HIGH SCHOOL EDITION
Journal of Student Research
Volume 12 Issue 3 (2023)

Examining Achievement Gaps Between Students in Kansas Public Schools

Advith Natarajan[1] and Deepesh Agarwal[2#]

[1]Manhattan High School
[2]Kansas State University
[#]Advisor

ABSTRACT

Disparities in academic performances across subgroups are widely recognized. This investigation examines these differences within Kansas public schools. Using the Kansas Department of Education state assessment scores published on the Kansas Report Card database, we compare various student subgroups on the basis of ethnicity, income, and literacy. Most notably, this inquiry uncovers novel findings regarding district size and student performance. This study is also the first to assess the impact of COVID-19 on Kansas students. Our findings indicate significant achievement gaps between minority and white students, self-paid lunch and free & reduced lunch students, and English Language Learner (ELL) and non-ELL students. These gaps in performance were greater in medium and large-sized school districts. We also analyzed the impact of the recent COVID-19 pandemic on academic performance across different student subgroups in grades 3-12. Significant decreases in performance were observed for students in grades 3-8 and a significant increase in performance for those in grades 10-12.

1. Introduction

The primary goal of public education is to provide all students with accessible and beneficial coaching. It aims to provide every student with opportunities to excel, irrespective of ethnicity, race, socioeconomic status, or any other differentiating factors. However, the system still falls short in terms of closing the achievement gap among various student subgroups.

Numerous studies have shown significant achievement gaps between high-income and low income students in the United States. Sean Reardon of Stanford University investigated the impact of increasing income inequality on student performance over a 50-year timeframe. His study concluded that the achievement gap between high and low-income families is 30 to 40 percent larger among children born in 2001 than 25 years earlier. Additionally, the black white achievement gap remains and research has indicated that Hispanic and Asian English Language Learner(ELL) students are falling further behind (Carnoy et al). Their work utilized data from the National Assessment of Education Progress to observe the shifts in performance across various subgroups over the past 15 years. In the state of Kansas, similar trends are seen. A released report by the Kansas Department of Education in 2017 revealed the performance gap between low and high-income students was significantly greater in larger schools (Carter). This research used Kansas State Assessment scores from the 2015-16 school year to observe the achievement gaps between subgroups based on income and school size. The report used the percentage of each subgroup that scored proficiently on the exam to draw conclusions about the nature of these performance gaps.

The public education landscape is ever-changing, making novel research initiatives like this one integral to understanding the socioeconomic dynamics of education. Our study investigates achievement gaps on various levels. Not only do we examine overall state assessment performance, but we observe how these metrics shift based on school

district sizes. We treat the impact of the recent COVID-19 pandemic as another variable in our analyses. This provides us an opportunity to analyze the impact of online learning on the performance of these student subgroups as well.

The purpose of this inquiry is to analyze and compare the performance of various subgroups of students in Kansas public schools. We want to determine where differences between these subgroups exist and attempt to answer why they occur. This research can be highly beneficial for public school districts across the state of Kansas, and possibly nationwide. Ultimately, these insights can be utilized by education professionals to provide under-served students with further resources that they need to succeed. This study is especially significant as the COVID-19 pandemic kept students away from normal school settings for almost two years. These objectives lead us to our main questions: Are differences found between opposing subgroups? Are these differences statistically significant? Are these differences smaller or larger with different-sized school districts? What impact did COVID-19 have on the performance of these subgroups?

2. Preliminaries

2.1 Description of student subgroups

For this study, we compared 3 groups of students, namely white students vs minority students, self-paid lunch students vs free & reduced lunch students, and non-English language learner (Non-ELL) students vs English language learner(ELL) students. Minority students encompassed African-American, Asian, Hispanic, Multi-Racial, and American Indian/Alaska Native students. As stated by the Kansas Department of Education (KSDE), free & reduced Lunch requirements are based on income and are obtained via application. Qualification for these programs is based on the income requirements outlined in Table 1.

Table 1: Income requirements for qualification for free & reduced lunch, as stated by the Kansas Department of Education (KSDE).

Household Size	Free Benefits	Reduced Price Benefits
1	17,667	25,142
2	23,803	33,874
3	29,939	42,606
4	36,075	51,338
5	42,211	60,070
6	48,347	68,802
7	54,483	77,534
8	60,619	86,266

An English language learner (ELL) is defined as a student whose primary language is not English, and whose English proficiency or lack thereof provides a barrier to successful learning. We chose these specific subgroups for various reasons. First, these subdivisions gave us the largest sample size in our dataset. Many of the other subgroups in the data were fairly inconsistent throughout the data (i.e. not enough data points). More importantly, however, our chosen subgroups best reflect the different demographics we wanted to explore in our research. Our analyses of white vs minority students gave us a comparison based on ethnicity, while our analyses of Self-paid vs Free & Reduced lunch students gave us a comparison based on income. And finally, our analyses of Non-ELL vs ELL students gave us a comparison based on literacy. For our district size-based comparison of these subgroups, we divided our dataset

into 3 categories based on population: Small districts with 0-10,000 students, Medium districts with 10,000-50,000 students, and Large districts with 50,000+ students.

