Extension Plan

Analysis of COVID-19 and Education in Essex County, New Jersey

Aaliyah Hänni

11/10/2021

Motivation

In Essex County, New Jersey effective March 18th, 2021 all schools in the State of New Jersey were closed for the remaining school year (source: [School responses in New Jersey to the coronavirus (COVID-19) pandemic](#)). The impacts of a school closure and of students who now have a 4 month gap in their education is still largely unknown.

For my analysis, I would like to explore how the COVID-19 pandemic has impacted the education system in Essex County, New Jersey. This is important to know for the people that were affected, to know how they stand compared to others, and also to educators who are going to be working with this group of students in the future. The results could be used to guide educators and education administrators in what the best decisions are when dealing with students who have been impacted, or if there is ever a similar future event that causes mass school closures.

Personally, I am interested in this topic because my two younger siblings were significantly affected by the introduction of COVID-19 into our community. One sibling struggled tremendously, where previously he was one of the best performing students in his grade. My other sibling, in middle school, actually thrived when the school moved to an online platform and found that he was able to learn more and be more productive. I thought it was interesting how this one event could impact two brothers so differently. This has motivated me to explore the overall impact of the COVID-19 pandemic on students.

Research questions and hypotheses

The research question that I would like to answer is:

How did the COVID-19 pandemic impact K-12 public school students in Essex County, New Jersey?

This is a very broad question and there are several ways of measuring the impact of the pandemic on the educational system. For this analysis, I have decided to focus on school enrollment, test scores, and graduation rates. Comparing the enrollment rates over the last three gives a sense of how many students are going to school. Looking at test scores allows us to gauge how students are performing and if their learning has been impacted. Graduation rates give a view of overall school completion, and if students are still completing on time. I plan to use these three metrics to assist in answering the high-level question of how students were impacted and get a sense of in what way.

Thus, I can break down my initial question into a summary of three more measurable questions that aim to answer the higher level question of impact:

1. What does the change in school enrollment for all grades look like when comparing the 2017-18, 2018-19 and 2019-20 school years, for all K-12 public schools in Essex County, New Jersey?
2. What is the difference in test scores when comparing the 2017-18, 2018-19 and 2019-20 school years, for all K-12 public schools in Essex County, New Jersey?
3. What is the change in 4 and 5-year school graduation rates for the last 4 years, for all public schools in Essex County, New Jersey?

Data used

Dataset Source: Official Site of the State of New Jersey

Dataset Link: https://rc.doe.state.nj.us/download

The datasets that I plan to use are contained in the New Jersey School Performance Report from the Official Site of the State of New Jersey for 2018-19 and 2019-20 school years, which can be sorted and filtered by the county and school level. These datasets are available for export in an excel file which has tabs for school enrollment trends, average test scores, and graduation rates. Each tab contains the County, School, and summary data for the year performance. In the enrollment trends tab, the data displays a total number of students enrolled in each grade. The test scores tab displays a column for the test type (ACT, PSAT, or SAT) , the school average, the state average, and the benchmark score. The graduation tab displays the 4 and 5 year graduation rate for the previous four school years.

Using the data from the last 3 school years (2017-2018, 2018-2019, 2019-2020), I will be able to extract these summary values into a final dataset that will be able to directly answer the questions outlined above, which will allow me to compare performances of the pandemic school year (2019-2020) to the previous two school years.

Given that we are looking at averages and total amounts for each school, and then culminating them together to display totals for the county in our final analysis, I don't believe there to be any ethical concerns with utilizing this dataset, especially for our analysis.

Unknowns and dependencies

There are several potential issues with this dataset that are noted, such as cancellation of tests due to COVID, or delays in graduation due to COVID in multiple school districts. This may impact the conclusions of the analysis, and will skew the results. It is unknown if some schools administered different policies when it came to how they handled the pandemic, so this could also be difficult to interpret when performing analysis.

Ideally this analysis would contain data from the 2020-2021 school year, to get a sense of the ongoing impact due to COVID-19 and to review how well the county is rebounding. Unfortunately, data is only publicly available up until the 2019-2020 school year.

Methodology

To tackle my research question of analyzing the impact of the COVID-19 pandemic on the public education system, I plan to utilize 3 key datasets: school enrollment, test scores, and graduation rates. Each of these datasets will then be analyzed and used to create a visualization.

