Effect of Government Spending on Standardized Test Scores

Project 1 - Team 1

Miao Er
Cheewee
Gilbert
Cathy

Problem Statement

Background

Following the new SAT format in 2016, you have been hired to study the trends of ACT and SAT participation and test scores from 2017 to 2019 and provide **budgetary recommendations** to the state government.

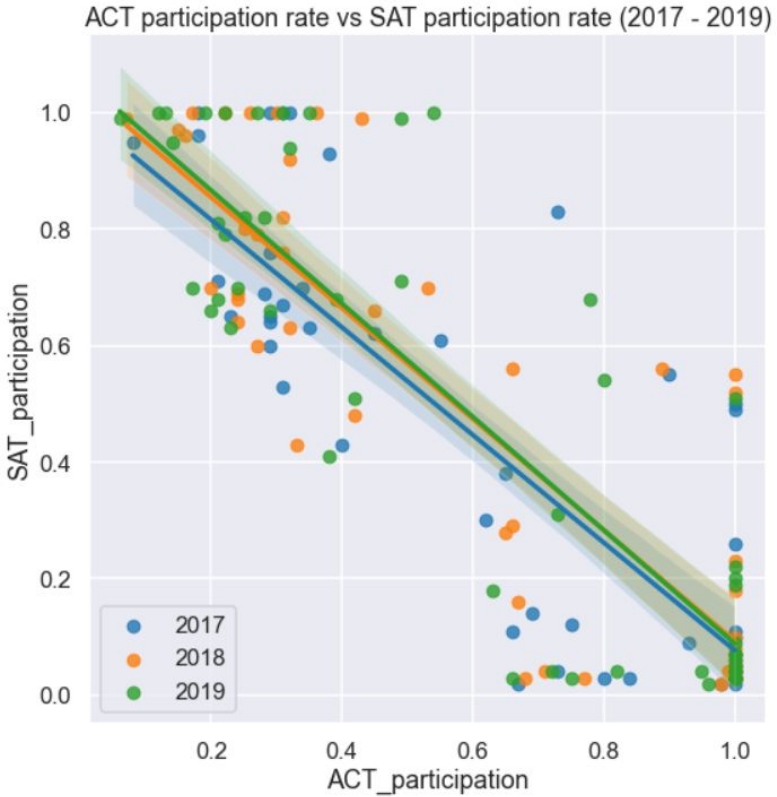
Problem Statement

This project aims to **investigate the impact of government spending** (and its various spending categories) on the standardized test (ACT and SAT) scores.

EDA and Visualisation

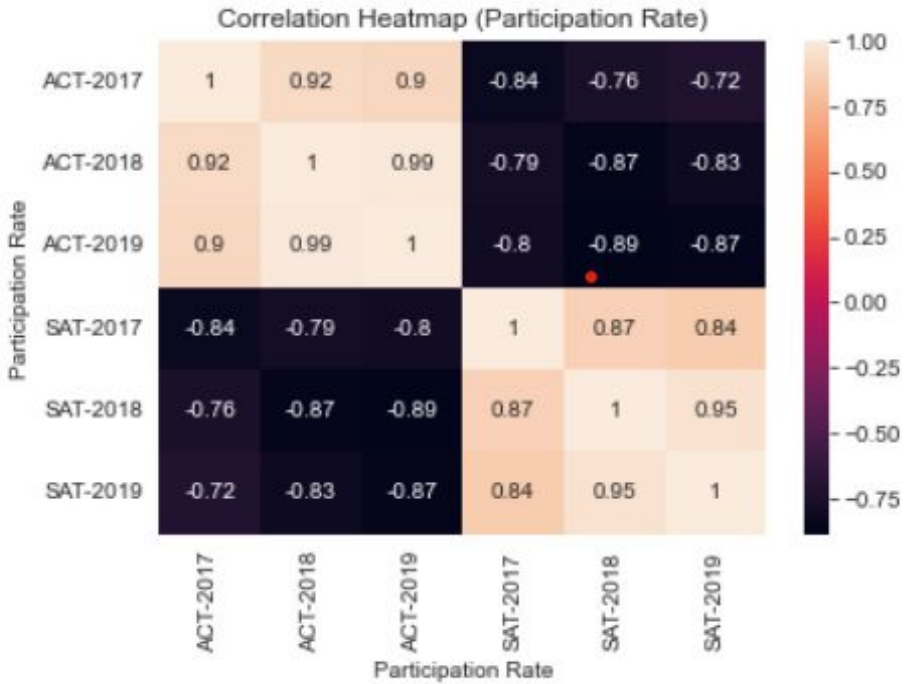
- Participation Rate
- Participation Rate vs Test Scores
- Government Spending

SAT and ACT Participation Rate



- There is an **inverse correlation** between the ACT and SAT test participation rate in each state.
- Some states prefer the ACT test while some prefer SAT test.
- The ACT test has a higher participation rate than the SAT test.

