Visualizing ELA and Math Assessment Trends in New York Schools, 2018-2019

John Bartlett, Ebo Edonu, Raya Young

Data Overview

- The New York State Education system contains 730+ districts; 4,400+ public schools; and 340+ charter schools.
- This information describes testing trends among grades 3-8 in the New York state public school system for 2018-2019.
- Besides location, the detailed dataset is also broken down by various student demographics including gender (equally split), ethnicity, and homelessness, for example.
- We imported the excel files and used the sqldf package access the information as a database. The detailed ELA scoring dataset contained 565k+ observations and 23 variables.

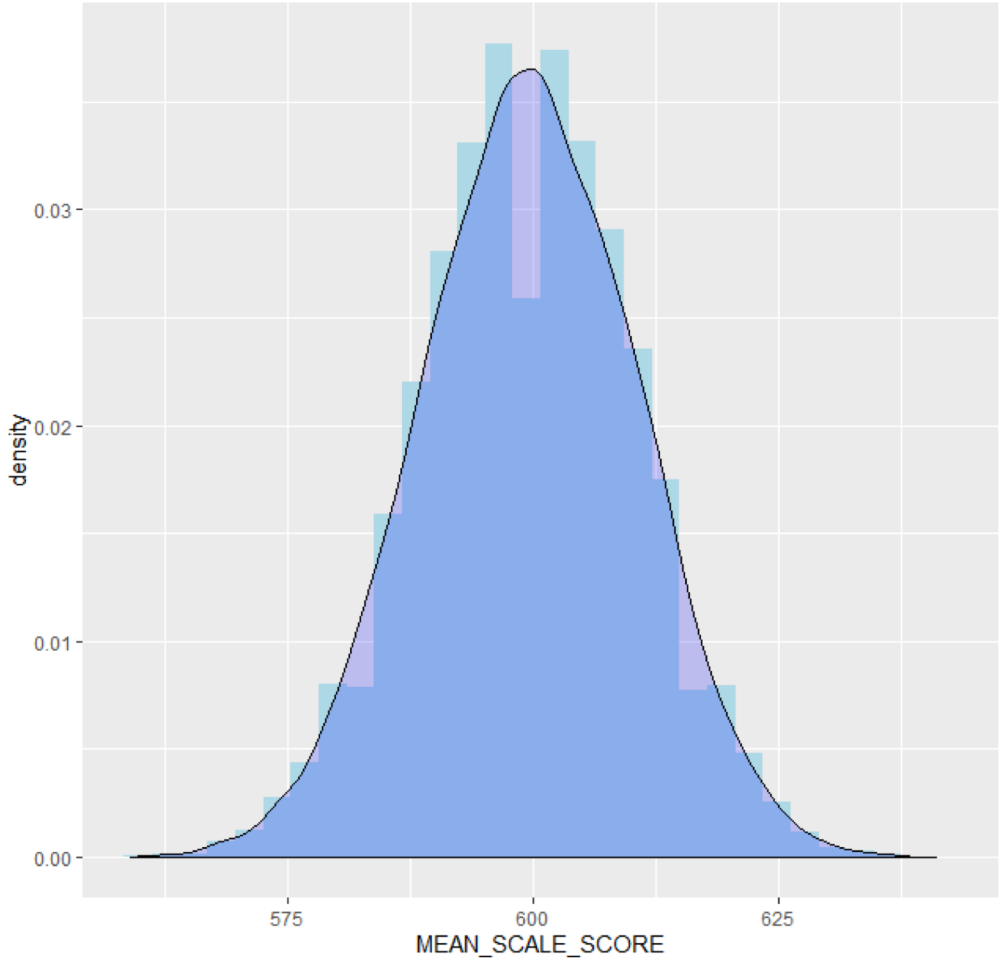
At a glance

Classes ‘tbl_df’, ‘tbl’ and 'data.frame': **565818 obs. of 23 variables:**

$ SY_END_DATE : POSIXct, format: "2019-06-30" "2019-06-30" "2019-06-30" ...
$ NRC_CODE : num NA NA NA NA NA NA NA NA NA NA ...
$ NRC_DESC : chr NA NA NA NA ...
$ COUNTY_CODE : logi NA NA NA NA NA NA ...
$ COUNTY_DESC : logi NA NA NA NA NA NA ...
$ BEDSCODE : num 0 0 0 0 0 0 0 0 0 0 ...
$ NAME : chr "STATEWIDE - ALL DISTRICTS AND CHARTERS" "STATEWIDE - ALL DISTRICTS AND CHARTERS" "STATEWIDE - ALL DISTRICTS AND CHARTERS" "STATEWIDE - ALL DISTRICTS AND CHARTERS" ...
$ ITEM_SUBJECT_AREA: chr "ELA" "ELA" "ELA" "ELA" ...
$ ITEM_DESC : chr "Grade 3 ELA" "Grade 3 ELA" "Grade 3 ELA" "Grade 3 ELA" ...
$ SUBGROUP_CODE : num 1 2 3 4 5 6 7 8 9 10 ...
$ SUBGROUP_NAME : chr "All Students" "Female" "Male" "American Indian or Alaska Native" ...
$ TOTAL_TESTED : num 172149 84744 87405 1216 29834 ...
$ L1_COUNT : chr "26358" "10645" "15713" "187" ...
$ L1_PCT : chr "0.15" "0.13" "0.18" "0.15" ...
$ L2_COUNT : chr "55837" "26224" "29613" "450" ...
$ L2_PCT : chr "0.32" "0.31" "0.34" "0.37" ...
$ L3_COUNT : chr "77771" "40751" "37020" "499" ...
$ L3_PCT : chr "0.45" "0.48" "0.42" "0.41" ...
$ L4_COUNT : chr "12183" "7124" "5059" "80" ...
$ L4_PCT : chr "7.0000000000000007E-2" "0.08" "0.06" "7.0000000000000007E-2" ...
$ L2-L4_PCT : chr "0.85" "0.87" "0.82" "0.85" ...
$ L3-L4_PCT : chr "0.52" "0.56000000000000005" "0.48" "0.48" ...
$ MEAN_SCALE_SCORE : chr "600" "602" "598" "599" ...

