Fiona Caretto

DATA 205

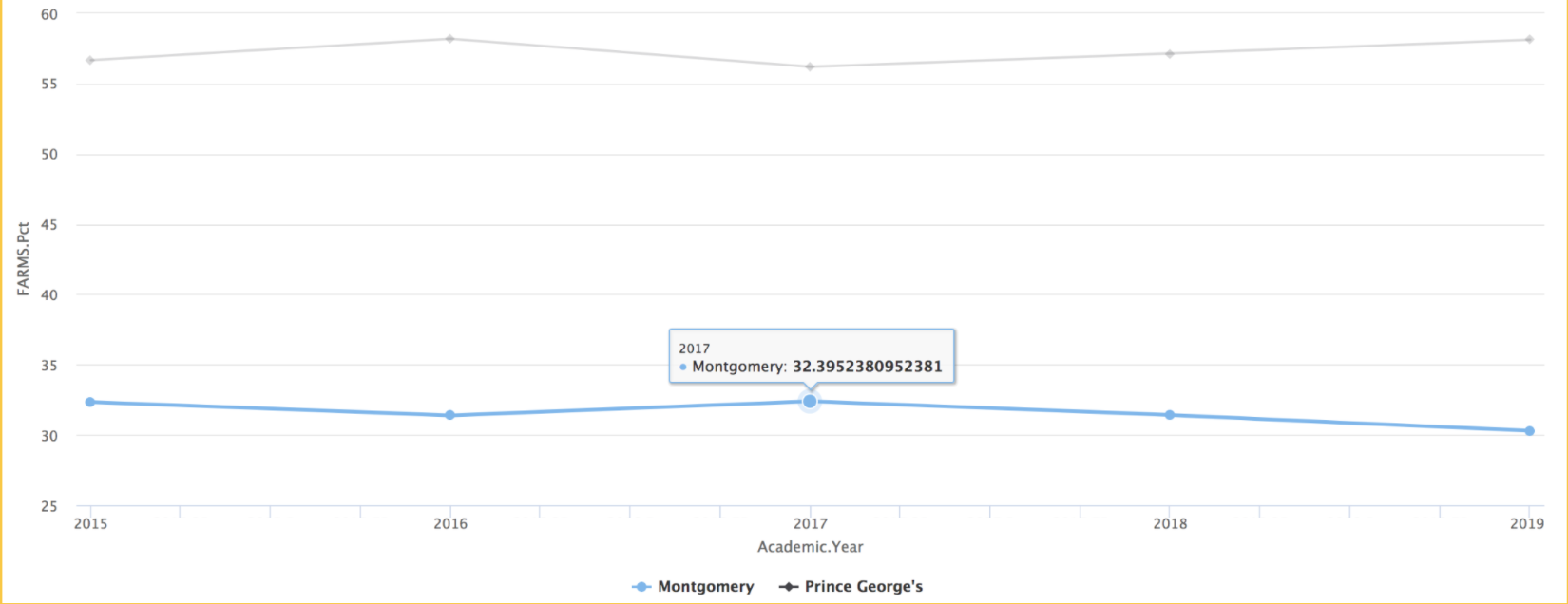
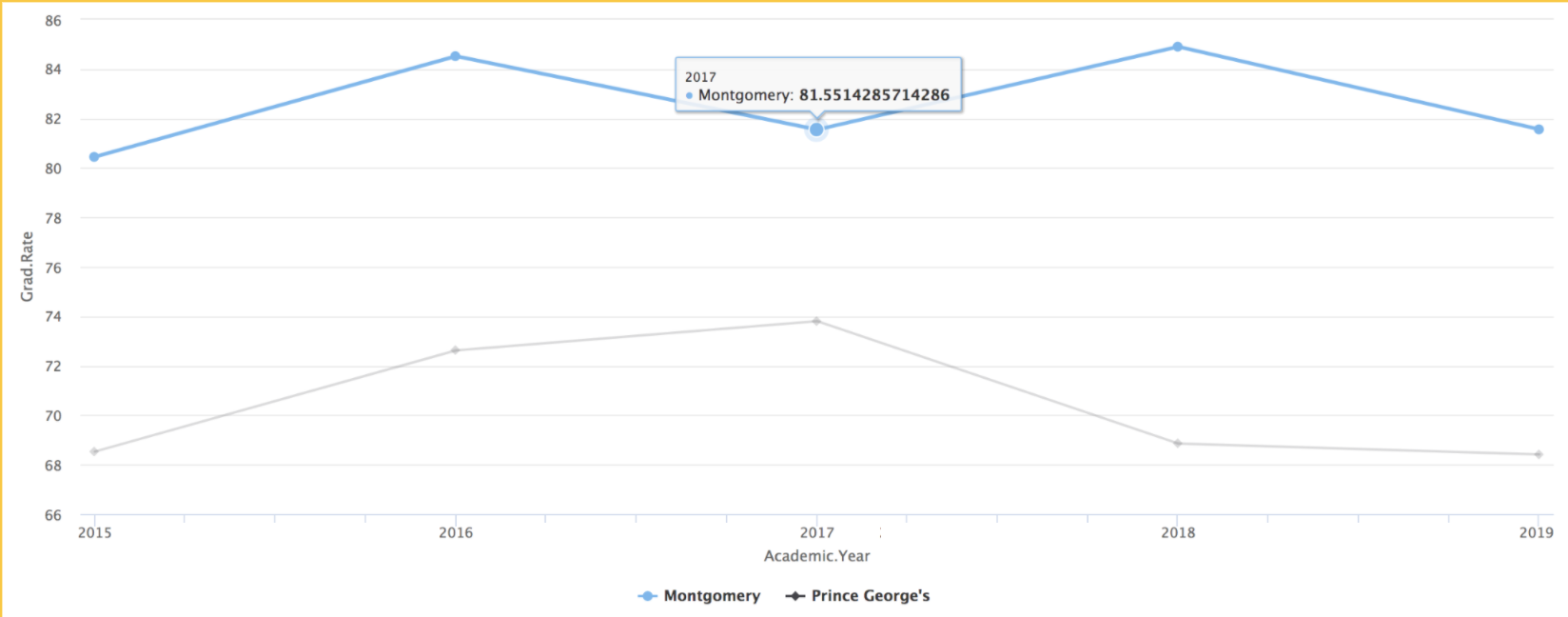
Spring 2021

