Final Research Paper

Grace Campos

Colorado State University Global

MIS581: Capstone – Business Intelligence and Data Analytics

Dr. Lisa Bryan

November 27, 2022

ABSTRACT

This work utilizes widely available academic data to show districts and schools how to build a tool that could support beginning of the year academic goal setting. In addition, the tool provides districts a more realistic projection of their end-of-year state test performance. In particular, this paper reviews diagnostic assessments, state test scores, and EVAAS projections to test commonly accepted beliefs in districts that drive outdated teaching and learning practices. Various tests were performed including one sample T-tests, ANOVA, and Kruskal-Wallis. Most importantly, this paper highlights the constant evolution of annual cut-score determination by grade level, and the importance of incorporating newly collected data. This practice not only broadens the dataset used in analytics, but also gives the district the opportunity to move away from national or state standards that are often not reflective of their unique demographics. Lastly, this work is meant to be a starting foundation for districts to use and show that the basic dataset requirements are already available or easily obtainable.

Keywords: cut-scores; student achievement; state test projection; data analytics

INTRODUCTION

Schools, private and public alike, are always seeking strategies to improve student competency and competitiveness. In the same vein, the state is also heavily invested to increase the potential of young people to earn a good living wage by the time they jump into the workforce. Ultimately, the goal is to educate students so that they become good citizens who are equipped with tools to become productive members of society. However, there are many steps that must be taken before any of it can happen. For instance, numerous factors contribute to the struggles many schools face in providing all the support a student requires before they graduate starting in elementary level. To better understand those factors, schools have adopted descriptive models. These models aim to better inform school personnel what is happening within the walls of schools and the educational culture as a whole. Realistically, these are not enough.

In addition, the state of Ohio has a “talent gap” problem. This means that the number of in-demand jobs that require postsecondary certification or higher are not being met by the working-age population. This is a serious problem, and if it continues, the state will lose ground resulting in long-term negative effects related to job creation and retention. In order to meet the needs of employers, Ohio must produce 1.3 million adults with highly qualified post-secondary certificates or degrees (ODHE, 2021). As one can imagine, secondary students who lag behind on academic standards do not help this case. Neither do underperforming elementary students. Therefore, finding a solution to the low achievement performance is dire. Though it may seem that the solution targets a specific age group, truthfully, it is not the case. Strategically, in order to attain this goal, one must step back and look at the achievement from elementary to grade 10 students.

OBJECTIVES

Plenty of studies have concluded that the most accurate predictor of student achievement is the extent to which the family is involved in his or her education (Henderson & Berla, 1994). Others claimed that students' academic resilience, regardless of their socio-economic background, is a strong factor that explains their academic success (Finn & Rock, 1997). The focus of this research is not to dispute or support these claims. Transparently, the researcher accepts and believes them. However, these are factors that are not within the locus of control of school administrators or teachers. The goal of this study is to provide districts and schools the tool that will support their academic goal setting without being inundated with indirect classroom variables.

To achieve this, the research will be divided into three groups or levels: kindergarten to 2nd grade, 3rd to 8th grade, and 9th to 10th grade. Different measures will be used including but not limited to diagnostic assessments, state tests, and Education Value-Added Assessment System (EVAAS) growth. Based on these findings, the tool will suggest achievement goals that are reflective of the district's student population. Furthermore, this tool will also be able to identify student subgroups that are struggling, who is performing well, and in which subjects. Thus, year after year, school achievement results will progressively get closer to the state's goal. On a related note, by projecting how students will perform on state assessments, the district is able to adjust the types of academic support teachers may need. In essence, there is a significant value in seeing results that are not favorable. In effect, knowing an undesired outcome ahead of time may lead to positive changes required to alter the future. Though the latter statement is outside of the scope of this research, it may open a new avenue for other studies to explore. To a point, interventions are only effective if they cater to the exact needs of the situation.

OVERVIEW OF STUDY

In a perfect world, barriers to attainment are removed and educational institutions, the community, and local businesses are willing to invest in the success of the students. Sadly, this is not the reality for struggling districts. Still, they are held to the same standards as well-funded schools. Nevertheless, the bottom line is that students must leave high school ready for college or careers. With this study, the researcher hopes to provide a simple yet impactful tool that schools can adjust and use based on their specific needs. By rethinking the way achievement goals are set, educators are not left feeling defeated which consequently results in students from all grade levels to simultaneously grow. There is no one size fits all, but understanding effectiveness of the available data can be the key to improve student academic performance.

RESEARCH QUESTIONS AND HYPOTHESES

The goal of this research is two-fold: define the effective cut scores and provide achievement projections. Additionally, it will provide the analytical backbone that other districts can use to perform similar studies based on their student demographics.

1. ***Increase academic achievement through effective grade-level cut scores using the historical datasets.***
    a. What is the maximum growth that is related to meeting typical end-of-year growth?
        i. H_0: Typical growth mean is less than or equal to 20 scale score points.
        ii. H_A: Typical growth mean is greater than 20 scale score points.

The current conception is that GMLSD students do not reach the average typical annual growth. By knowing the true average growth points, the district is able to set achievable cut scores and increase overall achievement performance.

    b. Does spending more time in EdTech programs aid in meeting typical end-of-year growth?
        i. H_0: Time spent makes no statistically significant difference to overall growth.
        ii. H_A: Time spent makes a statistically significant difference to overall growth.

Uncovering the time range that has a positive effect on growth may prove to be another avenue for educators to explore.

2. ***Project overall end-of-year performance using current school year's dataset.***
    a. Which groups are more likely to score proficient or higher on Spring state assessments?

i. H_0: New students are more likely to score proficient or higher.

ii. H_A: Students who have been with the district for at least two grade levels are more likely to score proficient or higher.

There is value in knowing if there is a "power of program" effect. On the other hand, if students are performing worse over the years, it might be worth revisiting the overall curriculum.

b. Do students below grade level grow?

i. H_0: Students below grade level do not show statistically significant growth in a single school year.

ii. H_A: Students below grade level do show statistically significant growth in a single school year.

If determined that below grade level students show significant growth, coupling this knowledge with Research Question 1 may produce very encouraging end-of-year proficiency performance.

LITERATURE REVIEW

Most schools focus on using the data to fulfill state and federal accountability rather than improving the overall teaching and learning practices. As a result, the benefits that guide data-informed decision-making seldom exist in districts based on their specific needs. Although plenty of research has shown that leveraging this resource can make it possible, it must be coupled with intentional strategies. A study conducted by Carlson et al. (1997) of 32 principals from various urban Chicago schools listed 13 strategies that led to the increase in student achievement. One of these strategies is setting clear goals and standards along with increased time on task. The data showed that students can flourish academically regardless of their economic background, race, and other challenges. The findings of this research support the author's claim that by establishing clear and achievable goals, both staff and students are set up for success. Specifically, a culture of achievement founded with high-performance standards provides educators the reflection and development needed to drive academic success.

There is a lot at stake with this information. The use of student achievement, especially state standardized tests, has become a primary indicator of schools and educators. It is not correct or fair, but unfortunately, it is a well-accepted behavior. This culture is hard to change so instead, efforts are better allocated to create better tools to alter the academic decline. Specifically, a tool that enables data to align, integrate, and support multiple components of the school system. The first step towards this effort, as Murray (2014) suggests, is to eliminate the confusion about how and when to use specific types of data. He also argues that data should be used to inform rather than drive decisions. This is very important to the author's research. For one, cut-scores are not meant to be set in stone; they are dynamic to reflect the systematic approaches of schools. Second, a sole focus on these numbers would not bring the desired compounding growth.

Academic analytical tools, from descriptive to prescriptive, are not new. There have been many efforts to even digitalize and analyze other variables that are believed to influence student outcomes which includes social media. Further, traditional business analytics has influenced many private EdTech companies to take advantage of the considerable benefits they bring. Still, and as Cazier et al. (2015) also points out, the benefits of analytics in education are still not fully realized. Although the scope of their thesis is beyond the goals of this research, it proves that much more needs to be done to capture valuable information to fully develop a predictive analytical tool. This is even more true if the goal includes a prescriptive purpose. On a more measurable note, the ability to identify students' strengths and weaknesses can be enough to build a predictive model along with key performance indicators.

As a final note, there is no arguing that data analytics has the potential to help at-risk students. However, there must be equal efforts to propel well-performing students, as well. This is why, though it may sound controversial, it is inevitable that students are divided into subgroups. For instance, as Finn &

Rock (1997) point out, pre-identified groups allow the accurate measurement in the variances of each test measurement. Also, grouping highlights unpopular facts. Namely, Black students are more transient than others which have been statistically proven to increase the likelihood of disengagement and drop-outs. Another predictor is class-size which, for under-funded districts, tends to be in the upper range. Afterall, it has been shown that schools with smaller class sizes encourage more student engagement which fosters positive behavior (Finn & Rock, 1997). This research, along with the three mentioned earlier, show that although studies of educational institutions are available, there is plenty more room to identify solutions that improve student achievement. Perhaps, instead of looking at things more broadly, research should be dialed down to individual districts and identify those unique needs.

RESEARCH DESIGN

Methodology

The research will be divided into three groups or levels: kindergarten to 2nd grade, 3rd to 8th grade, and 9th to 10th grade. Different quantitative measures will be used including but not limited to diagnostic assessments, state tests, and Education Value-Added Assessment System (EVAAS) growth. Due to its non-experimental design, the research will not perform complicated statistical analyses or blind studies. Additionally, the data required are readily available as all public schools must collect them to be compliant with the Educational Management Information System (EMIS). Essentially, this research will provide the needed navigation system that will allow districts to create SMART goals and harness the power of their own data. Note that it is anticipated that multiple cycles of data exploration and clean-up will occur during analysis.