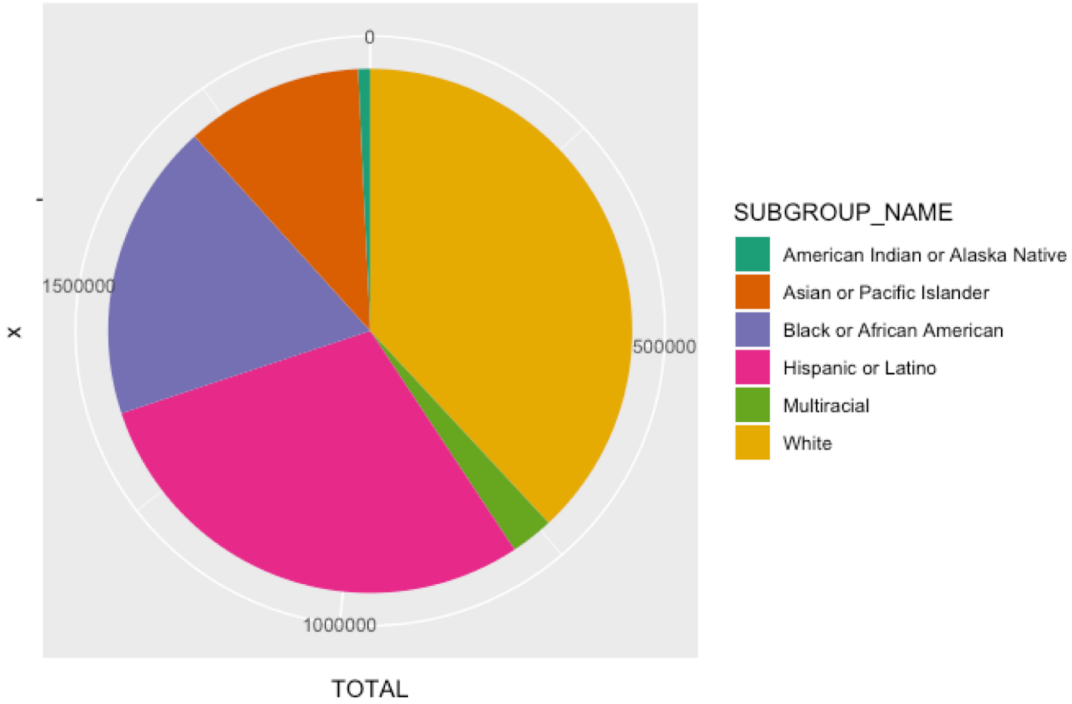
Score Overview

The mean scale score of both Math and English Language Arts exam scores closely resemble a normal Gaussian distribution.



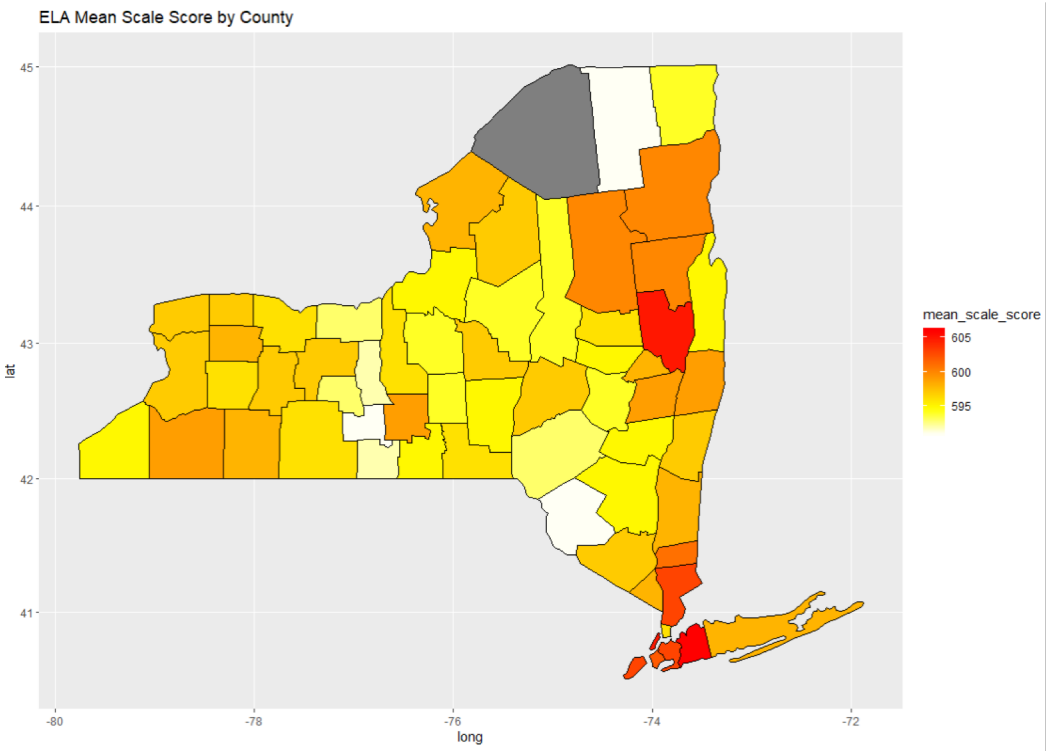
Subgroups by Race

Using ggplot2, this pi chart was generated to demonstrate the proportions of different subgroups that were represented in the data. White and Hispanic or Latino were the two largest groups, followed by Black or African American, then Asian or Pacific Islander.

