
Project 1: Analyzing SAT & ACT Scores

— By Rachel Koenig —

What are the SATs & ACTs?

These tests are used by colleges to measure a student's qualification for admittance. They are also often taken into consideration when being selected for a scholarship. The tests themselves are standardized, however, who has to take them is not and varies by state and college. Students are typically compared to the average, whether it be in their state or the national, so it is important to get a high score, especially in areas where there is a lot of competition.

According to collegevine.com, the SAT is a more of a general knowledge test, while the ACT focuses more on specific subjects.

Problem Statement

How can we increase SAT participation rates in states with high score like North Dakota? Should the tests be mandatory if 100% participation states have lower averages? If yes, should they be nationally standardized? Are they even helpful anymore since many colleges are dropping their requirement?

The Data

- We are given 4 sets of data:
 - 2017 ACT data included average scores by subject, the average composite score and participations rates for all 50 states, plus the District of Columbia.
 - 2018 ACT data only included the average composite score and participation rates.
 - 2017 SAT & 2018 SAT both included average scores by subject, average totals and participation rates for each state and District of Columbia.

Information was provided by the below sites.

<https://blog.prepscholar.com/act-scores-by-state-averages-highs-and-lows>

<https://blog.collegevine.com/here-are-the-average-sat-scores-by-state/>

<https://reports.collegeboard.org/sat-suite-program-results/state-results>

Clean Data Steps

1. First I had to check if the data was clean and complete.
 - a. Maryland had incorrect data for 2017 SAT Math and also 2017 ACT Science. I found the correct information on collegevine.com and prepscholar.com, respectively, and used the .replace method to fix.
2. Some of the numeric columns were typed as objects so I needed to change those to floats and integers next.
 - a. Before I could do that though, I had to use .replace again to get rid of a couple non-numeric values in the cells.
3. Next, I renamed the columns using .rename to make them all lowercase and more unique to prepare for merging datasets later on.

Clean Data Steps cont'd

4 . Before merging datasets, I had to make sure both dataframes had the same amount of rows. The `.shape` attribute showed me that ACT17 had an extra row called National. I used `.drop` to delete it.

5. Lastly, I merged the two dataframes with `pd.merge` and saved them to a new csv file with `.to_csv`.

6. Also had to do all the same steps on the 2018 data after importing, but with a few different errors to replace and drop.

7. Finally, I combined the 2018 ACT and SAT data and then merged that with the combined 2017 dataframe and voila! One giant dataframe!

States with Highest & Lowest Participation

- There were 13 states above the 75th percentile of participation in the 2017 SATs.
- There were 14 states with below 5% participation on the SATs in 2017
- There were 10 states with 90% or greater SAT participation in 2018.
- There were 19 states with 90% or greater ACT participation on 2018

	SAT 2017	SAT 2018	ACT 2017	ACT 2018
Highest	DC Michigan Connecticut Delaware	Colorado Connecticut Delaware Idaho Michigan	Alabama Arkansas Colorado North Dakota Mississippi	Minnesota Ohio Oklahoma South Carolina North Carolina
Lowest	Mississippi North Dakota Iowa	North Dakota Wyoming Wisconsin	Delaware Maine New Hamp Pennsylvania Rhode Island	Alaska California Colorado Connecticut Delaware

