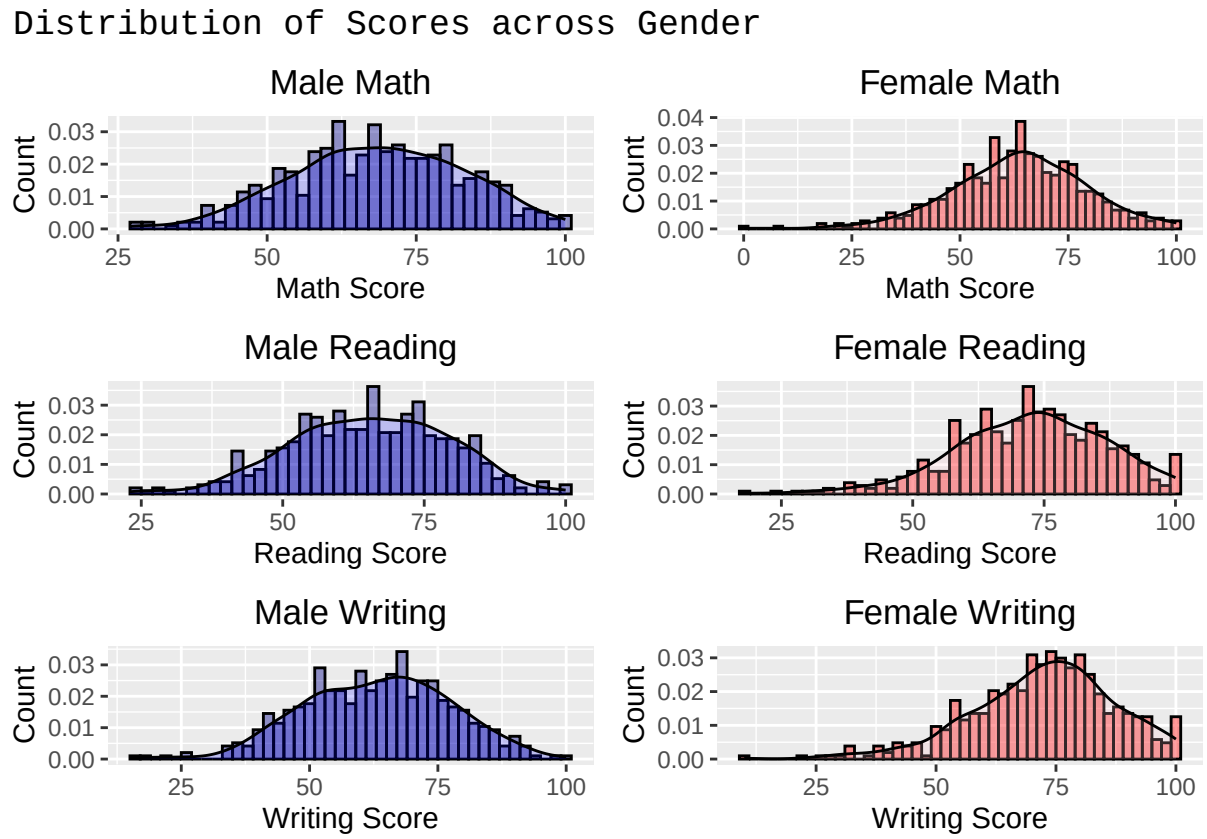
Predicting Student Performance

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.