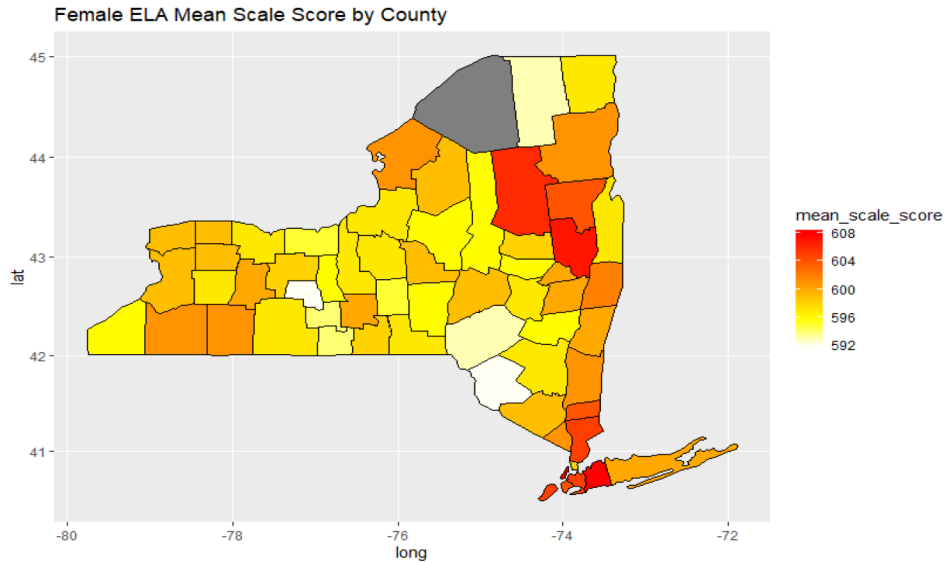


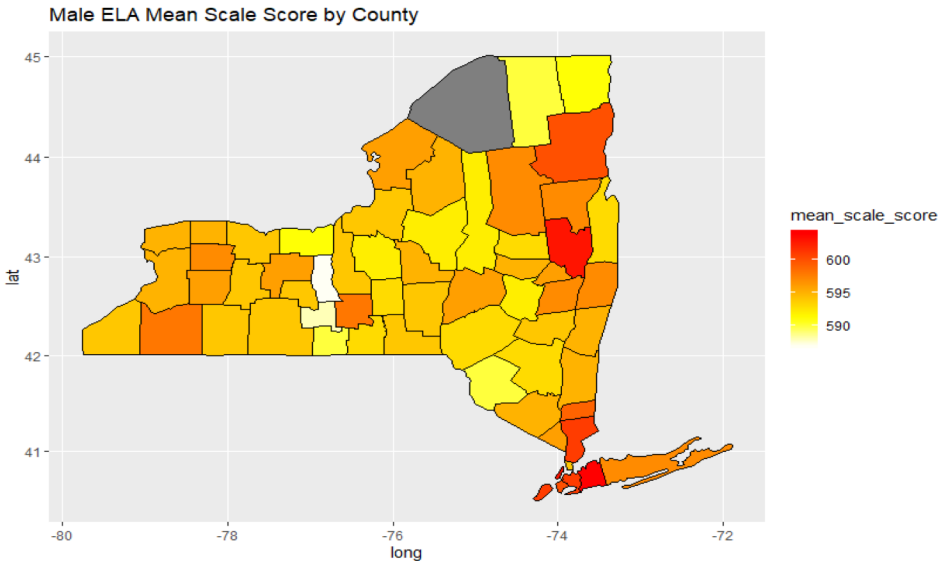
ELA Mean Scale Score by County

This color-coded map shows the ELA mean scale score by county, demonstrating trends in performance by geographic location. (***Grey county did not participate in the study.) The following slides gives a further breakdown of Mathematics scoring by gender for each county.

