

Parental Income and Graduation Rate among Colleges and Universities in the South

Introduction

For our final project, we seek to determine the extent to which parental income affects 4-year graduation rate across ten states in the Southeast U.S. We operationalize parental income as median parental income and 4-year graduation rate as the percentage of students who complete college or university within four years. We define the Southeast as including Alabama, Florida, Georgia, Arkansas, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, and Tennessee. We also look at different institution types in each state.

Overview of Data

I. Data Walk-through

To conduct this study, we joined four datasets. The first one is an *Opportunity Insights* dataset called “Preferred Estimates of Access and Mobility Rates by College.” We used this dataset to gather information on median parental income across colleges and universities throughout each state in the Southeast. Our second dataset is from *Data World* and is called ‘College Completion.’ We used this dataset to gather information on graduation rate within 4 years. Our third dataset is the *Census American Community Survey* data, which we used to gather the geometry of our 8 states. We also included a dataset from *Urban Institute* to gather information on racial composition, but we ultimately did not use these variables.

II. Data Manipulation & Tidying

First, we tried to join the datasets by state, but there was a problem: the state variable names were different across the datasets. Consequently, we renamed the state variable several times during our work to be able to join correctly. We also used the ‘state_abbrev_crosswalk’ dataset from Quiz 4 to match state abbreviations from our joined college dataset with the state names in the Census dataset. Additionally, we converted the graduation rate in our final dataset into numeric to be able to represent it accurately. Lastly, we filtered the data by state to only include colleges and universities from Southeast states. Having joined all the datasets, we created a dataset and rds file called ‘colleges’ that we used to make our R Shiny app.

Overview of App

Our R Shiny app features five tabs, three with interactive plots and the other two with interactive maps. The first tab (Figure 1) is a generalized scatterplot including all states from the Southeast with 4-year graduation rate on the x-axis and parental median income on the y-axis. This tab also features a checkbox input on the side panel that allows you to choose the states you want represented in the scatterplot.

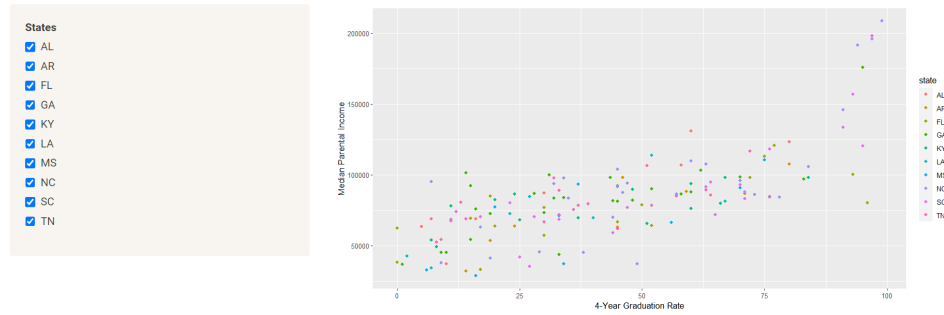


Figure 1: 1st panel which features a Scatterplot between Median Parental Income and 4-Year Graduation Rate across States in the Southeast with checkboxes on the side panel

The next tab (Figure 2) features a similar scatterplot with 4-year graduation rate on the x-axis and parental median on the y-axis, but in this tab, we faceted the plot by school type. Our options for school type are as follows: ‘For Profit,’ ‘Private More Selective,’ ‘Private Non-Selective,’ ‘Private Selective,’ ‘Public More Selective,’ ‘Public Non-Selective,’ and ‘Public Selective.’ Similar to the Scatterplot by State tab, we included checkboxes on the side panel to allow you to choose a combination of the school types to look at at one time.

