CollegeBoard SAT®

This is Year 2019.

We are a team of data scientists working for the SAT College Board.

Since the implementation of the new format in 2016, we've been tasked to analyze data and make recommendations to increase participation

SAT and ACT Data Analysis

DSI-18: Alicia, Mak, Sahaj, Mok

Agenda

1. **Why?** Background+Context
2. **How?** Datasets and cleaning
3. **What?** Key learnings
4. **So-what?** Recommendations

Why are we doing this analysis?

Background and context

- Standardized testing is an important part of the college application process for students
- In 2016, a new SAT format was introduced
- How does it affect participation rates in states? How can we increase participation?

Major changes coming to 2016 SAT test: Here's what, how and why

By Jamie Gumbrecht, CNN

Updated 1534 GMT (2334 HKT) March 6, 2014

The new SAT is A) a big change B) still important C) controversial D) all of the above

By Kelly Wallace, CNN

Updated 2053 GMT (0453 HKT) February 25, 2016

The New York Times

New, Reading-Heavy SAT Has Students Worried

Problem Statement

Recommend ways to increase
SAT participation in the US

How did we do it?

Datasets

SAT 2017, 2018 Datasets

- State
- Participation Rates
- Evidence-Based Reading & Writing
- Math
- Total Score

ACT 2017, 2018 Datasets

- State
- Participation Rates
- English
- Math
- Reading
- Science
- Composite Score

Datasets record the states' average results and participation rate

Data-cleaning

```
# Look at a summary of SAT 2017 data
sat_2017.describe()
```

	Evidence-Based Reading and Writing	Math	Total
count	51.000000	51.000000	51.000000
mean	569.117647	547.627451	1126.098039
std	45.666901	84.909119	92.494812
min	482.000000	52.000000	950.000000
25%	533.500000	522.000000	1055.500000
50%	559.000000	548.000000	1107.000000
75%	613.000000	599.000000	1212.000000
max	644.000000	651.000000	1295.000000

We did some cleaning of the result by examining our dataset.

E.g. For SAT, state Maryland, Math score - 52 to 524.

Data-cleaning

46	Vermont	29%	23.3	23.1	24.4	23.2	23.6
47	Virginia	29%	23.5	23.3	24.6	23.5	23.8
48	Washington	29%	20.9	21.9	22.1	22.0	21.9
49	West Virginia	69%	20.0	19.4	21.2	20.5	20.4
50	Wisconsin	100%	19.7	20.4	20.6	20.9	20.5
51	Wyoming	100%	19.4	19.8	20.8	20.6	20.2x

We did some cleaning of the result by examining our dataset.

E.g. For ACT, state Wyoming, Composite score - 20.2x to 20.2

Data-cleaning

```
# Check the data types for ACT 2017 Data
act_2017.dtypes
```

```
State            object
Participation    object
English         float64
Math            float64
Reading         float64
Science         float64
Composite        object
dtype: object
```

```
act_2017cln.dtypes
```

```
State            object
Participation   float64
English         float64
Math            float64
Reading         float64
Science         float64
Composite       float64
dtype: object
```

Data-cleaning

```
# Check the data types for SAT 2017 Data
sat_2017.dtypes
```

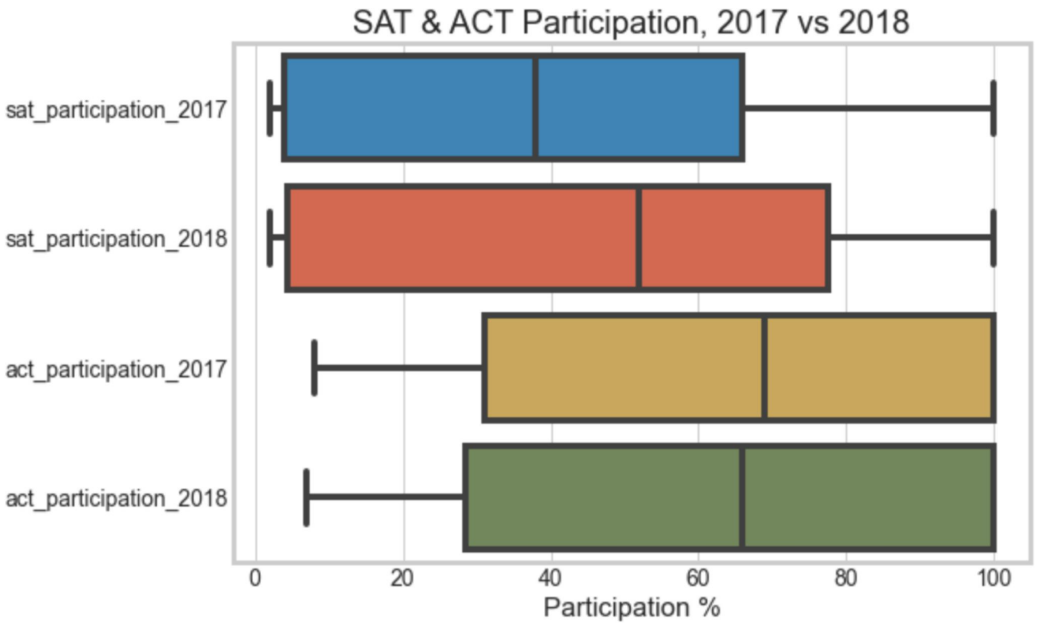
```
State                                  object
Participation                          object
Evidence-Based Reading and Writing      int64
Math                                    int64
Total                                   int64
dtype: object
```

```
sat_2017cln.dtypes
```

```
State                                  object
Participation                         float64
Evidence-Based Reading and Writing      int64
Math                                    int64
Total                                   int64
dtype: object
```

What did we learn?

