Closing the Gap

Methodology of Analysis of F&P Reading Proficiency Scores

METHODOLOGY

The main tools used were Microsoft Excel and Python. Three data frames were used to conduct the initial analysis. Two data frames were created from the original file, DF2 and DF3. DF2 excluded all null values. This reduced the data set to 268 entries from the original 376. DF3 included mean imputation to replace all missing values, consequently the data frame retained the original 376 entries. The original data frame was used to test how deletion of null values and imputation of null values affected summary statistics.

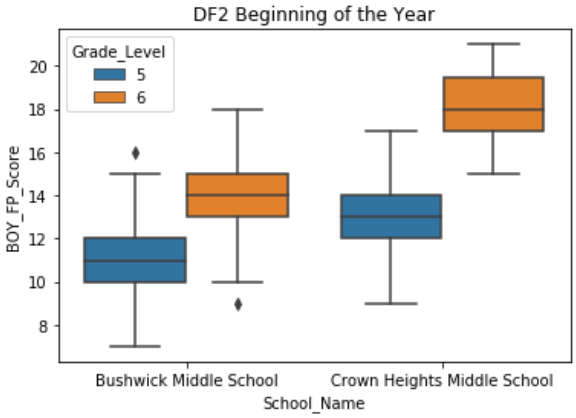
The primary analysis and findings are based on DF2.

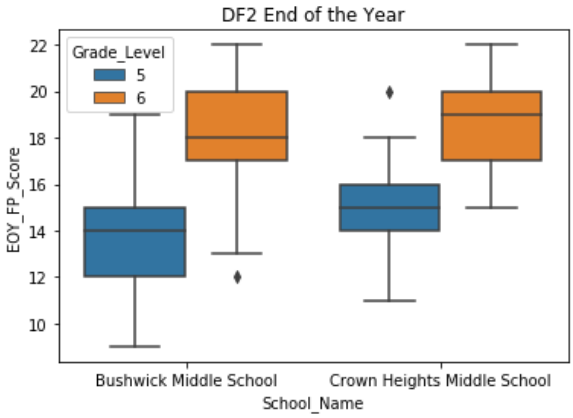
Subsets were created from the DF2 data set to allow for comparative analysis between school performance, grade performance and retention of proficiency from beginning of the school year to the end of the school year.

Final Data Frame: DF2

SCHOOL NAME	5TH GRADE	6TH GRADE
Crown Heights Middle School	73	15
Bushwick Middle School	86	94

ANALYSIS





The above box plots examine the variation of the values for each school by grade for reading proficiency scores for the beginning of the year and end of the year. The interquartile ranges at the beginning of the year are not very large indicating a concentration of scores between a smaller range. For example, a majority of the scores for Bushwick 5th graders at the beginning of the year fall between 10 and 12 and for Crown Heights between 12 and 14. The distribution of the ranges also suggests that Crown Heights outperforms Bushwick in reading proficiency at the beginning of the year.

The end of the year box plots varies more in range. For 5th grade Bushwick EOY scores a significant portion of the data fall below the median whereas the inverse is true for Bushwick EOY scores for 6th graders indicating the data is somewhat skewed. There also seems to be significant jump in results between Bushwick EOY 5th grade scores and EOY 6th grades scores with the latter showing a higher concentration of higher test scores.

The following distribution plots better visualize the data distribution and if/where they overlap.