Methods

In order to answer the first research goal, a combination of a one-sample T-test and Wilcoxon test will be used. By comparing the observed mean with the known or hypothesized value, the district can certainly determine if there is enough evidence to dispute the unscientifically held idea. Based on the results, suggestions of typical growth adjustment may be required. For research question 2, the Wilcoxon test compares two paired groups and determines if they are different from each other in a statistically significant manner (Elliott & Woodward, 2015). Specifically, the Wilcoxon test will highlight if minutes spent using curated lessons from EdTech programs work better.

For the second research goal, a Kruskal-Wallis test will be used. The Kruskal-Wallis test is an extension of the Wilcoxon test which is applicable to two or more independent samples (Elliott & Woodward, 2015). By comparing the independent variable, *group_id*, based on previous and current Reading and Math percentiles, the district will know which group is most likely to improve. Lastly, an analysis of variance (ANOVA) test effectively compares three or more means across groups or compares three or more dependent samples (SAS Institute, 2017). Since only a specific group will be tested here, comparing their means side-by-side provides insight on predicting end-of-year performance. In other words, students who perform poorly on diagnostic assessments, when compared to their peers of the same level, may show a tertiary-subgroup that are likely to be reachable without inserting too much effort.

Limitations

Although it may seem that the effects of the pandemic have passed, the disruption it has caused still looms greatly. For instance, remote learning and the cancellation of state assessments has resulted in non-continuous data. This is not unique to this project. Additionally, school districts often change their diagnostic programs, which in this case, resulted in obtaining only two years' worth of data. One of which is during a school year that was still affected by remote learning. In a different light, since the data used are purely academic and noting the results in literature reviews, there are aspects of the results that could be easily challenged. For example, one could argue that it is more important to include students' social and emotional learning data than minutes on a platform. However, due to the limitations of the information available, the academic variable is preferred.

Ethical and Legal Considerations

In 1974, Family Educational Rights and Privacy Act (FERPA) was enacted and its main purpose was to oversee the privacy, discharge, and accuracy of educational records and to ensure that administrators and instructors protect student information at all costs (Toglia, 2007). In the K-12 public school sector, FERPA has a significant importance to both inside and outside of the classroom as it provides explicit regulations regarding the release of students' educational records. Documents containing these records can include a student's date and place of birth, parental information, grades, test scores, courses taken, demographics, disciplinary records, attendance, medical records, and so on. Moreover, these personally identifiable information (PII) can exist in both physical and electronic forms. Regardless of the format, the release and access of these records must be approved by the parent or guardian of the student. In this instance, because the data consists of test records, permission was acquired through the district. A legal document highlighting misuse and unauthorized release of PII was signed by the researcher.

In addition to the legal ramifications, ethical policies must also be followed. Gajjar (2013) suggests the following five core principles for research ethics: cite intellectual property properly, be conscious of multiple roles, follow informed-consent rules, respect confidentiality and privacy, and consider other ethical obligations from other researchers. Specifically in this research, since there will be a comparison to the existing cut scores provided by EdTech programs, proper citing will be included. Moreover, regardless of the results, the employability and reputations of educators must not be compromised. In other words, the factors considered in this research will not be directly tied to a specific individual. In regards to students, there will be reasonable efforts to avoid deducing that groups should be treated inconsistently. For example, although some students may require more academic focus and one-on-one tutoring, these efforts should not take away from those who also perform well. Equal treatment must be provided though they may look differently.

TRANSITIONAL CONCLUSION

In this first draft, introduction, objectives, research question and hypotheses, literature reviews, and research design were provided. So far, the foundation of the study has been theoretical and with only the support of previous studies. However, the major point of this study is yet to be revealed. While the data is poised for statistical analyses, no actual testing has begun. Therefore, the author was cautious to make a strong claim on any of the sections mentioned above. By virtue, the study is not tainted by known information and the reader should feel confident that the study has not been altered to skew towards the author's claims. This study is meant to be educational to the reader as much as it is for the researcher. With that said, should the null hypothesis be rejected, the author suggests additional tests instead of accepting the alternative hypothesis. Standards goal setting, especially in education, is complicated and should not be deduced to a single person's conclusion. In other words, there is nothing more detrimental to growth than being complacent. Therefore, as the results come in, the report shall evolve only to communicate those findings and nothing more. Suggestions for future studies and/or adjustments to the study may be included.

FINDINGS

Recall that the purpose of this research was to build a tool that will help districts calculate effective cut scores and provide projections. Therefore, the statistical results are based on the entire student population as many districts do not set different goals per grade level based on subgroups. However, the Tableau dashboard will allow the user to navigate through the different cut scores for the subgroups, if desired. With that said, the first step is to determine the typical annual growth per grade level. Since many administrators and teachers believe, and perpetuated by iReady trainers, that students who grow 20 points from Fall to Spring administration would be on-track to proficiency by the time they take the Spring administration, the researcher wanted to test this theory. Table 1 shows the t-test results with Figure 1 as supporting evidence. Evidently, younger students, specifically grades kindergarten through 4th grade, have the highest observable growth. However, starting 5th grade, the growth tapers off to less than 20 points. Obviously, this is concerning and should be addressed but knowing that younger students have the ability to grow rapidly is very encouraging. Consequently, since there seems to be a band where the null hypothesis is true and where it is false, accepting it is not straightforward. For grades kindergarten through 4th, the null hypothesis should be rejected and noted that students show much higher growth points. On the other hand, 5th through 8th seem to accept the null hypothesis but should note why that is. The researcher recommends further analyses on this specific grade band.

Table 1

One-sample t-test results for Math and Reading by grade level

Math Diagnostic	**Typical Growth**	**Frequency/ Peaks**	**95% CL Mean**	**t Value**	**Pr > |t|**
Grade K	31.8	87.2%	30.7 - 30.9	156.39	<.0001
Grade 1	28.8	71.4%	30.3 - 30.6	139.07	<.0001
Grade 2	25.8 28.8	41.9% 53.1%	27.2 - 27.4	151.32	<.0001
Grade 3	26 27	36.4% 36.1%	27.1 - 27.3	189.53	<.0001
Grade 4	23.1 24	66.7% 31.8%	23.2 - 23.3	200.83	<.0001
Grade 5	18.2 19.9	65.1% 32.7%	18.5 - 18.6	-52.63	<.0001
Grade 6	14 15	47.9% 42.5%	14.3 - 14.4	-381.77	<.0001
Grade 7	12 12.9	31.2% 66.4%	12.6 - 12.7	-633.46	<.0001
Grade 8	9 12	33.6% 54.8%	10.7 - 10.8	-292.09	<.0001
Reading Diagnostic	**Typical Growth**	**Frequency/ Peaks**	**95% CL Mean**	**t Value**	**Pr > |t|**
Grade K	49.2	80.7%	47.9 - 48.1	611.64	<.0001

Grade 1	49	76.2%	48.8 - 49.1	401.08	<.0001
Grade 2	39.8 44.3	40.4% 43.7%	38.8 - 39.4	132.12	<.0001
Grade 3	26 33	19.7% 35%	28.6 - 29.2	64.61	<.0001
Grade 4	20 28	38.2% 29.8%	21.8 - 22.2	18.70	<.0001
Grade 5	16 20	29.3% 28.2%	18.3 - 18.8	-11.88	<.0001
Grade 6	12.4 18.8	23.9% 40.8%	14.1 - 14.5	-53.26	<.0001
Grade 7	16.8	54.2%	13.0 - 13.5	-64.80	<.0001
Grade 8	18	57.3%	13.2 - 13.7	-51.41	<.0001

Figure 1

Summary statistics of Math diagnostic growth by grade level

Subject=Math

Student_Grade	N Obs	Variable	Mean	Std Dev	Median	Std Error	Mode	Skewness	Kurtosis
0	2043	Typical_Growth Stretch_Growth	30.8203622 38.7121880	3.1273237 0.9174461	32.0000000 39.0000000	0.0691893 0.0202977	32.0000000 39.0000000	-2.3655228 -3.5213196	3.8600829 11.2611877
1	1914	Typical_Growth Stretch_Growth	30.4555904 41.7429467	3.2892231 8.6124531	29.0000000 37.0000000	0.0751835 0.1968593	29.0000000 37.0000000	0.6533830 1.1840978	0.0714197 -0.5489143
2	1890	Typical_Growth Stretch_Growth	27.3402116 42.2677249	2.1087737 6.1359458	29.0000000 48.0000000	0.0485064 0.1411402	29.0000000 48.0000000	-1.6299657 -0.1552003	4.0688476 -1.9033097
3	1927	Typical_Growth Stretch_Growth	27.2060197 42.4042553	1.6690342 7.7640465	27.0000000 43.0000000	0.0380211 0.1768672	26.0000000 35.0000000	0.5021709 0.5880628	0.3162500 -1.0102998
4	1807	Typical_Growth Stretch_Growth	23.2606530 39.7548423	0.6901588 5.8965986	23.0000000 41.0000000	0.0162357 0.1387147	23.0000000 34.0000000	-2.9638248 -0.0846982	17.8785722 -1.1512541
5	1813	Typical_Growth Stretch_Growth	18.5626034 34.7357970	1.1628314 5.0223807	18.0000000 35.0000000	0.0273098 0.1179535	18.0000000 31.0000000	-0.8244074 -0.1971995	3.1769541 -0.2910509
6	1870	Typical_Growth Stretch_Growth	14.3294118 30.4941176	0.6423115 4.3148824	14.0000000 30.0000000	0.0148534 0.0997811	14.0000000 35.0000000	-0.4302299 -0.2732495	-0.7014026 -1.2415683
7	1902	Typical_Growth Stretch_Growth	12.6498423 28.3664564	0.5060367 4.8872387	13.0000000 33.0000000	0.0116032 0.1120620	13.0000000 33.0000000	-0.9548895 -0.1651344	-0.3637570 -1.8584295
8	1960	Typical_Growth Stretch_Growth	10.7653061 26.9489796	1.3997084 4.5055067	12.0000000 31.0000000	0.0316162 0.1017690	12.0000000 31.0000000	-0.3108767 -0.2407601	-1.7974378 -1.8700375

