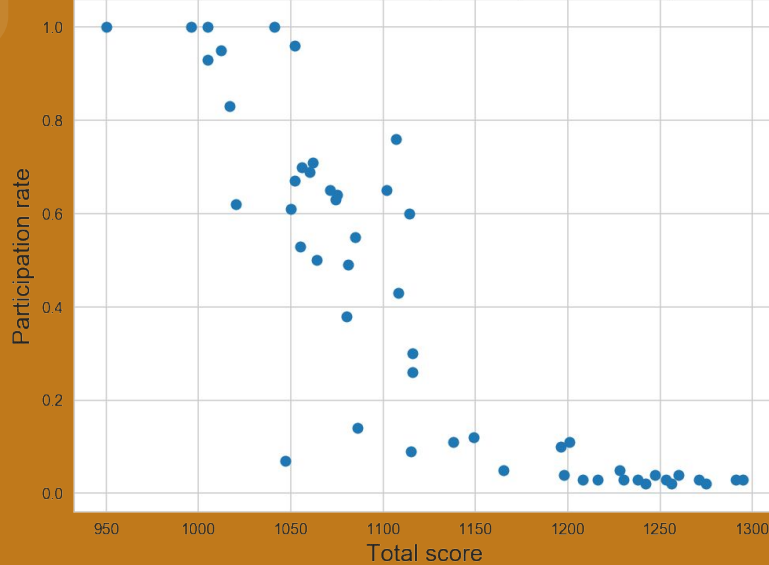
SAT and ACT Participation Rates 2017 and 2018



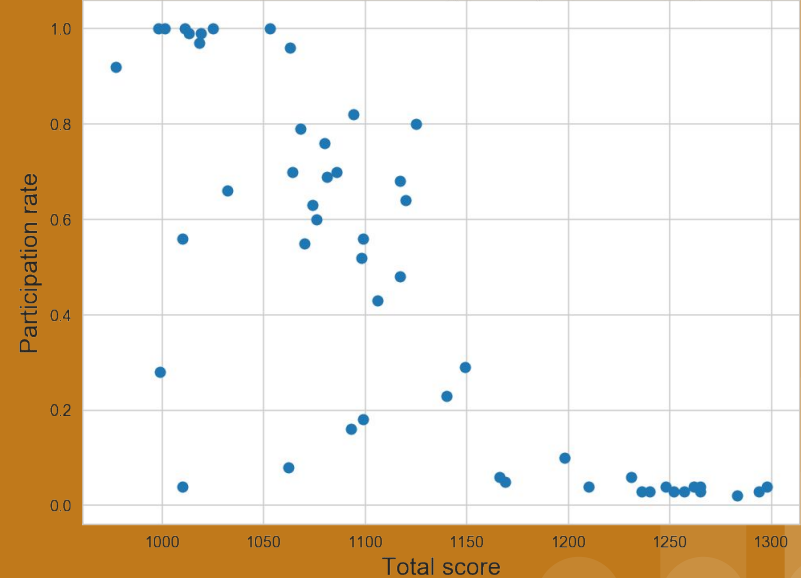
- Boxplots can help further describe the behavior of the histograms
- SAT participation rate 25-50-75 quartiles are around the lower end of participation rates
- Has longer top whiskers indicating outliers are high participation rates
- ACT participation rate 25-50-75 quartiles are around the higher end of participation rates
- Has longer bottom whiskers indicating outliers are low participation rates

Continued

SAT2017 total score and participation comparsion



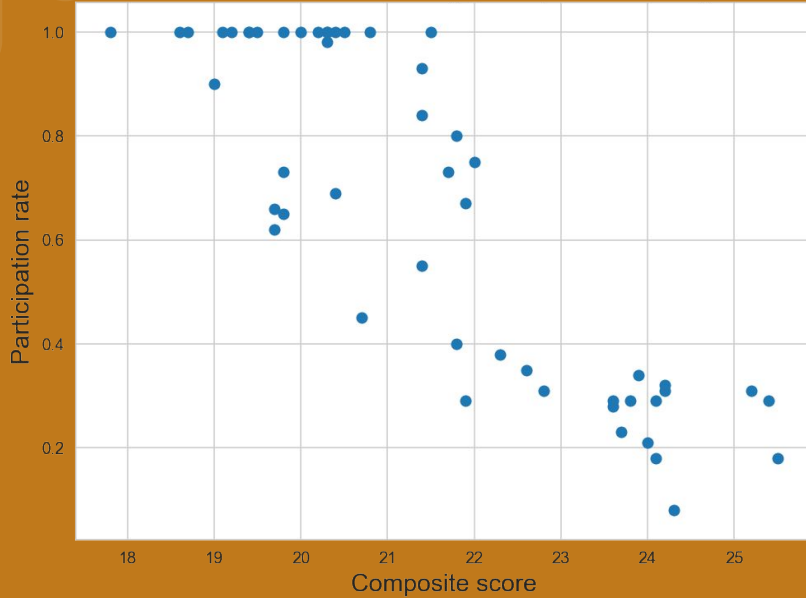
SAT2018 total score and participation comparsion