Correlation Heatmap (Participation Rate)



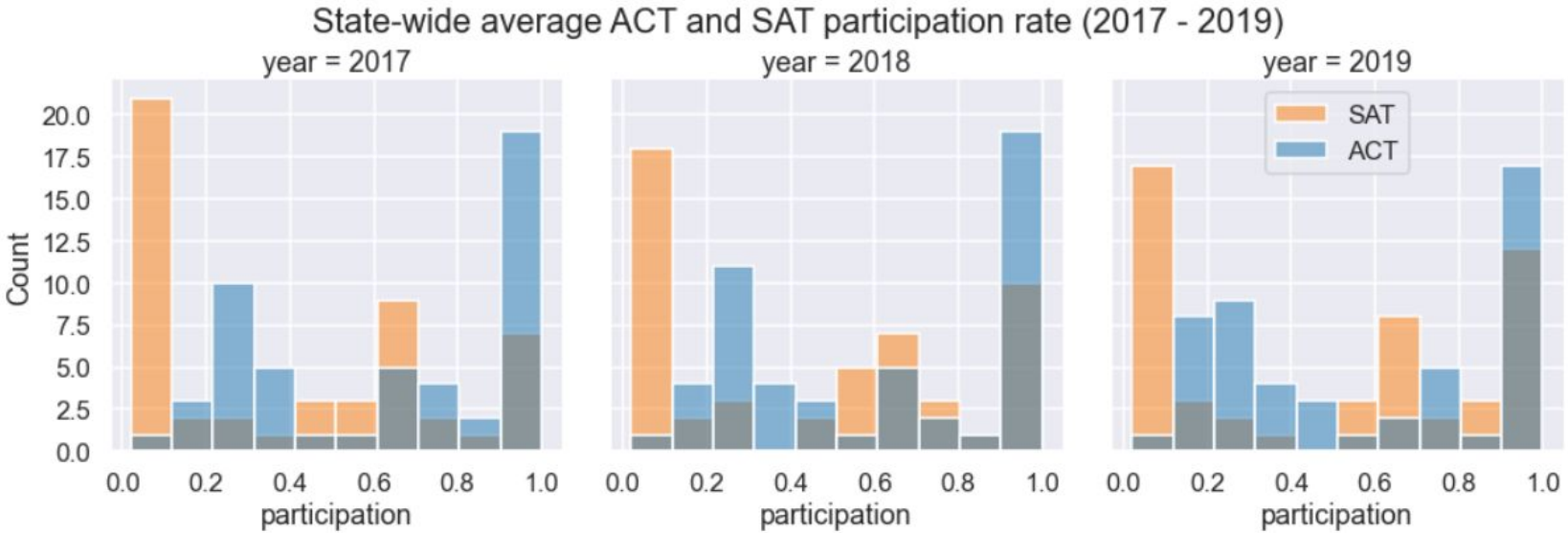
Positive correlation (> 80% correlation):

- The states that have high participation rate will **maintain high participation rate**, and vice-versa

Negative Correlation (-70% correlation):

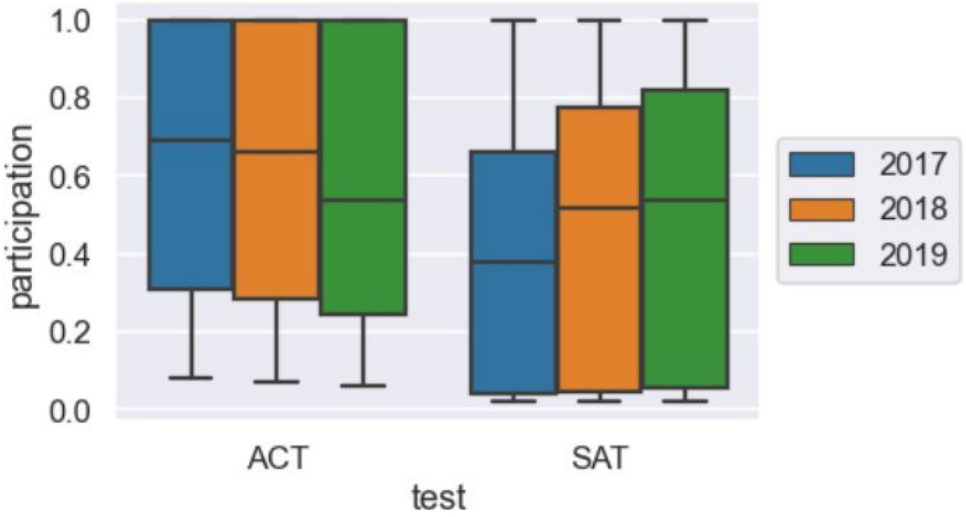
- The states that have high ACT participation rate, would have a low SAT participation rate, and vice-versa

Bimodal distribution in the participation for both tests



- There are less states with full ACT participation rate (1.0) at latter years.
- There are less states with very low SAT participation rate (0.0 - 0.10) at latter years.
- This could be due to some states converting away from mandatory ACT participation.

