# High School Student Performance In Exams

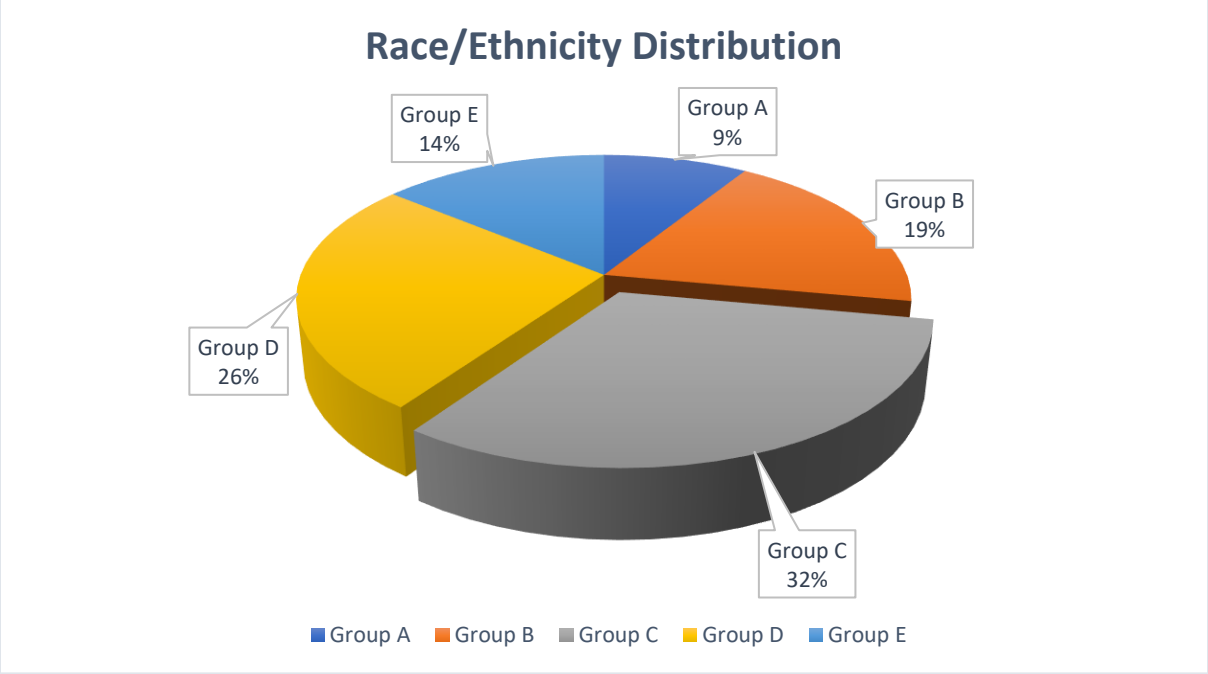
The dataset was extracted from Kaggle. By studying this dataset, the main goal is to understand the influence of the parents' background, race, gender, etc. on students' performance in three exams.

Main Considerations:

1. How parent's education distributed in different races?
2. How parent's education will affect student's performance in exams?
3. How race/ethnicity will affect student's performance in exams?
4. Does gender have different influence on scores in math/reading/writing respectively?
5. Do courses correlated to each other?

First, let's have a look at the race distribution of the students in the high school.

In this dataset, most of the students are from Group C which consist 32% of the total students. The second most students are from Group D which consists 26% of the total students. Group A has the least students in the high school.