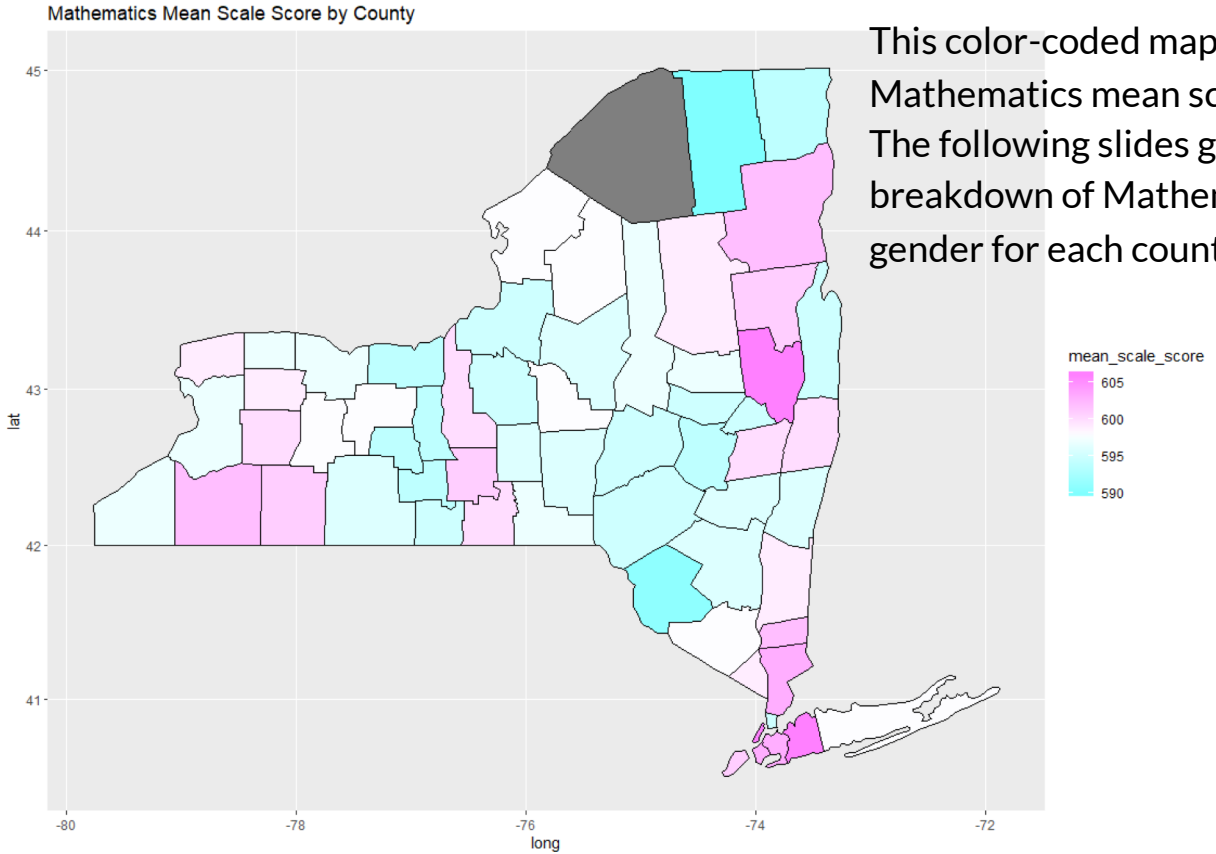


Mean Scale Score by County; by Gender



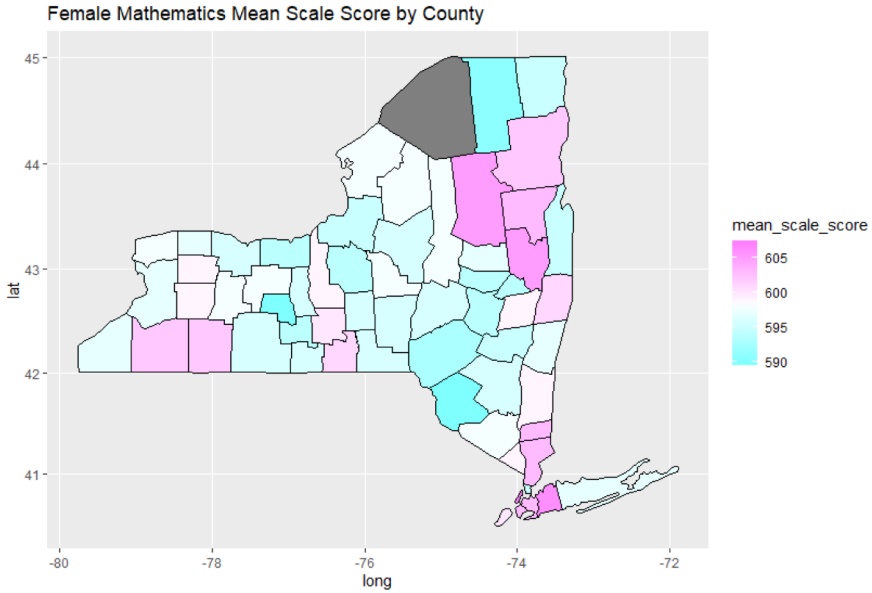


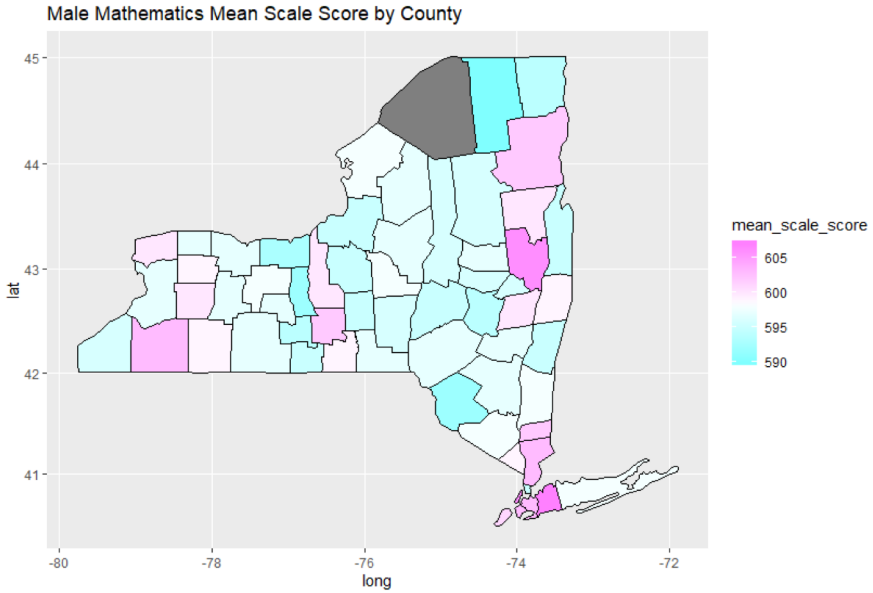
Mathematics Mean Scale Score by County



This color-coded map shows the Mathematics mean scale score by county. The following slides gives a further breakdown of Mathematics scoring by gender for each county.

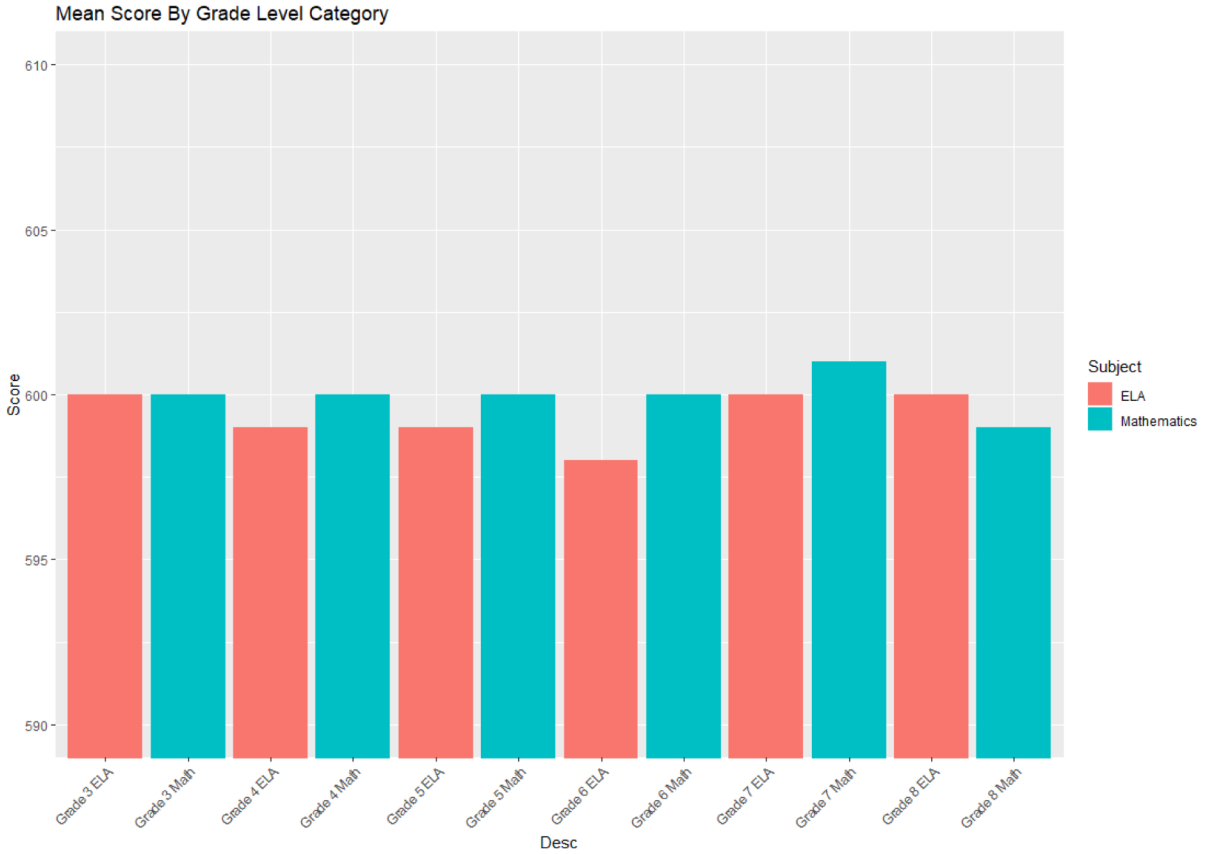
Mathematics Mean Scale Score by County; by