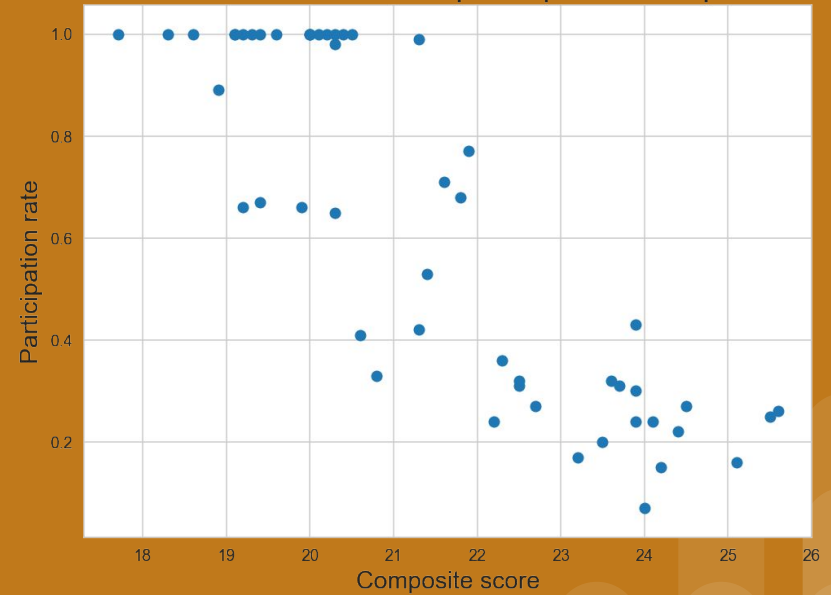
-We observe negative correlation between SAT participation rates and total scores

Continued

ACT2017 total score and participation comparison



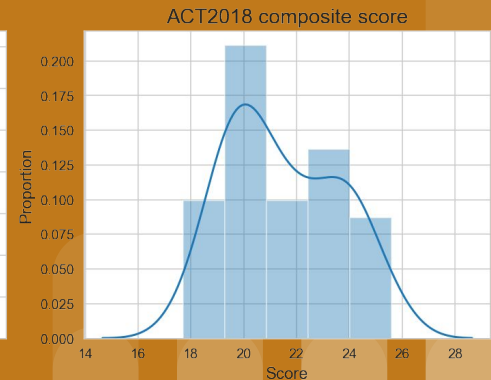
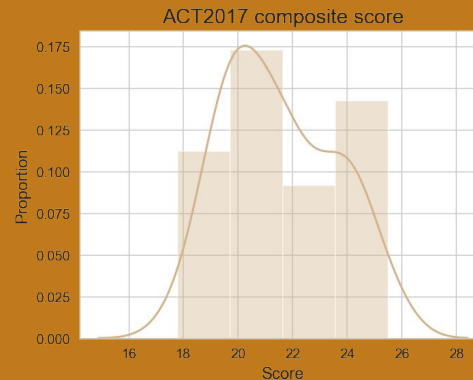
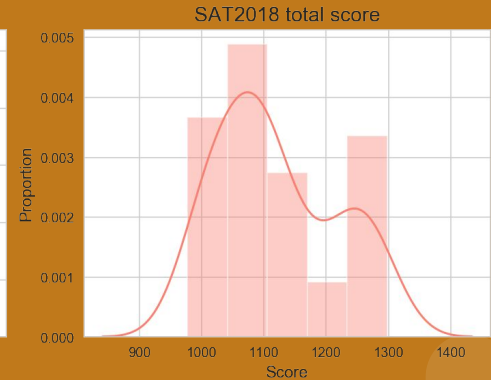
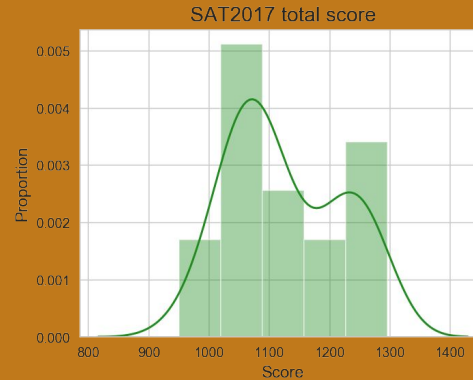
ACT2018 total score and participation comparison



