

Predicting Student Performance

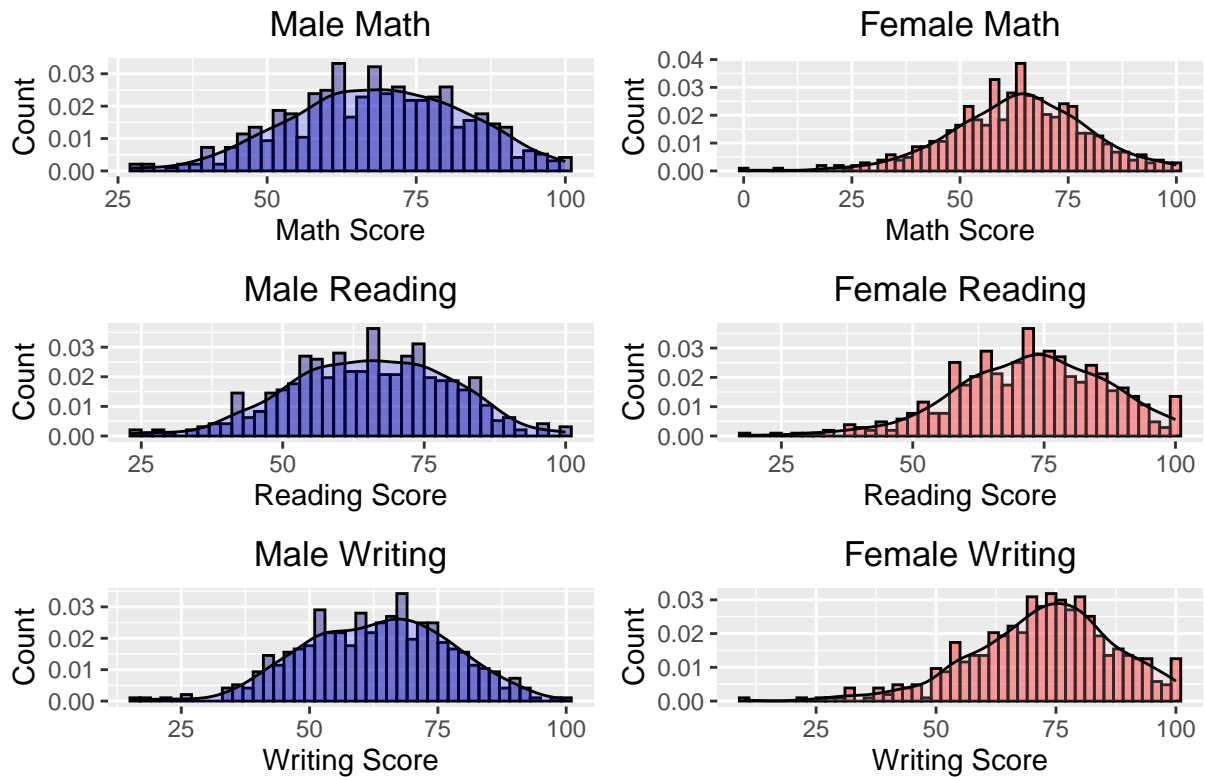
John Russell

2023-01-04

The goal of this project is to gain insight on classroom performance. We will begin with a series of visualizations and conclude with a predictive model.

We will first examine the distribution of subject specific scores across gender.