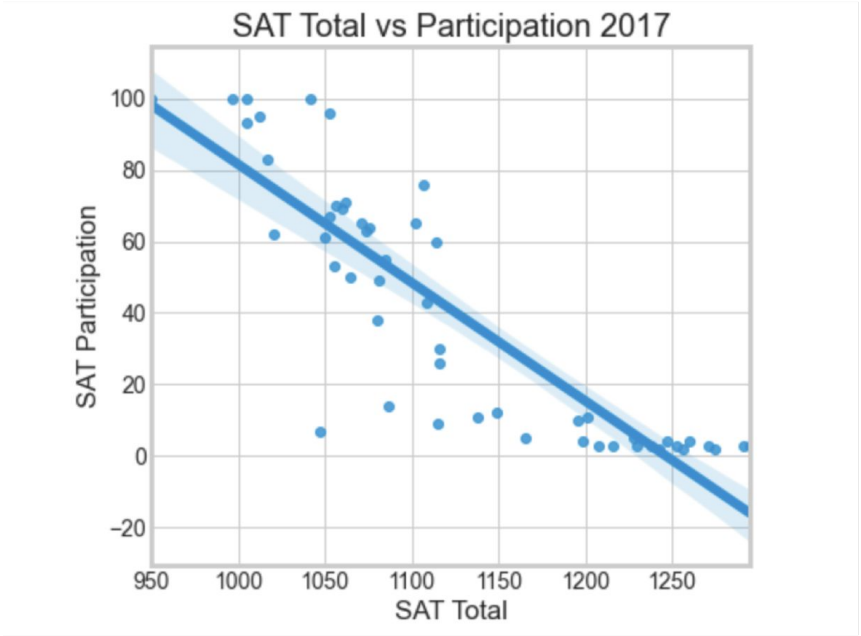
We can see a clear difference between SAT and ACT participation rates across the USA



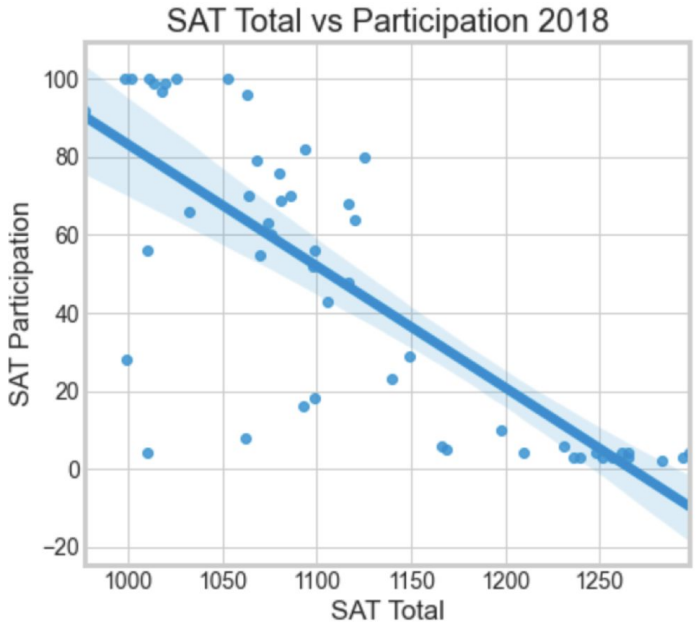
From 2017 - 2018:

- ACT has higher average participation% than SAT across the US
- 6% increase in mean SAT participation
- 4% decrease in mean ACT participation

There appears to be an inverse correlation between average State scores and participation, and this relationship can be attributed to selection bias*



