GA Project 1

Standardized Test Analysis

Problem Statement

- The ACT and SAT are both **widely-recognised** in the admission process.
  - Adequate **representations of 'intelligence'**.
- However, does **each test** per se **accurately** portray the level of 'intelligence' of candidates?

In theory…

- A candidate or group of people with a **higher SAT score(s)** should have a **higher ACT score(s)** as well. Vice-versa.

Testing the theory

- Compare the data on a ‘States’ level
  - No access to individual candidates’ data
- Study the movements by comparing the difference in test scores Y-O-Y

Data Dictionary

Feature	Type	Dataset	Description
state	object	ACT/SAT	The State where recorded scores are from
total_act2017	float	ACT	Average total ACT score from the year 2017
total_act2018	float	ACT	Average total ACT score from the year 2018
total_act2019	float	ACT	Average total ACT score from the year 2019
total_sat2017	float	SAT	Average total SAT score from the year 2017
total_sat2018	float	SAT	Average total SAT score from the year 2018
total_sat2019	float	SAT	Average total SAT score from the year 2019
act_diff1718	float	ACT	Difference of ACT score for years 2017 and 2018
act_diff1819	float	ACT	Difference of ACT score for years 2018 and 2019
sat_diff1718	float	SAT	Difference of SAT score for years 2017 and 2018
sat_diff1819	float	SAT	Difference of SAT score for years 2018 and 2019

Summary Statistics

Key points:

1. Difference in mean ≠ 0
   a. close to 0
   b. size of each state was not disclosed
2. SD of tests relatively constant

	act_diff1718	act_diff1819	sat_diff1718	sat_diff1819
count	42	42	42	42
mean	-0.1190	-0.0452	4.9524	-9.0000
std	0.4244	0.2890	19.6319	12.2076

	total_act2017	total_act2018	total_act2019	total_sat2017	total_sat2018	total_sat2019
count	42	42	42	42	42	42
mean	21.633	21.514	21.469	1,134.095	1,139.048	1,130.048
std	2.066	2.148	2.242	92.392	91.612	91.321

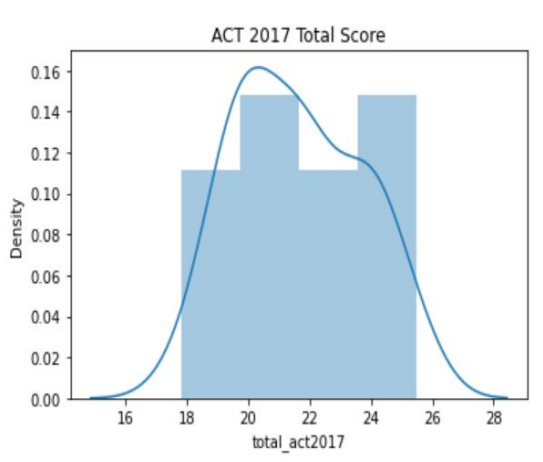
Caveats

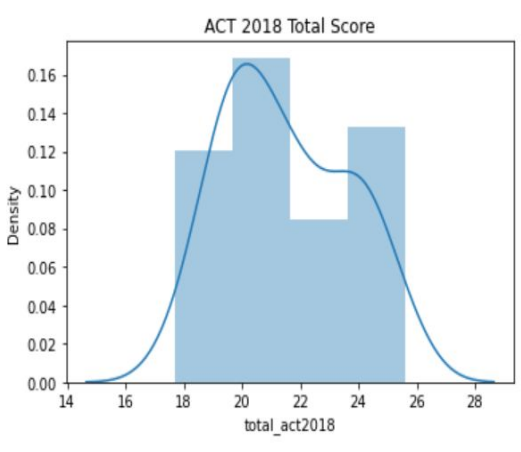
- Scores not adjusted
  - No breakdown of State's sizes
- Final sample size used: 42
  - Relatively small but each sample is an average of a larger sample
- Y-O-Y value differences compared at its sign, not percentages/values
  - Due to nature of scores (floor/ceiling)
  - Lack of access to grading distribution and matrix

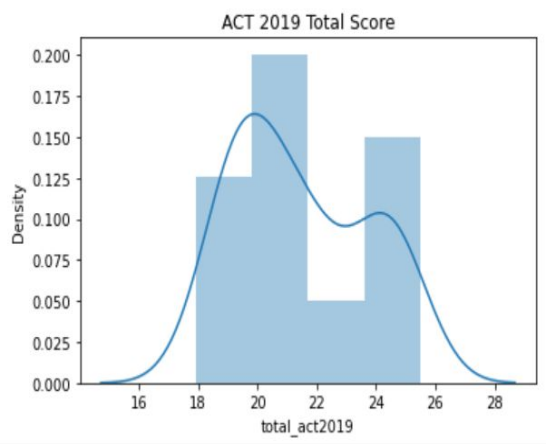
Assumptions

- ACT & SAT have **internal scaling** to produce a **similar distribution** yearly