Housing 101: Low-Income Housing and its Relationship to Educational Outcomes

For this project, I looked at the relationship between low-income housing and academic success in Montgomery County, Maryland. Montgomery County is known for having one of the best public school systems in the country, as well as one of the most diverse. The county is also known for having the country’s oldest and largest affordable housing program. As research on these topics suggest, having access to affordable housing positively impacts educational outcomes for students from low-income families, so I wanted to see if this appeared to be the case in Montgomery County.

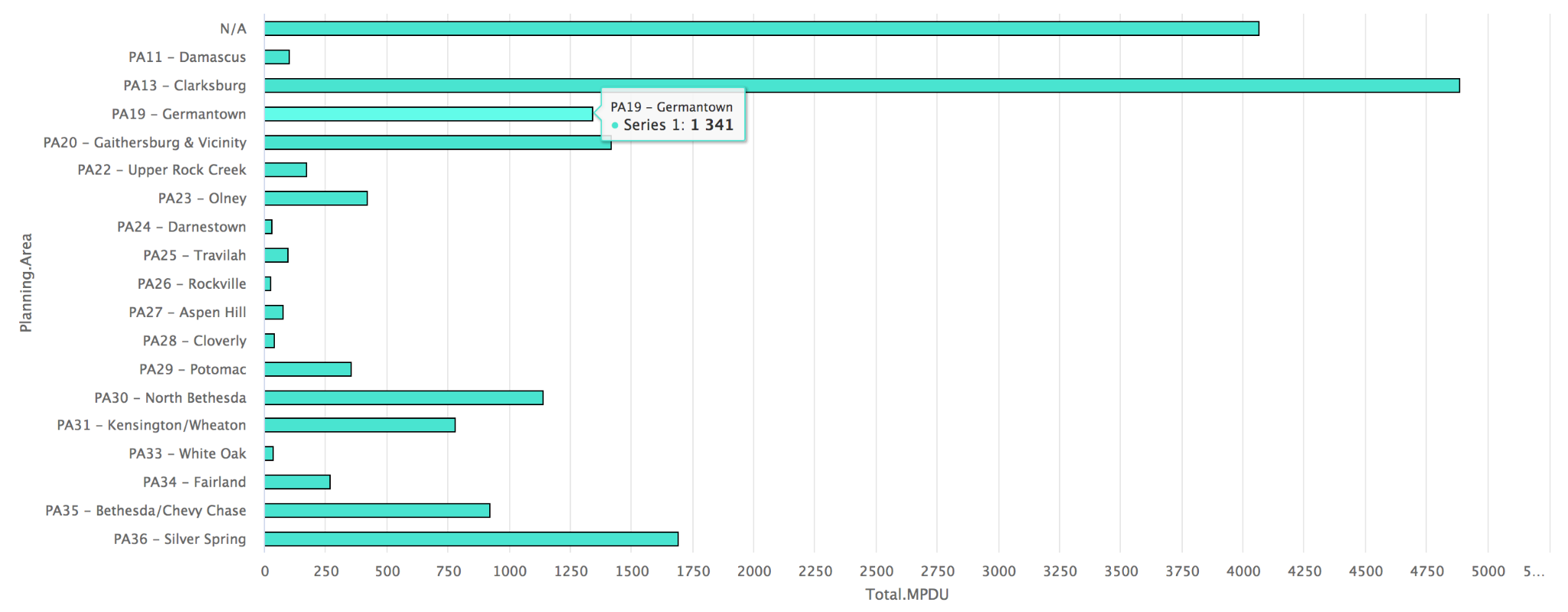
In total, I used five public datasets for my project. Two of these datasets are from dataMontgomery: “Moderately Priced Dwelling Units” and “Housing Licensing and Registration.” The “Moderately Priced Dwelling Units” dataset provides data on the number of moderately priced dwelling units (MPDUs) available out of the total constructed units across various “planning areas,” which mostly match up with various cities/towns in Montgomery County. I was interested in exploring the planning area, total units, and total MPDU variables. The “Housing Licensing and Registration” provides data on housing built in Montgomery County including details such as location, structure type (i.e., Condominium in Garden Style Property, Townhouse Single Family), number of units, and ownership type (i.e., Housing Opportunities Community, Trust, LLC). I was interested in exploring the ownership type, property zip code, and unit count variables. An important thing to note is that one of the responsibilities of the Housing Opportunities Community of Montgomery County is to acquire MPDUs for rental to lower-income families. This is where the connection is made between these two datasets from dataMontgomery.