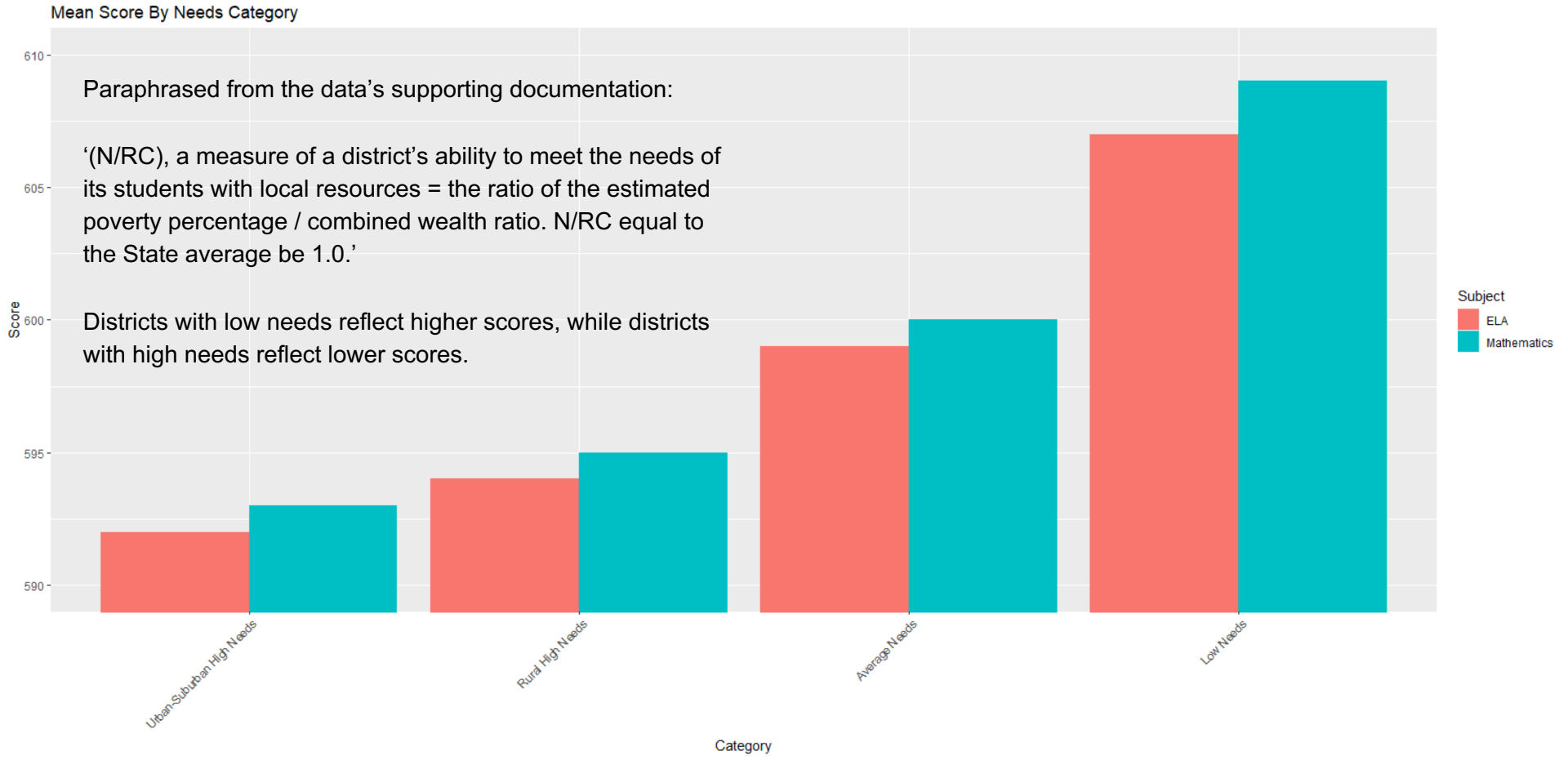


Mean Scale Score by Grade Level

This visualization shows mean scores by grade, in both ELA and mathematics assessments. There appears to be little variation present and seems to imply the success of the program. In general, performance on ELA is slightly less than Mathematics for most grades with Grade 6 ELA needing the most improvement, but scoring is high for both areas.

