Standardized Test Performance of U.S. Schools

Shankar Prabhu

July 2020

In this section, we will take a look at how the testing statistics in reading & math of U.S schools have changed over the past 30 years.

```
data = read.csv(
  "C:/Users/shank/Documents/Data_Science_Projects/education/states_all_extended.csv",
  header = TRUE, sep = ",")
```

```
STATE = data$STATE
YEAR = data$YEAR
ENROLL = data$ENROLL
test_nums = data[grep('..._._._MATHEMATICS$|..._.._._MATHEMATICS$|..._._._READING$|..._.._._READING$', n

test_scores = cbind(STATE, YEAR, ENROLL, test_nums)

summary(test_scores)
```

```
##         STATE           YEAR          ENROLL        G04_A_A_READING
##  ALABAMA   :  33   Min.   :1986   Min.   :  43866   Min.   :179.0
##  ALASKA    :  33   1st Qu.:1994   1st Qu.: 264514   1st Qu.:215.0
##  ARIZONA   :  33   Median :2002   Median : 649934   Median :220.0
##  ARKANSAS  :  33   Mean   :2002   Mean   : 917542   Mean   :218.8
##  CALIFORNIA:  33   3rd Qu.:2010   3rd Qu.:1010532   3rd Qu.:224.0
##  COLORADO  :  33   Max.   :2019   Max.   :6307022   Max.   :237.0
##  (Other)   :1517                  NA's   :491       NA's   :1065
##  G04_A_A_MATHEMATICS G04_A_M_READING G04_A_M_MATHEMATICS G04_A_F_READING
##  Min.   :192.0       Min.   :174.0   Min.   :191.0       Min.   :183.0
##  1st Qu.:232.0       1st Qu.:211.0   1st Qu.:233.0       1st Qu.:218.2
##  Median :238.0       Median :217.0   Median :240.0       Median :224.0
##  Mean   :236.3       Mean   :215.3   Mean   :237.3       Mean   :222.4
##  3rd Qu.:242.0       3rd Qu.:221.0   3rd Qu.:244.0       3rd Qu.:227.0
##  Max.   :253.0       Max.   :234.0   Max.   :255.0       Max.   :239.0
##  NA's   :1150        NA's   :1065    NA's   :1150        NA's   :1065
##  G04_A_F_MATHEMATICS G04_WH_A_READING G04_WH_A_MATHEMATICS G04_BL_A_READING
##  Min.   :192.0       Min.   :214.0    Min.   :232.0        Min.   :188.0
##  1st Qu.:231.0       1st Qu.:226.0    1st Qu.:245.0        1st Qu.:200.0
##  Median :237.0       Median :229.0    Median :248.0        Median :204.0
##  Mean   :235.3       Mean   :230.1    Mean   :248.3        Mean   :204.6
##  3rd Qu.:241.0       3rd Qu.:233.0    3rd Qu.:251.0        3rd Qu.:209.0
##  Max.   :253.0       Max.   :260.0    Max.   :276.0        Max.   :228.0
##  NA's   :1150        NA's   :1450     NA's   :1450         NA's   :1489
##  G04_BL_A_MATHEMATICS G04_HI_A_READING G04_HI_A_MATHEMATICS G04_AS_A_READING
```

```
##  Min.   :211.0        Min.   :196.0    Min.   :218.0        Min.   :199.0
##  1st Qu.:219.0        1st Qu.:204.0    1st Qu.:226.2        1st Qu.:228.0
##  Median :224.0        Median :209.0    Median :230.0        Median :236.0
##  Mean   :223.3        Mean   :209.0    Mean   :230.3        Mean   :235.8
##  3rd Qu.:227.0        3rd Qu.:212.8    3rd Qu.:234.0        3rd Qu.:244.0
##  Max.   :241.0        Max.   :230.0    Max.   :246.0        Max.   :255.0
##  NA's   :1486         NA's   :1465     NA's   :1465         NA's   :1551
##  G04_AS_A_MATHEMATICS G04_AM_A_READING G04_AM_A_MATHEMATICS G04_HP_A_READING
##  Min.   :227.0        Min.   :169.0    Min.   :208.0        Min.   :184.0
##  1st Qu.:252.8        1st Qu.:191.8    1st Qu.:218.5        1st Qu.:197.8
##  Median :258.5        Median :198.0    Median :222.0        Median :202.0
##  Mean   :257.8        Mean   :197.0    Mean   :222.5        Mean   :202.9
##  3rd Qu.:264.0        3rd Qu.:204.0    3rd Qu.:227.0        3rd Qu.:212.0
##  Max.   :275.0        Max.   :223.0    Max.   :238.0        Max.   :216.0
##  NA's   :1547         NA's   :1651     NA's   :1652         NA's   :1699
##  G04_HP_A_MATHEMATICS G04_TR_A_READING G04_TR_A_MATHEMATICS G08_A_A_READING
##  Min.   :212.0        Min.   :203.0    Min.   :228.0        Min.   :236.0
##  1st Qu.:222.5        1st Qu.:220.0    1st Qu.:238.0        1st Qu.:260.0
##  Median :226.0        Median :224.0    Median :242.0        Median :265.0
##  Mean   :226.1        Mean   :224.3    Mean   :242.4        Mean   :263.6
##  3rd Qu.:230.5        3rd Qu.:228.0    3rd Qu.:247.0        3rd Qu.:268.0
##  Max.   :236.0        Max.   :253.0    Max.   :259.0        Max.   :280.0
##  NA's   :1700         NA's   :1532     NA's   :1532         NA's   :1153
##  G08_A_A_MATHEMATICS G08_A_M_READING G08_A_M_MATHEMATICS G08_A_F_READING
##  Min.   :231.0       Min.   :229.0   Min.   :230.0       Min.   :241.0
##  1st Qu.:272.0       1st Qu.:254.0   1st Qu.:273.0       1st Qu.:265.0
##  Median :280.5       Median :260.0   Median :281.0       Median :270.0
##  Mean   :278.3       Mean   :258.4   Mean   :278.7       Mean   :268.9
##  3rd Qu.:286.0       3rd Qu.:263.0   3rd Qu.:286.0       3rd Qu.:273.0
##  Max.   :301.0       Max.   :276.0   Max.   :301.0       Max.   :286.0
##  NA's   :1113        NA's   :1153    NA's   :1113        NA's   :1153
##  G08_A_F_MATHEMATICS G08_WH_A_READING G08_WH_A_MATHEMATICS G08_BL_A_READING
##  Min.   :233.0       Min.   :256.0    Min.   :272.0        Min.   :235.0
##  1st Qu.:272.0       1st Qu.:270.0    1st Qu.:288.0        1st Qu.:244.0
##  Median :280.0       Median :272.0    Median :292.0        Median :247.0
##  Mean   :277.8       Mean   :272.9    Mean   :291.8        Mean   :247.6
##  3rd Qu.:285.0       3rd Qu.:275.0    3rd Qu.:295.0        3rd Qu.:251.0
##  Max.   :300.0       Max.   :300.0    Max.   :323.0        Max.   :273.0
##  NA's   :1113        NA's   :1450     NA's   :1450         NA's   :1493
##  G08_BL_A_MATHEMATICS G08_HI_A_READING G08_HI_A_MATHEMATICS G08_AS_A_READING
##  Min.   :248.0        Min.   :239.0    Min.   :255          Min.   :257.0
##  1st Qu.:256.0        1st Qu.:251.0    1st Qu.:266          1st Qu.:274.0
##  Median :260.0        Median :255.0    Median :270          Median :281.0
##  Mean   :260.4        Mean   :254.8    Mean   :270          Mean   :279.8
##  3rd Qu.:264.0        3rd Qu.:258.0    3rd Qu.:274          3rd Qu.:286.0
##  Max.   :277.0        Max.   :277.0    Max.   :287          Max.   :298.0
##  NA's   :1494         NA's   :1469     NA's   :1467         NA's   :1562
##  G08_AS_A_MATHEMATICS G08_AM_A_READING G08_AM_A_MATHEMATICS G08_HP_A_READING
##  Min.   :277.0        Min.   :229.0    Min.   :240.0        Min.   :234.0
##  1st Qu.:301.0        1st Qu.:242.0    1st Qu.:257.8        1st Qu.:244.2
##  Median :309.0        Median :247.0    Median :261.0        Median :248.5
##  Mean   :306.9        Mean   :246.7    Mean   :261.0        Mean   :248.3
##  3rd Qu.:314.0        3rd Qu.:252.0    3rd Qu.:265.0        3rd Qu.:253.5
##  Max.   :332.0        Max.   :261.0    Max.   :275.0        Max.   :259.0
```

```
##  NA's   :1558         NA's   :1654     NA's   :1655         NA's   :1701
##  G08_HP_A_MATHEMATICS G08_TR_A_READING G08_TR_A_MATHEMATICS
##  Min.   :250.0        Min.   :249.0    Min.   :266.0
##  1st Qu.:258.0        1st Qu.:263.0    1st Qu.:279.0
##  Median :266.0        Median :267.0    Median :284.0
##  Mean   :264.8        Mean   :268.1    Mean   :284.1
##  3rd Qu.:269.0        3rd Qu.:272.0    3rd Qu.:289.0
##  Max.   :276.0        Max.   :291.0    Max.   :305.0
##  NA's   :1702         NA's   :1574     NA's   :1570
```