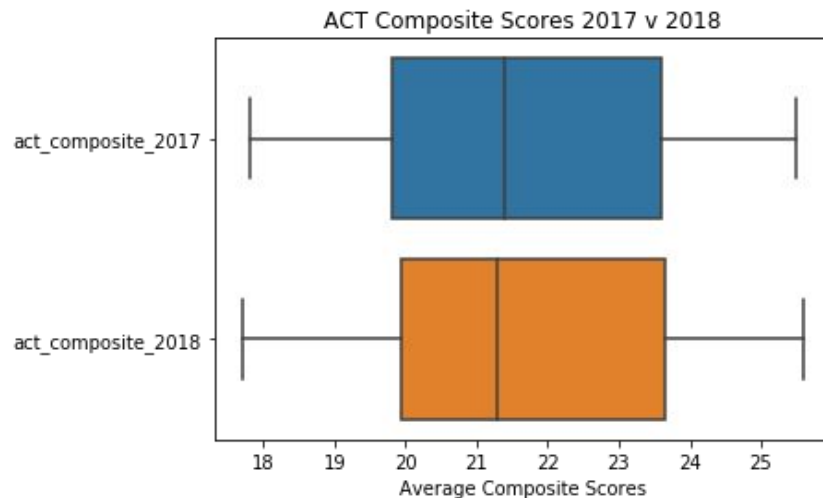
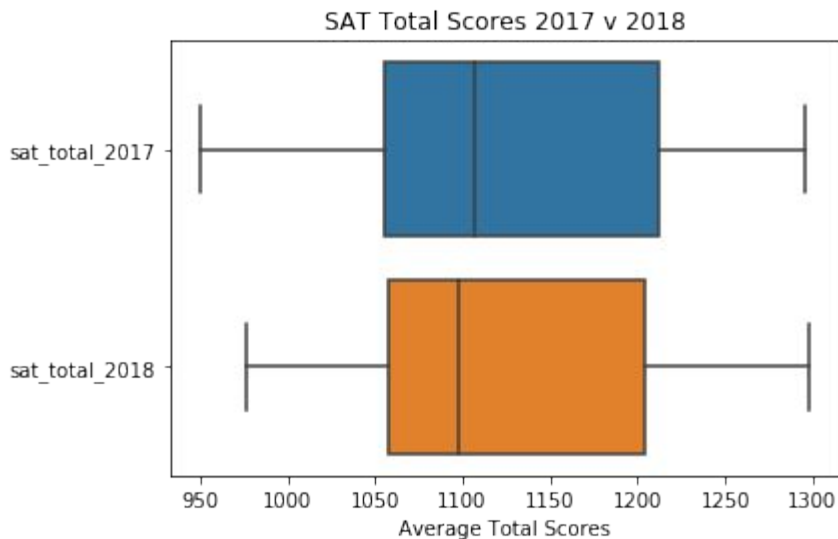
Top 5 Highest & Lowest Total Scores

	SAT 2017	SAT 2018	ACT 2017	ACT 2018
Highest	Minnesota Wisconsin Iowa Missouri Kansas	Minnesota Wisconsin North Dakota Iowa Kansas	New Hampshire Massachusetts Connecticut Maine DC	Connecticut Massachusetts New Hampshire New York Michigan
Lowest	DC Delaware Idaho Michigan Maine	DC Delaware West Virginia Idaho Utah	Nevada Mississippi S. Carolina N. Carolina Hawaii	Nevada South Carolina Mississippi Hawaii Alabama