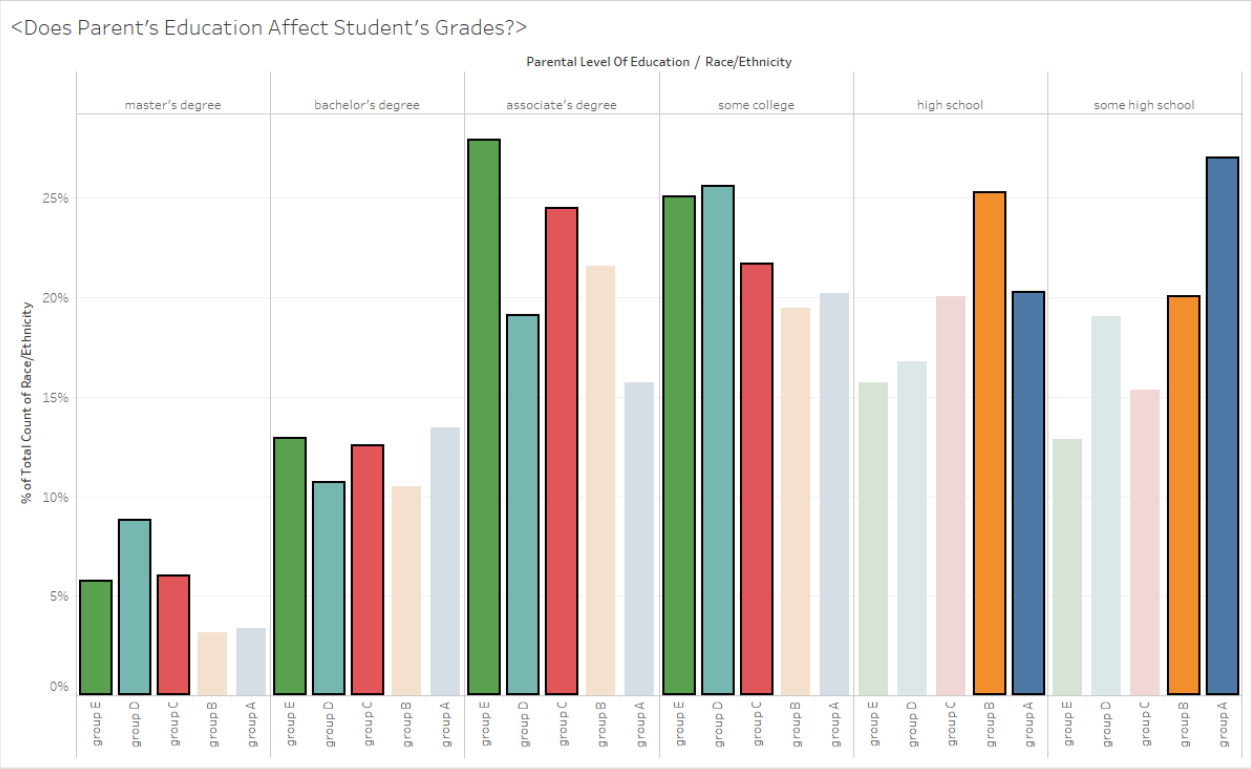

Made in Excel

Let's do more analysis.

Source from Kaggle https://www.kaggle.com/spscientist/students-performance-in-exams

In this graph, we want to find some patterns between race and parental level of education.

We found when the parent held a master's degree, most of them are from Group D, C, E. We can also find that if the parental education is higher than 'some college', most of them are from Group C, Group D and Group E. However, if parental education is 'high school' or 'some high school', most of them are from Group A and Group B.