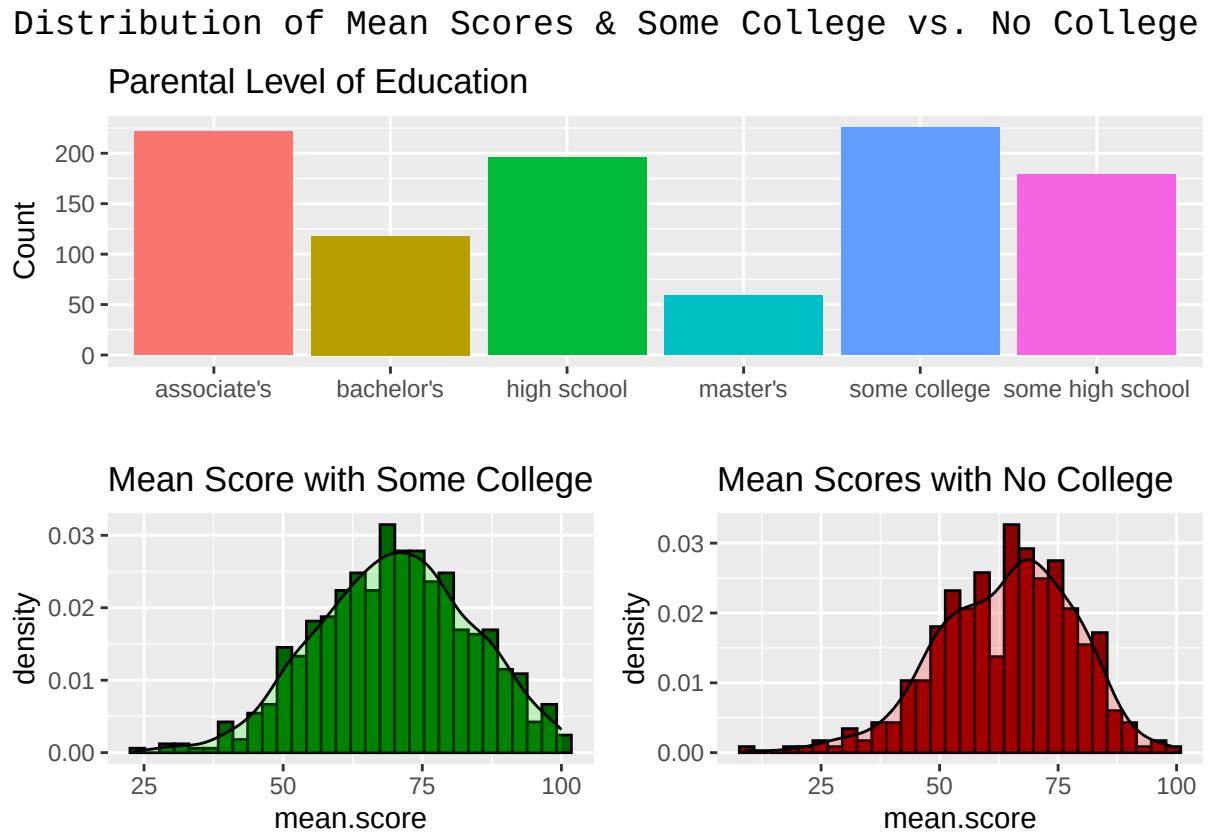
Mean Scores

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

We can see that male's tend to score higher in Math while female's tend to score higher in Reading and Writing. We see that the Reading and Writing histograms for the female group are skewed to the right. The 'why' is an interesting question in and of itself.

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

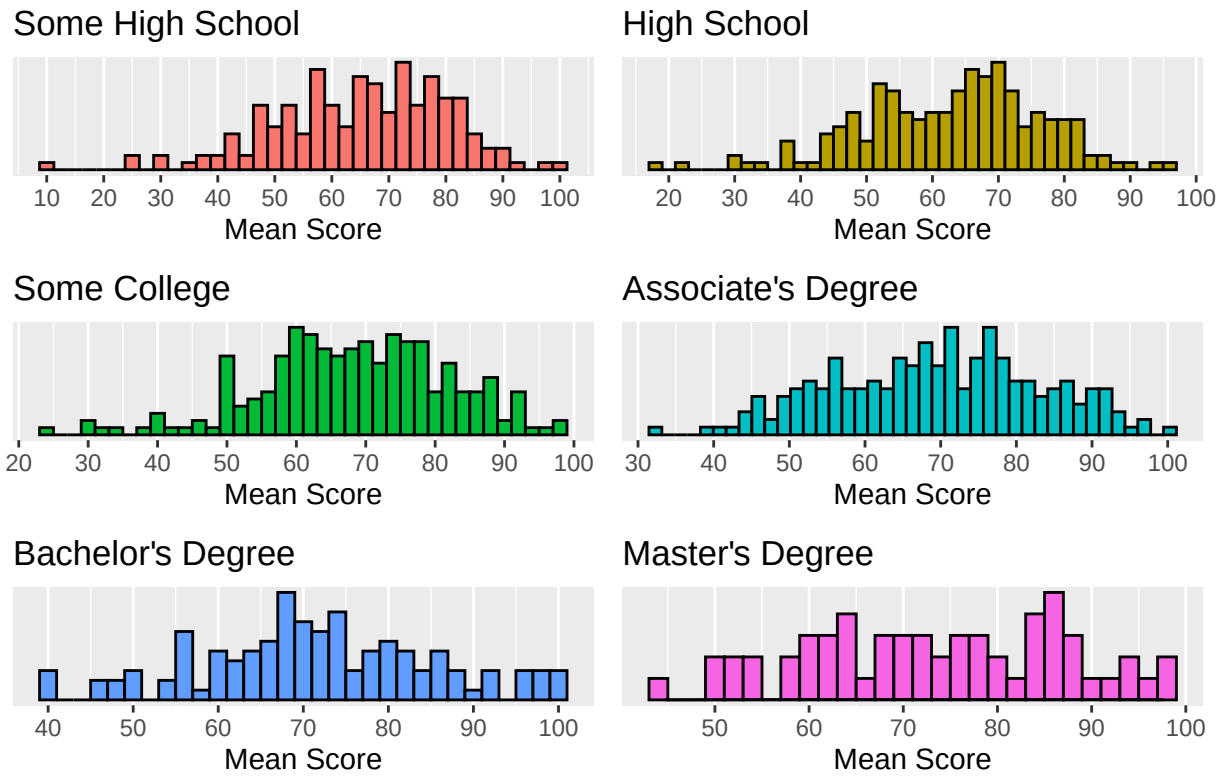
parental.level.of.education	n	percent
associate's	222	22.2
bachelor's	118	11.8
high school	196	19.6
master's	59	5.9
some college	226	22.6
some high school	179	17.9



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



Overlay Mean Scores across Parental Education Level