2.2 Statistical Tools and Techniques

To perform our comparisons, we used hypothesis testing. Hypothesis testing is a form of statistical inference that uses data to draw conclusions about different populations. In hypothesis testing, an initial claim is made about the population. This is known as the null hypothesis, denoted by H0. An alternative hypothesis (HA), which is opposite to the claim of the null hypothesis, is then made. Therefore, after analyzing our sample data, rejecting our null hypothesis would prove our alternative hypothesis to be true. This would be classified as a statistically significant change. In the case of our research, we use independent, two-sided t-tests to observe whether gaps between our comparable subgroups were significant. We also use Hedges' G as an indicator of practical magnitude for the gap between subgroups. Hedges' G is a measure of effect size. Effect size tells us how one group differs from another. This metric gives us a more practical measure of the difference between subgroups, which is especially helpful due to our astronomically small p-values.

3. Dataset Description

To conduct our analyses, we utilized the Kansas Report Card database, available publicly on the KSDE report card website. We used the KSDE state-assessment scores –published annually for every public school district in the state of Kansas– as our measure of performance. The Kansas State assessments are administered to all Kansas students in grades 3-12. These exams test students' proficiency in Math, English Language Arts (ELA), and Science. We used a collection of 6 datasets (one for each year 2016-2022, excluding 2020 due to COVID-19) to perform our comparisons.

Assessment scores are measured on a 4-level scale. The Kansas Assessment Program (KAP) classifies these levels as the following:

- Level 1: A student at Level 1 shows a limited ability to understand and use the skills and knowledge needed for postsecondary readiness.
- Level 2: A student at Level 2 shows a basic ability to understand and use the skills and knowledge needed for postsecondary readiness.
- Level 3: A student at Level 3 shows an effective ability to understand and use the skills and knowledge needed for postsecondary readiness.
- Level 4: A student at Level 4 shows an excellent ability to understand and use the skills and knowledge needed for postsecondary readiness.

In our dataset, scores were categorized by school district, student subgroup, and grade. The scores were initially displayed as percentages of each level. However, we took the expected value for each level in order to obtain the mean score for that specific grade and subgroup. Due to this structure, we weren't able to obtain a precise n value for our hypothesis tests. Instead, we classified one "datapoint" as one row of data, corresponding to a specific grade and subgroup (See appendix A). As aforementioned, the data was also separated by test subject (Math, English, and Science). However, we analyzed performance holistically and disregarded the grade level and subject in our comparisons for each subgroup. In other words, we analyzed performance as an overall metric of grade level and subjects together. We did this because the KSDE State assessment is an examination that tailors to each grade level's respective knowledge.

4. Experimental Study

To apply our statistical methods, we used data science libraries in Python. We used the Pandas library to organize our data, Scipy library to perform independent t-tests and Matplotlib library for data visualizations. We also used an online calculator to find Hedges' g values for our district size-based comparisons.

4.1 Baseline Comparisons

The baseline comparisons highlight state-wide differences in performance between the subgroups detailed previously. We performed two-sided, independent t-tests to observe whether the performance gaps between subgroups were significant. In this case, the null hypothesis states that the performance of subgroup A is equal to the performance of subgroup B. Therefore, the alternative hypothesis establishes that the performance of subgroup A is not equal to the performance of subgroup B. Note that these comparisons are statistically valid, as displayed by the Quantile-Quantile (Q-Q) normality plot in Figure 1.

Table 2: T-Statistics and p-values for each subgroup comparison by year during 2016-2022. (Null Hypothesis: White = Minority means that the mean performance of students from White and Minority subgroups is same; Alternate Hypothesis: White = Minority means that the mean performance of students from White and Minority subgroups is not same. Similar explanations can be written for all other comparison groups.)

Test ID	Year	T-statistic	p-value
Comparison Groups: White vs. Minority Null Hypothesis: White = Minority Alternate Hypothesis: White /= Minority			
1	2016	10.47	2.71e-11
2	2017	10.96	7.40e-12
3	2018	10.59	1.74e-11
4	2019	9.92	5.71e-11
5	2021	9.64	2.24e-10
6	2022	5.28	3.17e-06
Comparison Groups: ELL vs. non-ELL Null Hypothesis: ELL = non-ELL Alternate Hypothesis: ELL /= non-ELL			
7	2016	7.78	1.11e-08
8	2017	7.45	2.77e-08
9	2018	10.07	4.91e-11
10	2019	10.61	1.53e-11
11	2021	10.86	6.74e-12
12	2022	12.26	8.77e-15
Comparison Groups: self-paid vs. free & reduced lunch Null Hypothesis: self-paid lunch = free & reduced lunch Alternate Hypothesis: self-paid lunch /= free & reduced lunch			
13	2016	8.67	1.31e-09
14	2017	8.87	7.33e-10
15	2018	9.25	3.41e-10
16	2019	9.01	6.48e-10
17	2021	7.74	1.72e-08
18	2022	9.58	2.42e-11