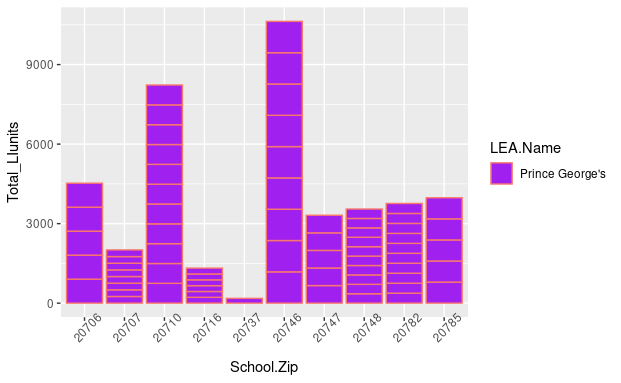
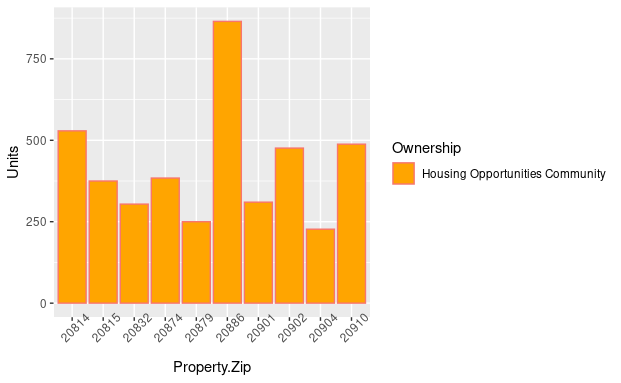
I used two datasets from the Maryland State Department of Education’s “Report Card” website: “Adjusted Cohort Graduation Rate” and “Students Receiving Special Services.” These datasets are published for every academic year, so I used the datasets from a five-year span (2015-2019) to have a decent amount of data for observation. This data was not reported for 2020 or 2021, which is why I did not include those in my project. As some of my datasets were not very large, I decided to also look at the data for Prince George’s County Public Schools (PGCPS), another large and diverse county in Maryland, in addition to Montgomery County Public Schools (MCPS). The variables I was interested in for the “Adjusted Cohort Graduation Rate” dataset were graduation rate, county, and academic year. I also chose to focus on the “4-year adjusted cohort graduation rate” specifically in this dataset, which is the percentage of high school freshmen who graduate within four years of starting 9th grade, adding in any students who later transfer into the cohort and subtracting any who leave. The “Students Receiving Special Services” dataset provides data on various special services provided by Maryland public schools, such as Free and Reduced Price Meals (FARMS) and ADA/Section 504. I was interested in looking at the FARMS percentage, county, and academic year variables. FARMS is often used to measure school need. While this variable alone is not a sufficient measure of disadvantaged schools/students, for the purposes of my project I chose to see what the FARMS data could tell us. Below are two line graphs depicting the trends in FARMS (top) and Graduation Rates (bottom) in each county from 2015-2019.