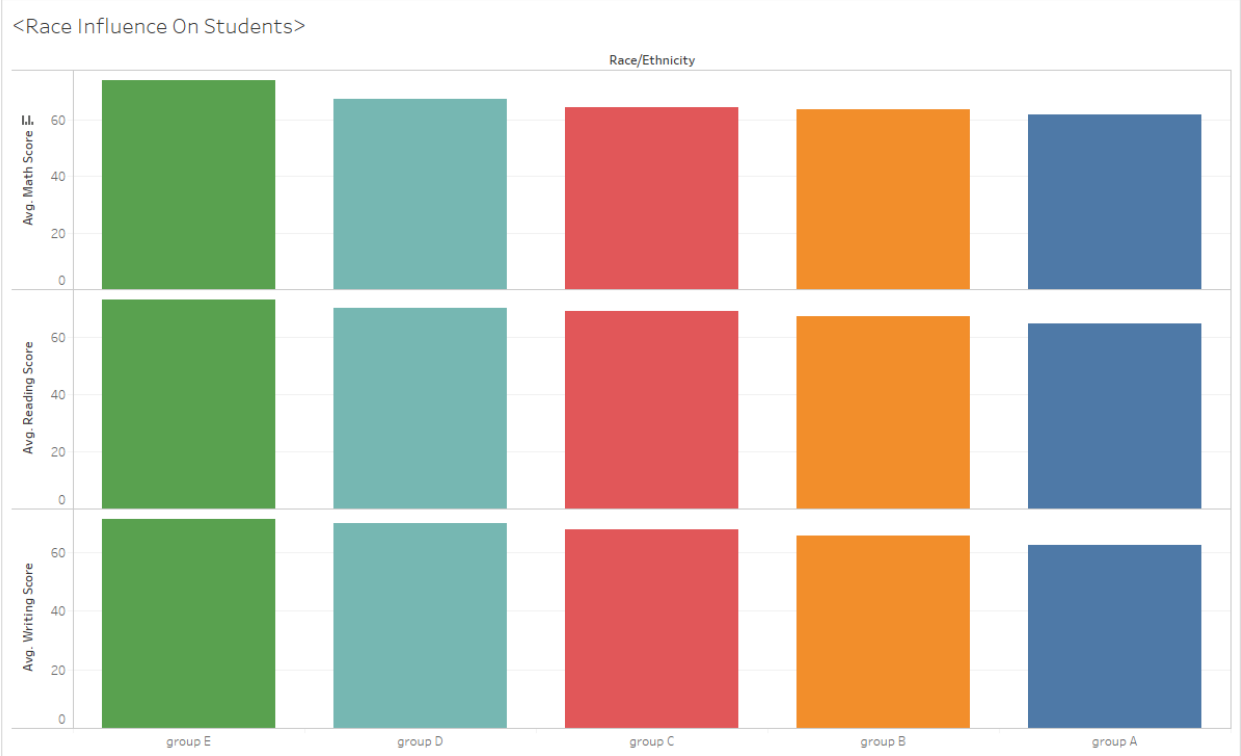



Made in Tableau

Let's dig deep to see if parent's education and race will affect student's grades.

As we found the parental education distribution among the 5 different race groups, we found parents from Group C, D, E obtained more higher education, Group A and B obtained more lower education. This education level directly affected the children's performance in the three exams. Students from Group C, D, and E achieved higher grades than students from Group A and B.

Source from Kaggle https://www.kaggle.com/spscientist/students-performance-in-exams

