STAT 2600 Final Project

Zachary Carifa

5 May 2025

Contents

Ask a Question

Between 2009 and 2010, Colorado standardized math test scores across Colorado were particularly low. To improve these marks, the Colorado Department of Education (CDE) employed an increased focus on math curriculum for beginning high school students. As a result, Colorado exhibited an increase in their statewide math scores. As part of a statewide evaluation, the CDE is interested in whether the increased emphasis on the math curriculum improved test results. In particular, the CDE would like to know if math rates increased due to the changes made. And if so, did an increased performance in math come at the cost of the reading and writing exams? The CDE would also like to know if the reading and writing rates were associated with math. And, how accurate of a predictor are each of the other two subjects at predicting math performance? The answers to these questions will provide valuable insight into statewide test performance, as well as evaluate the effects of the modified math curriculum. The questions inform whether the CDE's implemented changes had any positive influence on math scores, and if so, an associated negative effect on reading and writing scores. Also, determining if there is a correlation between math and the other two subjects allows CDE administrators to interpret current scores, and better predict future scores.

Acquire the Data

All data for this report comes from CDE-provided public access to schools' test results via the Colorado Information Marketplace API. It contains the 2009 and 2010 standardized test results for all Colorado 9th graders in math, reading, and writing. First, all necessary data is imported in three stages via the API endpoint, leaving us with a table for each subject. After combining the three subjects row by row, we are left with a single table of 1,911 rows. To simplify the dataset to only what is necessary, we must wrangle the data. Limiting our table to schools with a minimum of 31 students in 2009 and 2010 and not numbered zero, and only selecting the columns representing counts, subjects, and school numbers. We are left with a data frame, `schools_all`, containing a column for `year`, `school number`, `subject`, and each performance category(`noscore`, `unsatisfactory`, `partial`, `proficient, advanced`). Each observation now represents a unique combination of year, school number, and test subject.

```
## Rows: 1,488
## Columns: 8
## $ year           <dbl> 2009, 2009, 2009, 2010, 2010, 2010, 2009, 2009, 2009, 2~
## $ school_no      <dbl> 10, 10, 10, 10, 10, 10, 15, 15, 15, 15, 15, 15, 24, 24,~
## $ subject        <fct> MATH, READING, WRITING, MATH, READING, WRITING, MATH, R~
## $ noscore        <dbl> 0, 1, 0, 4, 2, 3, 0, 0, 0, 0, 0, 0, 18, 17, 16, 12, 9, ~
## $ unsatisfactory <dbl> 351, 141, 68, 313, 113, 83, 37, 8, 2, 32, 9, 9, 287, 76~
## $ partial        <dbl> 98, 195, 339, 139, 222, 348, 36, 20, 43, 48, 23, 56, 81~
## $ proficient     <dbl> 34, 151, 81, 51, 174, 78, 20, 66, 44, 31, 86, 50, 32, 1~
## $ advanced       <dbl> 5, 0, 0, 6, 1, 0, 4, 3, 8, 10, 3, 6, 4, 1, 2, 7, 1, 3, ~
```

Finally, to prepare ourselves for analysis, we randomly sample 120 schools from the population data set, resulting in a unique sample of `schools_all`, and `my_schools`.

```
## Rows: 720
## Columns: 8
## $ year           <dbl> 2009, 2009, 2009, 2010, 2010, 2010, 2009, 2009, 2009, 2~
## $ school_no      <dbl> 10, 10, 10, 10, 10, 10, 24, 24, 24, 24, 24, 24, 40, 40,~
## $ subject        <fct> MATH, READING, WRITING, MATH, READING, WRITING, MATH, R~
## $ noscore        <dbl> 0, 1, 0, 4, 2, 3, 18, 17, 16, 12, 9, 9, 1, 1, 1, 1, 1, ~
## $ unsatisfactory <dbl> 351, 141, 68, 313, 113, 83, 287, 76, 47, 280, 106, 59, ~
## $ partial        <dbl> 98, 195, 339, 139, 222, 348, 81, 180, 289, 134, 207, 33~
## $ proficient     <dbl> 34, 151, 81, 51, 174, 78, 32, 148, 68, 60, 172, 93, 4, ~
## $ advanced       <dbl> 5, 0, 0, 6, 1, 0, 4, 1, 2, 7, 1, 3, 1, 0, 0, 3, 1, 1, 1~
```