Figure 2

Summary statistics of Reading diagnostic growth by grade level

Subject=Reading

Student_Grade	N Obs	Variable	Mean	Std Dev	Median	Std Error	Mode	Skewness	Kurtosis
0	2044	Typical_Growth	47.9995108	2.0696392	49.0000000	0.0457777	49.0000000	-1.5794543	0.6268431
		Stretch_Growth	66.1516634	2.7614386	67.0000000	0.0610794	67.0000000	-3.0669750	19.9971918
1	1963	Typical_Growth	48.9704534	3.2002713	49.0000000	0.0722315	49.0000000	-1.9241536	6.8445574
		Stretch_Growth	69.2175242	11.7327555	67.0000000	0.2648133	67.0000000	1.0303042	1.6546593
2	1932	Typical_Growth	39.1123188	6.3583081	39.0000000	0.1446565	44.0000000	-1.5590977	1.5276444
		Stretch_Growth	62.5569358	17.6053276	53.0000000	0.4005351	81.0000000	-0.2766089	-1.0950806
3	1970	Typical_Growth	28.8918782	6.1085479	33.0000000	0.1376274	33.0000000	-0.4785147	-1.1015322
		Stretch_Growth	53.7263959	17.7615480	63.0000000	0.4001729	63.0000000	-0.1087053	-1.0956816
4	1866	Typical_Growth	22.0219721	4.6714063	20.0000000	0.1081414	20.0000000	-0.2199932	-0.5511773
		Stretch_Growth	44.0235798	13.9297156	36.0000000	0.3224679	36.0000000	0.0511398	-1.2916928
5	1866	Typical_Growth	18.5739550	5.1864070	20.0000000	0.1200634	16.0000000	-0.1091688	-0.4788768
		Stretch_Growth	40.7202572	14.2431579	47.0000000	0.3297239	30.0000000	0.1729636	-1.3803384
6	1887	Typical_Growth	14.3184950	4.6335253	14.0000000	0.1066659	19.0000000	-0.6433204	-0.3941928
		Stretch_Growth	37.4610493	12.7184806	38.0000000	0.2927853	51.0000000	-0.2115388	-1.4388138
7	1880	Typical_Growth	13.2537234	4.5143007	17.0000000	0.1041146	17.0000000	-0.7211151	-0.9139055
		Stretch_Growth	39.1861702	12.7851789	50.0000000	0.2948682	50.0000000	-0.5759878	-1.2900007
8	1985	Typical_Growth	13.4891688	5.6423128	18.0000000	0.1266418	18.0000000	-0.6791289	-1.2001187
		Stretch_Growth	39.1168766	13.4676076	50.0000000	0.3022805	50.0000000	-0.6257194	-1.2729079

The second hypothesis tests if spending more time with EdTech programs lead to higher growth. Figure 3 shows the total minutes spent based on subject and meeting or not meeting typical growth. Overall, students spent nearly thrice as much in Math but still did not meet their annual typical growth. A similar case in Reading where students who spent nearly twice on the platform still did not meet their goals. In contrast, students who met their annual typical growth did not spend much time on the platform.

Figure 3

Total minutes spent in a year by subject

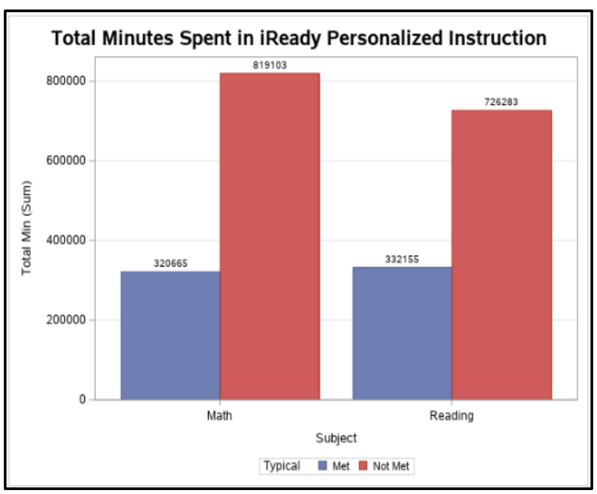


Figure 4

Kruskal-Wallis test of total minutes based on group - met or not met the annual growth goal

Wilcoxon Scores (Rank Sums) for Variable TotalMin
Classified by Variable Typical

Typical	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Not Met	5447	21134634.5	21167042.0	90546.4901	3880.05039
Met	2324	9063471.5	9031064.0	90546.4901	3899.94471

Average scores were used for ties.

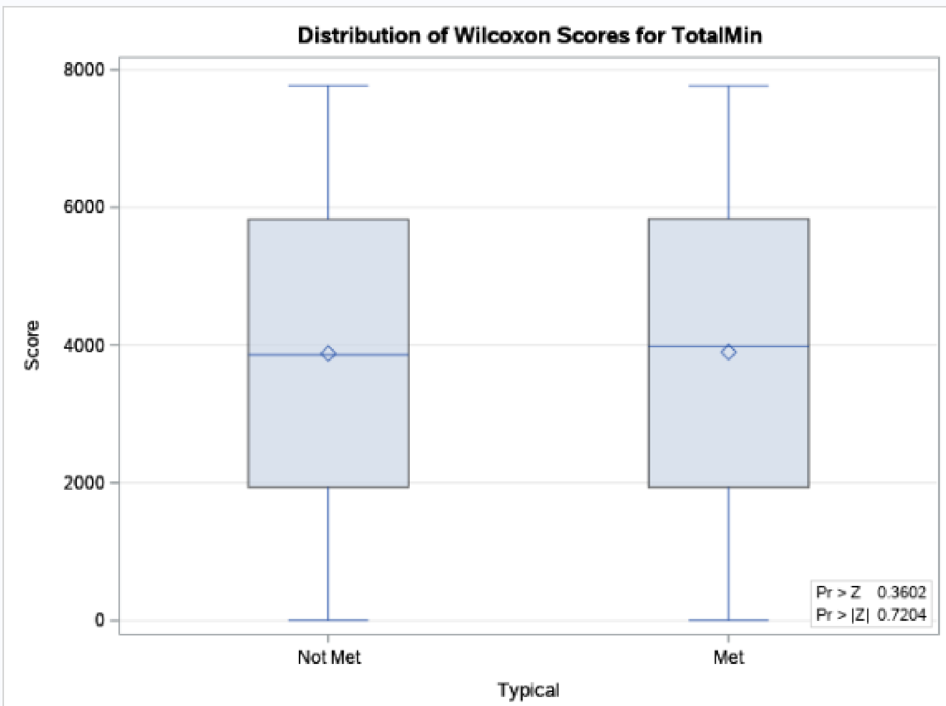
Wilcoxon Two-Sample Test

Statistic	Z	Pr > Z	Pr > |Z|	t Approximation Pr > Z	t Approximation Pr > |Z|
9063472	0.3579	0.3602	0.7204	0.3602	0.7204

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	DF	Pr > ChiSq
0.1281	1	0.7204

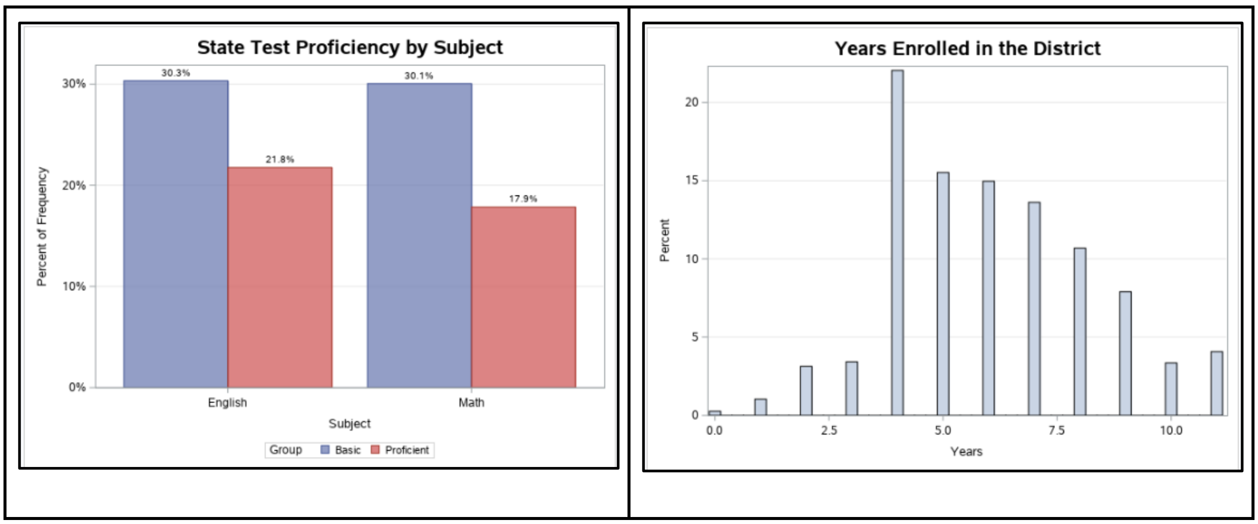


Based on the results shown in Figure 4, there seems to be significant evidence to accept the null hypothesis. In other words, spending time on the platform (suggested 45-60 minutes per week) does not lead to better achievement performance. Perhaps the money is better spent on hiring an additional intervention specialist (IS) so that they can conduct small group in-person sessions with students who are not meeting their growth goals.

The third hypothesis is based on projected performance at the end of the year. Figure 5 shows the state test proficiency by subject for all tested grades and the distribution of years the students have been in the district. This shows that students are likely to stay in the district. It also shows that the district proficiency in English and Math are lower than expected. The purpose of showing these graphs is to give the reader the broader view, in separate buckets, that 1) students are underperforming, but 2) the district is not experiencing a high transition. The latter is very important as most urban and underperforming districts tend to have high attrition rates with the byproduct of environmental inconsistency for students.