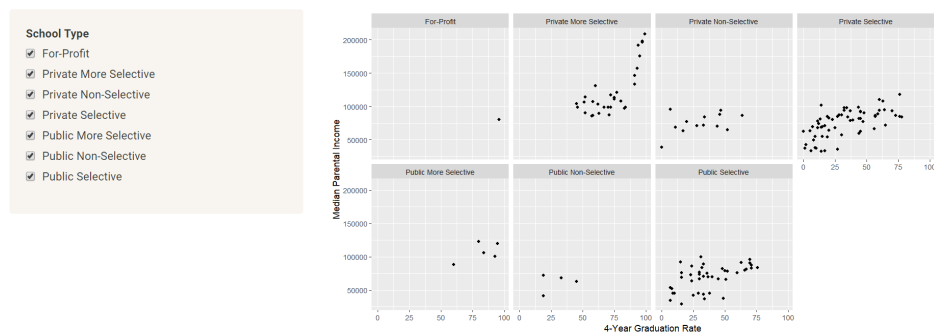


Figure 2: 2nd tab which features our Scatterplot faceted by School type with checkboxes on the side panel

In the third tab (Figure 3), we chose to make boxplots with graduation rate on the x-axis and parental income on the y-axis. For this plot we choose to have a drop down bar for our interactive side panel, allowing you to view one state at a time. However, this feature is limited as it does not allow us to view a boxplot of all the states combined.

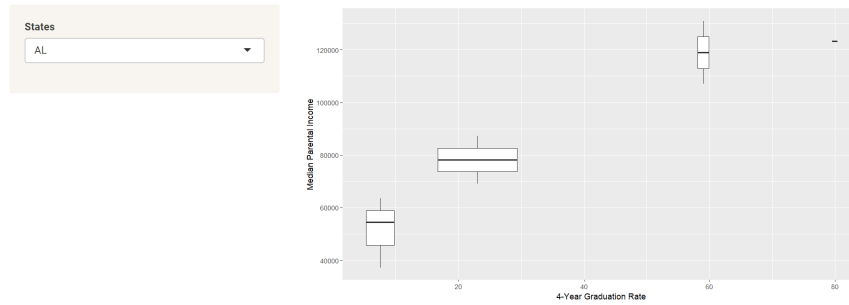


Figure 3: 3rd tab which features our Boxplot between Median Parental Income and 4-Year Graduation Rate across States in the Southeast with the drop down bar on the side panel

The fourth tab (Figure 4) is our first map which shows median parental income by state. This tab features checkboxes on the side panel that allows you to pick which states you would like to view on the map. The fifth tab (Figure 5) features our second map which displays graduation rate by state. In this tab, we also include checkboxes on the side panel that allows you to filter the map by state.

Parental Income Map

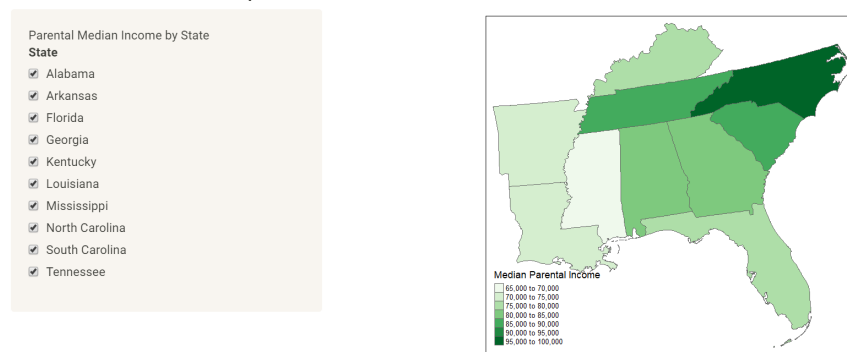


Figure 4: 4th tab which features our first map of Parental Income by state with state checkboxes on the side panel

Graduation Rate Map

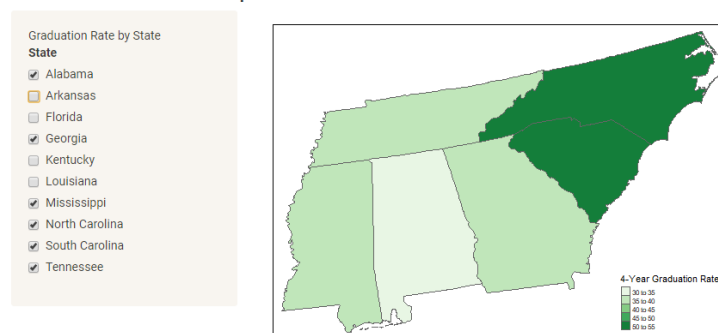


Figure 5: 5th tab which features our second map of Graduation Rate by state with state checkboxes on the side panel