Analyze the Data

To determine whether the CDE's implemented changes to the 9th-grade curriculum affected performance in any way, we calculate the difference in the pass rate for each subject at every school. This is done by summing the students who scored proficient or advanced and dividing that by the total number of tested students for each year. We then subtract the 2009 passing rates from the 2010 rates to find our difference. For each subject, we have a data frame of four columns: `school_no`, `rate2009`, `rate2010`, and `diff`.

```
## Rows: 120
## Columns: 4
## $ school_no <dbl> 10, 24, 40, 146, 203, 298, 361, 378, 432, 604, 640, 664, 812~
## $ rate2009  <dbl> 0.1659836, 0.1658768, 0.2083333, 0.9285714, 0.1000000, 0.762~
## $ rate2010  <dbl> 0.15234375, 0.19393939, 0.13846154, 0.83050847, 0.11627907, ~
## $ diff      <dbl> -0.013639857, 0.028062617, -0.069871795, -0.098062954, 0.016~
```

Figure 1 depicts the bootstrap sampling distribution for the mean difference in writing pass rate from 2009 to 2010, along with 95% confidence bounds. We can see the *zero* is not within the confidence interval, showing that writing scores worsened from 2009 to 2010. In particular, we are 95% confident that the passing rate for writing decreased by between 2.2 and 4.9 percent for an average Colorado 9th grader.

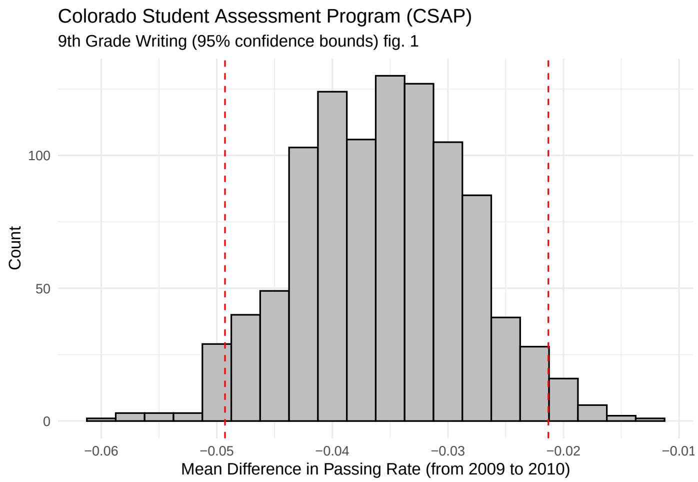


Figure 2 depicts the bootstrap sampling distribution for the mean difference in reading pass rate from 2009 to 2010, along with 95% confidence bounds. We can see the *zero* is within the confidence interval, showing that reading scores did not significantly change from 2009 to 2010. Because *zero* is within the confidence interval, we can reasonably infer any change in passing rate.