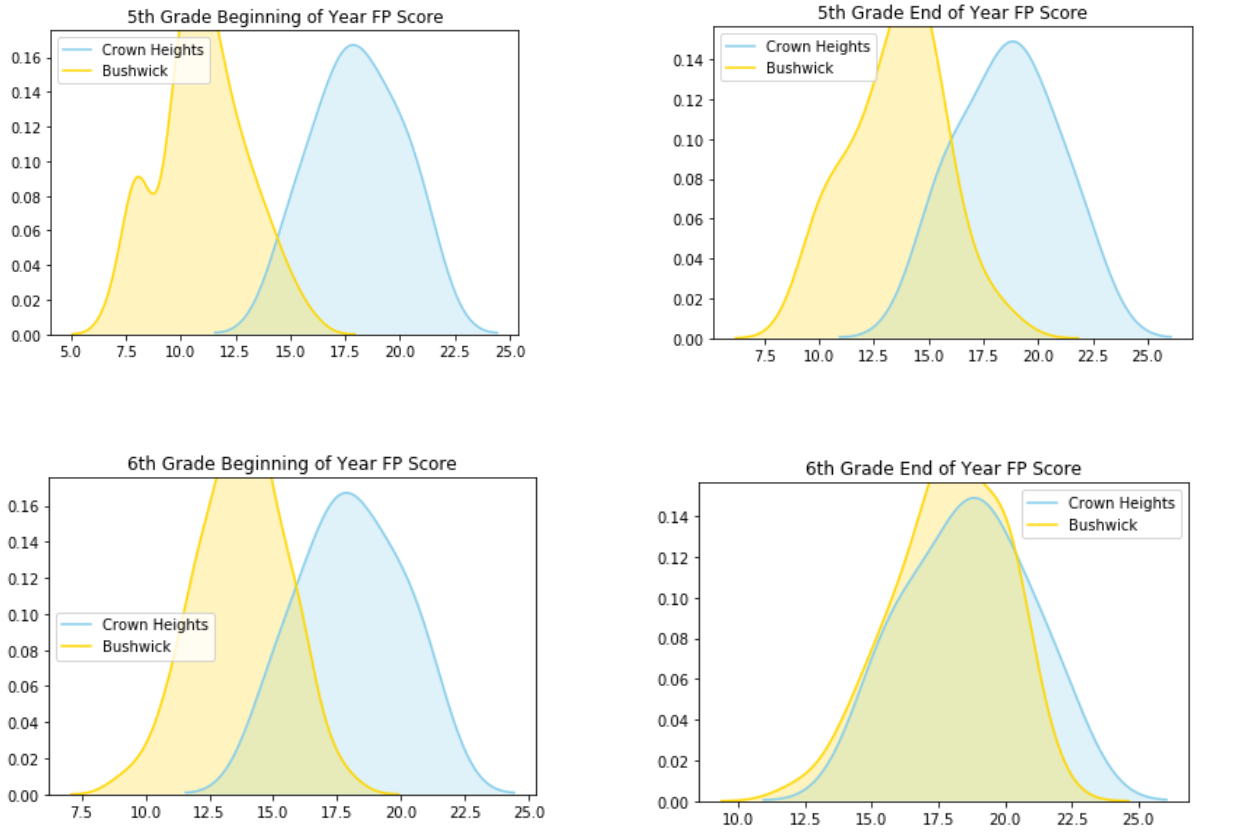


Figure 3-6. The above distribution plots further highlight the difference in scores between Bushwick and Crown Heights.

The distribution is fairly normal except for Bushwick 5th grade BOY scores. There appear to be two values that occur frequently in the data suggesting that incoming 5th graders at Bushwick have very different capabilities in terms of reading proficiency. Given where the two points peak it would appear that a significant portion of the students are either at proficient or above or below proficient to remedial.

To better understand where the scores fall in terms of proficiency two crosstabs, one for BOY and one for EOY were generated.

School_Name	BOY Grade_Level	Advanced	Below Proficient	Proficient	Remedial	All
Bushwick Middle School	5	10	38	22	16	86
	6	15	30	40	9	94
Crown Heights Middle School	5	32	8	30	3	73
	6	13	0	2	0	15
All		70	76	94	28	268

School_Name	EOY Grade_Level	Advanced	Below Proficient	Proficient	Remedial	All
Bushwick Middle School	5	13	21	35	17	86
	6	58	11	23	2	94
Crown Heights Middle School	5	26	15	29	3	73
	6	11	1	3	0	15
All		108	48	90	22	268

Comparing the two crosstabs gives further insight into student proficiency and improvement throughout the year. While Crown Heights has higher BOY scores that doesn't translate into higher EOY scores. 5th graders at Crown Heights who were proficient or higher decreased by 11.6% from 85% at the beginning of the year to 75% at the end of the year. However, Bushwick 5th graders who were proficient or higher saw a 51% increase from 37% at the beginning of the year to 56% at the end of the year.

Scores for 6th graders show a similar trajectory with more Bushwick 6th graders starting at proficient or higher and ending the year at proficient or higher.

Bushwick maintains a positive trajectory in reading proficiency, steadily increasing reading scores from beginning of the school year to the end of the year, from 5th grade to 6th grade.

Summary: Crown Heights Middle School

Though Crown Heights has higher reading scores they do not maintain the momentum throughout the year. There is a gradual but steady decline among those that are at proficient or higher during the school year. There is a slight bump at the beginning of each year, but it does not provide a foolproof buffer to keep proficiency levels from decreasing and deficiency levels increasing.

Summary: Bushwick Middle School

Despite starting with lower reading proficiency scores Bushwick does a better job of maintaining proficiency among students that are already proficient and increasing proficiency among students who are deficient.

Overall Progress Toward Reading Proficiency

For now, both schools appear to be on track to close the reading gap. 61% of students begin the school year at proficient or higher and 74% end the year at proficient or higher.

However, Bushwick is doing a better job at closing the achievement gap for below proficient and remedial students. Of the 71 5th and 6th grade Bushwick students who ended the year at advanced proficiency 86% began the year at proficient or lower and 48% began the year at below proficient or lower. Of the 41 5th and 6th grade Crown Heights students who ended the year at advanced proficiency 44% started that year at proficient or lower and 17% started at below proficient. It is possible that should the negative trajectory increase the achievement gap could widen at Crown Heights.

Additional Notes

The number of values for Crown Heights 6th graders was drastically reduced (from 94 to 15) when the null values were deleted. While mean imputation was initially used (DF3) to fill in the missing values given that over 80% of the data was missing it was decided that dropping the null values would better preserve the integrity of the analysis. DF2 was used to conduct the remaining analysis to avoid bias mean imputation would introduce and because it could not adequately be determined if the missing values were missing completely at random or simply missing at random.