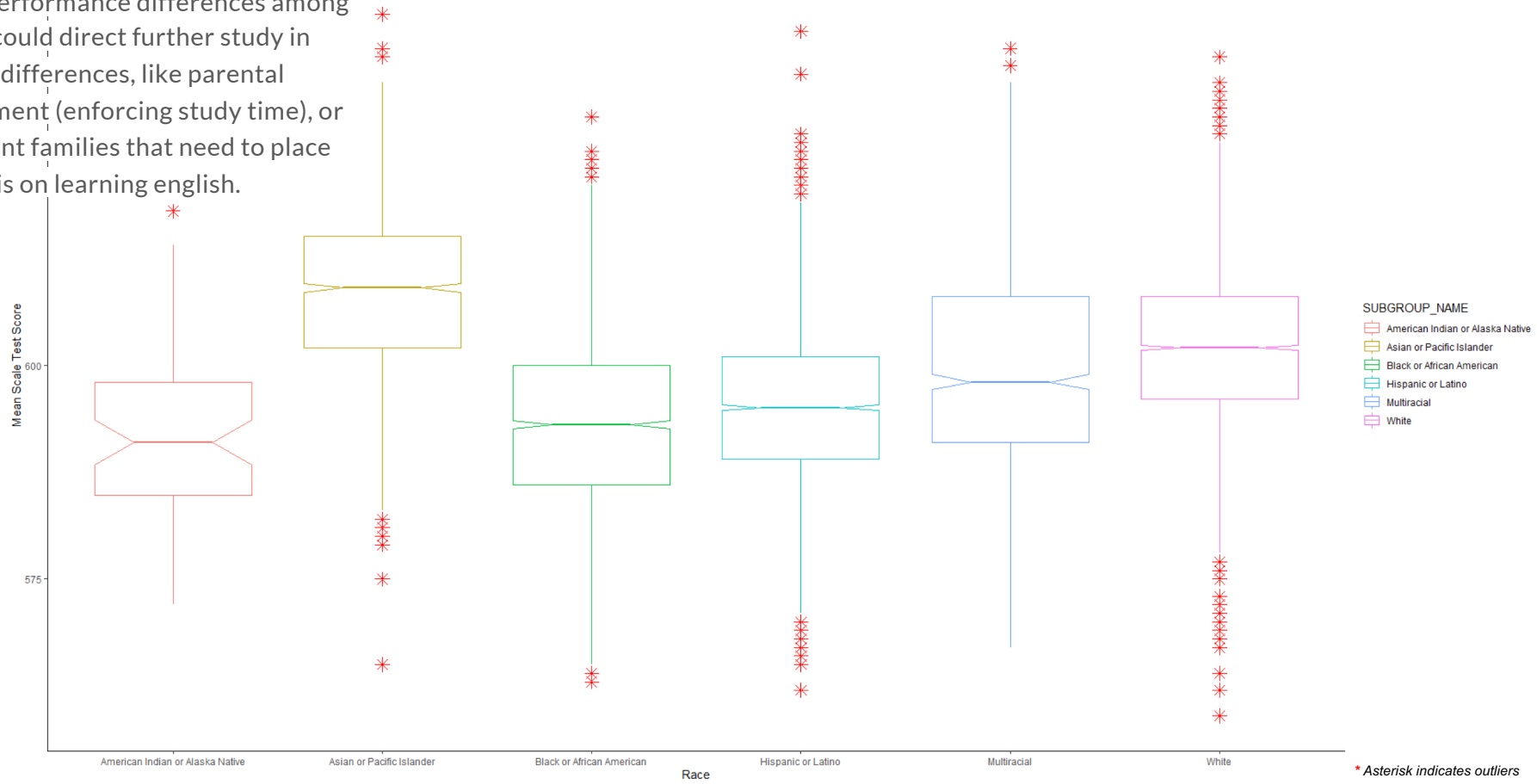


Mean Score by Need/Resources Capacity



English - Mean Scale Score by Race

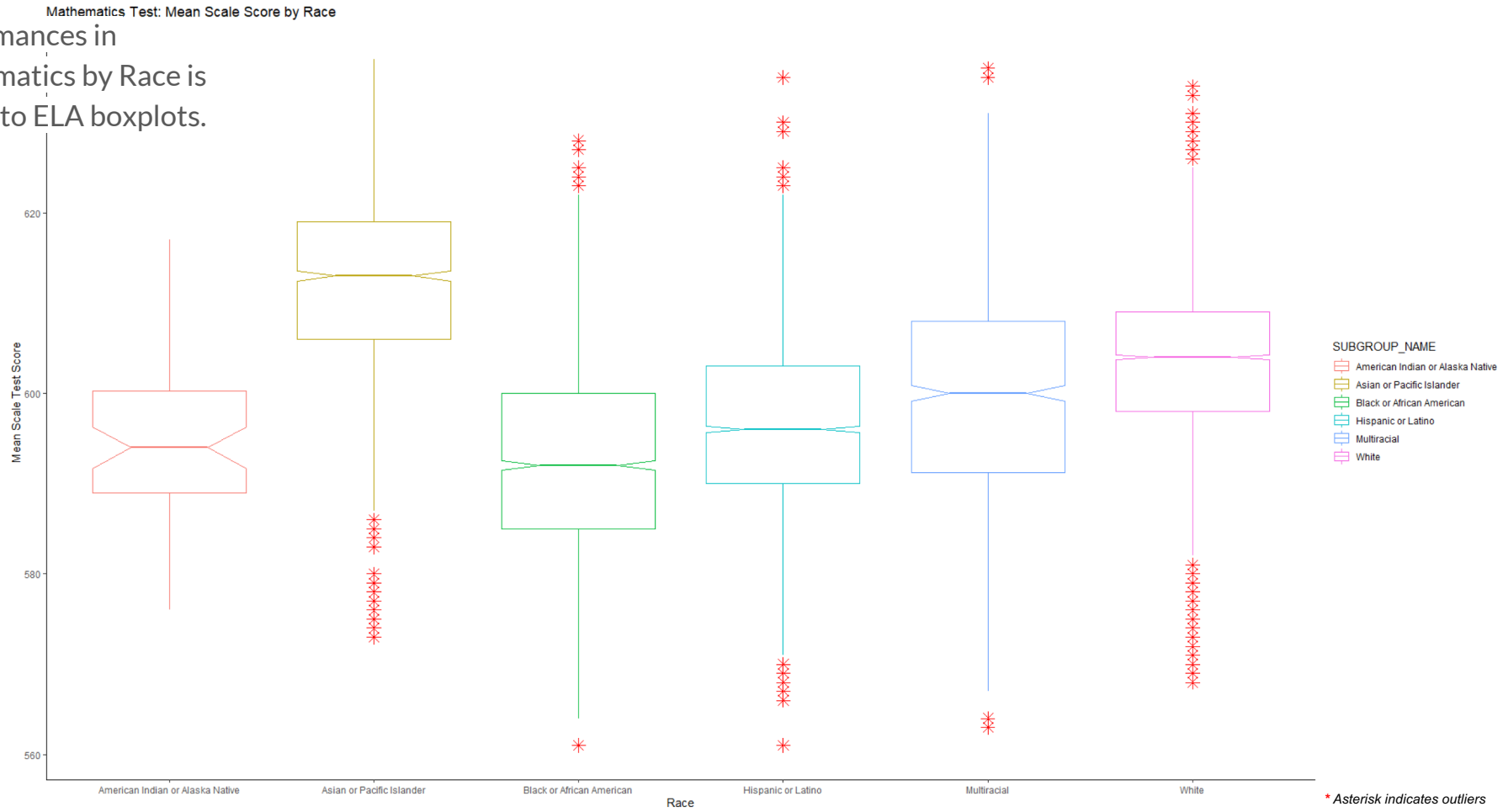
These boxplots show a breakdown of ELA assessment by race. Asian and Pacific Islanders are well above the mean. Performance differences among groups could direct further study in cultural differences, like parental involvement (enforcing study time), or immigrant families that need to place emphasis on learning english.



* *Asterisk indicates outliers*

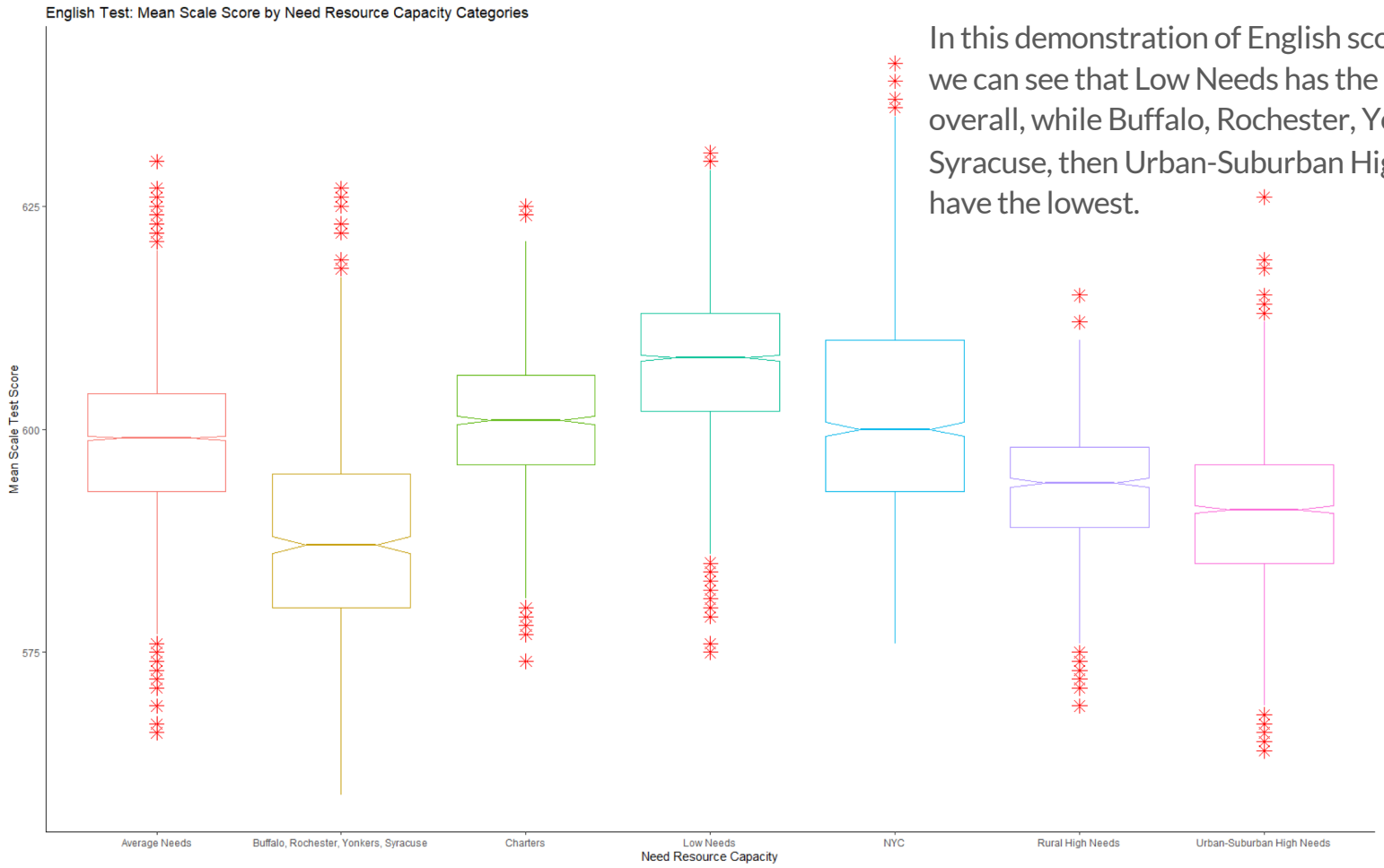
Mathematics Test - Mean Scale Score by Race