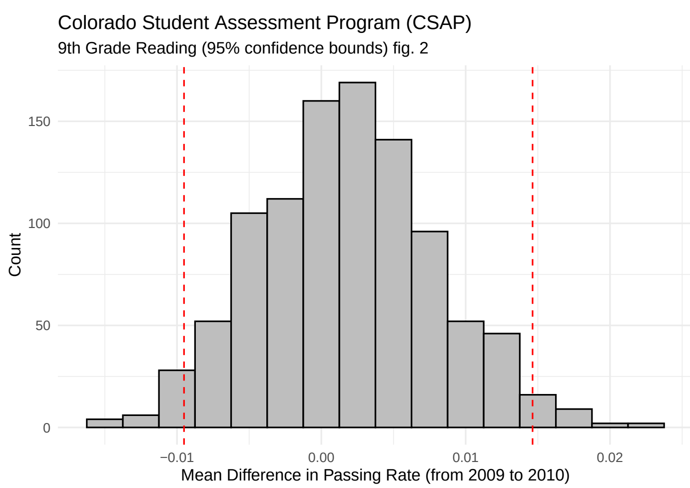
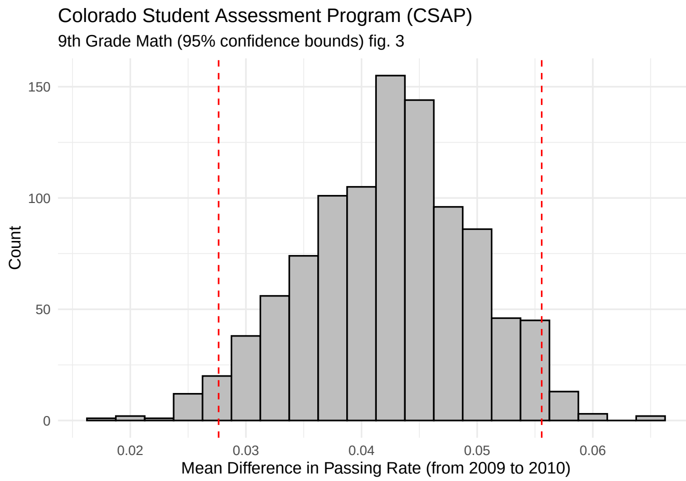
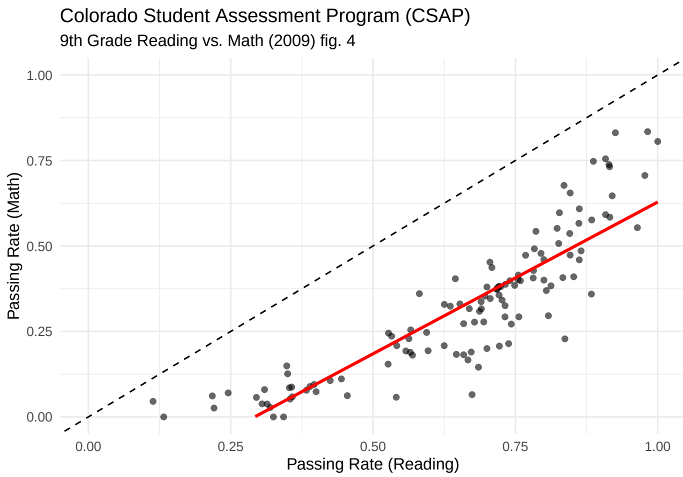


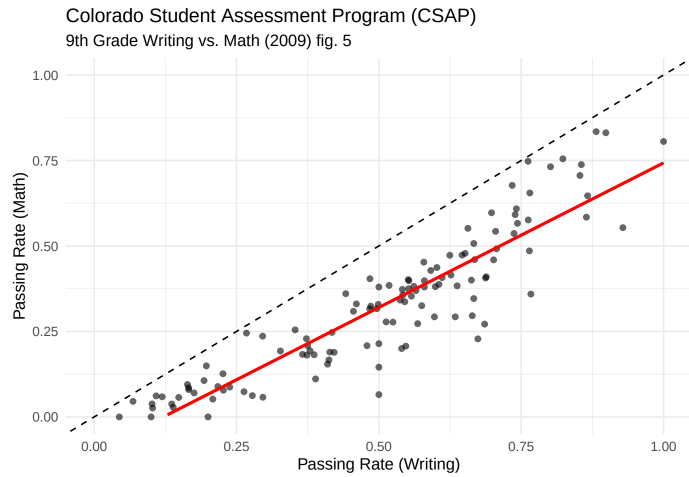
Figure 3 depicts the bootstrap sampling distribution for the mean difference in math pass rate from 2009 to 2010, along with 95% confidence bounds. We can see the *zero* is not within the confidence interval, showing that math scores improved from 2009 to 2010. In particular, we are 95% confident that the pass rate for math increased by between 2.7 and 5.6 percent for an average Colorado 9th grader.