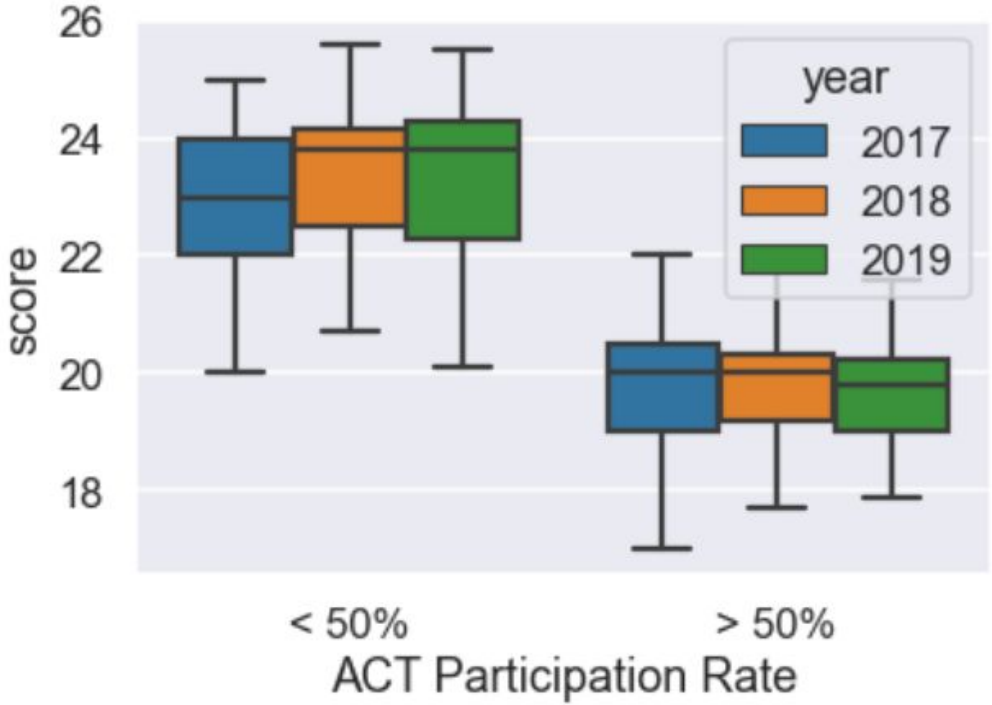
Participation Rate of two tests



- The year-on-year participation of both tests are quite **consistent**

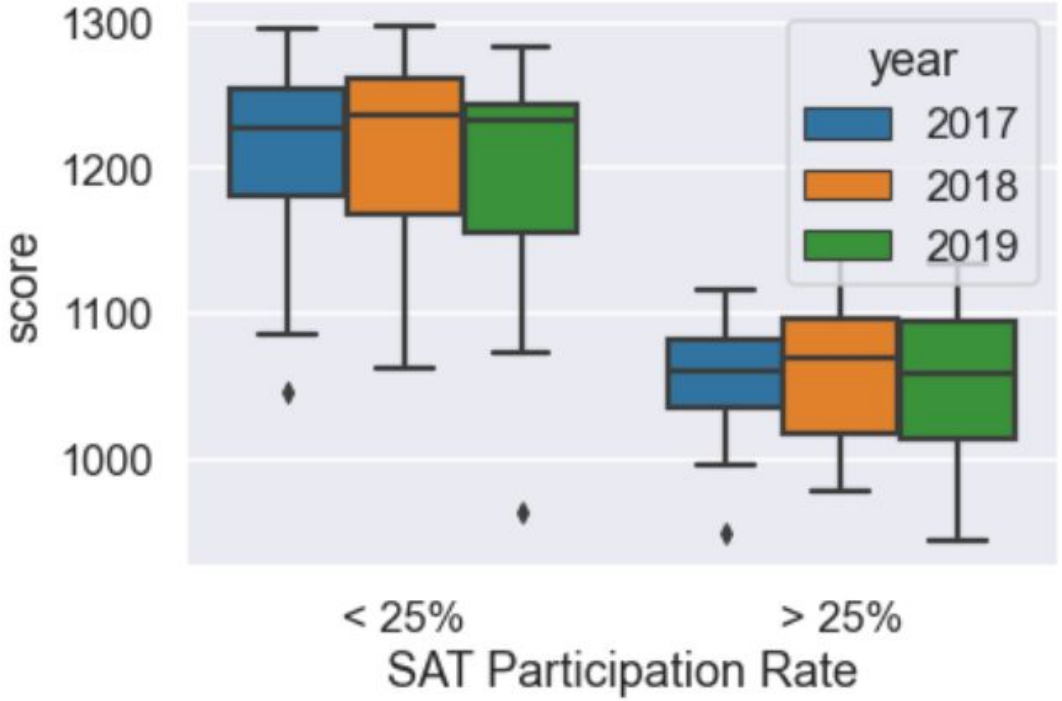
- ACT is having lower year-on-year participation rate, while the opposite is true for SAT