Performances in Mathematics by Race is similar to ELA boxplots.



** Asterisk indicates outliers*

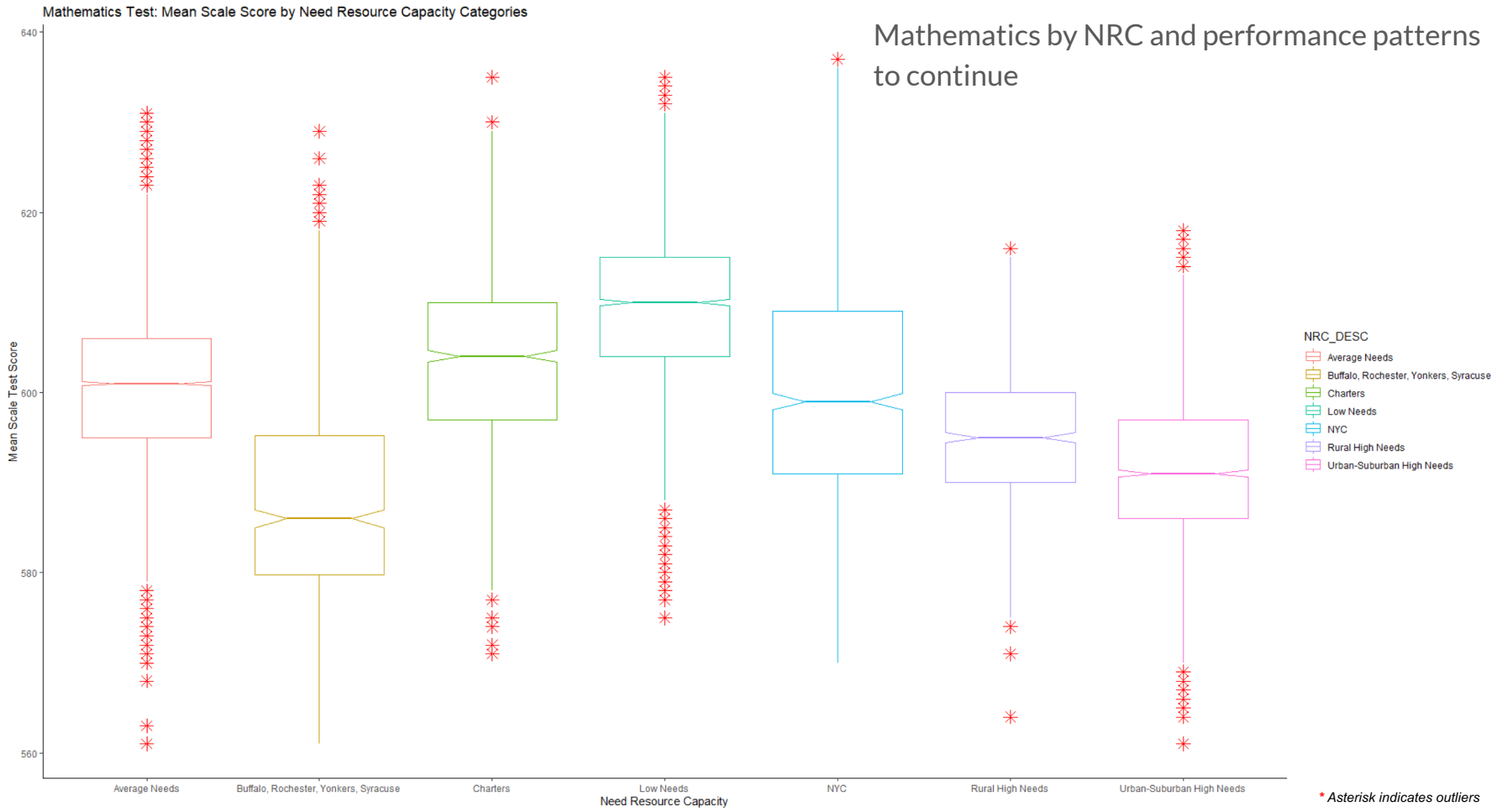
English - Mean Scale Score by NRC



In this demonstration of English scoring by NRC, we can see that Low Needs has the highest scoring overall, while Buffalo, Rochester, Yonkers, Syracuse, then Urban-Suburban High Needs and have the lowest.