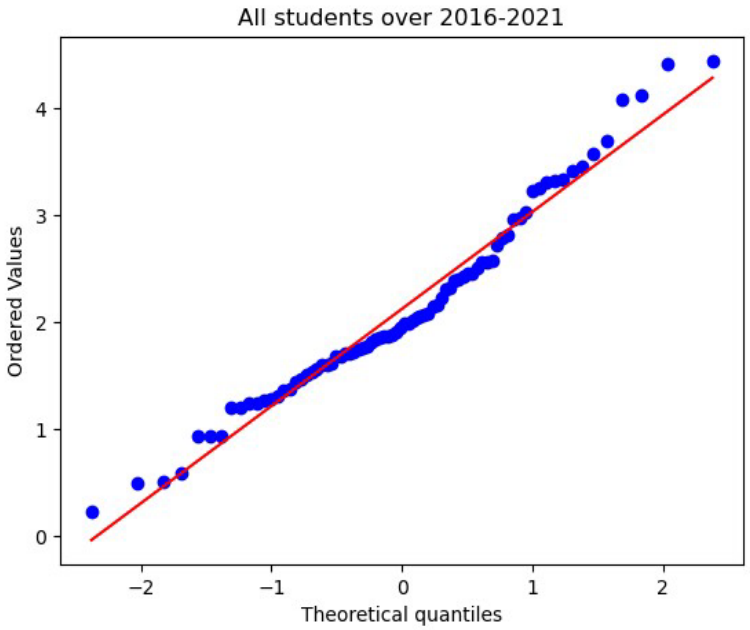

Figure 1: Q-Q Plot

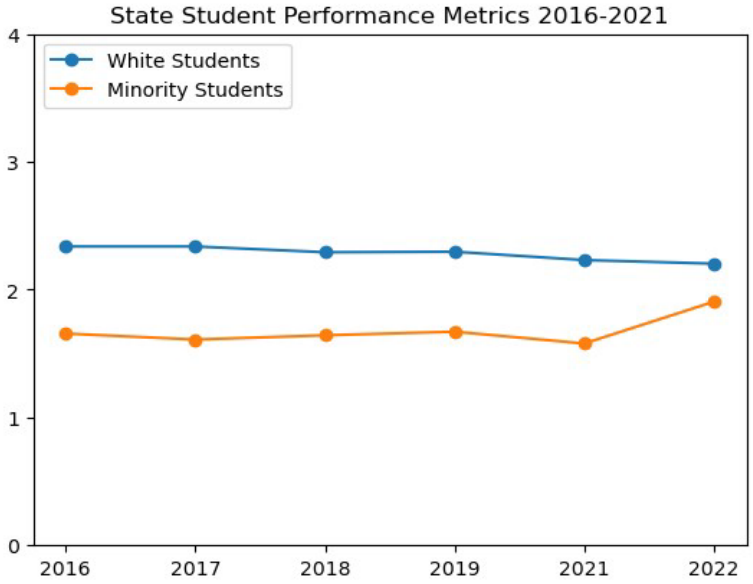

Figure 2: White vs Minority

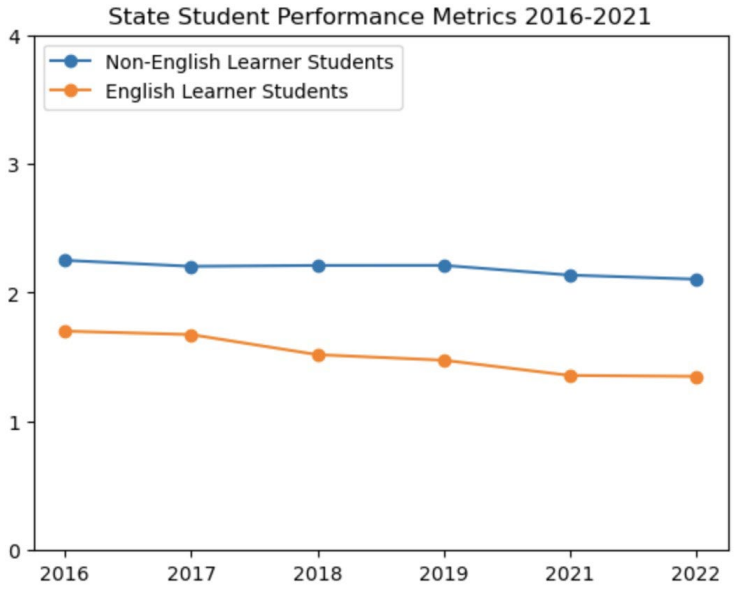

Figure 3: ELL vs Non-ELL students

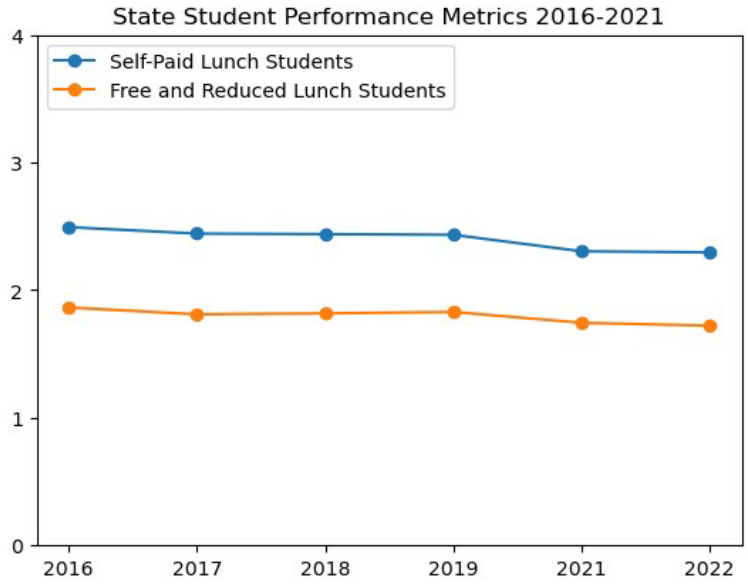

Figure 4: Self-Paid Lunch vs Free & Reduced Lunch

As shown in Table 2, the baseline comparisons showed significant gaps in performance between subgroups throughout all six years. Minority students have consistently scored significantly lower than their white counterparts (Figure 2). ELL students have scored significantly poorer than non-ELL students (Figure 3). Free & reduced lunch students have scored significantly poorer than self-paid lunch students (Figure 4).