School Enrollment

The school enrollment datasets for the previous three school years, which contains the data of the total number of students enrolled by grade for each school, organized by county. For my analysis, I plan to aggregate the total number of enrolled students for each grade in the district. Using the resulting data, I will create a visualization to show the total county school enrollment for each year, segregated by grade, to get a sense of how the different grades were impacted.
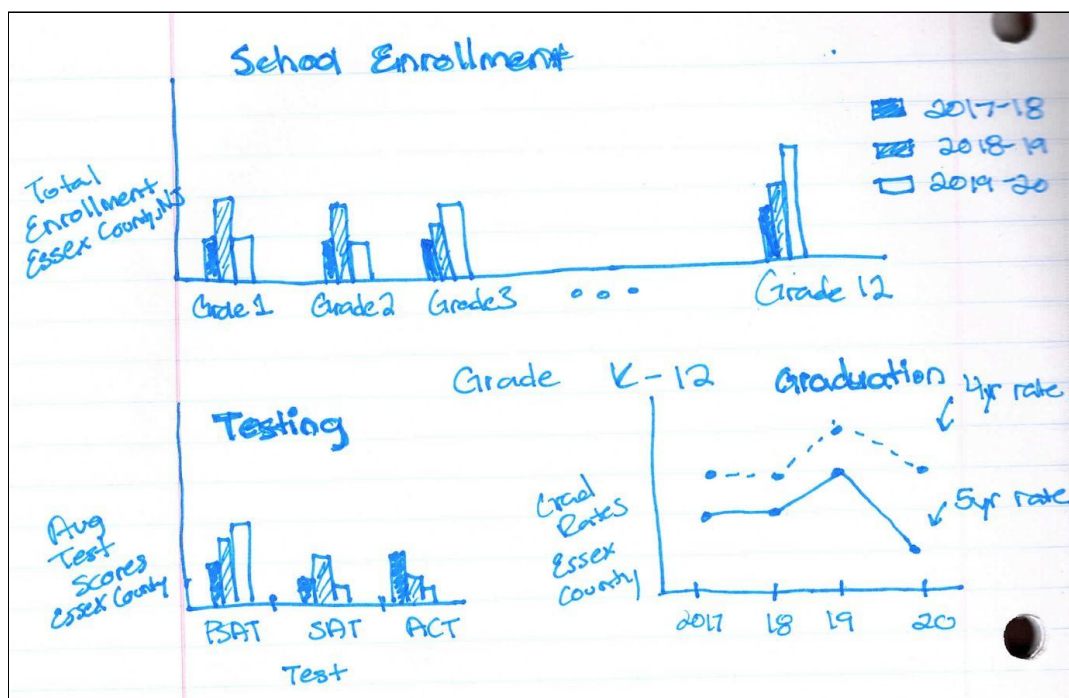
Test Scores

The test scores datasets contain the ACT, SAT, & PSAT average test scores for the previous three school years, for each school. It also contains the state averages for each test. I will use these datasets to aggregate the average test scores for Essex County for each of the three tests, for each school year. Using the resulting data, I will create a bar plot of the average county test scores for each year, segregated by the three different tests.

Graduation Rates

The graduation rates dataset contains the historical data for the 4-year and 5-year graduation rate for the previous five school years, for each school. It also contains the annual target and data about if that target was met. I will use this dataset to compute the average graduation rate, for both 4-year and 5-year, for Essex County for each school year. Using the resulting dataset, I will plot the 4 and 5-year graduation rates for Essex County over the last four years.

All three measurements will be combined into a single interactive dashboard which can be used to analyze the 'how' question. A dashboard summarizing key education metrics over time, will allow the user to explore how schools were impacted by the resulting school closures due to the COVID-19 pandemic. In order to explore the impacts of COVID-19 on the educational system, I plan to create multiple visualizations that can be filtered and interacted with by the user. The dashboard will be created using Tableau, which allows me to combine multiple datasets and create a variety of visualizations. Below is a rough draft mockup of what this visualization will look like.

Timeline to completion

To be completed by **11/11/21:**

- [ ] Gather the required data for analysis
- [ ] Design the final dashboard layout
- [ ] List out fields needed in final dataset that will be used to create dashboard

To be completed by **11/18/21**:

- [ ] Perform data cleaning by isolating needed fields, handling NULL values
- [ ] Perform feature engineering of data, aggregate values for the county for each school year
- [ ] Merge together all datasets into a final dataset
- [ ] Create an excel file of final dataset to use for dashboard

To be completed by **11/25/21**:

- [ ] Import dataset into Tableau
- [ ] Create previously designed visualizations
- [ ] Migrate visualizations into a final dashboard

To be completed by **12/2/21**:

- [ ] Finalize dashboard: update titles, labels, fonts, colors for consistency and clarity
- [ ] Add intuitive instructions where needed and cite sources for data used
- [ ] Publish dashboard to Tableau public

Cited Sources

https://ballotpedia.org/School_responses_in_New_Jersey_to_the_coronavirus_(COVID-19)_pandemic#Timeline_by_school_year

https://rc.doe.state.nj.us/