ACT Total Score 2017 to 2019



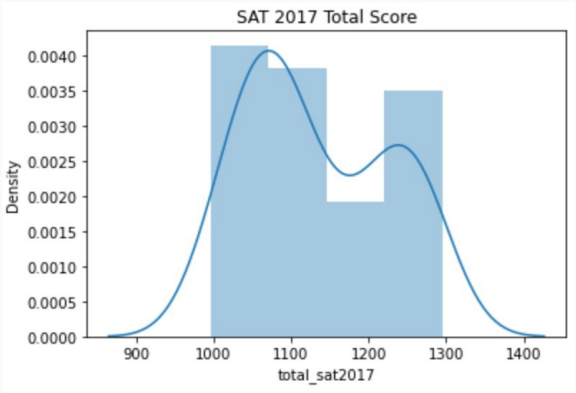


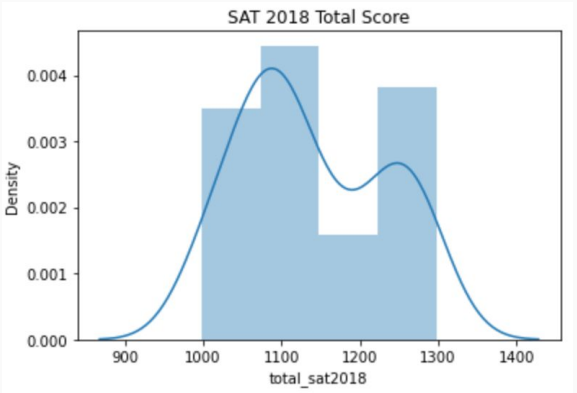


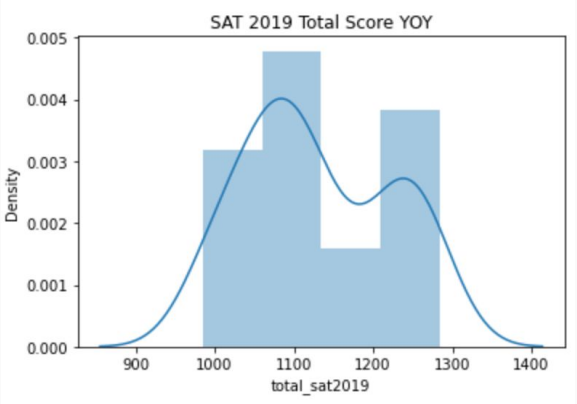
Assumptions

- ACT & SAT have **internal scaling** to produce a **similar distribution** yearly