ACT Participation vs ACT Scores



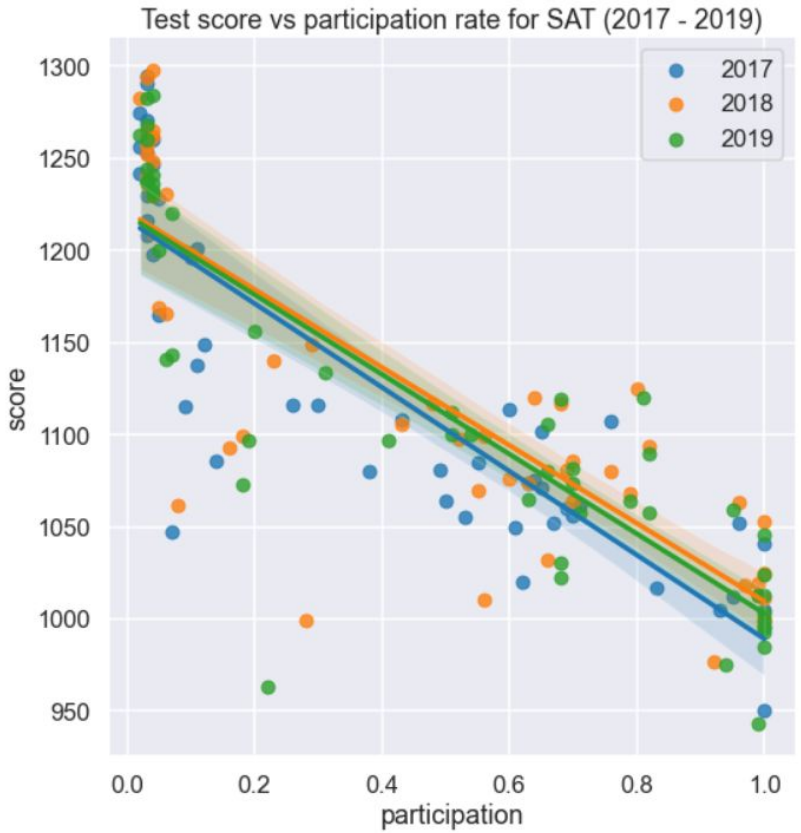
- The boxplots of ACT scores shows that states with a ACT participation rate lower than 50% generally **perform better** than those higher than 50%.

SAT Participation vs SAT Scores

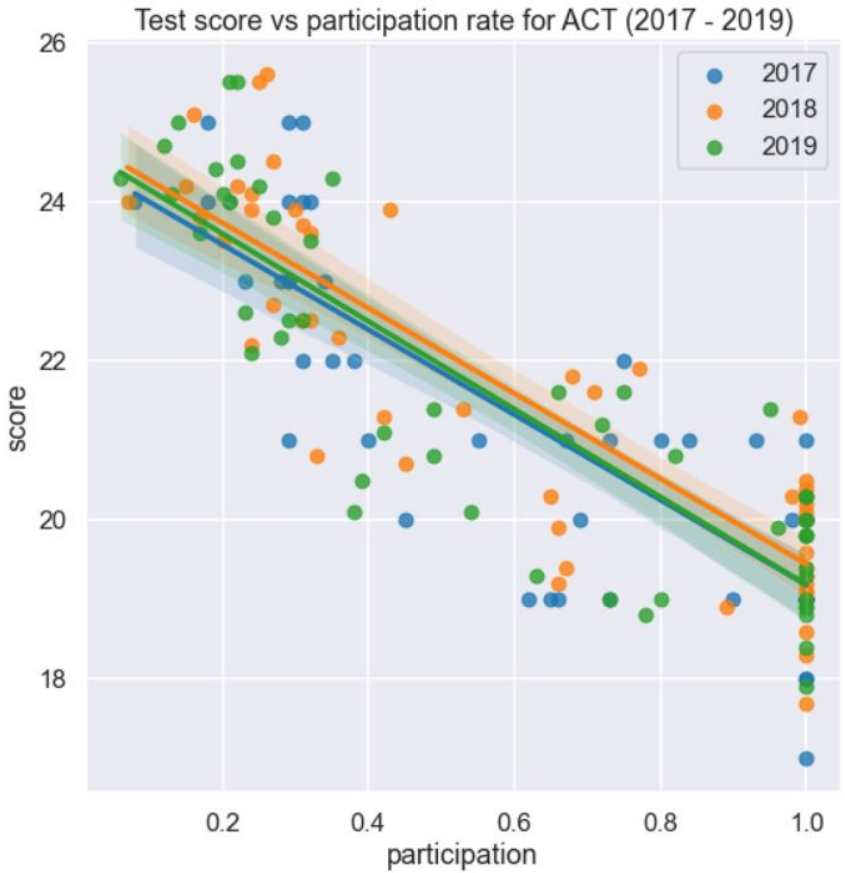


- Similarly with the boxplots of SAT scores for the two groups of states based on their SAT participation rate, we can see that in general states with a lower SAT participation rate show a higher score.