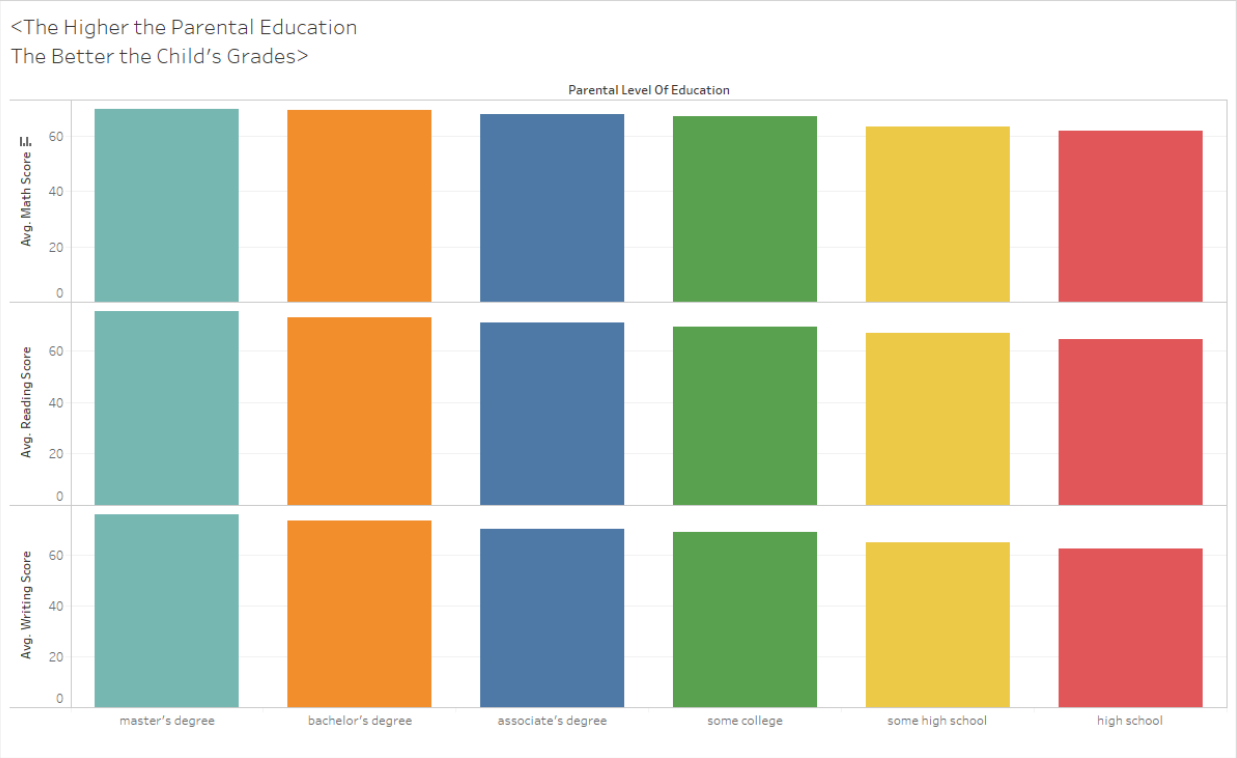



Made in Tableau

Then let's have a look at the relationship between parental education and exam performance.

We can find that if the parent held a master's degree, then his/her children will gain better grades in each of the exams. The lower the parental education, the worse the child's grades will be. This phenomenon exists in each of the exams.

Source from Kaggle https://www.kaggle.com/spscientist/students-performance-in-exams