Figure 5

Five-year state test results and respective years students have been in the district



To determine if new students are more likely to score proficient on the state test, Kruskal-Wallis test was performed. Figures 6.a and 6.b show there does not appear to be a significant difference between students who have been with the district long enough and those who are new. This could be accepted in a good or bad way; it is good that the district is not 'failing' the students by letting them lag behind, and it is bad that the district is not pulling them up either. Simply put, neither group is likely to perform better on the state assessment. In summary, it is better to reject the null hypothesis but the alternative should not be accepted either. Perhaps a smaller research, studied in consecutive years, should be performed separately.

Figure 6.a

Kruskal-Wallis test to determine if years affect state performance

Wilcoxon Scores (Rank Sums) for Variable Years Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Basic	24906	510635096	513686250	1169661.82	20502.4932
Proficient	16343	340125530	337074375	1169661.82	20811.6949
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr > Z	Pr > |Z|	Pr > Z	Pr > |Z|
3.4013E8	2.6086	0.0045	0.0091	0.0045	0.0091
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
6.8047	1	0.0091

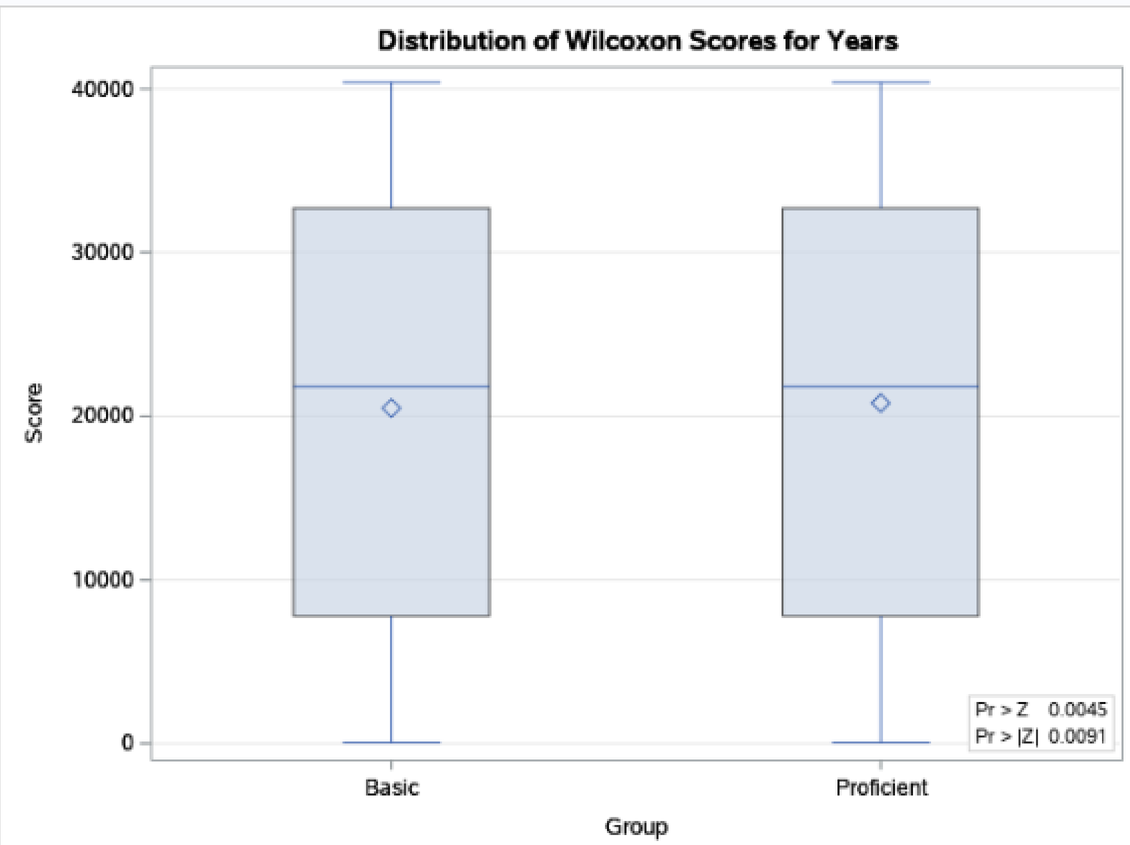


Figure 6.b

Continued - Kruskal-Wallis test to determine if years affect state performance

Median Scores (Number of Points Above Median) for Variable Years
Classified by Variable Group

Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Basic	24906	12498.7359	12452.6981	46.400753	0.501836
Proficient	16343	8125.2641	8171.3019	46.400753	0.497171

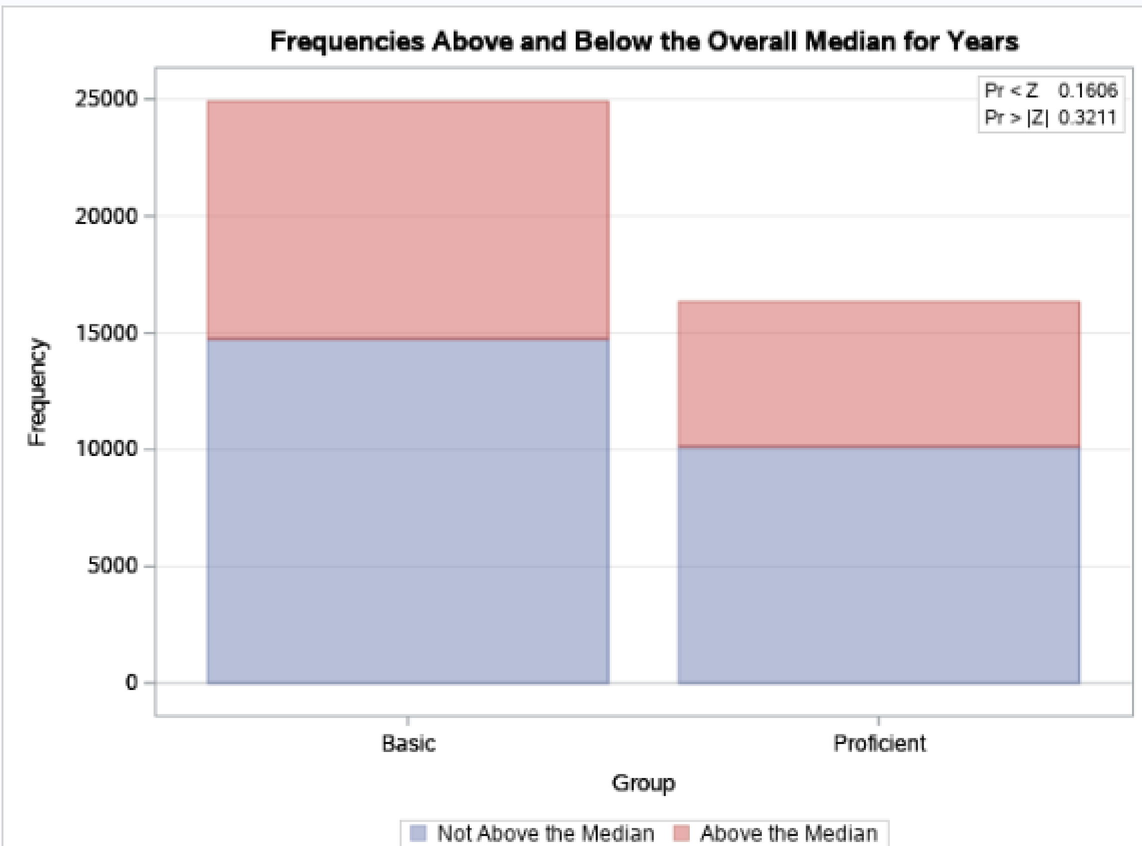
Average scores were used for ties.

Median Two-Sample Test

Statistic	Z	Pr < Z	Pr > |Z|
8125.264	-0.9922	0.1606	0.3211

Median One-Way Analysis

Chi-Square	DF	Pr > ChiSq
0.9844	1	0.3211



The last hypothesis tests if students below grade level (performance level 1 and 2) grow significantly compared to their counterparts (performance levels 3+). Based on the ANOVA test, only

Group 2 showed significant growth compared to the proficient group (see Figure 7). This was both encouraging and discouraging. It was encouraging that there was a high chance that Group 2 grew enough to move to Group 3 but Group 1 was so far behind that the educational gap seemed to be worsening. As such, similar to the first hypothesis, the answer to this question was not definitive for the entire group. Specifically, for Group 1, students did not show significant growth and therefore the null hypothesis failed to be rejected. However, Group 2 showed evidence that there was growth and therefore the null hypothesis was rejected. The author suggests additional studies with other variables to define what led the groups to have different conclusions.