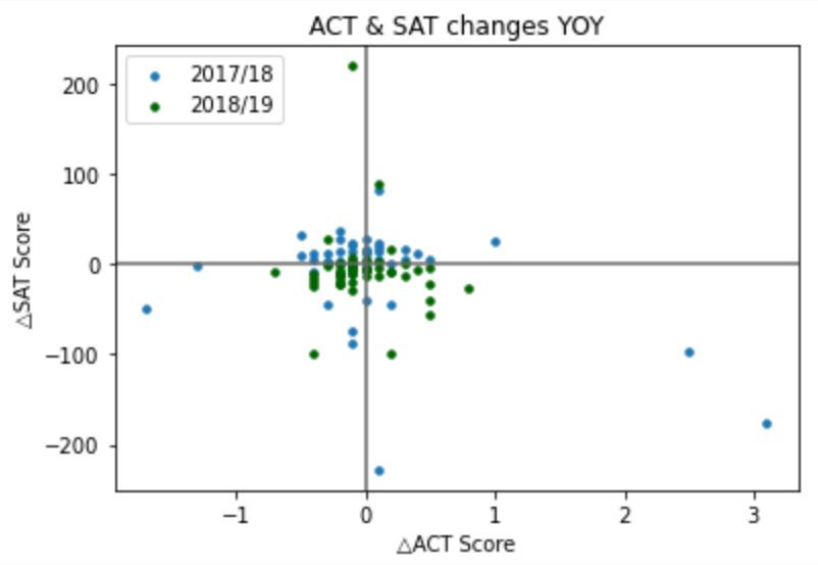
SAT Total Score 2017 to 2019







ACT & SAT changes Y-O-Y



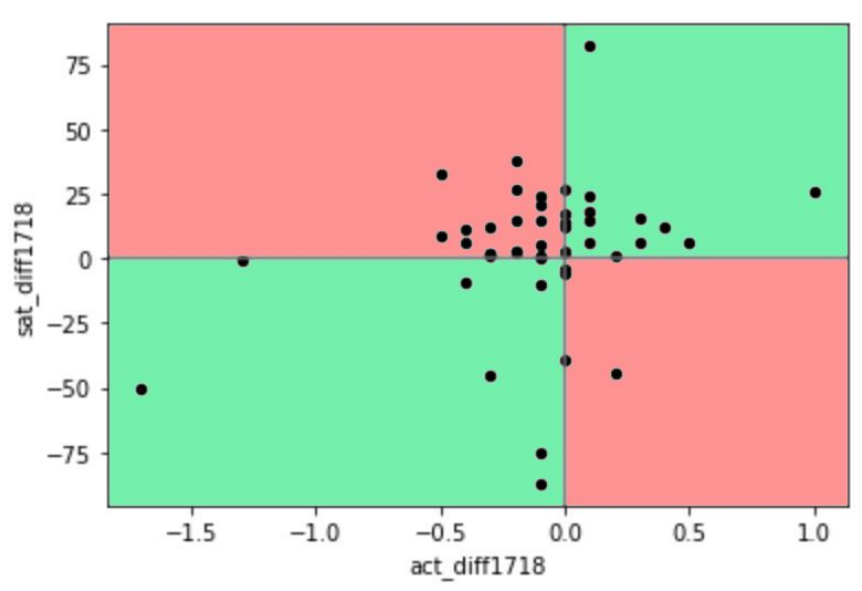
Observation:

- No noticeable relationship
- Outliers

Action:

- Separate the years to avoid unwanted ‘noise’
- Remove outliers

ACT & SAT 2017/18 changes Y-O-Y

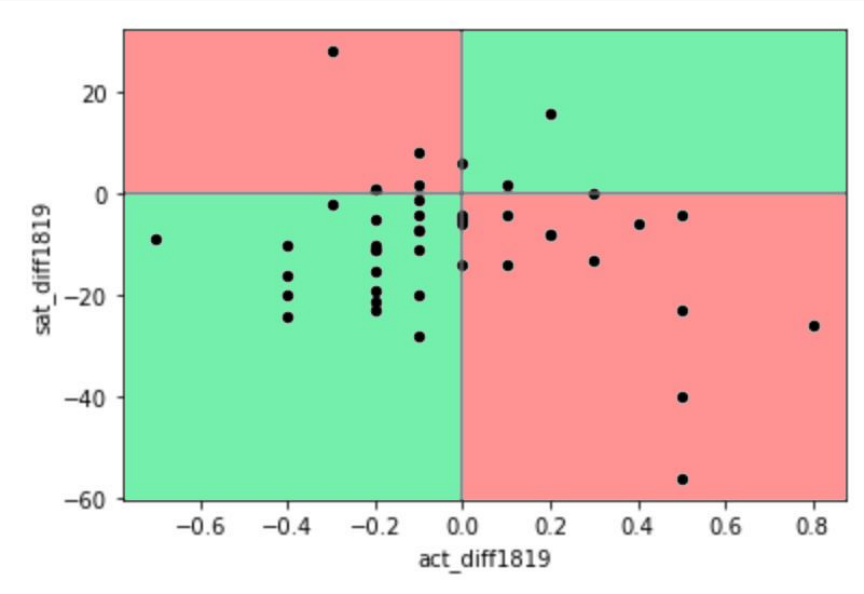


- Large cluster of states that ↑ SAT but ↓ ACT scores

17	10
5	1

Expected	
Unexpected	

ACT & SAT 2018/19 changes Y-O-Y



- More expected results but still no clear relationship
- A few extreme data points in unexpected area

5	2
20	9

Expected	
Unexpected	

Future Improvements

1. Source for the **number of candidates** for each State
2. Look for datasets with actual test **scores directly tied to candidates** rather than using a proxy (States)
3. Identify **relationship** of college intake **scores** with ACT & SAT scores to adjust appropriately for reasonable data extrapolation

Conclusion

- The ACT and SAT individually may not accurately represent a candidate's abilities
  - Significant amount of data points refuted the theory
    - Years 2017/18: 42.86%
    - Years 2018/19: 33.33%

Recommendations

1. Academic Institutions can require both tests scores collectively
    a. Gain a more accurate representation of a candidate's capabilities
    b. Avoid loophole: candidates can take both tests and apply with the better test score only
2. Consider other matrices/methods to assess candidates