4.2 Digging Deeper

The aim of these comparisons was to go beyond the surface-level achievement gaps summarized above. We wanted to observe how these performance disparities changed based on school district size. Once again, we used independent, two-sided t-tests to compare the subgroups. We also utilized Hedges' G in order to gauge the practical magnitude of

the gaps between subgroups based on school district size. Similar to the baseline comparisons, the null hypothesis states that the performance of subgroup A is equal to the performance of subgroup B. Therefore, the alternative hypothesis states that the performance of subgroup A is not equal to the performance of subgroup B. We also considered our Hedge's g value for each size-based comparison in order to observe which district size category had the greatest practical difference.

Table 3: T-statistics, p-values, and Hedges' G for all district sizes and population subgroups comparison by year during 2016-2022.

Year	Small Schools				Medium Schools				Large Schools			
	Test ID	T-statistic	p-value	Magnitude (Hedges G)	Test ID	T-statistic	p-value	Magnitude (Hedges G)	Test ID	T-statistic	p-value	Magnitude (Hedges G)
Comparison Groups: White vs. Minority Null Hypothesis: White = Minority; Alternate Hypothesis: White /= Minority												
2016	19	27.10	3.33e-149	0.63	25	19.25	1.56e-74	0.77	31	12.95	1.94e-31	1.03
2017	20	31.69	3.48e-198	0.73	26	21.47	2.54e-90	0.86	32	13.82	8.26e-35	1.10
2018	21	21.81	1.05e-95	0.70	27	16.25	1.77e-51	0.93	33	10.64	8.19e-21	1.15
2019	22	26.51	2.19e-138	0.77	28	20.01	7.77e-76	0.98	34	13.98	1.14e-32	1.31
2021	23	28.73	2.67e-161	0.81	29	18.79	1.31e-67	0.95	35	12.84	1.83e-28	1.27
2022	24	26.24	2.03e-138	0.72	30	16.22	2.62e-53	0.72	36	8.01	4.67e-14	0.66
Comparison Groups: ELL vs. non-ELL Null Hypothesis: ELL = non-ELL; Alternate Hypothesis: ELL /= non-ELL												
2016	37	27.35	1.06e-118	1.04	43	17.92	4.18e-63	1.01	49	11.12	2.13e-25	1.05
2017	38	26.42	9.29e-116	0.99	44	19.84	2.18e-76	1.07	50	10.12	9.18e-22	0.95
2018	39	25.75	4.60e-82	1.49	45	21.51	1.18e-73	1.75	51	12.64	9.93e-28	1.69
2019	40	31.88	1.84e-107	1.77	46	30.94	8.30e-129	2.25	52	19.62	7.73e-53	2.34
2021	41	30.99	1.89e-105	1.75	47	29.68	3.52e-123	2.15	53	18.94	1.26e-49	2.26
2022	42	30.05	2.54e-99	1.67	48	29.04	9.02e-123	2.05	54	18.13	5.93e-47	2.16
Comparison Groups: self-paid vs. free & reduced lunch Null Hypothesis: self-paid lunch = free & reduced lunch; Alternate Hypothesis: self-paid lunch /= free & reduced lunch												
2016	55	63.52	0	1.16	61	31.46	8.97e-169	1.57	67	15.02	3.07e-40	1.41
2017	56	63.30	0	1.15	62	32.30	6.52e-176	1.61	68	16.57	8.12e-47	1.56
2018	57	43.98	0	1.13	63	24.81	1.18e-100	1.75	69	12.71	2.49e-27	1.69
2019	58	50.83	0	1.17	64	28.01	5.34e-127	1.77	70	16.07	1.39e-39	1.92
2021	59	50.60	0	1.14	65	23.71	1.27e-97	1.49	71	12.01	2.32e-26	1.43
2022	60	47.01	0	1.06	66	22.03	4.21e-87	1.39	72	13.96	1.08e-32	1.66