EDA Learnings

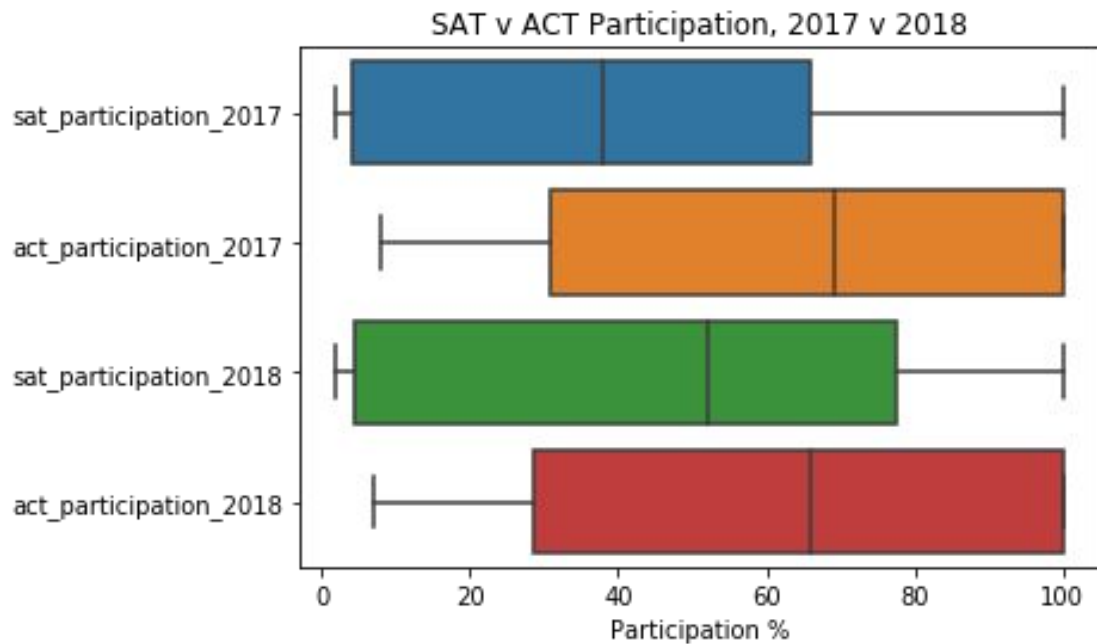
- Connecticut, Delaware and Michigan had 100% SAT participation both 2017 and 2018. DC had 100% participation in 2017 but went down to 92% in 2018. Their total scores went up when participation went down. Colorado had the biggest jump from only 11% participation in 2017 to 100% in 2018, but score dropped about 200 points with the increase in participation.
- Colorado had the opposite spike for ACTs with 100% participation in 2017 and only 30% in 2018 and composite score went up. Nebraska and Ohio went up to 100% but the score didn't change very much.
- Only 3 states, Florida, Georgia and Hawaii, had participation over 50% for both SAT and ACT tests in 2017. North and South Carolina joined them in 2018. Total and Composite scores for all were below average.

Box Plot Visualizations for Score



I used box plots to compare year to year Total SAT scores and Composite ACT scores. They show very little change.

Box Plot Visualization for Participation



This boxplot shows overall participation for the SATs was lower than the ACTs both years & all of the distributions are skewed.

Colorado

- Colorado is interesting to me because the SAT participation grew from 11% to 100% from 2017 to 2018 when the tests become mandatory (<https://blog.prepscholar.com/which-states-require-the-sat>). Average total score dropped from 1200 to 1025.
- Consequently, the ACT numbers reversed. It's participation dropped from 100% to 30% which caused the average composite score to increase from 20.3 to 23.9.
- The required test for Colorado students changed due to "hard-fought testing reform legislation" and to align with the core curriculum.
- <https://www.chalkbeat.org/posts/co/2015/12/23/goodbye-act-hello-sat-a-significant-change-for-colorado-high-schoolers/>

North Dakota

- According to our data, North Dakota had the lowest SAT participation rates both years, but one of the highest scores.
- The ACT test is mandatory for high school juniors to take in North Dakota and free to take. <https://blog.prepscholar.com/which-states-require-the-act-full-list-and-advice>
- Even though the ACT scores for ND are lower than the national average, the fact that the test is free provides many opportunities to low income kids who may not have been able to take it otherwise. https://www.nd.gov/dpi/SchoolStaff/assessment/act_workkeys/

District of Columbia

District of Columbia intrigued me because of its 100% SAT participation but extremely low SAT scores. I was surprised to find out the test is actually not mandatory there but it is free so that is great that juniors are taking advantage of the opportunity.

<https://blog.prepscholar.com/which-states-require-the-sat>

Key Takeaway

I think it is important to note correlation between states having a high participation for one test and a very low for the other. It seems we can deduce that few students want to take both tests so in states where one is mandatory, the other is very low.

North Dakota SAT participation is the lowest in the country. If we made that test free, but not mandatory, I think the number of students taking it would definitely increase.

I would have liked to dive deeper into the necessity of the tests. They seem like an ancient and inaccurate tool. Especially with the recent college admissions scandals in the news, I think it would be important to figure out the percent of fraudulent high scoring tests last year and focus on fixing that problem first.