-We also observe negative correlation between ACT participation rates and composite scores

Descriptive and Inferential Statistics

- Graphs shown earlier on indicates data is biased towards extremes of participation rates
- Negative correlation is observed
- The central limit theorem states that as n gets bigger and is at least 30 and data is randomly distributed, the distribution would look like a normal distribution. However, this is not the case
- We can see perhaps participation is not all random and bias exists.
- Observe the SAT total and ACT composite scores distribution plots
- Kernel density estimators are non parameter descriptions of the probability density function
- They do not form the normal distribution



Continued

-Conducted hypothesis testing

Scenario 1

-null hypothesis is population mean of SAT participation rate 2017 is equal to ACT participation rate 2017

-Alternative hypothesis is population mean of SAT participation rate 2017 is not equal as ACT participation rate 2017

-p value obtained is 0.00024

-Assuming alpha is 0.05, reject the null hypothesis

Scenario 2

-null hypothesis is population mean of SAT participation rate 2018 is equal to ACT participation rate 2018

-Alternative hypothesis is population mean of SAT participation rate 2018 is not equal as ACT participation rate 2018

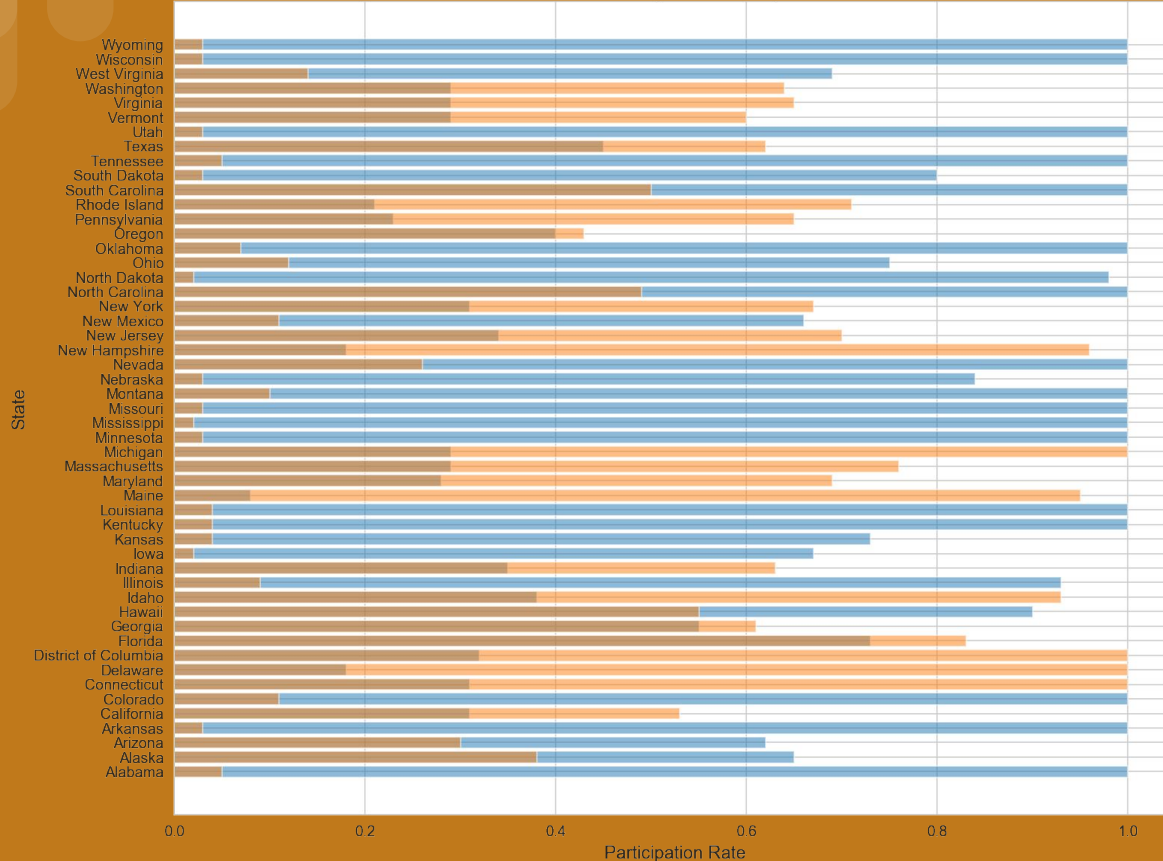
-p value obtained is 0.02684

-Assuming alpha is 0.05, reject the null hypothesis

-Though we carried out hypothesis testing, it is not reliable as data is assumed to not be randomly distributed

Outside Research

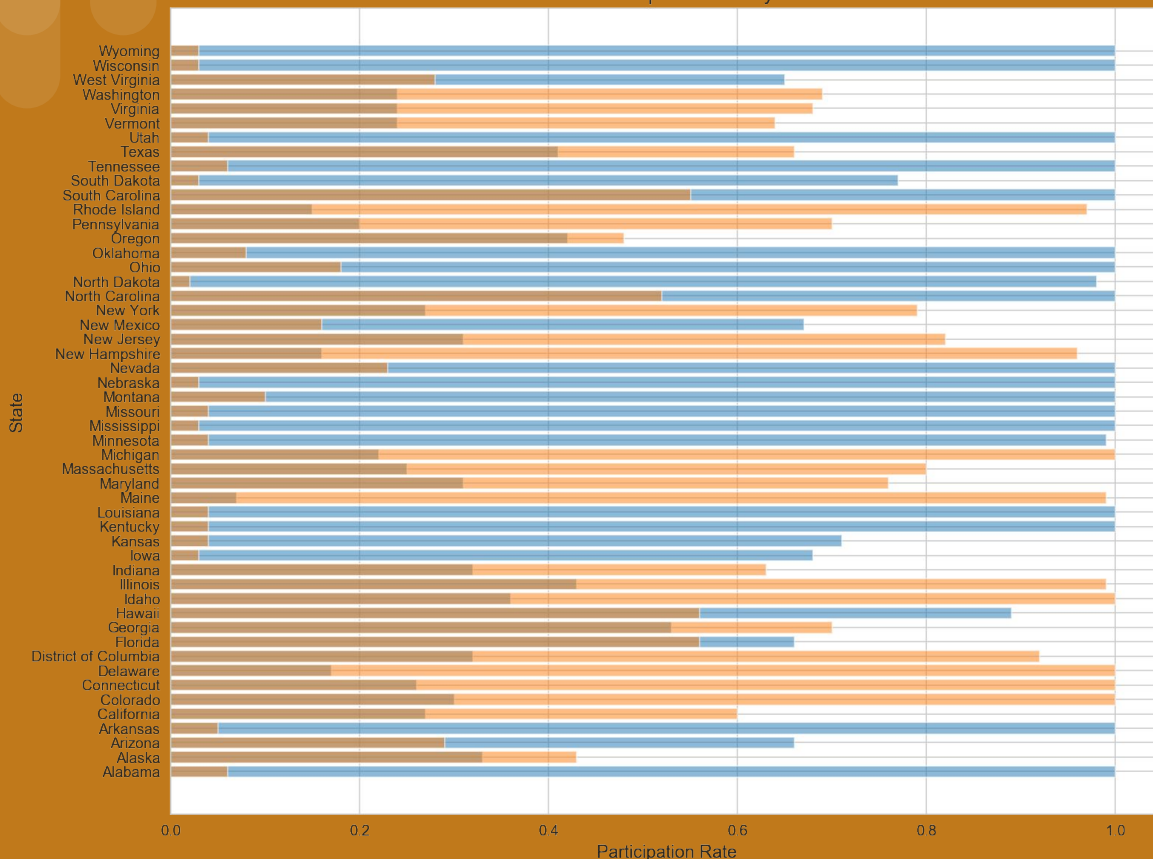
SAT and ACT Participation rate by state 2017



- Bar chart shows participation rates for 2017
- Orange is SAT
- Blue is ACT
- Brown is overlapping

Continued

SAT and ACT Participation rate by state 2018



-Bar chart shows participation rates for 2018

-Orange is SAT

-Blue is ACT

-Brown is overlapping

-**By the bar chart we can see that Alaska, Colorado and Illinois display interesting participation rate changes from 2017 to 2018**

-More explained in the following slide

Continued

- "PR" is participation rate

- Alaska's PR for ACT was higher than SAT in 2017 at 65% vs 38% respectively but in 2018 it did a flipside and SAT PR overtook the ACT PR at 43% vs 33% respectively as seen in the table above. The Anchorage daily news reported that as of 2017, high school students in Alaska are not required to take either the SAT, ACT or Workkeys which is a job skills test in order to graduate. This could be part of the explanation why Alaska doesn't see full or near full PR for either SAT or ACT.