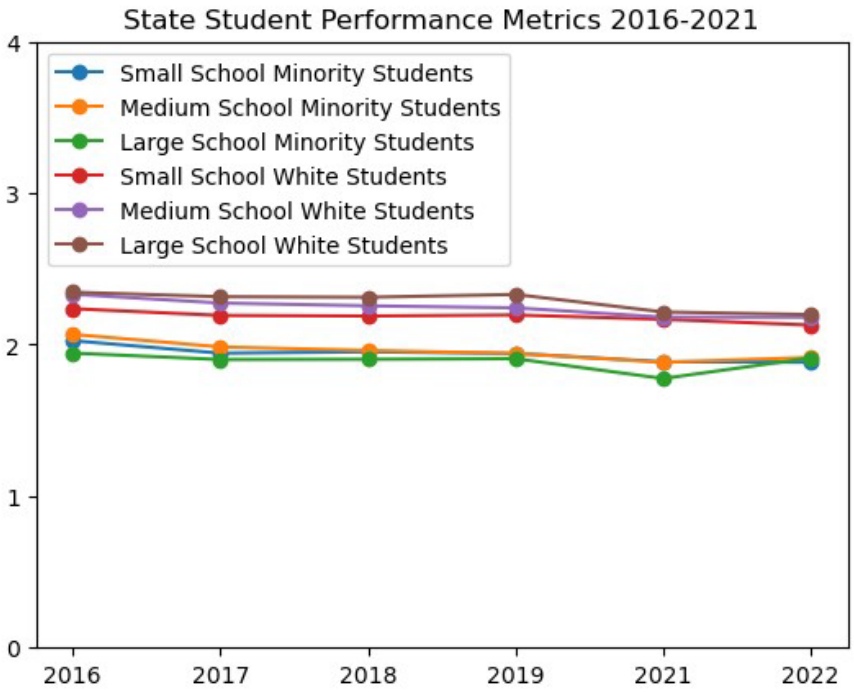

Figure 5: White Students vs Minority Students by District Size

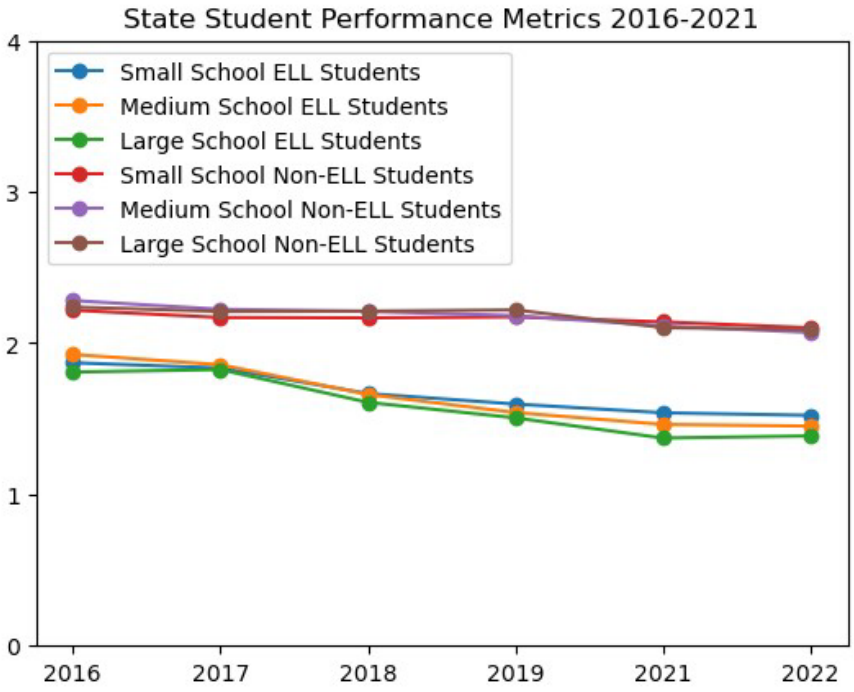

Figure 6: ELL Students vs Non-ELL Students by District Size

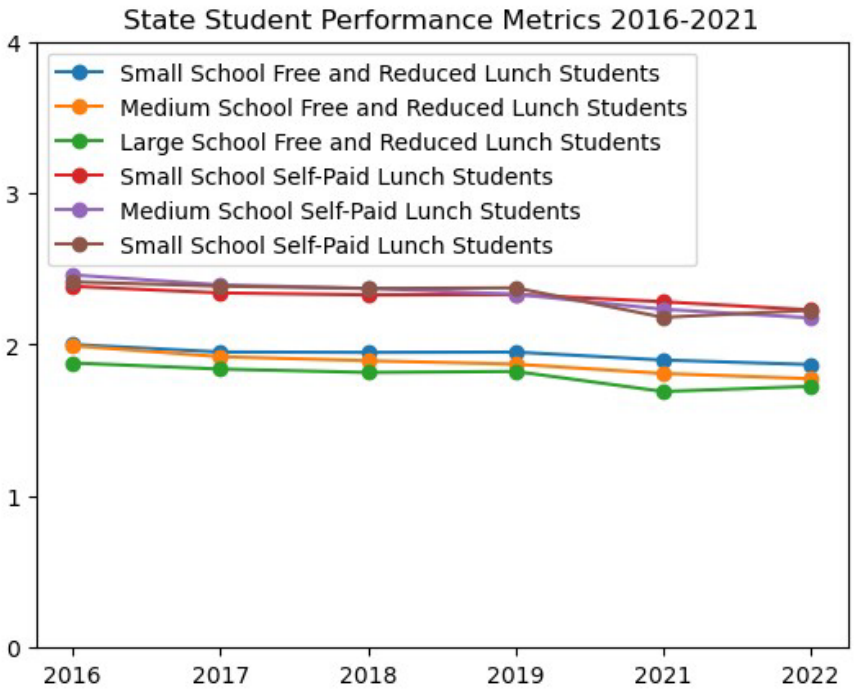

Figure 7: Self-Paid Lunch Students vs Free & Reduced Lunch Students by District Size

These comparisons examined performance based on district size. As shown in Table 3, the gaps in performance were larger in medium and large-sized school districts (10,000+ students) as compared to small school districts (0-10,000 students). This trend was consistent for white v minority (Figure 5), Ell v non-ELL (Figure 6), and self-paid v free & reduced lunch (Figure 7) comparisons.

4.3 Impact of COVID-19

The COVID-19 pandemic kept students away from normal in-person learning for almost two years. This resulted in the Kansas State Assessments not being administered in 2020. The purpose of these comparisons is to observe how students' performance shifted as a result of this gap year. We analyze these performance dynamics by grade level to find which grades experienced the greatest drop-off or increase in achievement. This could tell us what impact online learning had on students. This means that in order to compare the same group of students from 2019 to 2021, we would have to compare 3rd graders in 2019 to 5th graders in 2021 (4th to 6th, 5th to 7th, etc.). This would allow us to compare roughly the same groups of students and observe how their performance shifted. We applied this technique across all of our subgroups.