Distribution of Scores across Gender



Mean Scores

Gender	Math	Reading	Writing
Male	68.73	65.47	63.31
Female	63.63	72.61	72.47

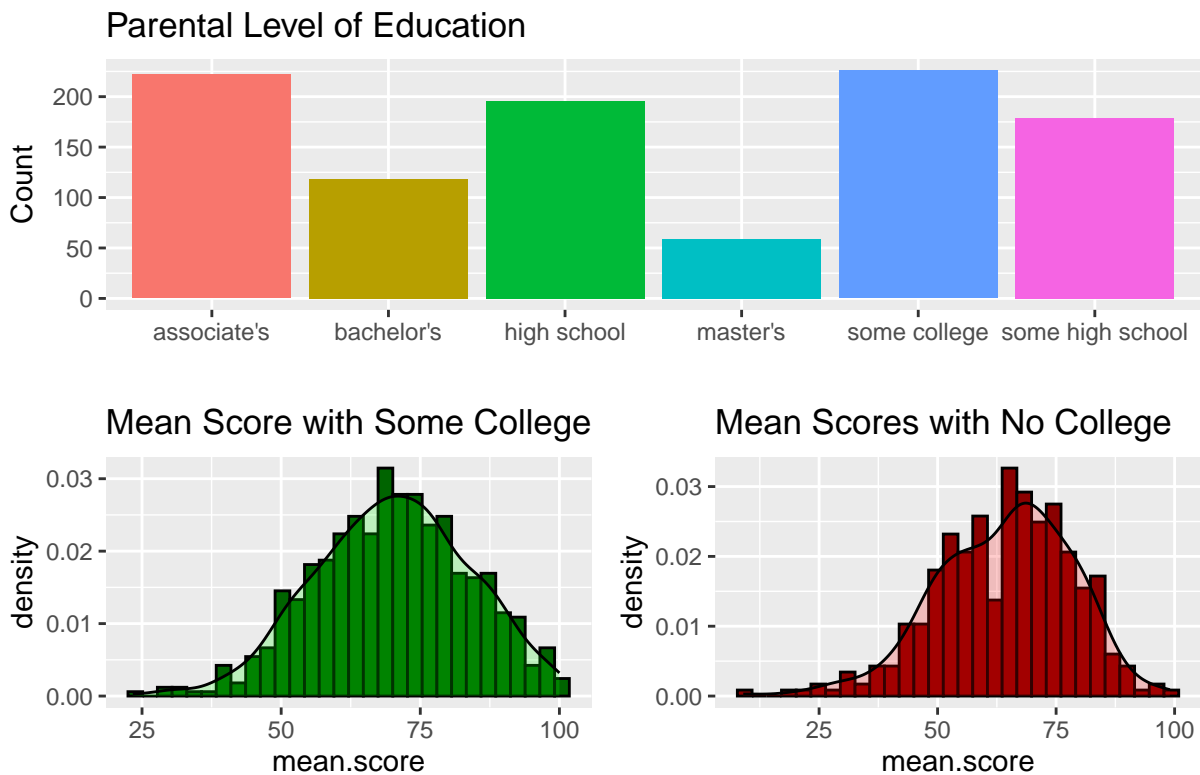
There are some interesting conclusion we can make:

- Male's tend to score higher in Math.
- Female's tend to score higher in Reading and Writing.
- The distribution for the male group is relatively normal while the distribution for the female group is skewed to the right.

Next, we want to look at the effect of parental level of education on subject scores.

parental.level.of.education	n	percent
associate's	222	22.2
bachelor's	118	11.8
high school	196	19.6
master's	59	5.9
some college	226	22.6
some high school	179	17.9

Distribution of Mean Scores & Some College vs. No College

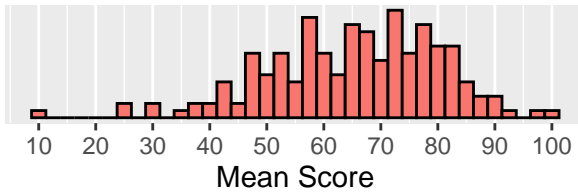


There are a few things to notice:

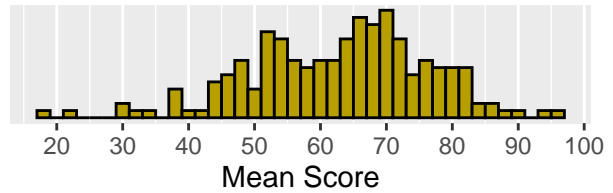
- Most students have parents who achieved some level of college (65.7%).
- The median score for students whose parents achieved some level of college is 70.3%.
- The median score for kids whose parents did not reach college is 65.7%.