Figure 7

Comparing actual_growth across performance level using the ANOVA test

ANOVA Test run

The ANOVA Procedure

Dependent Variable: Actual_Growth

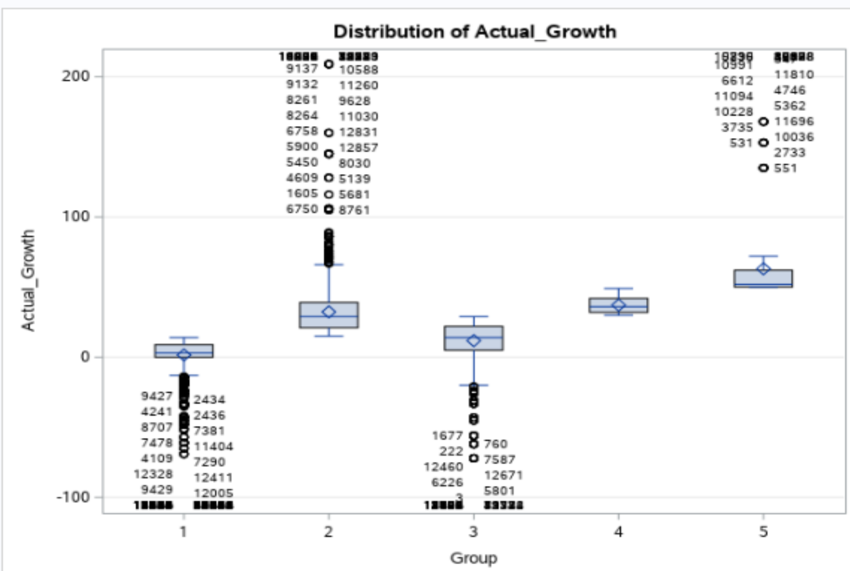
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2854678.410	713669.603	3318.71	<.0001
Error	13107	2818584.214	215.044		
Corrected Total	13111	5673262.625			

R-Square	Coeff Var	Root MSE	Actual_Growth Mean
0.503181	70.55530	14.66439	20.78424

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Group	4	2854678.410	713669.603	3318.71	<.0001

ANOVA Test run

The ANOVA Procedure



CONCLUSION

Evident by the results above, accepting or rejecting the null hypotheses are not always straightforward. There are instances where it is more likely that the null hypothesis is true without contest while others depend on the circumstances. Although this may seem inconclusive, this only proves that it is not easy setting up guidelines for educational success, so much more if it deals with statistics. On the other hand, a continued effort to better understand student performance is necessary. By constantly creating hypotheses and testing them, the district is much more likely to find what works and what does not within the walls of their schools. In this example, there was clear evidence that spending more screen time is not helpful and high attrition does not contribute to the district's low performance. Therefore, minds should shift towards the variables that do.

As a final thought and to tie-up goals 1 and 2, the following Tables show proposed cut scores for determining students who are on-track to score proficiency on Spring assessments. With each administration, the district will be able to project the percentage of students who are on-track to score proficient or higher on the state assessment. Note that grades K-2 do not take state assessments, however, since they are foundational grade levels, it is better to include them and set their respective cut scores early to establish continuity. The cut scores are based on the Typical Growth results in Table 1 and may be compared to iReady's Diagnostic Placement chart (see Figure 7 in Appendix). Note that 3rd Grade must achieve a ≥510 scale score before the year ends, or they may not be promoted based on state law (ODE, 2022).

Table 2

Increase achievement with effective grade-level cut scores based on diagnostic results

On-Track in Reading	Fall	Winter	Spring
Kindergarten	≥346	≥370	≥394
1st Grade	≥394	≥419	≥443

2nd Grade	≥443	≥462	≥481
3rd Grade	≥481	≥495	≥510
4th Grade	≥510	≥520	≥530
5th Grade	≥530	≥538	≥546
6th Grade	≥546	≥556	≥565
7th Grade	≥565	≥574	≥582
8th Grade	≥582	≥591	≥600
On-Track in Math	**Fall**	**Winter**	**Spring**
Kindergarten	≥336	≥352	≥368
1st Grade	≥368	≥383	≥397
2nd Grade	≥397	≥412	≥426
3rd Grade	≥426	≥439	≥452
4th Grade	≥452	≥464	≥475
5th Grade	≥475	≥484	≥493
6th Grade	≥493	≥500	≥507
7th Grade	≥507	≥514	≥520
8th Grade	≥520	≥526	≥532

Next, the researcher proposes the following cut score ("22-23 Goal Projection") instead of the 22-23 Goal. It is not advisable to change district goals mid-year, and for this reason, the proposal will be hidden until the end of year when the actual results from Spring 2023 are available. This will stand as a test if the analysis of this research is sound. Lastly, the 22-23 Projection is related to the cut scores above which reflects how students are currently performing this school year. However, since students tend to perform the worst on Fall administrations, the researcher plans to recalculate this proposal after Winter administration.

Table 3

Project overall end-of-year performance

	Subject	**18-19 Baseline**	**20-21 Results**	**21-22 Actual**	**22-23 Goal (Current)**	***22-23 Goal Projection***
Grade 3	Reading	52%	23%	45%	67%	*42%*
	Math	60%	27%	40%	60%	*28%*
Grade 4	Reading	60%	41%	54%	67%	*50%*
	Math	63%	35%	45%	63%	*20%*
Grade 5	Reading	63%	48%	56%	64%	*52%*
	Math	49%	20%	35%	50%	*12%*
Grade 6	Reading	39%	33%	45%	57%	*58%*
	Math	46%	20%	25%	46%	*13%*
Grade 7	Reading	56%	38%	48%	58%	*52%*
	Math	46%	14%	17%	46%	*9%*
Grade 8	Reading	34%	29%	36%	42%	*43%*
	Math	39%	23%	23%	39%	*7%*
Grade 9/10	Algebra I	27%	20%	20%	27%	*10%*
	English II	35%	44%	39%	44%	*44%*

As a final note for this section, some of the data used are limited to a single year because of the effects of remote learning and state test cancellations. Therefore, as the year goes by, the author plans to revisit this study and include additional data to strengthen its findings. In addition, at the end of the year, the author plans to revisit the cut scores and projections proposed here. By comparing the actual achievement growth and results, the author will know if the calculations are on the right track. This is an evolving project and the author plans to recalibrate the tool as needed especially when the projection seems to be way lower than the current goals. Perhaps there will be a downward trend that administrators are not willing to look at, and it is shown in the data. Time will tell.

RECOMMENDATIONS

The steps taken to produce the desired results in this research are considered standard in the analytics point of view. However, keep in mind that the target audience here are those who are not statisticians or have the funding to use the tools provided. For instance, SAS is not a free program and neither is Tableau. Tableau Public is not recommended as it does not have the right guardrails to protect sensitive data. Therefore, the author recommends using more accessible tools such as Sheets (Google) or Excel. Another option is R but this requires the user to know the R language. Regardless of the method chosen, the steps taken here should build the foundation to create effective cut scores and project EOY performance. On the analytics end, the author recommends revisiting the ANOVA test. There seems to be plenty of outliers though they were not present in other datasets. Another is diving deeper into the scale score distribution (see Figures 9 and 10 in the Appendix). In most cases, symmetric distributions are desired but, in this case, it is much more preferable to have a left skewed distribution as it indicates that more students are scoring proficiently.

REFERENCES

Carlson, K. G., Shagle-Shah, S., & Ramirez, M. D. (1999). Leave no child behind: a baker's dozen strategies to increase academic achievement. *Chicago Schools Academic Accountability Council*. https://eric.ed.gov/?id=ED436615

Cazier, J., Cech, T., & Spaulding, T. (2015). *Applying business analytic methods to improve organizational performance in the public school system*. CORE. https://core.ac.uk/download/pdf/301365696.pdf

Cody, R. (2019). *A Gentle Introduction to Statistics Using SAS Studio.* Cary, NC: SAS Institute Inc. https://platform.virdocs.com/r/s/0/doc/1316035/sp/178880722

Curriculum Assessments. (2022). *Using iReady Diagnostic for Ohio's Third Grade Reading Guarantee 2022-2023.* https://i-readycentral.com/articles/i-ready-diagnostic-approved-by-the-ohio-department-of-education-doe-for-the-2022-2023-school-year/

Elliott, A. C., & Woodward, W. A. (2015). *SAS Essentials: Mastering SAS for Data Analytics* (2nd ed.). John Wiley & Sons. ISBN-13: 978-1119042167

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of applied psychology*, *82*(2), 221.

Gajjar, D. (2013). Ethical consideration in research. *Education*, *2*(7), 8-15.

Henderson, A. T., & Berla, N. (1994). *A new generation of evidence: The family is critical to student achievement.* ERIC. https://files.eric.ed.gov/fulltext/ED375968.pdf

Kelleher, C., & Wagener, T. (2011). Ten guidelines for effective data visualization in scientific publications. *Environmental Modeling & Software*, *26*(6), 822-827.

Marr, B. (2015). *Big Data: Using SMART Big Data, Analytics, and Metrics to Make Better Decisions and Improve Performance* (1st ed.). Wiley. ISBN-13: 978-1118965832

Murray, J. (2014). Critical issues facing school leaders concerning data-informed decision-making. *Professional Educator*, *38*(1), 14-22.

Ohio Department of Education (ODE). (2022, October 14). *Parent Resource Toolkit*. Testing. https://education.ohio.gov/Topics/Testing/Testing-Forms-Rules-and-Committees/Parent-Resource-Toolkit#OST

Ohio Department of Higher Education (ODHE). (2021). *Ohio Attainment Annual Report.* https://www.ohiohighered.org/sites/default/files/uploads/attainment/Attainment-Report_2021_FINAL.pdf

SAS. (2022). *SAS EVAAS for K-12*. https://www.sas.com/en_us/software/evaas.html#:~:text=Our%20Education%20Visualization%20and%20Analytics,students'%20future%20needs%20and%20goals.

SAS Institute (2017). *Advanced predictive modeling using SAS Enterprise Miner*. Cary, NC: SAS Institute Inc. ISBN-13: 978-1-63526-115-8

Shaffer, J. (2016, April 20). 5 tips on designing colorblind-friendly visualizations. *Tableau*. https://www.tableau.com/about/blog/examining-data-viz-rules-dont-use-red-green-together

Toglia, T. V. (2007). How does FERPA affect you? *Tech Directions, 67*(2).