Table 4: Comparison in performance of student subgroups before and after COVID-19. (Null Hypothesis: There is no impact of COVID-19 on the performance of students; Alternate Hypothesis: There is a significant impact of COVID-19 on the performance of students.)

Grade Level Comparisons	**T-statistic**	**p-value**	**Significant increase/decrease**	**T-statistic**	**p-value**	**Significant increase/decrease**
	Population subgroup: White students			Population subgroup: Minority students		
3 vs. 5	8.01	2.75e-15	Decrease	17.89	1.80e-61	Decrease
4 vs. 6	16.59	1.77e-54	Decrease	26.96	2.54e-100	Decrease
5 vs. 7	16.57	6.41e-56	Decrease	21.92	3.09e-78	Decrease
6 vs. 8	13.49	2.37e-38	Decrease	19.68	4.23e-72	Decrease
8 vs. 10	6.74	2.44e-11	Decrease	5.00	7.44e-07	Decrease
10vs. 12	-14.43	7.26e-43	Increase	-6.31	7.38e-10	Increase
	Population subgroup: ELL students			Population subgroup: non-ELL students		
3 vs. 5	5.88	2.55e-08	Decrease	8.63	1.99e-17	Decrease
4 vs. 6	14.91	2.61e-29	Decrease	17.84	7.55e-62	Decrease
5 vs. 7	8.09	2.73e-13	Decrease	17.45	1.95e-61	Decrease
6 vs. 8	6.81	1.66e-09	Decrease	14.07	2.35e-41	Decrease
8 vs. 10	5.48	2.86e-07	Decrease	6.62	5.50e-11	Decrease
10vs. 12	-6.67	1.18e-09	Increase	-14.05	4.32e-41	Increase
	Population subgroup: Students with self-paid lunch			Population subgroup: Students with free & reduced lunch		
3 vs. 5	9.90	4.59e-22	Decrease	8.12	1.45e-15	Decrease
4 vs. 6	16.95	1.55e-55	Decrease	21.41	1.23e-80	Decrease
5 vs. 7	18.15	1.66e-64	Decrease	14.22	5.37e-42	Decrease
6 vs. 8	14.32	9.44e-42	Decrease	12.91	8.40e-35	Decrease
8 vs. 10	9.09	4.69e-19	Decrease	7.47	2.25e-13	Decrease
10 vs. 12	-12.69	1.06e-33	Increase	-14.75	9.37e-43	Increase