Correlation between Participation and Score



- Participation rate are negatively correlated to score in both SAT and ACT
- High participation rate are associated with low score
- Low participation rate are associated with high score



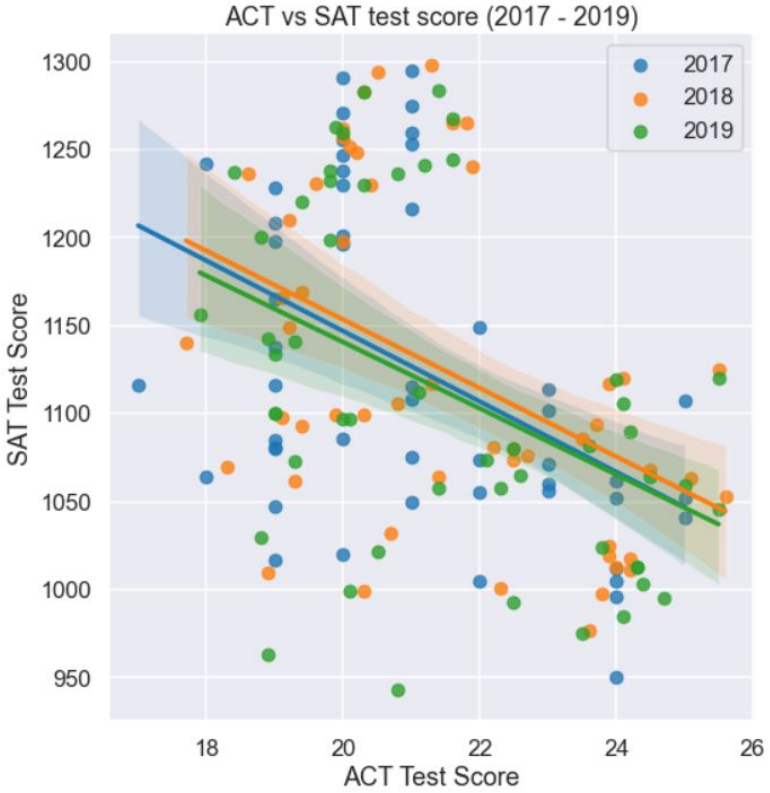
- Potential reason: ‘Self-selection bias’
- High performing students will likely be taking the tests
- States with high participation rate inevitably include low-performing students

Correlation between **ACT and SAT test score**

- **Negative correlation** between ACT & SAT (**?!**)
  - States with ‘smarter’ students should perform equally well in both tests

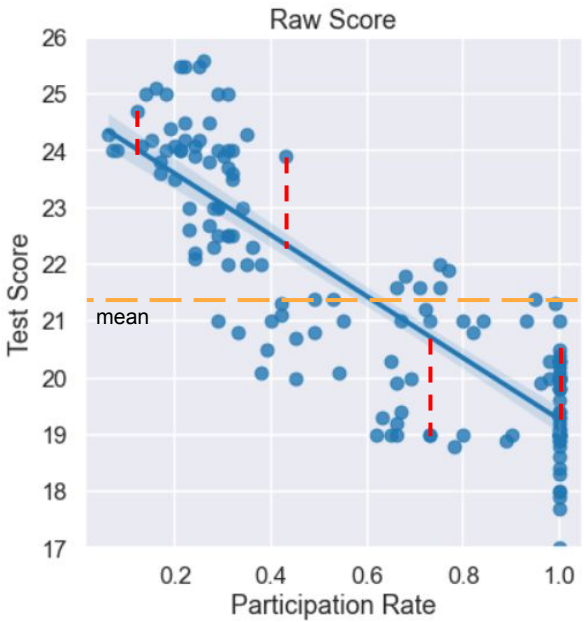
- Might be due to negative correlation between **test score and participation rate**
  - e.g.: states with **high SAT participation** has **low ACT participation**
  - **High SAT participation** results in **low SAT score**
  - **Low ACT participation** results in **high ACT score**
  - The state will have (relatively) low SAT score and high ACT score

- The **test scores** needs to be **adjusted for participation rate** before making further analysis