Appendix

Figure 8

iReady Reading and Math Diagnostic Placement

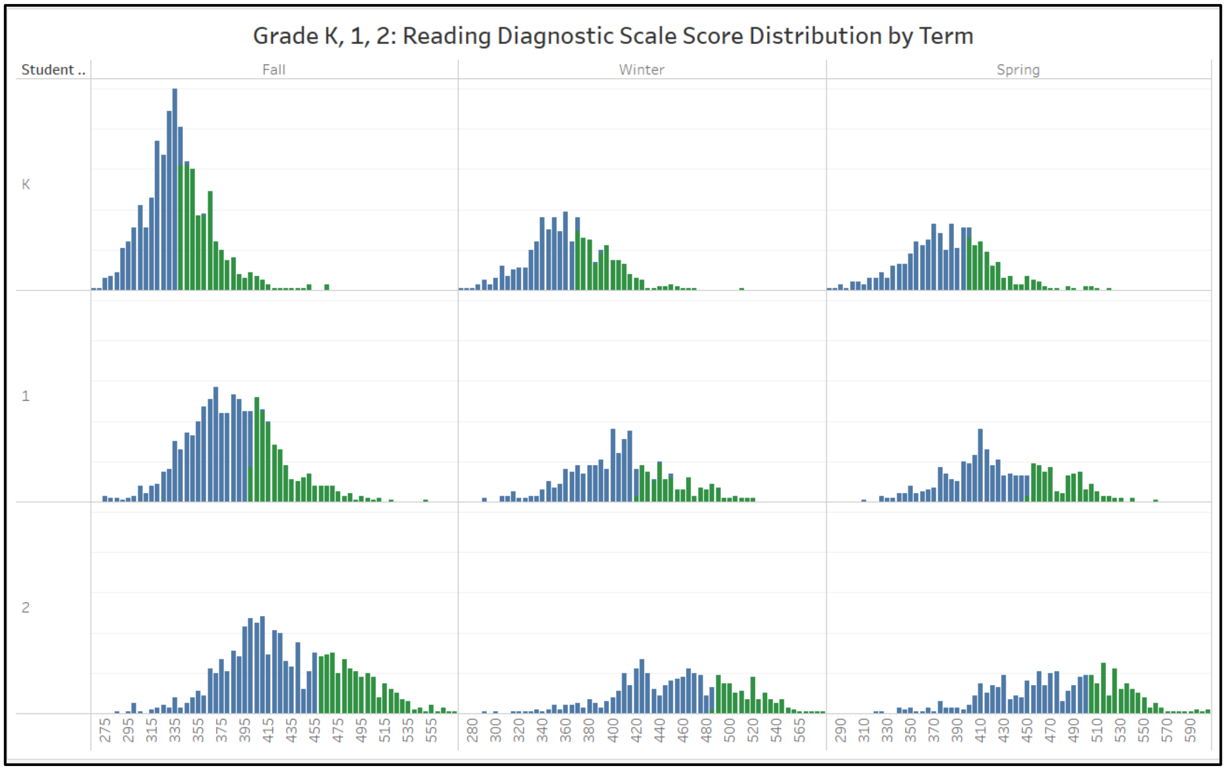
		iREADY Reading Diagnostic Placement		
On-Grade Ranges		Fall 22-23	Winter 22-23	Spring 22-23
Grade K	Below	≤346	≤370	≤394
	At	347 - 366	371 - 403	395 - 423
	Above	≥367	≥404	≥424
Grade 1	Below	≤402	≤423	≤453
	At	403 -428	424 - 470	454 - 496
	Above	≥429	≥471	≥497
Grade 2	Below	≤459	≤488	≤504
	At	460 - 504	489 - 527	505 - 540
	Above	≥505	≥528	≥541
Grade 3	Below	≤509	≤521	≤533
	At	510 - 540	522 - 557	534 -570
	Above	≥541	≥558	≥571
Grade 4	Below	≤532	≤547	≤557
	At	533–570	548–584	558–594
	Above	≥571	≥585	≥595
Grade 5	Below	≤556	≤568	≤577
	At	557–595	569–608	578–618
	Above	≥596	≥609	≥619

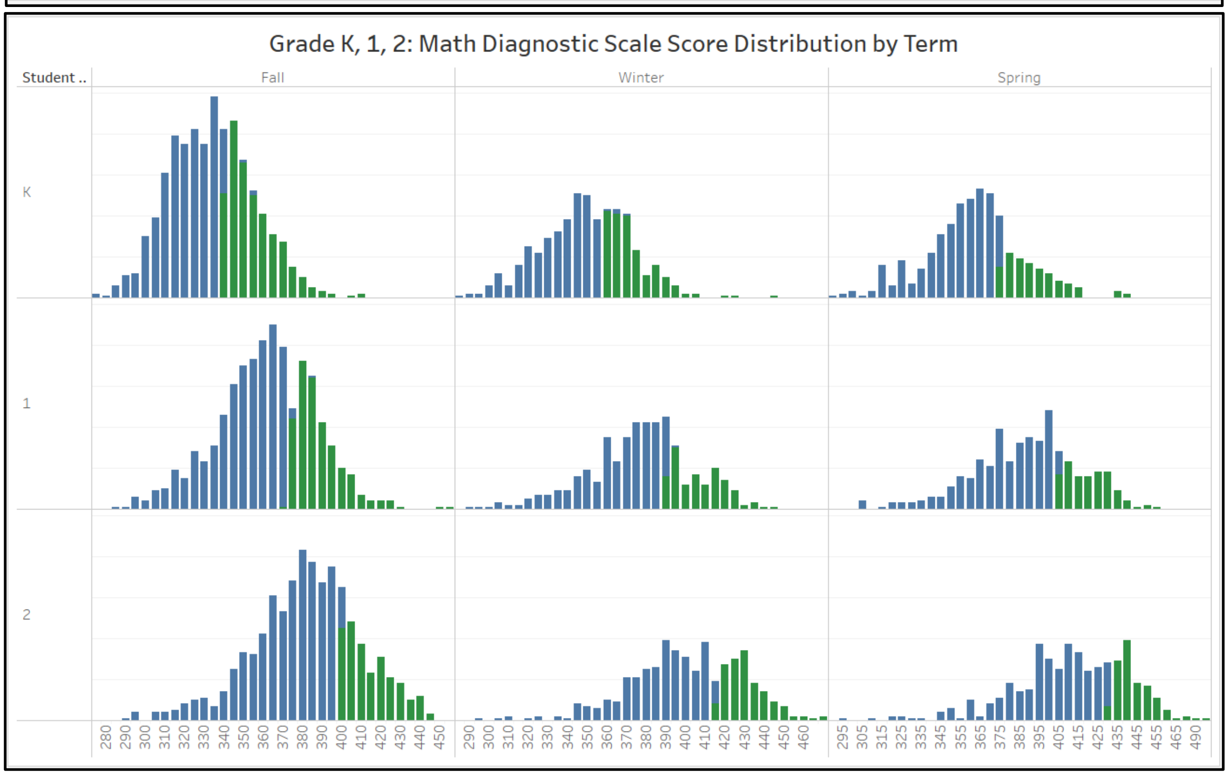
		iREADY Math Diagnostic Placement		
On-Grade Ranges		Fall 22-23	Winter 22-23	Spring 22-23
Grade K	Below	≤341	≤360	≤377
	At	342–360	361–379	378–395
	Above	≥361	≥380	≥396
Grade 1	Below	≤375	≤392	≤407
	At	376–392	393–412	408–429
	Above	≥393	≥413	≥430
Grade 2	Below	≤402	≤417	≤433
	At	403–422	418–441	434–454
	Above	≥423	≥442	≥455
Grade 3	Below	≤427	≤444	≤458
	At	428–447	445–461	459–479
	Above	≥448	≥462	≥480
Grade 4	Below	≤451	≤465	≤487
	At	452–473	466–485	488–501
	Above	≥474	≥486	≥502
Grade 5	Below	≤470	≤480	≤490
	At	471–491	481–501	491–513
	Above	≥492	≥502	≥514

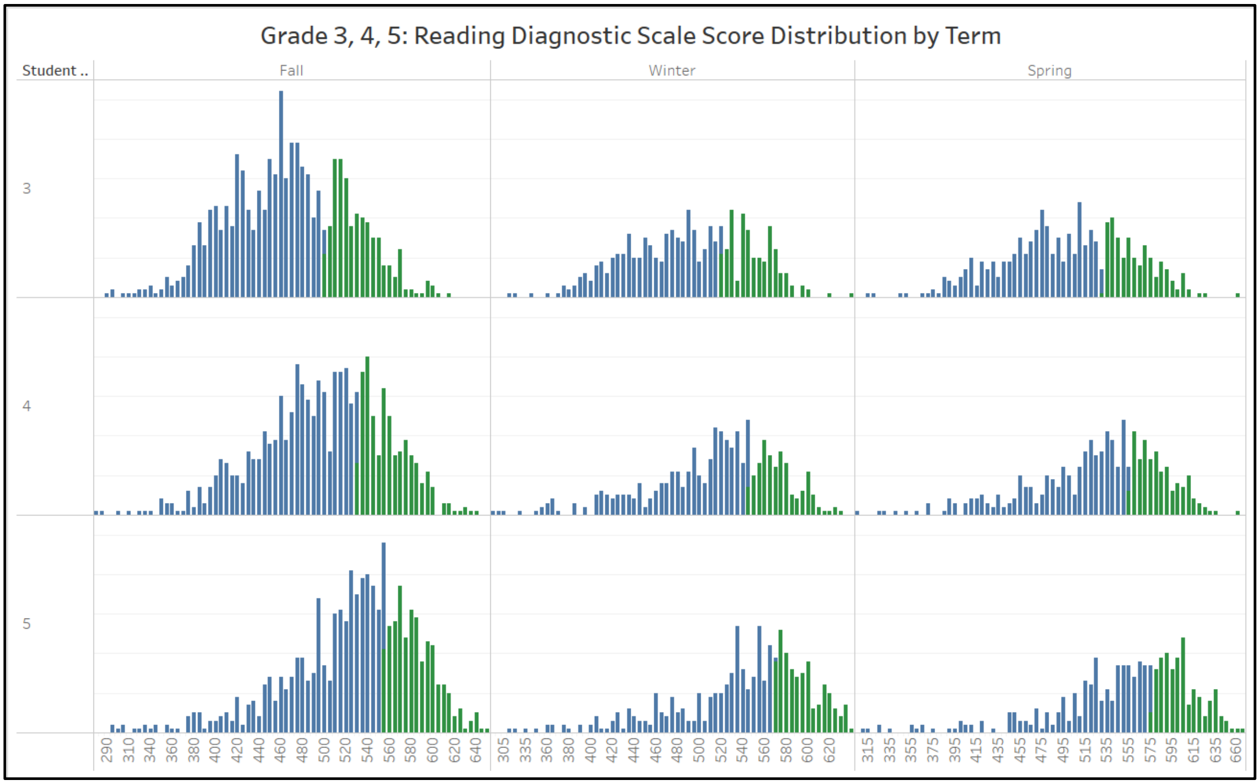
Note. *iReady Diagnostic Placement. From Curriculum Associates Guidance Brief, by Curriculum Associates. Copyright 2022 by Curriculum Associates.*