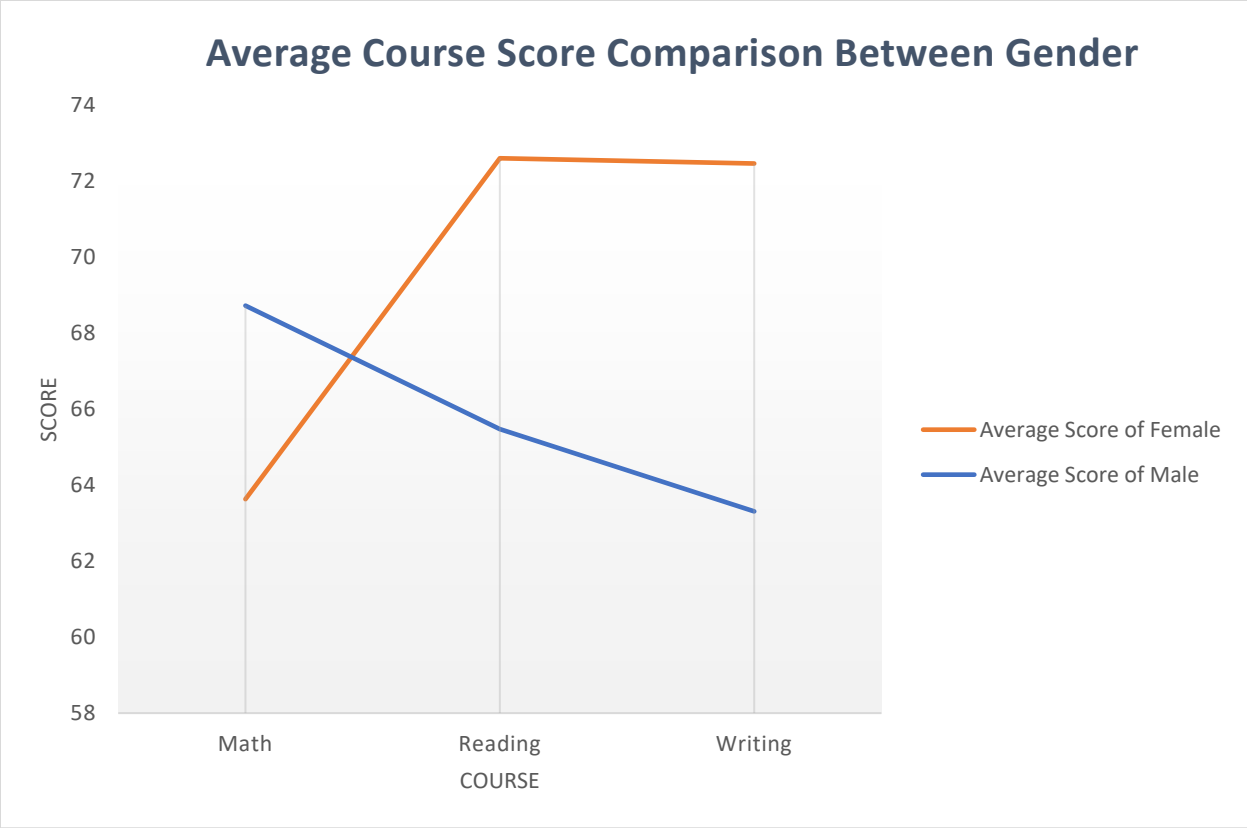

Made in Tableau

Thus, we found that parental education and race/ethnicity do affect student performance in the exams.

Let's do more analysis to see if gender will affect exam grades.

We calculated the average scores of the three exams for female and male respectively. In the graph below, it's obvious that male had a better performance in the math exam, however, female had better performances in both the reading exam and writing exam.

Source from Kaggle https://www.kaggle.com/spscientist/students-performance-in-exams