Together, these results show that math scores improved significantly, while writing scores declined, and reading scores remained relatively stable. All in all, we can see that while the increased emphasis on the math curriculum served its purpose in increasing math test scores, it negatively impacted writing test scores and had no observable effect on reading.

To determine whether the passing rates for reading and writing are correlated to the passing rates in math, we must create two linear regression models: `math_reading_model` and `math_writing_model`, mapping math to each respective subject. These models are trained on the passing rates in each subject for each school in 2009. Below, we can visualize the association between math and each subject in 2009.





For both plots, we can visualize a strong correlation between passing rates in math and reading, as well as math and writing. The slope of the best-fit line (red) in both graphs appears close to 1 (dashed black line), but slightly more shallow. With our linear model, we confirm the visual findings and obtain the following regression equations:

- $Math = 0.888 * Reading - 0.259$
- $Math = 0.846 * Writing - 0.103$

The negative intercepts suggest that students do better in reading and writing than they do in math, and a slope less than one signifies that this difference in performance only gets larger as the passing rates for reading and writing get larger, in both graphs. So as writing and reading performance increase, the difference from performance in math also increases.

By using our model to predict the 2010 data, we can determine its accuracy. And in doing so, determine if reading and writing performance are viable predictors for performance in math. Computing the root mean squared error (RMSE) of our predicted scores from the true 2010 results, we find:

```
## # A tibble: 1 x 1
##   RMSE_Writing
##          <dbl>
## 1        0.100
```

```
## # A tibble: 1 x 1
##   RMSE_Writing
##          <dbl>
## 1       0.0949
```

These RMSEs indicate that our models are very strong. It means if we know a school's passing rate for writing or reading, we can predict the passing rate for math to be within 0.1 and 0.95 percentage points, respectively.

Advise on Results

The analysis provides a lot of valuable insight that can heavily inform the CDE's understanding of student performance in all subjects, and influence the modification of standardized testing in the future. Based on the error of both our models, as well as the plots of math performance versus each of the other subjects, it is obvious that they are correlated. This tells us that if students do better in reading or writing, they are very likely to do better in math. However, we are also able to see that performance in reading and writing is better than math overall. This is based on the negative intercepts. From our bootstrap resample, we were able to visualize the impact an increased focus on the math curriculum had on each subject's performance. It resulted in a significant improvement in statewide math scores while writing scores got significantly worse (reading scores were unaffected). Although this adjustment to the math curriculum improved math performance and did not harm reading, it negatively impacted writing scores. There is clear success in math performance, however further work must be done to improve performance in all subjects and provide equal focus. Luckily, we are now able to predict future test scores fairly accurately, which will serve as an invaluable tool for the CDE to use in the future. Despite the profound results we gathered, the analysis has its limitations. With a sample size of 120 schools, we may not have been able to capture heavy influences within our analysis. Perhaps a region of Colorado of also 120 schools exhibited opposite results to ours, which strongly changed our data. And without implementing demographics within our analysis, we may be neglecting powerful influences on student performance.

Answer the Question

Our analysis proved to be fruitful in answering the questions posed. Due to the implementation of an increased math focus, our data clearly shows that math performance improved significantly, by nearly 5%, from 2009 to 2010. However, this success came at the expense of writing performance, whose scores experienced a significant decline (Reading was unaffected). We also found that math performance can be well predicted by performance in reading and writing. Our linear regression models based on 2009 reading and writing scores estimated 2010 math pass rates with an average error of 6.9% and 7.4%, respectively. Further analysis could come from including more variables within the data. By adding in variables like income, sex, or race, we can further complexify and improve our model. The applications of this analysis are endless and limitless to subject and grade level. If implemented, it could positively impact test scores of students of all ages, at any school.