** Asterisk indicates outliers*

Mathematics - Mean Scale Score by NRC

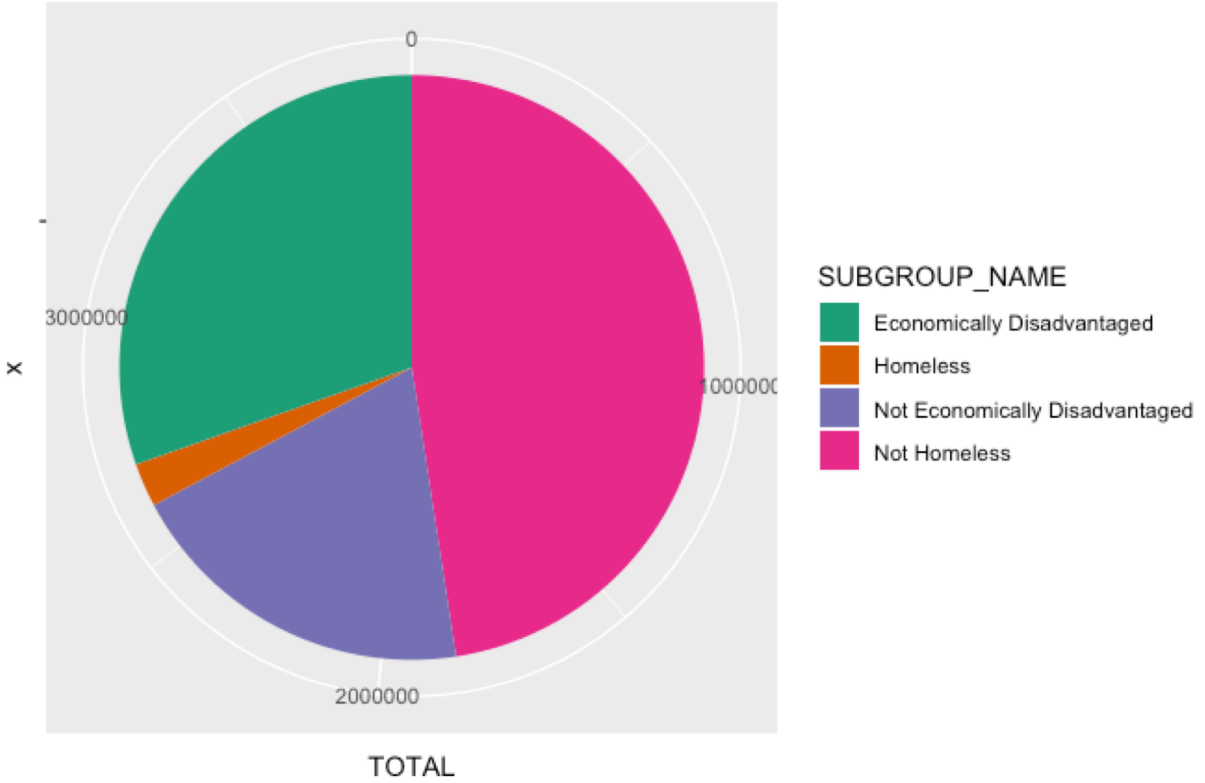


Mathematics by NRC and performance patterns to continue

** Asterisk indicates outliers*

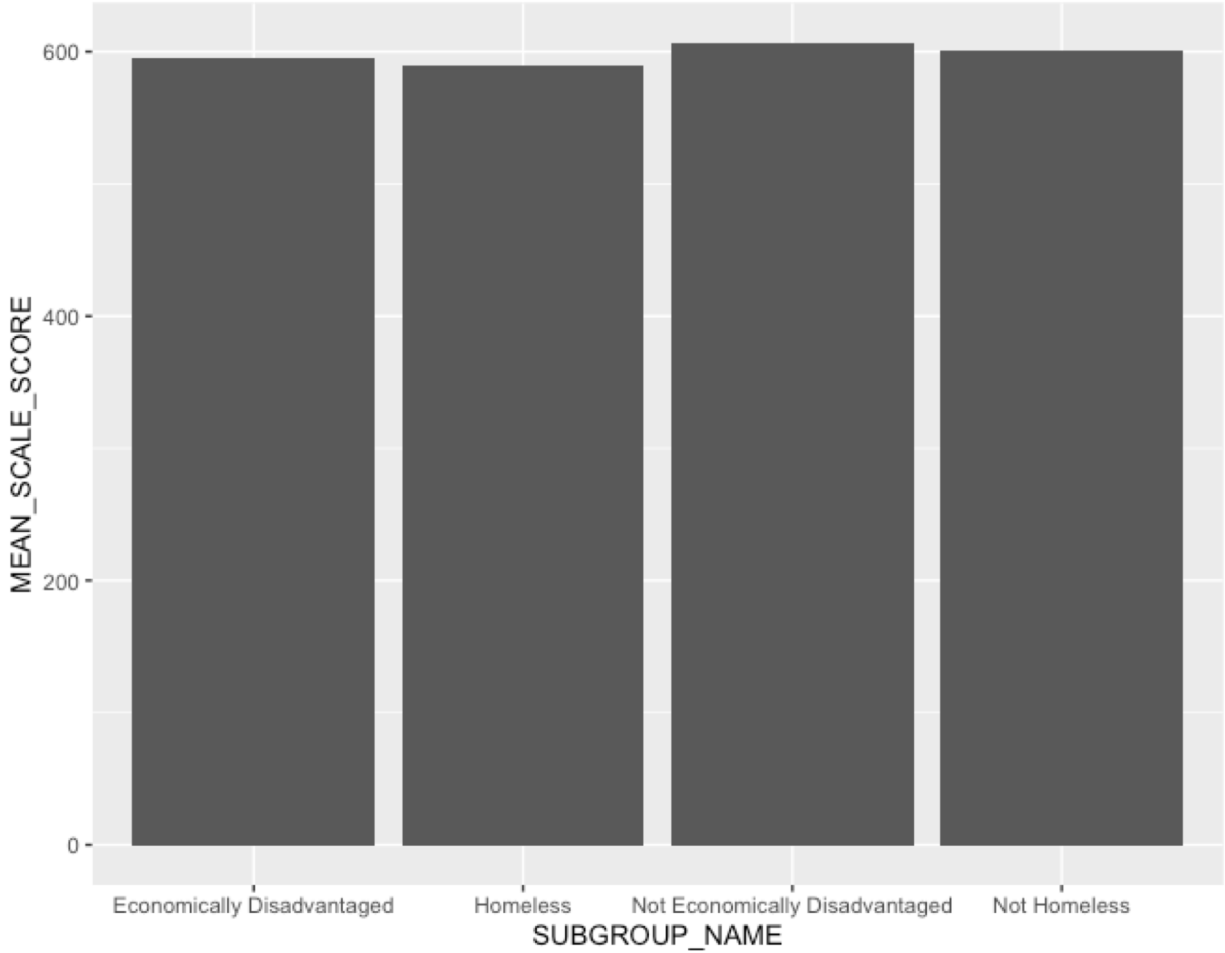
Subgroups, Home Status

This pi chart shows the proportion of homeless and economically disadvantaged children in the sample.



Performance, Home Status

This chart shows the compared mean performance of those groups with differing home and economic status. Homeless scored the lowest at 590, while Not Homeless scored at 600. Other categories Disadvantaged and Not Disadvantaged scored at 595 and 606, respectively. Mode revealed almost identical findings.



Conclusion

The data set we used for the project contained several dimensions of information, so it took some exploring to discover which features we wanted to focus on. We found interesting patterns among socioeconomic status and test performance, as well as performance variations by race and gender. Linear analysis showed clumping around residuals, and predictive modeling showed strange results. There were no considerable differences in performance between grades, which could be seen as a measure of success by the program.