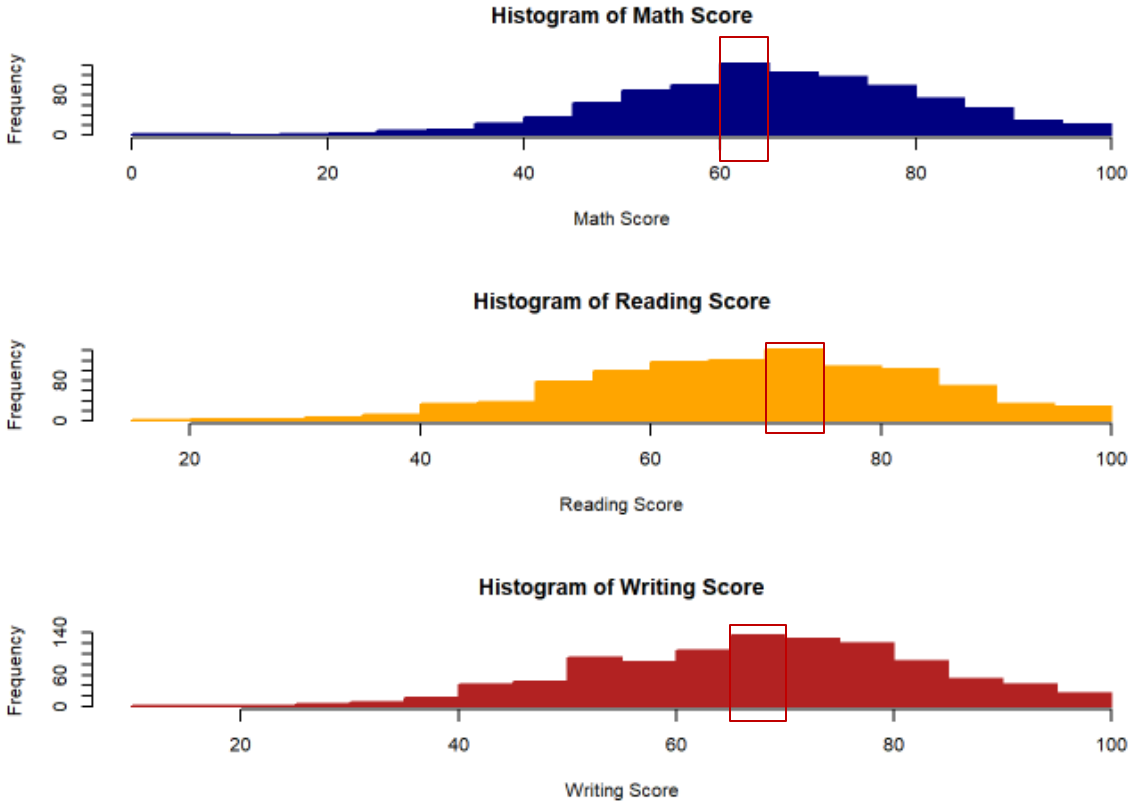



Made in Excel

However, is there any correlation among the three exams?

Let's first have a look at the grades distribution of the three exams. The three exams all have a left skewed distribution of the grades. Therefore, we can conclude, all of the mean of the three exams are lower than the median of the three exams. For math exam, most of the students obtained their scores between 40 and 90, even most of the students got scores from 60 to 65. For reading and writing exams, they have similar distributions, most of the scores are from 50 to 85. The highest occurrence happened from 70-75 for reading and 65-70 for writing exam.

Source from Kaggle https://www.kaggle.com/spscientist/students-performance-in-exams