- Colorado saw full ACT PR at 100% in 2017 while its SAT PR was at a meager 11% but in 2018 it's SAT PR shot up to 100% while ACT PR went down to 30%. Colorado made a switch from ACT to SAT in 2017 for its high school students. Due to feedback of lack of preparation for the sudden change, high school students were allowed to take the ACT then. It was found by the state selection committee that the College Board's Practice SAT, given to 10th-graders, aligned better with the high school Colorado Academic Standards

- Illinois also showed high ACT PR in 2017 at 93% while its SAT PR was 9% but in 2018 it was the opposite and SAT PR was at 99% while its ACT PR was 43%. The Chicago Tribune reported that college board won the tender for SAT to be the standardized test and dropped the ACT after 15 years as it was decided the SAT was better aligned with standards for what students should know in Illinois.

Conclusions and Recommendations

What has been done is to:

- Filter out SAT participation rates of 90% and above (Those that are under state policies and already high participation rates)

- Filter out ACT participation rates of 90% and above (Those that are under state policies and already high participation rates)

What we understand so far:

- Sampling bias, partly due to state policies implementing mandatory SAT or ACT

- States with high SAT participation rates had low ACT participation rates and vice versa

- Negative correlation between SAT total score and participation rate, so states with lower participation rates might be a better choice

- Coastal states to have high SAT participation rates so can consider looking at other states

-States with significant total score improvement and not too high participation rates would be Alaska, Arizona and Florida. However, Arizona saw relatively high ACT participation rates for both years and Florida's ACT participation rate overtook SAT in 2018. Alaska would then be a recommendation as both SAT and ACT participation rates are somewhat similar and in 2018 the SAT participation rate overtook ACT.

Conclusion

We have seen that data becomes a lot trickier to analyze when sampling bias is involved. But through filtering and visualization certainly helps to paint a better picture. From observing the bar charts, I would not recommend College board to direct their resources to states that have seen high ACT participation rates in both years as their rates tend to be dominant and SAT participation rate minimally fluctuates in those states. We should also look at states besides those with high or near full participation rates as they might have already implemented policies on making SATs mandatory for high school students and also it might be more worthwhile investing resources to a state with lower participation rate to maximize nurturing and getting the students and other stakeholders such as school faculty accustomed to SAT.

Therefore, Alaska would seem like a suitable choice for College board to allocate resources to increase the participation rate. Alaska does not have any mandatory policies on SAT or ACT but the SAT participation rate overtook the ACT participation rate at 43% vs 33% respectively in 2018 suggesting that an increasing number of Alaskan high school students see the SAT as a way to get into university on their own accord. It's reported out of 7796 high school students in 2018, 3334 chose to take SAT which represents the 43% so it would be great to increase the figure and significantly lead ahead of ACT participation rates. Also noticed is the low 12% of SAT participants in Alaska using fee waiver and it would be an incentive to encourage more high school students to take SAT by giving awareness about the fee waivers. It is an option to recommend for Alaska to implement state policy on making SAT mandatory.

We do have to be aware though that we see negative correlation between high SAT participation rate and total scores. Possibly high school students who are part of the low participation rate and scoring higher are those that already have in mind the universities they would like to apply to. I would recommend as well that resources should also be focused on continuously improving test preparation access and quality such as Khan Academy currently teaming up with College board and have these easily and readily available for Alaskan high school students. This can help to lower the impact of the participation rate and total score negative correlation.

The End!

Thank you!

References:

<https://www.adn.com/alaska-news/education/2016/06/30/students-no-longer-need-national-tests-to-graduate/>

<https://www.denverpost.com/2017/03/06/colorado-juniors-sat-college-exam/>

<https://www.chicagotribune.com/news/ct-illinois-chooses-sat-met-20160211-story.html>

<https://reports.collegeboard.org/pdf/2018-alaska-sat-suite-assessments-annual-report.pdf>