Results and Analysis

I. All States in the Southeast

Overall, we find a positive correlation between median parental income and 4-year graduation rate across the ten states in the Southeast:

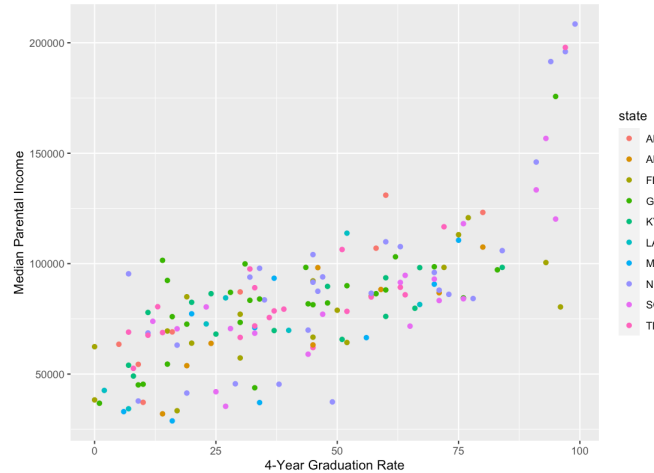


Figure 6: Scatterplot between Median Parental Income and 4-Year Graduation Rate across States in the Southeast

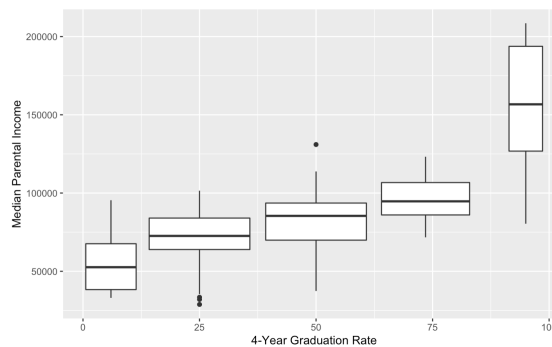


Figure 7: Boxplot between Median Parental Income and 4-Year Graduation Rate across States in the Southeast

According to Figures 6 and 7, as median parental income increases, so does 4-year graduation rate, suggesting that median parental income has a positive impact on graduation rate in the Southeast at large. There are a number of plausible explanations for this relationship. To begin, parents who are wealthier can afford more academic resources like tutoring which could improve their students' academic success, helping them graduate within 4 years. Moreover, students whose parents have more money likely do not have as much pressure to receive a job outside of school, as their parents can afford basic necessities such as food and clothing. In consequence, such students can dedicate more time toward academics, contributing to more academic success and thus higher graduation rates. Another possible explanation is that wealthier parents can send their children to more prestigious schools, which tend to be more expensive. Such institutions often have higher graduation rates, as they tend to have more resources that contribute to academic achievement. Of course, there are numerous other explanations behind the correlation

between median parental income and 4-year graduation rate in the Southeast, but we posit that the ones listed above are of particular importance.

II. *Individual States in the Southeast*

Although there is a positive correlation between our variables collectively, there is considerable variation within individual states. For one, we found that in Georgia, there is a smaller correlation between median parental income and 4-year graduation rate:

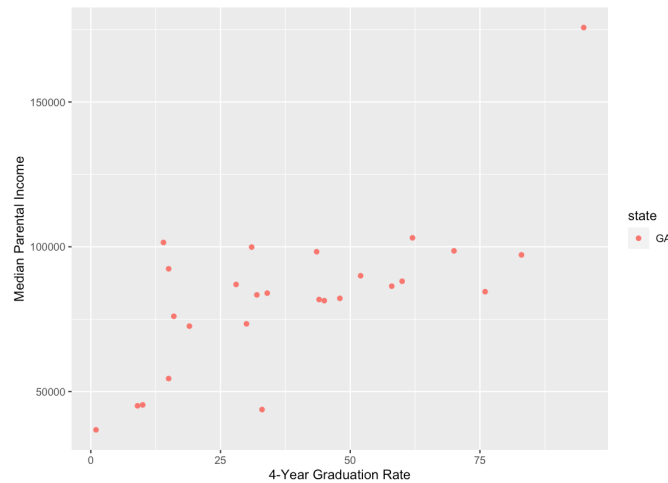


Figure 8: Scatterplot between Median Parental Income and 4-Year Graduation Rate in Georgia

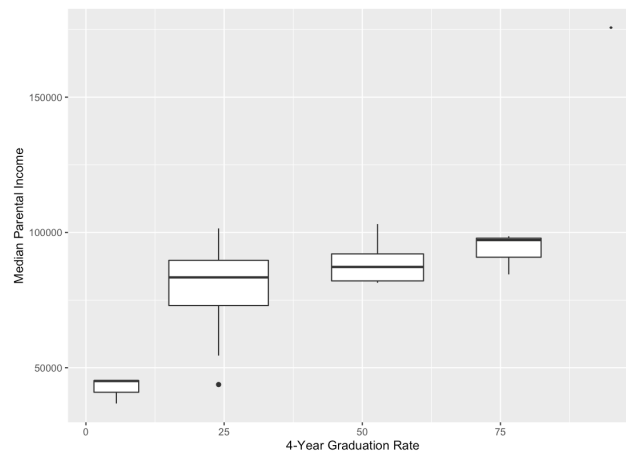


Figure 9: Boxplot between Median Parental Income and 4-Year Graduation Rate in Georgia

Looking at Figures 8 and 9, we can see that although there is a slight increase in graduation rate as parental income increases, there appears to be a smaller correlation between these variables. In Figure 8, we can see that the slope of the points is less steep, indicating a smaller coefficient. Although there is a jump between the first boxplot in Figure 9 to the subsequent boxplots, there does not seem to be a significant increase in graduation rate as parental income increases thereafter.

On the other hand, there appears to be a particularly strong correlation between parental median income and 4-year graduation rate in South Carolina:

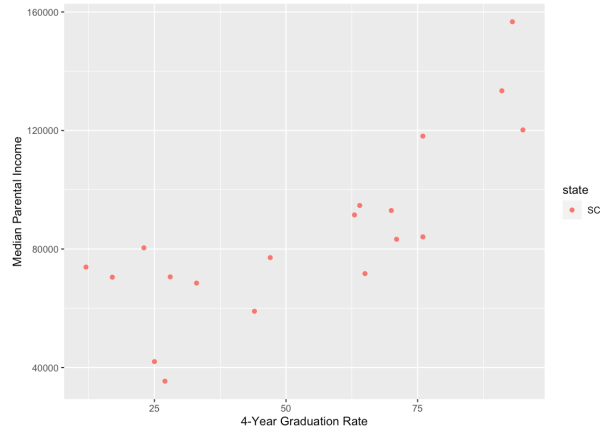


Figure 10: Scatterplot between Median Parental Income and 4-Year Graduation Rate in South Carolina

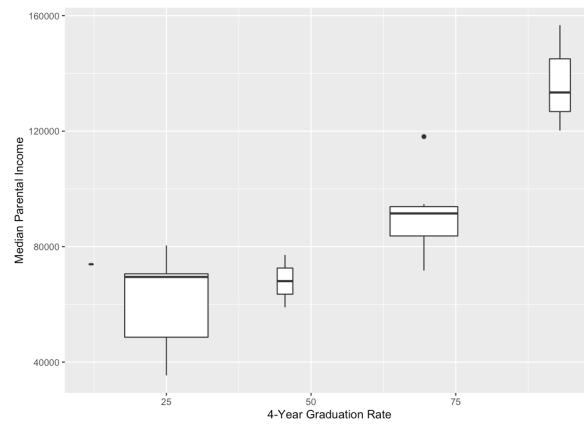


Figure 11: Boxplot between Median Parental Income and 4-Year Graduation Rate in South Carolina

According to Figures 10 and 11, as median parental income increases in South Carolina, there is also a considerable increase in graduation rate. There are several potential explanations behind the variation in the relationships between our variables according to different states. A primary explanation could be that there are various levels of government funding going toward colleges and universities within each state, affecting the quality of education and therefore graduation rates.