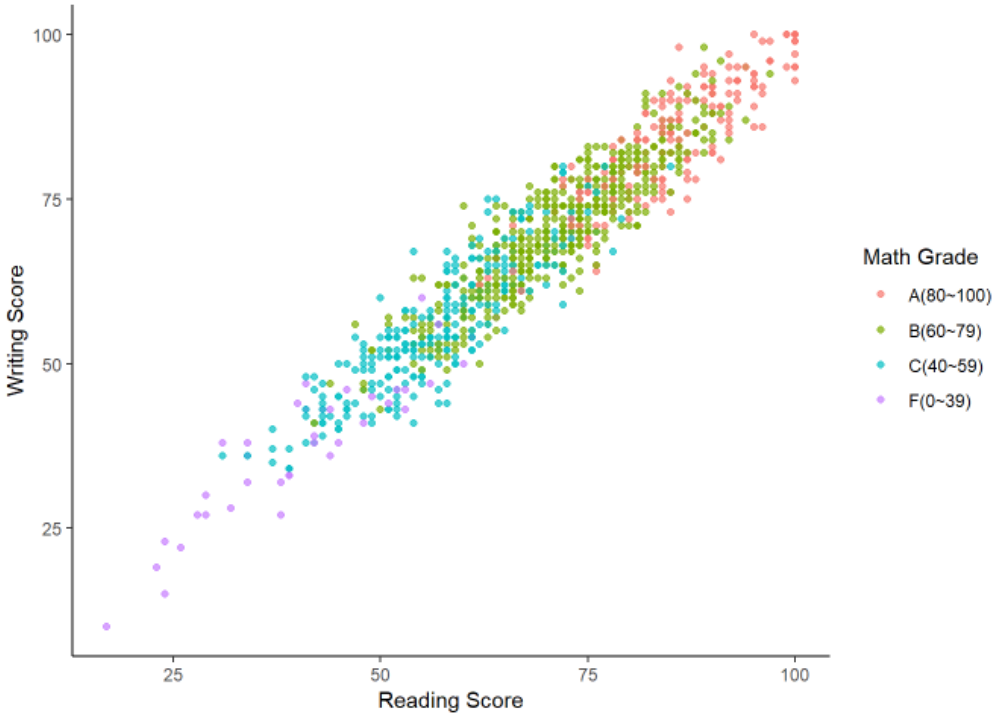

Made in R Studio

Then let's have a look at if the three exams have correlations.

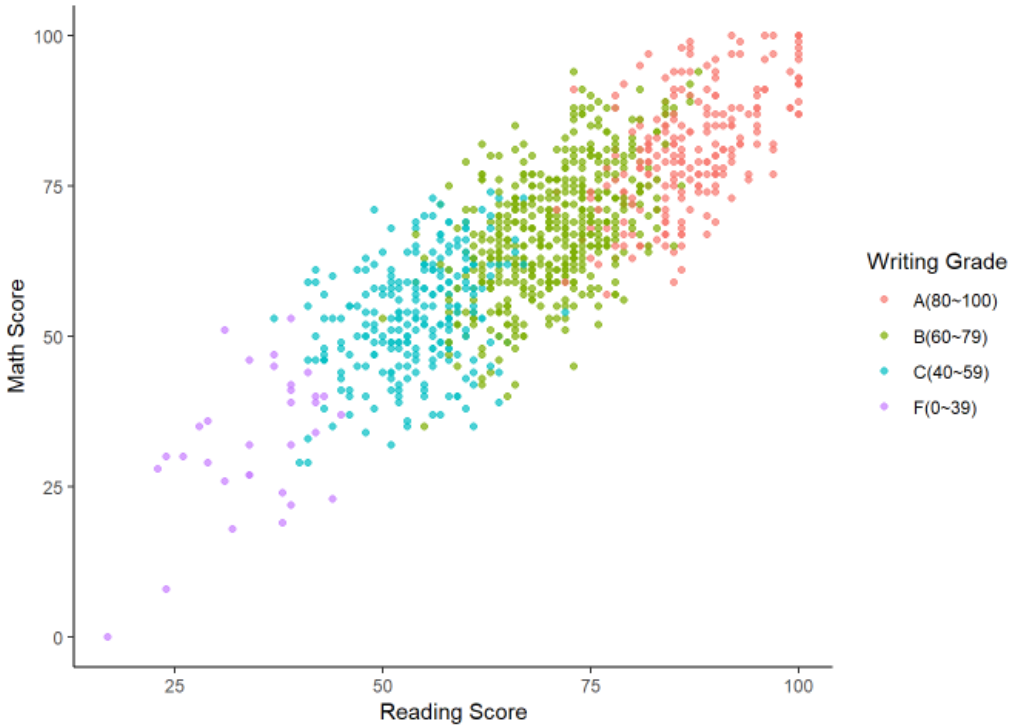
We can find that if a student got a high score in both reading exam and writing exam, some of the students also could get high scores in math exam, however, there are also many students they got a lower grade even if they got high scores in both reading and writing exams. The scores are randomly distributed, we can not separate them clearly.

However, look at the second picture, we can find if a student got a high score as well as a high score in math, then in most of the cases, the students got a high score in the writing exam. Therefore, we can see there is a positive relationship between the reading score and writing score. But we cannot conclude if they can get high math score according to their reading and writing scores.