Lastly, since I wanted to compare my housing variables across Montgomery County and Prince George’s County, I used the “Low-Income Housing Tax Credit” dataset from the Department of Housing and Urban Development (HUD). This dataset provides data on low-income housing units and projects around the country including the project location, the number of low-income units, and the year the project was placed in service. I looked at the data for Montgomery County and Prince George’s County, specifically pulling the zip codes of the high schools within those counties. I used the “FIPS2010” variable--the unique census 2010 tract ID which I was able to pull out the state and county codes from and filter based on county. I also looked at the project zip code and total low-income (LI) units variables.

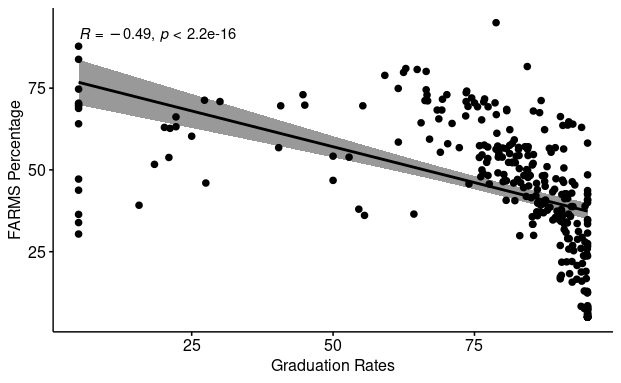
I wanted to get a picture of the low-income housing in Montgomery and Prince George’s counties. As you can see from the following bar charts, the amount of low-income housing units is not evenly distributed across the counties. Some trends I saw in Montgomery County were that Montgomery Village, Kensington, North Bethesda, Bethesda, Chevy Chase, Silver Spring, and Clarksburg were among the areas with the most MPDUs/housing units owned by the Housing Opportunities Community.





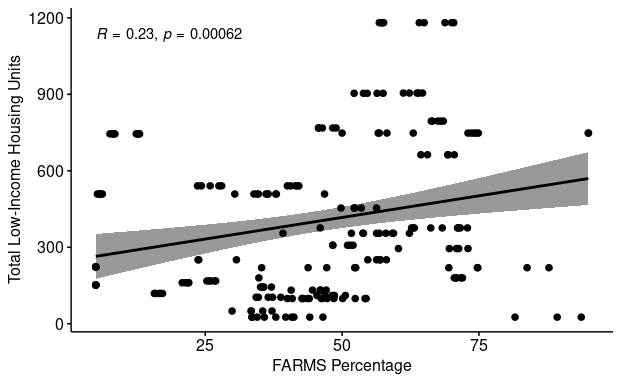
These results are important because research shows that economically integrated housing improves academic success, and the government could use this data to help plan where to construct affordable housing units to improve the economic integration of housing.

Next, I looked at the relationship between FARMS and graduation rates. I performed a Pearson’s correlation test and found that these variables had a statistically significant, moderate, negative correlation, with a correlation coefficient of -0.49 (as shown in the scatter plot with regression line below).



These results weren’t too surprising to me. School funds are secured through local property taxes. If a school is located in a lower income area, then that school will receive fewer funds. The amount of students who receive free and reduced price meals can serve as a somewhat reliable predictor of a school’s community income level. It is reasonable to conclude that schools with more students who receive free and reduced price meals will be located in lower income areas and will therefore have fewer funds. Schools with fewer resources suffer academically, which could include lower graduation rates. These results are important because these variables represent the two primary needs kids are seeking to have fulfilled when they go to school: to be fed/kept safe and to be educated. This data shows that MCPS is struggling to meet both of those needs simultaneously. The Montgomery County government can therefore use this data to pinpoint the schools that need more resources from the government (i.e., Title I funding).