III. Different Institution Types

Finally, the relationship between median parental income and 4-year graduation rate varies depending on the type of institution:

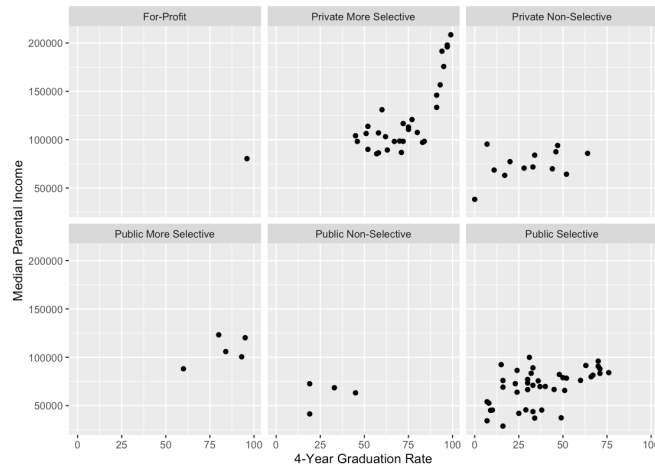


Figure 12: Scatterplots between Median Parental Income and 4-Year Graduation Rate across Different Institution Types

As Figure 12 shows, the correlation between median parental income and 4-year graduation rate varies according to institution type. Unfortunately, for For-Profit, Public More Selective, and Public Non-Selective schools, we do not have as many observations, likely skewing our results. However, there appears to be stronger and more positive relationships between our variables in more selective schools. This finding makes sense, as more selective schools often have better resources, facilitating academic success and hence graduation rates.

Limitations and Suggestions

Due to the time constraints on our project, there are several areas in which we would ideally improve our research design. For starters, we would want to look at parental median income and graduation rates in individual CZs or counties rather than states at large, as doing so would give us a better understanding of the relationship between our variables. By generalizing our findings to entire states, we are glossing over certain nuances, such as the fact that parental median income and graduation rates might vary based upon whether an area is urban or rural. If we had instead incorporated individual CZs, our maps would be more informative as well. Furthermore, we would ideally include more observations per each state, as some states have fewer college and university data points.

Importantly, we would also like to control for more variables such as race, gender, quality of institutions, prior educational attainment, etc. We contend that variables like these are critical to consider when ascertaining the relationship between parental median income and 4-year graduation rate. For instance, someone who is a racial minority could possibly face more barriers to graduation even if their parents are wealthier due to institutional racism.

In terms of the app we produced, there are several changes we would suggest. For one, our maps on graduation rate and parental income are in separate tabs in our app, but we would ideally want to place them next to each other so that they are easier to compare. In addition, we would want to make our state labels consistent in each tab so that all of them either include state

abbreviations or full state names. Lastly, our boxplot tab in the app only displays boxplots for individual states, but it would be more informative if there was an option to create a boxplot that captures all states in the Southeast.

Sources

Opportunity Insights Dataset:

- <https://opportunityinsights.org/paper/mobilityreportcards/>
 - (Preferred Estimates of Access and Mobility Rates by College)

Data World Dataset:

- <https://data.world/databeats/college-completion>
 - (College Completion)

Census Dataset:

- <https://www.census.gov/programs-surveys/acs/news/data-releases/2021/release.html>
 - (2017-2021 ACS 5-year Estimates)

Urban Institute Dataset:

- <https://datacatalog.urban.org/dataset/racial-and-ethnic-representativeness-us-postsecondary-education-institutions/resource>
 - (4-year College Representativeness)