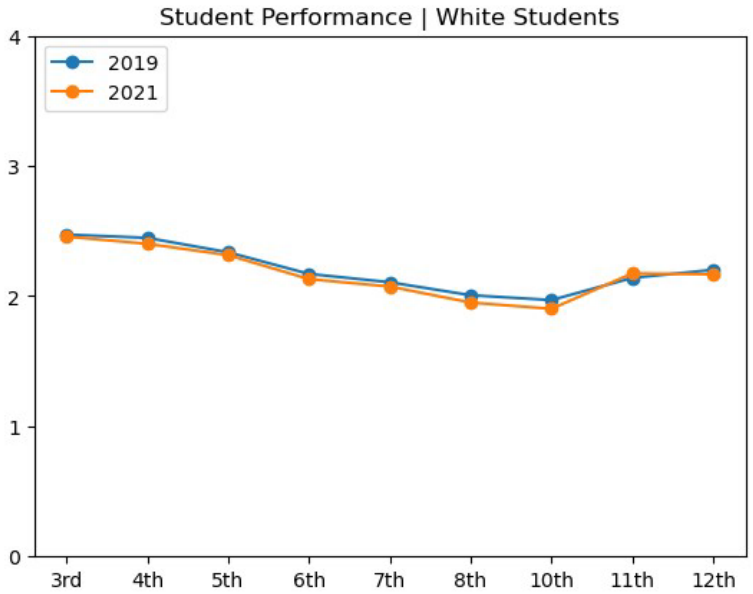

Figure 8: White Students

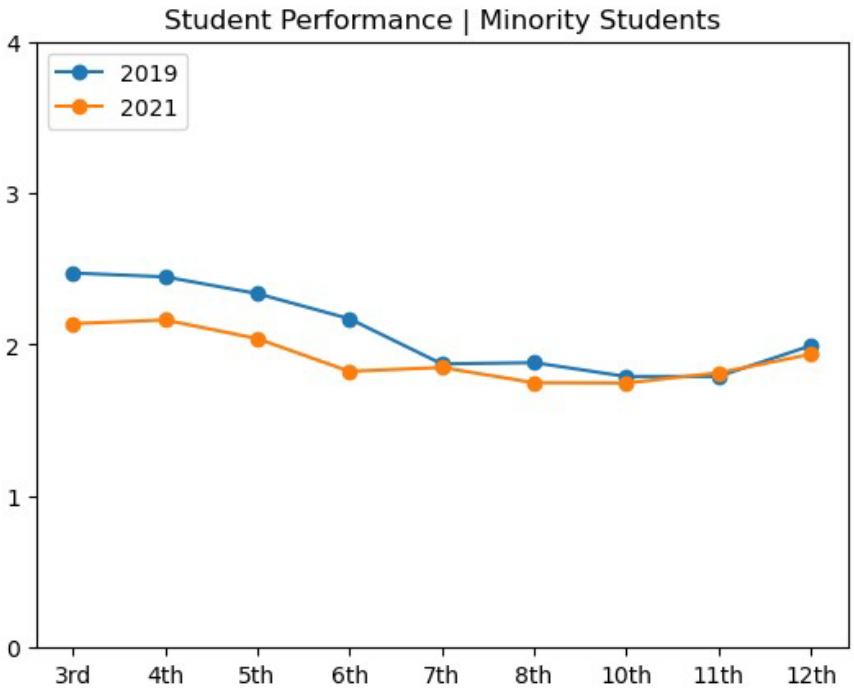

Figure 9: Minority Students

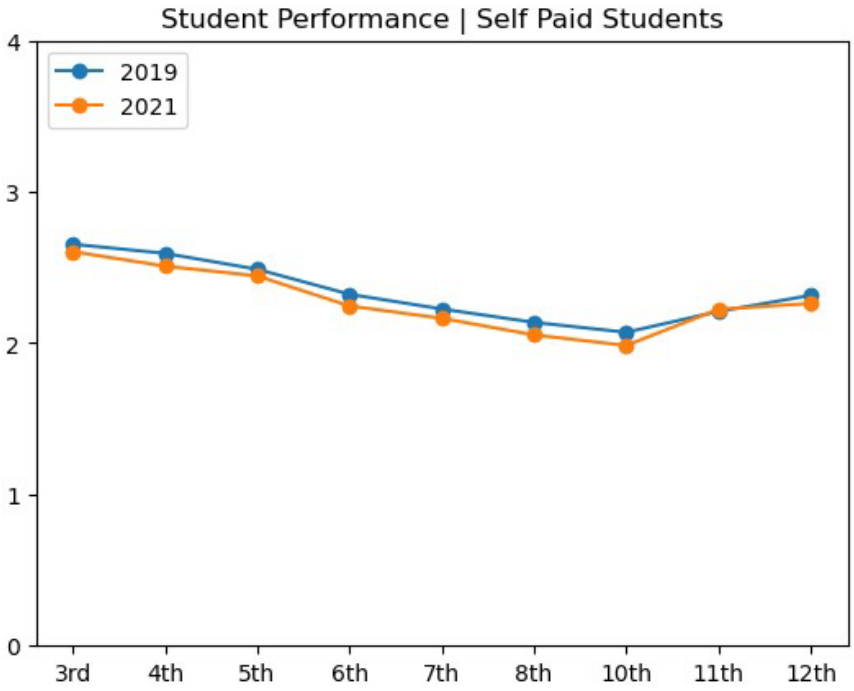

Figure 10: Self-Paid Lunch Students

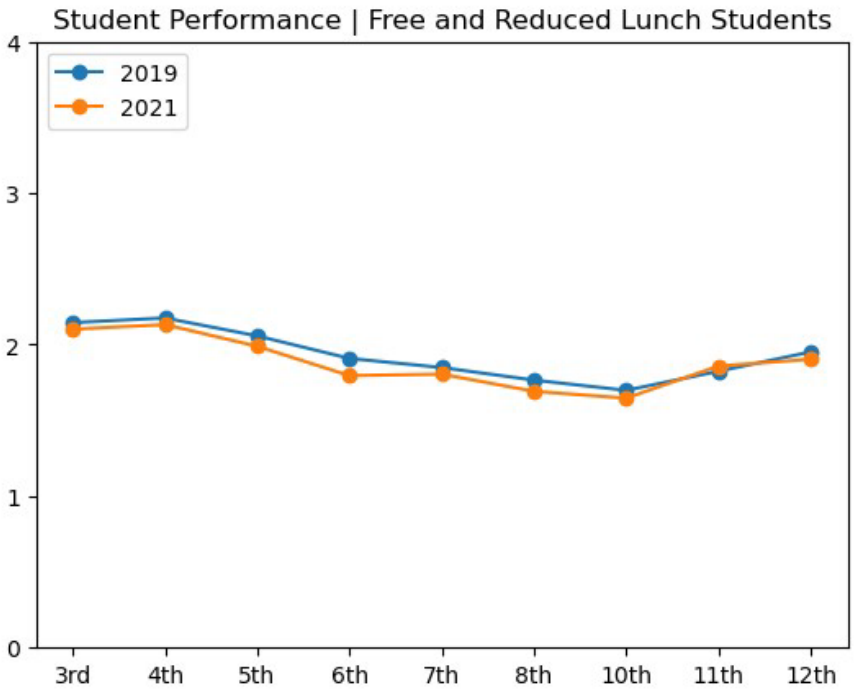

Figure 11: Free & Reduced Lunch Students

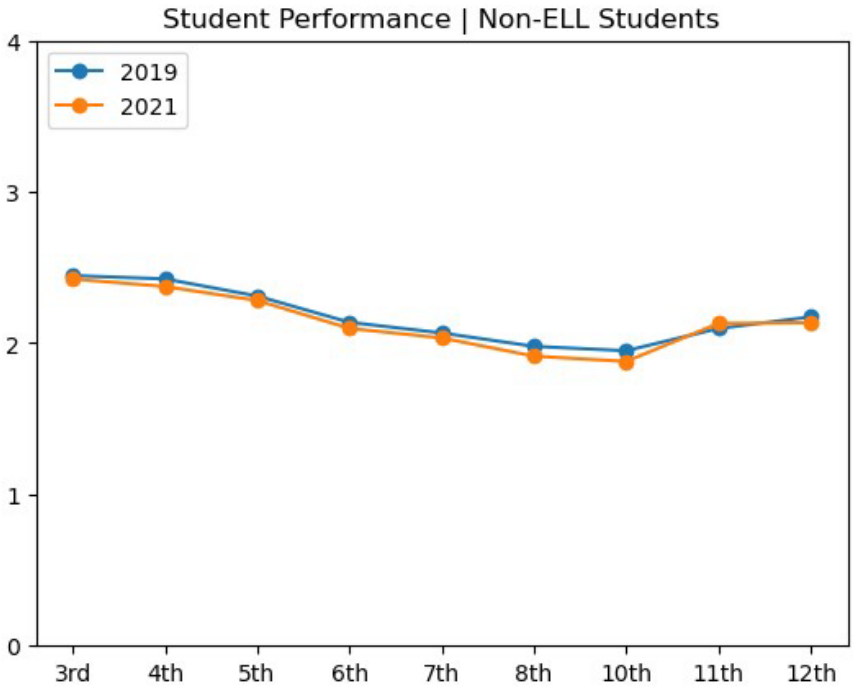

Figure 12: Non-ELL Students

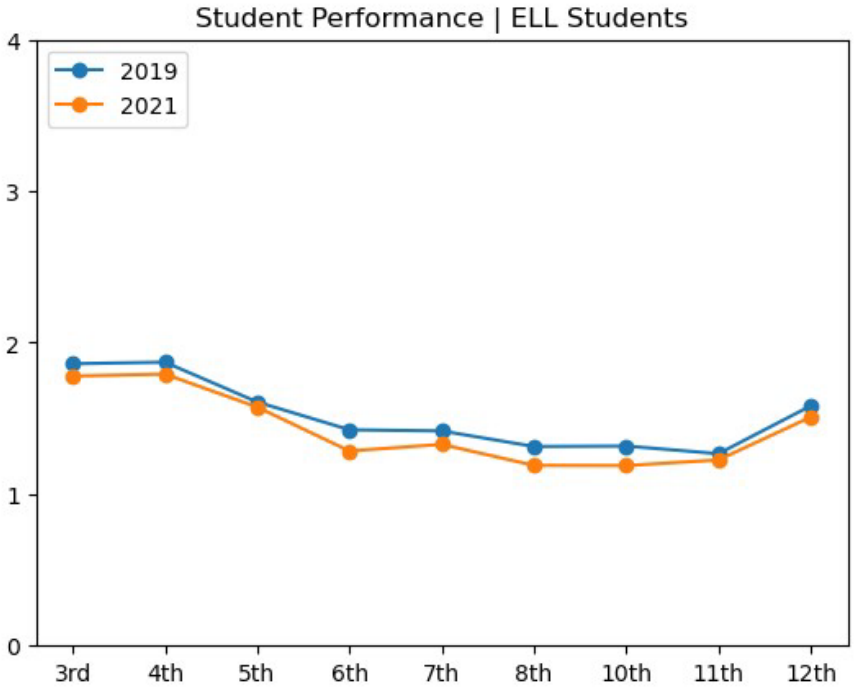

Figure 13: ELL Students

The impact of COVID-19 on student performance was also profound. As delineated in Table 4, across all subgroups, we observed statistically significant decreases in performance for students in grades 3-8. However, from grades 10-12, we actually observed a statistically significant increase in performance between pre-COVID-19 vs post-COVID-19 scores. The most statistically significant decrease in performance from grades 3-8 was seen in Minority students (Figure 9), while the most statistically significant increase in performance from grades 10-12 was seen in White students (Figure 8).

5. Discussion

Analyzing 6 years of student performance data in Kansas public schools highlighted many disparities in student achievement. As detailed in the baseline comparisons above, significant changes were observed between all of our subgroup comparisons. These results reflect previous research on achievement gaps, namely that of low-income and minority students (Carnoy). This study also compared these subgroups based on district size. We suspect these increased gaps in medium and large school districts are due to the lack of one-on-one instruction in schools with a greater number of students. This leaves many students deprived of proper tutoring and access to learning resources to help them excel in the classroom. A significant impact of COVID-19 was also observed. We believe