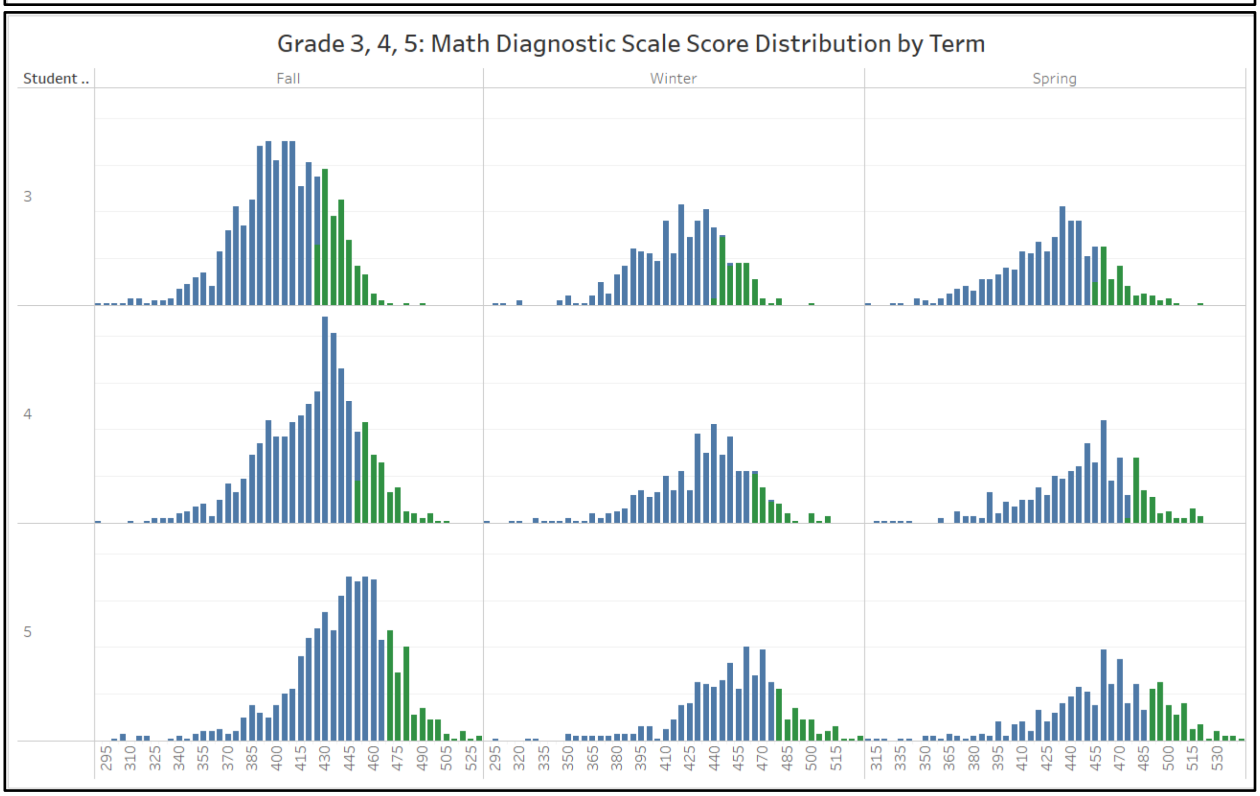
Figure 9

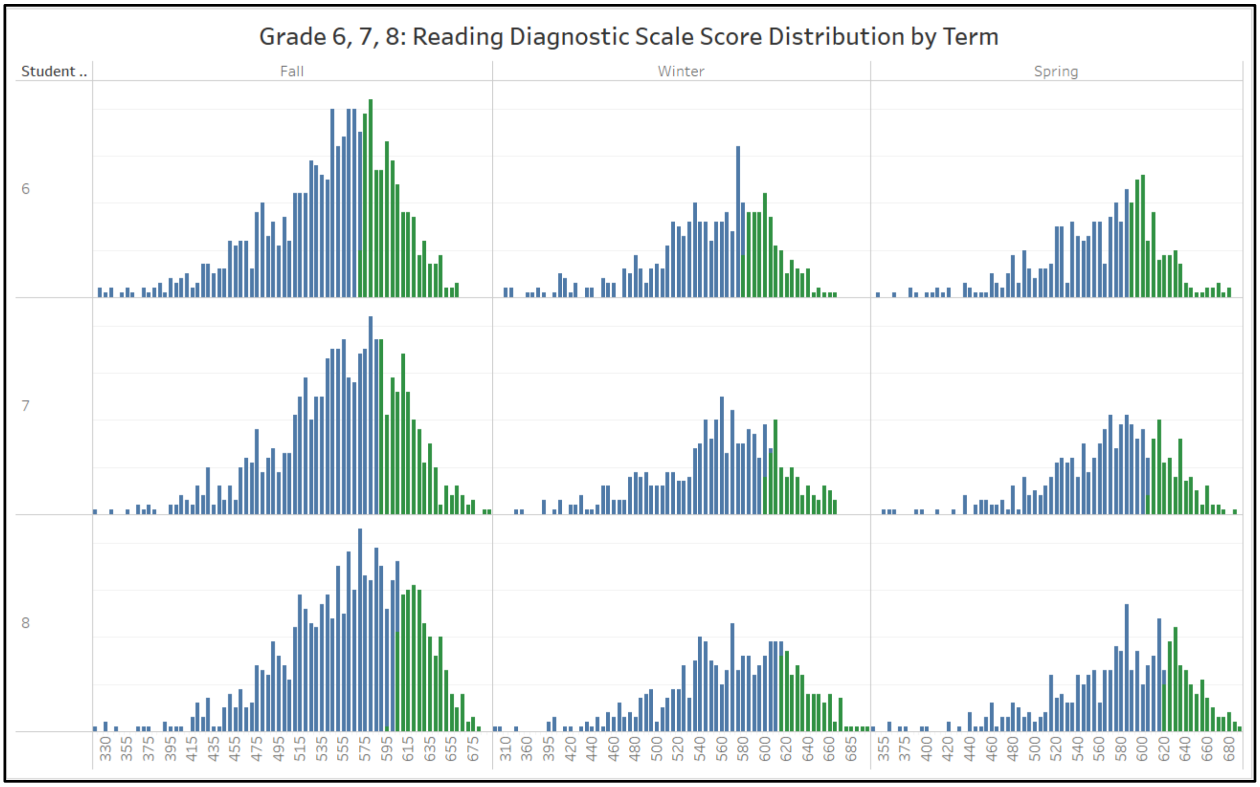
Grades kindergarten through 8th Math and Reading diagnostic scale score distribution






Grade 3, 4, 5: Reading Diagnostic Scale Score Distribution by Term
Student ..
Fall
Winter
Spring
3
4
5
290 310 340 360 380 400 420 440 460 480 500 520 540 560 580 600 620 640
305 335 360 380 400 420 440 460 480 500 520 540 560 580 600 620
315 335 355 375 395 415 435 455 475 495 515 535 555 575 595 615 635 660


Grade 3, 4, 5: Math Diagnostic Scale Score Distribution by Term
Student ..
Fall
Winter
Spring
3
4
5
295 310 325 340 355 370 385 400 415 430 445 460 475 490 505 525
295 320 335 350 365 380 395 410 425 440 455 470 485 500 515
315 335 350 365 380 395 410 425 440 455 470 485 500 515 530


Grade 6, 7, 8: Reading Diagnostic Scale Score Distribution by Term
Student ..
Fall
Winter
Spring
6
7
8
330 355 375 395 415 435 455 475 495 515 535 555 575 595 615 635 655 675
310 360 395 420 440 460 480 500 520 540 560 580 600 620 640 660 685
355 375 400 420 440 460 480 500 520 540 560 580 600 620 640 660 680