Source from Kaggle https://www.kaggle.com/spscientist/students-performance-in-exams



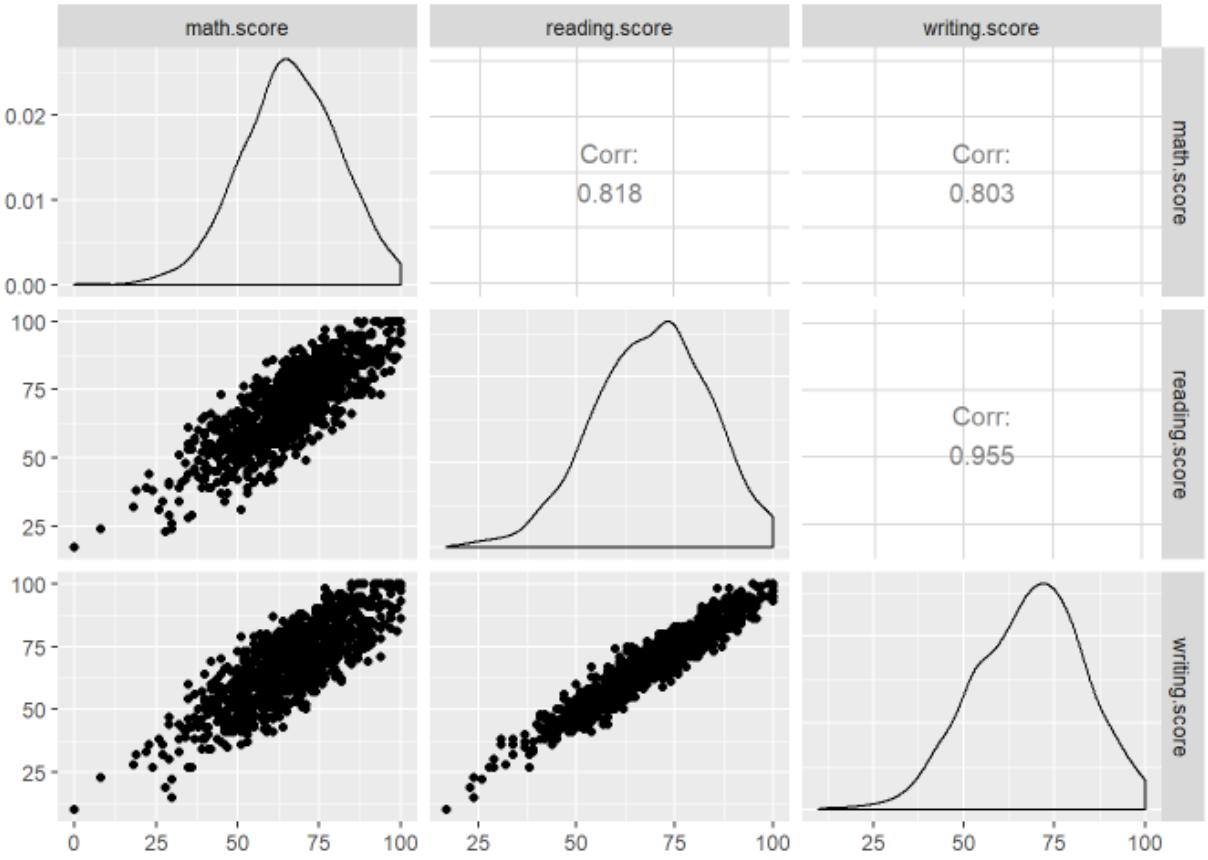

Made in R Studio




Made in R Studio

Source from Kaggle https://www.kaggle.com/spscientist/students-performance-in-exams

Here is a picture to illustrate the relationship between each of the two exam scores. As we found in the last two pictures, reading score and writing score has a high correlation. Math score also has high correlation with reading score or writing score, but not as high as reading to writing. And we cannot make a conclusion to say students with high score in reading and writing must have high score in math. We need to check the score for specific students. The reason should be reading and writing are under the same subjects, but math is under different thinking structure.




Made in R Studio

Thank you for reading the student performance analysis report.

Source from Kaggle https://www.kaggle.com/spscientist/students-performance-in-exams