the drop off in performance seen in grades 3-8 is due to the lack of academic intervention during hybrid and online schooling. We think the increase in performance seen from grades 10-12 was most likely due to the majority of these students having a larger base level of knowledge than younger students. We think this gave them a stronger ability to retain mastery across the subjects during hybrid and online schooling. We cannot exclude the possibility of error arising from our varied sample sizes of subgroups and lack of a specific n-value. Although we believe these limitations have not affected the core outcomes of this research, it would be beneficial to obtain these controls for future work.

6. Conclusion

Understanding the shortcomings in the current public education system is critical to improving opportunities for students of all backgrounds. The performance trends statistically detailed here pave the way for educational administrators to find more ways to address achievement gaps in their school districts. Moreover, schools can provide further resources for these under-served students to help them achieve greater success. Moreover, this 6-year interval of focus (2016-2022) can be compared to past decades of student performance to analyze long-term trends in student achievement. Ultimately, these findings allow for further research focused on lessening achievement gaps in the State of Kansas and beyond.

Works Cited

Carnoy, Martin. Garcia, Emma. "Five key trends in U.S. student performance" *Policy Institute*, 12 Jan. 2017.

Carter, Ted. Garcia, Emma. "The Impact of Poverty and School Size on the 201516 Kansas State Assessment Results," *Kansas Association of School Boards*, Mar 2017.

Reardon, Sean. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations," *Stanford University*, 2011.