Distribution of Mean Scores across Parental Education Level

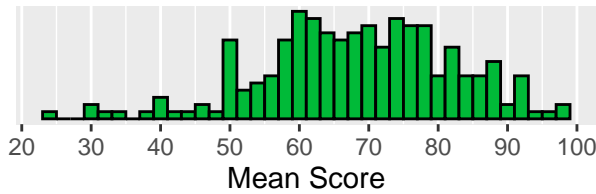
Some High School



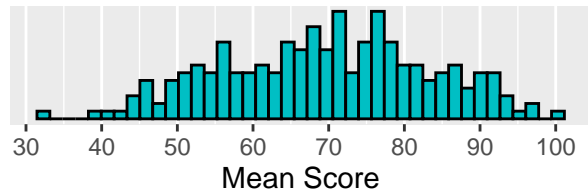
High School



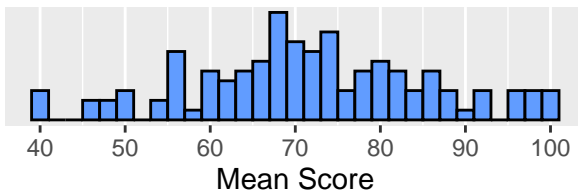
Some College



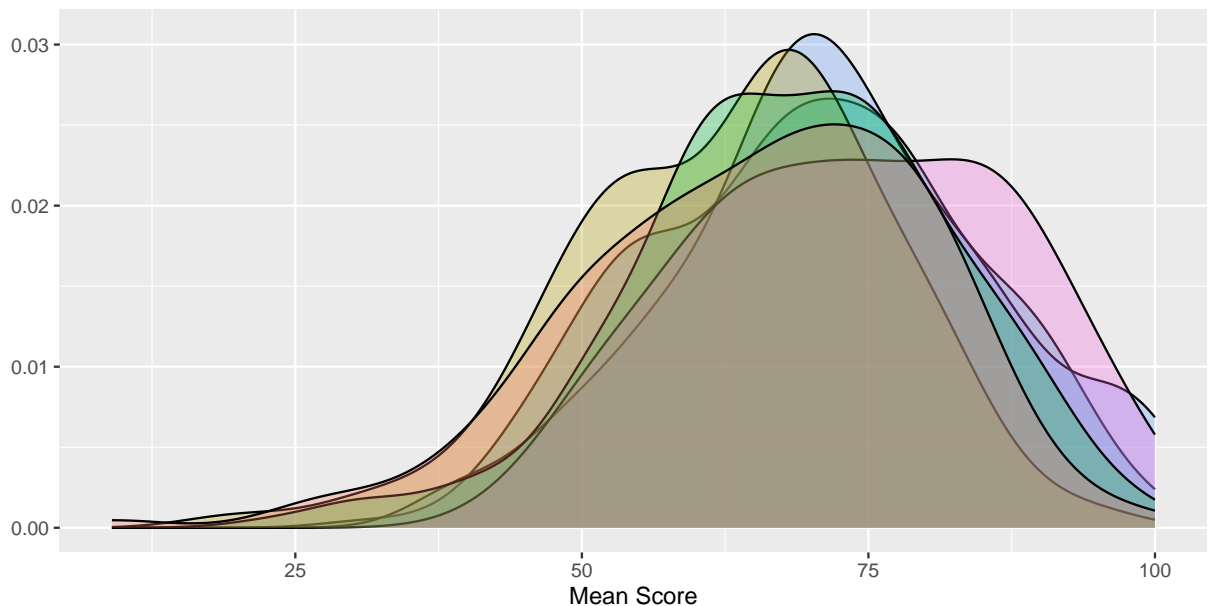
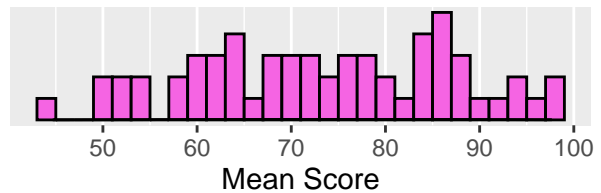
Associate's Degree



Bachelor's Degree



Master's Degree



We see that student's whose parents achieved beyond a bachelor's degree are more likely to score beyond 80%. Furthermore, student's whose parents did not attend college are more likely to score below 60%. Otherwise, it appears to be difficult to differentiate between different levels of parental education.

Next, we will look at how race affects subject specific exam scores.