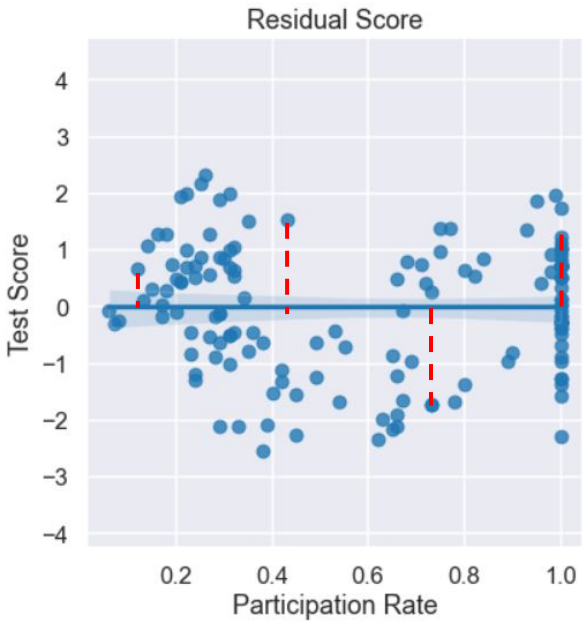


Adjusting the ACT and SAT test score for participation rate

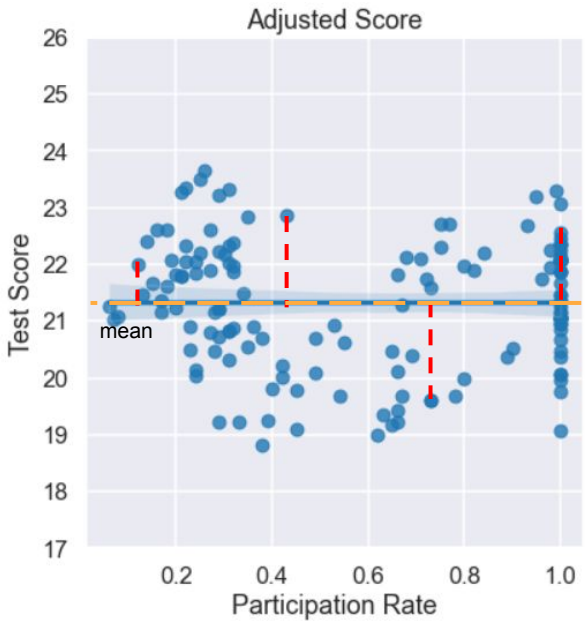
Step 1:
Obtain linear regression line



Step 2:
Obtain 'residual score'

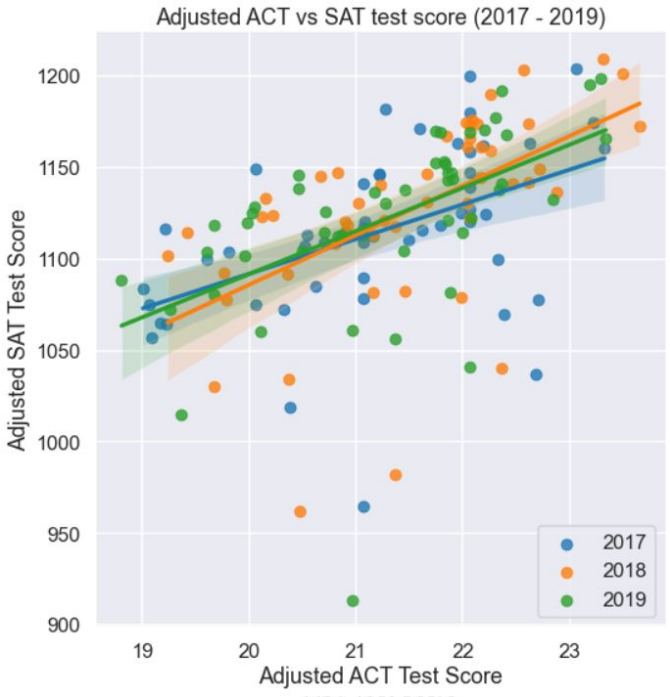


Step 3:
Shift up by the mean score

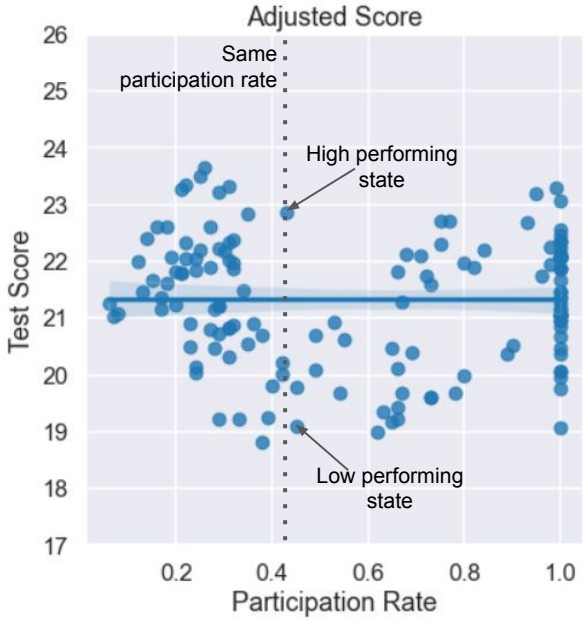


The **adjusted test score** is now **independent** of the **participation rate**

Adjusted ACT and SAT test score



Positive correlation between ***adjusted*** **ACT and SAT scores**



There are still **variations in test scores** for diff. states

There are **various factors** behind this variation in score:

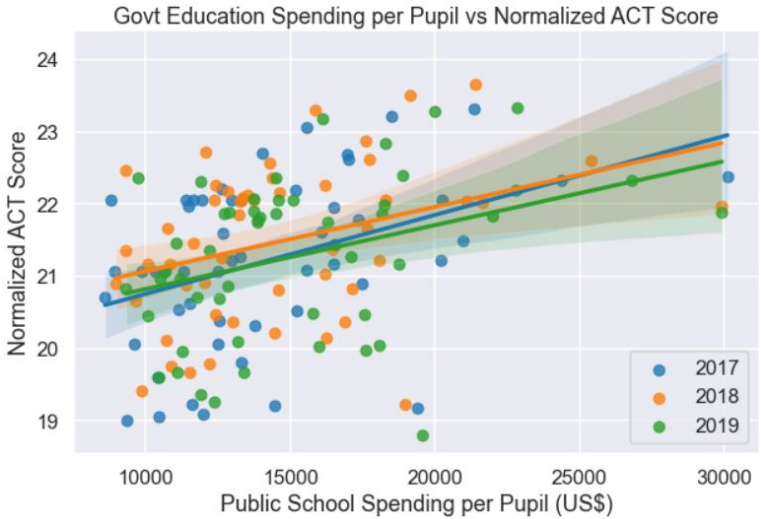
- Economic status / income
- State poverty rate
- Parents' education
- **Government education spending**