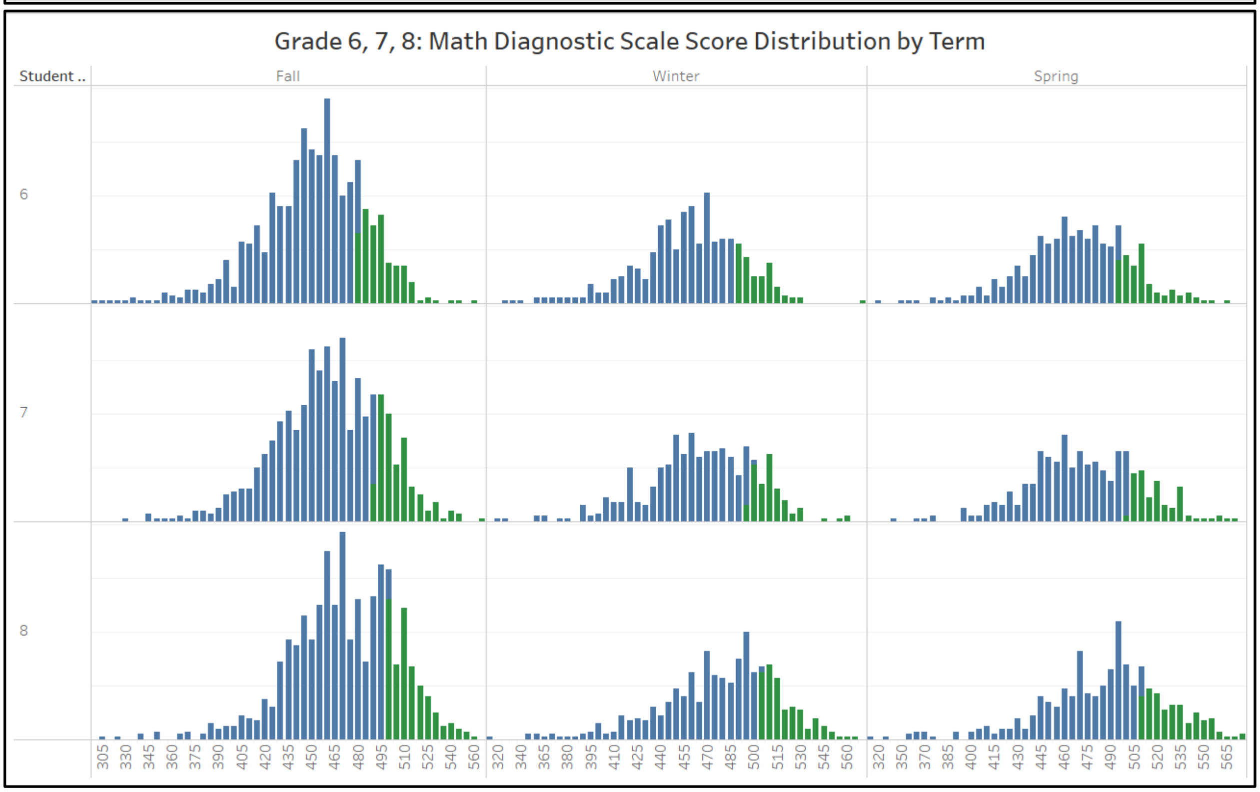
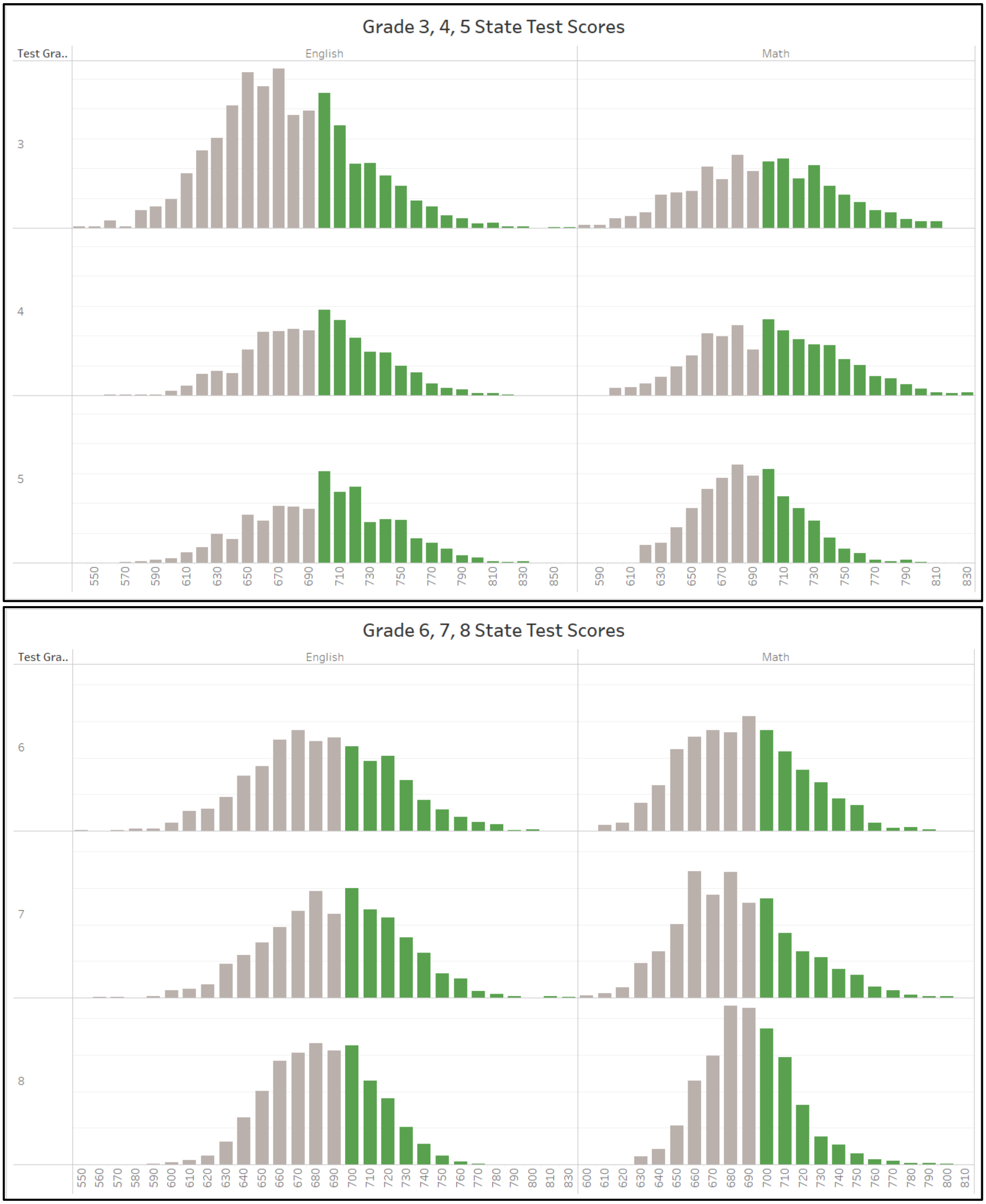
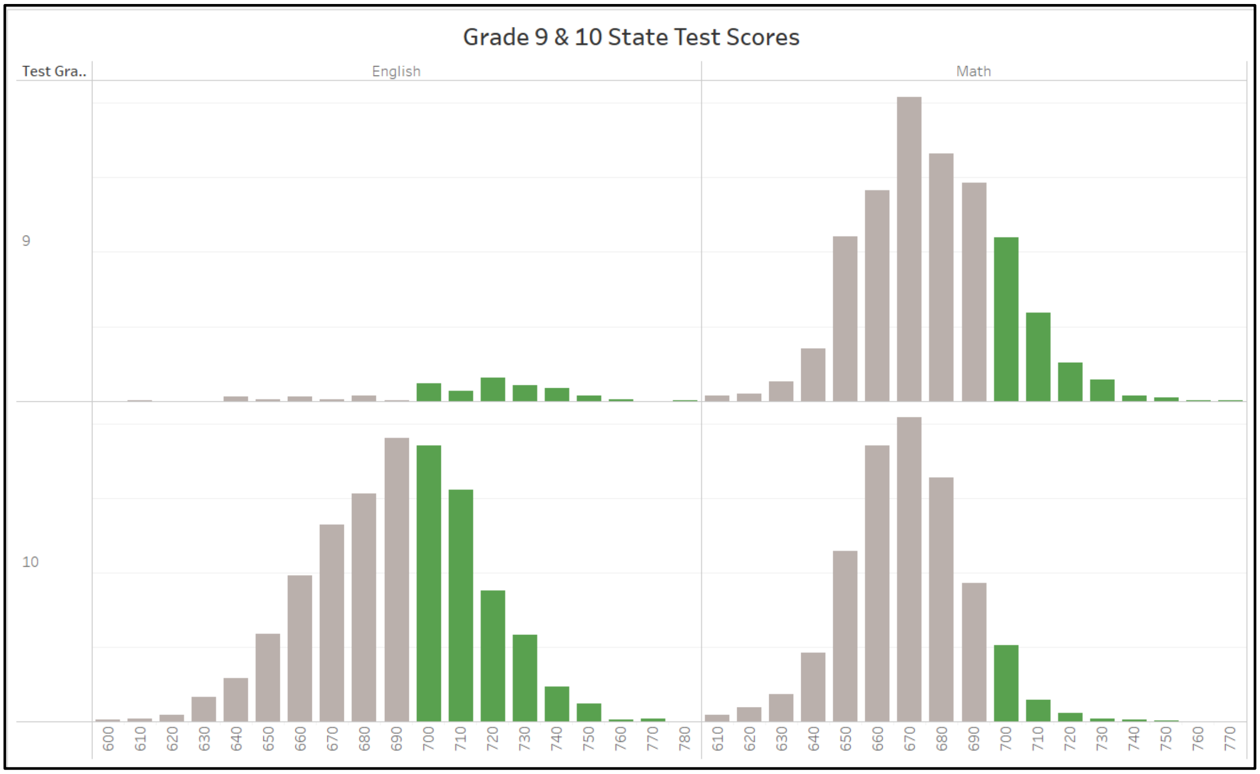

Grade 6, 7, 8: Math Diagnostic Scale Score Distribution by Term
Student ..
Fall
Winter
Spring
6
7
8
305 330 345 360 375 390 405 420 435 450 465 480 495 510 525 540 560
320 340 365 380 395 410 425 440 455 470 485 500 515 530 545 560
320 350 370 385 400 415 430 445 460 475 490 505 520 535 550 565

Figure 10

Grades 3rd to 10th Math and Reading state test scale score distribution





Grade 9 & 10 State Test Scores
Test Gra..
English
Math
9
10
600 610 620 630 640 650 660 670 680 690 700 710 720 730 740 750 760 770 780
610 620 630 640 650 660 670 680 690 700 710 720 730 740 750 760 770