Based on the summary, it still looks like we will have a lot of missing data. Similar to the reasoning used in the enrollment_demos exploration file, we will have to go ahead and set these missing values to 0. Unfortunately, this means losing out on some interesting trend information.

First, let's take a look at how overall reading and math scores have changed over time. To do this, we can create a dataset which shows the average math and reading scores in the state over time. This could be prone to some adjustment errors if enrollment demographics shift over grade levels, but earlier analysis revealed this is not the case so we should be good to go.

```
test_scores[is.na(test_scores)] = 0
test_scores$READING <- rowSums(data[grep('..._._._READING$|..._.._._READING', names(data))])
test_scores$MATHEMATICS <- rowSums(data[grep('..._._._MATHEMATICS$|..._.._._MATHEMATICS$', names(data))]

total_scores_by_year = setNames(aggregate(cbind(test_scores$MATHEMATICS,
                                                test_scores$READING,
                                                test_scores$ENROLL)
                                          , by=list(YEAR=test_scores$YEAR),
                                          FUN = sum, na.rm = TRUE), c("YEAR", "Avg. Math Score",
                                                                      "Avg. Reading Scores",
                                                                      "Total Enrollment"))
summary(total_scores_by_year)
```

```
##       YEAR      Avg. Math Score Avg. Reading Scores Total Enrollment
##  Min.   :1986   Min.   :    0   Min.   :   0        Min.   :       0
##  1st Qu.:1994   1st Qu.:    0   1st Qu.:   0        1st Qu.:       0
##  Median :2002   Median :    0   Median :   0        Median :46432782
##  Mean   :2002   Mean   : 1244   Mean   :1153        Mean   :34032451
##  3rd Qu.:2010   3rd Qu.:    0   3rd Qu.:   0        3rd Qu.:48266209
##  Max.   :2019   Max.   :10321   Max.   :9515        Max.   :48571827
```

Looks like there is a lot of missing data in the scores section, and the only year for which their is really good data is 2013. All other years either have no testing data or limited data to where it might not be usable.

So, unfortunately, we will not be able to look at test scores over time, we should confine our analysis to 2013 to see what other information we are able to gleam.