I also looked at the relationship between FARMS and low-income housing. I performed another Pearson’s correlation test here and found that they had a statistically significant, weak, positive correlation, with a correlation coefficient of 0.23 (as shown in the scatter plot with regression line below).



Initially, I was surprised by these findings because I had predicted that if there are more students on FARMS, then there would be more low-income housing in the area, as both variables reflect low-income status. As mentioned previously, research shows that access to affordable housing improves educational outcomes for children from low-income families. Using FARMS as a measure of financial need, if there is not a strong correlation between students on FARMS and amount of low-income housing, perhaps this is a contributing factor to the lower graduation rates associated with higher percentage of students on FARMS. The government could use this data to pinpoint where in the county has the most need for MPDUs/low-income housing units, to support the education of students from low-income families.

Looking forward, I think the “Moderately Priced Dwelling Units” and “Housing Licensing and Registration” datasets from dataMontgomery capture very important and useful data. However, I believe they could be improved and I think these improvements would be of great benefit to the county. I suggest a dataset that combines the variables of both of these datasets, but provides more consistent and complete data. Similar to HUD’s “Low-Income Housing Tax Credit” dataset, this dataset should capture the number of housing projects/units being built in Montgomery County, the funding source (“Ownership Type”), the number of low-income units, the project location (using a more specific variable than “Planning Area,” like in the “Moderately Priced Dwelling Units” dataset, such as address, city, or zip code, like in the “Housing Licensing and Registration” dataset). While working on my project, I found some of the data provided in these datasets difficult to compare with similar data from other sources/parts of the country. Data quality is important because incomplete data produces unreliable results. Governments, communities, and businesses rely on these findings to make serious decisions for their residents and clients. Improving data quality should be a priority for all data analysts so that those using this data can make well-informed, educated decisions.

According to Montgomery County’s Department of Housing and Community Affairs website, the Moderately Priced Housing Law establishes certain benchmarks for the availability of affordable public housing. According to this law, between 12.5% and 15% of the houses in new subdivisions of 20 or more units must be moderately priced dwelling units. Further, 40% of these MPDUs must be offered to the Housing Opportunities Commission and other non-profit housing agencies for use by low and moderate income families. Additionally, participants must apply to this housing program and are selected via a random selection drawing. Further amplifying the importance of this issue is the fact that the need for housing greatly outweighs the availability of units. Based on this information, I believe a dataset that captures quality data on the amount of MPDUs available in the county (perhaps per housing selection period), the amount of applicants, the amount of applicants selected each period, and the location of these units (using a variable such as address or zip code, to provide more specific data) would be of great utility to the Montgomery County government.

Another new dataset I would suggest is, again, based off of the research that shows that access to affordable housing increases educational outcomes for children from low-income families. It would be helpful to collect data on this and provide it through dataMontgomery. Variables that would be beneficial to capture in this dataset could include the poverty level of families with children who attend Montgomery County Public Schools, the amount of affordable housing units per school district, the amount of students in the school district who live in affordable housing, and a measure of academic success (such as grades, test scores, and/or graduation rates). This dataset could be used to see if academic success does improve for children from low-income families when they are provided with affordable housing, and the government could then use this to help plan construction of affordable housing units within the county and increase the overall success of MCPS and its students.

You can find my R Markdown file, as well as links to all of these datasets, and my PowerPoint presentation on my Github repository located here: <https://github.com/fcaretto/DATA-205-Capstone-Project>. All of the datasets were read live into the Markdown file through the URLs, so my code is reproducible for those interested. Additionally, I would like to acknowledge all of those who have helped me with my capstone project and throughout the Data Science Certificate Program at Montgomery College. To Professor Saidi, Professor Mohamed, Professor Iapalucci, my classmates, Victoria Lewis, dataMontgomery, and the Montgomery County Government--I can’t thank you enough for your support and guidance. I have learned so much from all of my professors, and from conversations with classmates. I thoroughly enjoyed each of my courses and especially this capstone project. I recommend the Data Science Certificate program at Montgomery College for anyone with an interest in data science, no matter their experience level. I came in having never written a single line of code, and now I’m finishing the program with a breadth of new knowledge, a capstone project, and a Github portfolio that I can be proud of. I hope this project can provide some helpful insight into the topics of low-income housing and education.