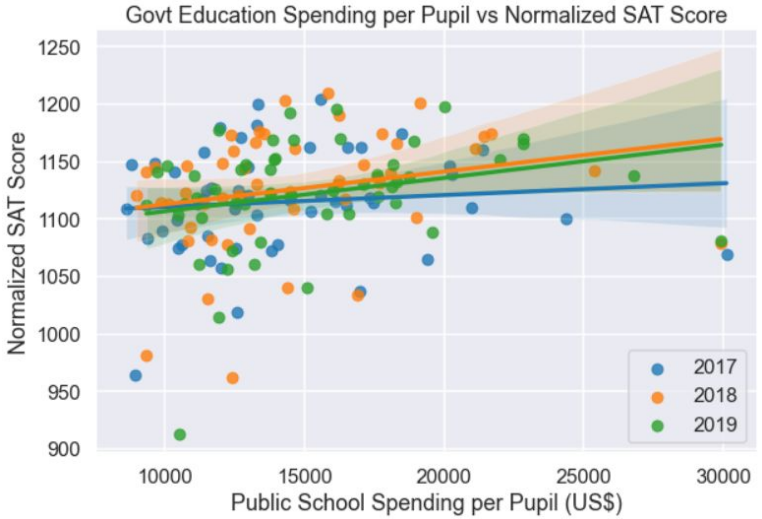
Effect of **Government Spending** on **Test Scores**

- The US Dept. of Education releases data relating to **government spending on K-12 education**
  - The data is split by **state** and by **spending category**
  - The spending is presented in **US$ spent per pupil**
  - Data source: https://nces.ed.gov/programs/digest/d19/tables/dt19_236.75.asp

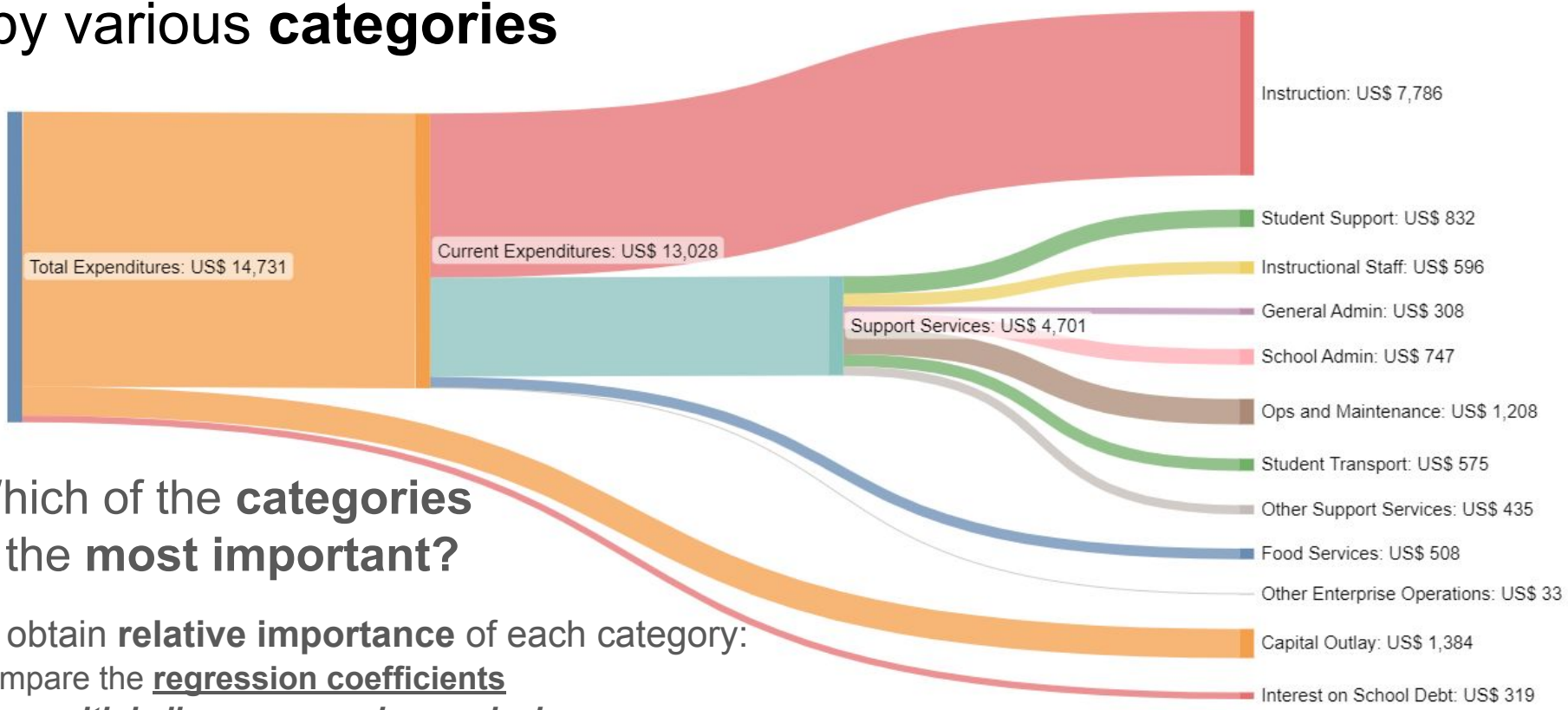
- Government spending is **positively correlated** to the adjusted ACT and SAT test scores