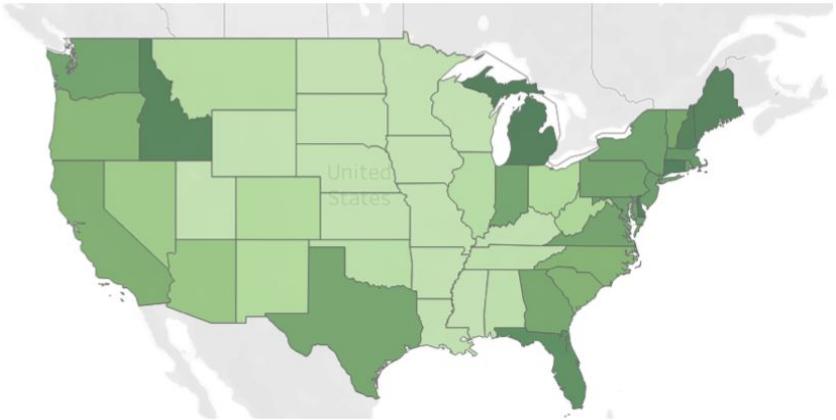
r = -0.868, p = 0.0



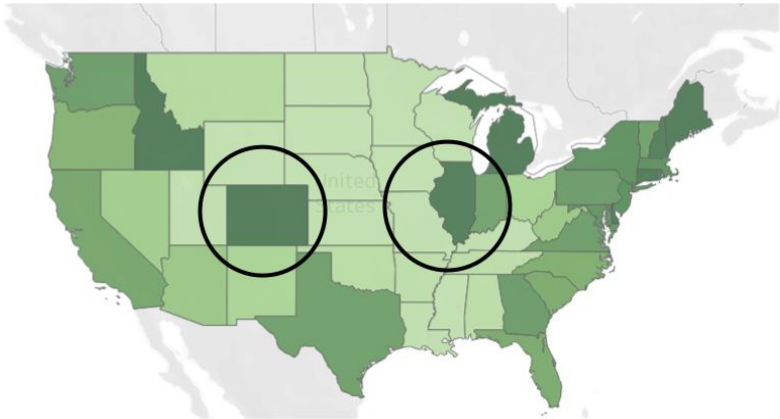
r = -0.786, p = 0.0

*Source:
https://eml.berkeley.edu/~jrothst/publications/clark-rothstein-schanzenbach-EER2009.pdf

As a result of SAT being made mandatory*, Colorado and Illinois both saw ~90% increases in SAT participation from 2017 to 2018



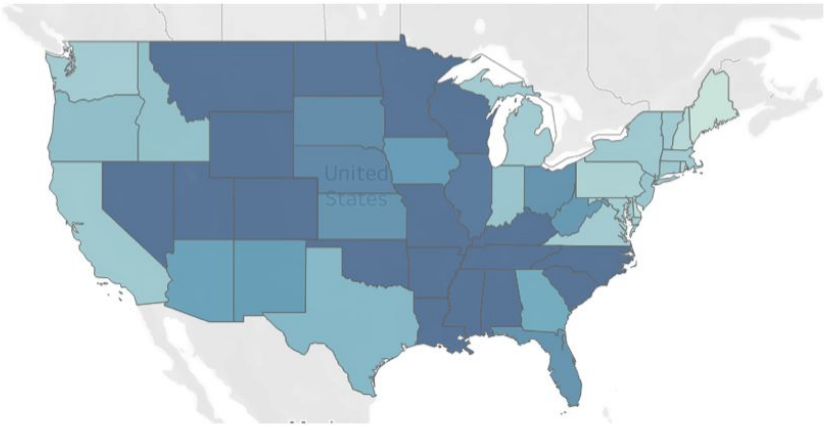
SAT Participation 2017



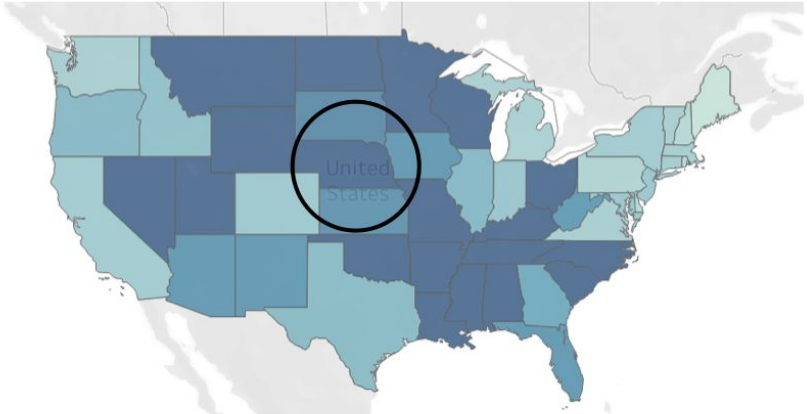
SAT Participation 2018

*Sources:
https://www.testive.com/colorado-sat-change-2017/
https://www.chicagotribune.com/news/ct-illinois-chooses-sat-met-20160211-story.html

Similarly, Ohio has also seen a 25% increase in ACT participation after it was made mandatory* in 2017



ACT Participation 2017



ACT Participation 2018

*Source:
https://www.daytondailynews.com/news/historically-low-act-scores-red-flag-for-our-country/djfx9Urp719WyEaMfykyxL/

With average SAT participation of 3% in 2018, we have identified South Dakota and Iowa as the states that represent the greatest opportunity for growth

state	sat_participation_2017	sat_participation_2018	act_participation_2017	act_participation_2018
Iowa	2.0	3.0	67.0	68.0
South Dakota	3.0	3.0	80.0	77.0

What do we do now?

Key Insight

It is evident that **state policy** is the factor driving participation rates for both the SATs and the ACTs.

Recommendation

In most states, a state board of education is in charge of educational policymaking.

It is recommended that the College Board **actively market the SAT towards state board of educations**.

Specifically, the states **Iowa and South Dakota** have been identified as high-potential targets.

Questions

Interesting Correlations?