- Which **spending category** has the **most impact** on the **test scores**?





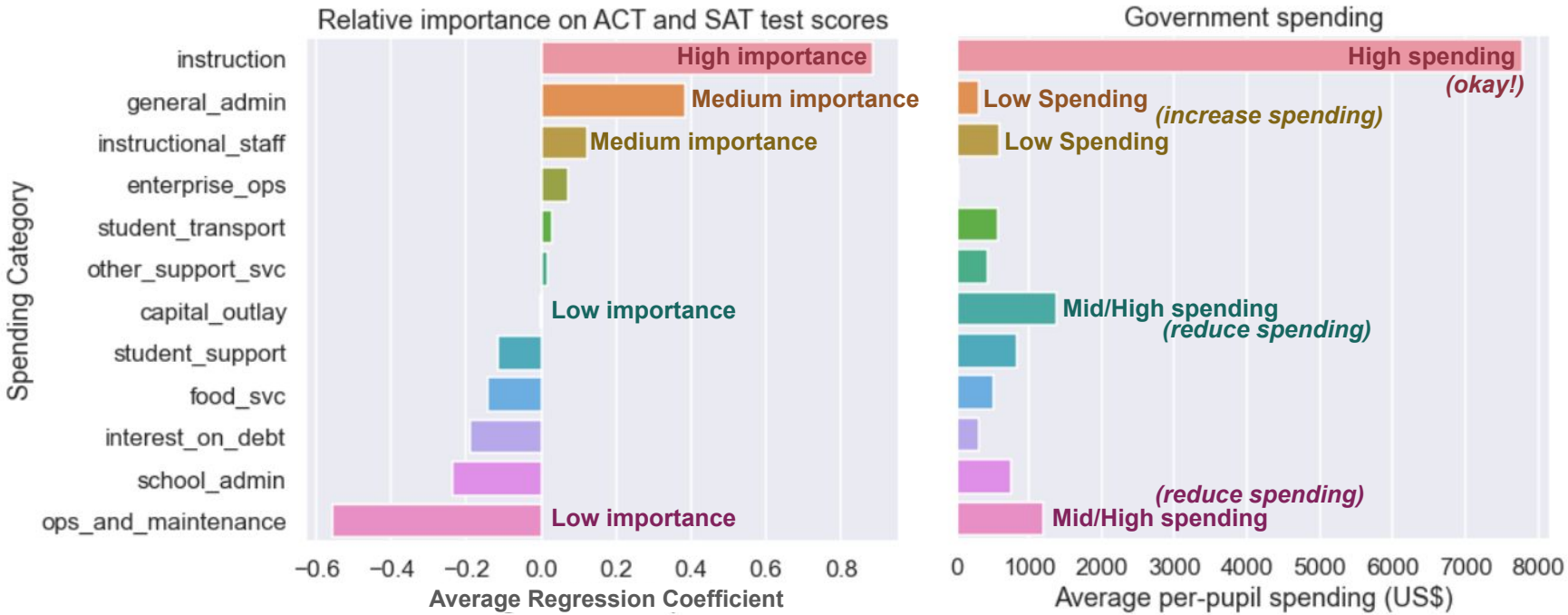
Government **Education Spending** by various **categories**



Which of the **categories** is the **most important?**

To obtain **relative importance** of each category:
Compare the **regression coefficients**
from ***multiple linear regression analysis***
on ***normalized*** features and variables

Relative Importance of Spending Categories



Recommendations can be made based on the **relative importance** of various spending categories and the **actual government spending** on those category

Summary & Conclusion

- There is an **inverse correlation** between the ACT and SAT **test participation rate** in each state.
- There is an **inverse correlation** between the **test score and test participation rate**.
- The distribution of ACT and SAT **test scores** are **bimodal** in nature.
- Using the raw scores there is an inverse correlation between the ACT and SAT test scores.
- The ACT and SAT scores are **adjusted** through linear regression analysis.
- Using the **adjusted scores**, there is a **positive correlation** between the ACT and SAT **test scores**.
- The **relative importance** of each spending category along with the actual government spending, can be used to obtain **recommendations** for future government spendings.

Future Works

Adjust the dollar values to 2019 US$ (to eliminate effect of inflation)

Conduct hypothesis testing and perform regression only on statistically significant features

Study the correlation / impact of socio-economic factors (race, income, etc.) on the standardized test scores