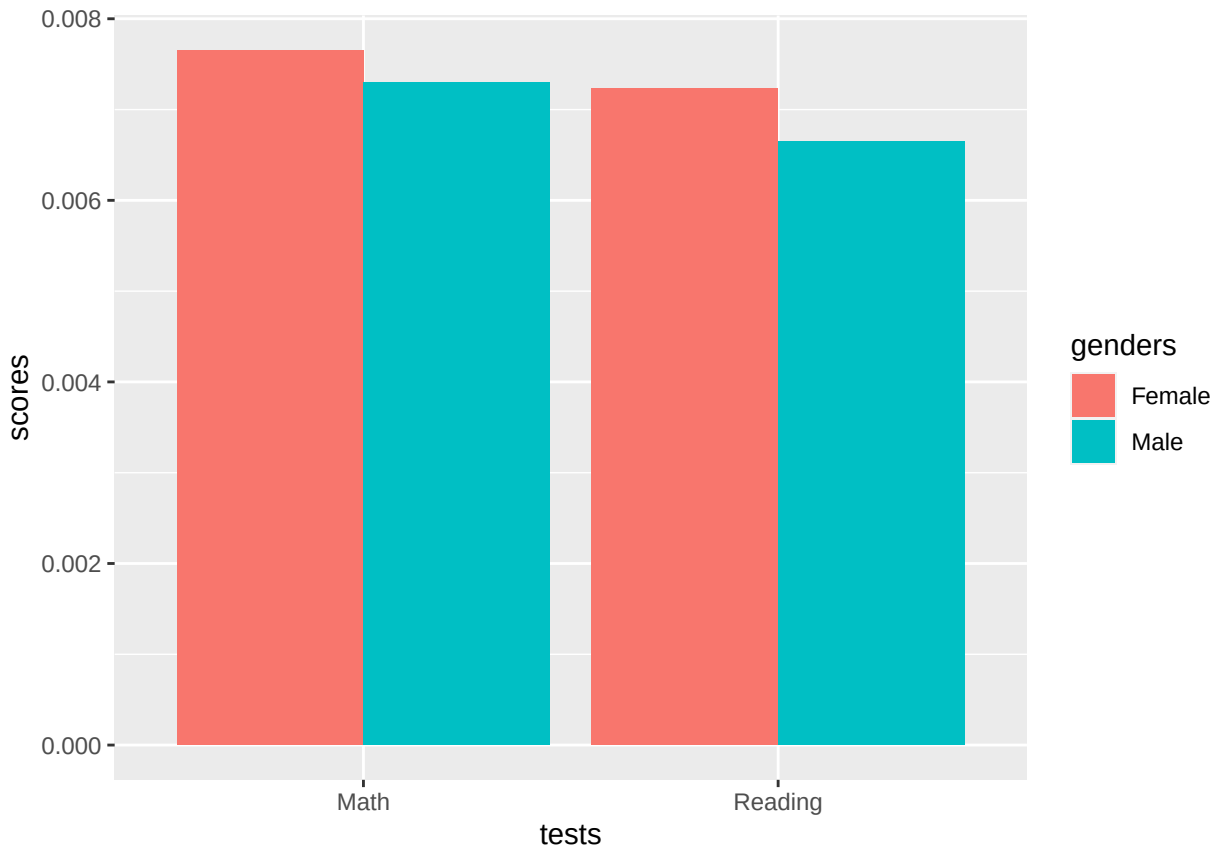
```
data = read.csv(
  "C:/Users/shank/Documents/Data_Science_Projects/education/states_all_extended.csv",
  header = TRUE, sep = ",")

year_2013 = data[data$YEAR == 2013,]
year_2013[is.na(year_2013)] = 0

# we should have columns for enrollment by race/gender and then columns for total reading scores and to
```

```
AM_4 = year_2013$G04_AM_F + year_2013$G04_AM_M
AS_4 = year_2013$G04_AS_F + year_2013$G04_AS_M
HI_4 = year_2013$G04_HI_F + year_2013$G04_HI_M
BL_4 = year_2013$G04_BL_F + year_2013$G04_BL_M
WH_4 = year_2013$G04_WH_F + year_2013$G04_WH_M
HP_4 = year_2013$G04_HP_F + year_2013$G04_HP_M
TR_4 = year_2013$G04_TR_F + year_2013$G04_TR_M
FEM_4 = year_2013$G04_AM_F + year_2013$G04_AS_F +
  year_2013$G04_HI_F + year_2013$G04_BL_F +
  year_2013$G04_WH_F + year_2013$G04_HP_F +
  year_2013$G04_TR_F
MAL_4 = year_2013$G04_AM_M + year_2013$G04_AS_M +
  year_2013$G04_HI_M + year_2013$G04_BL_M +
  year_2013$G04_WH_M + year_2013$G04_HP_M +
  year_2013$G04_TR_M

AM_8 = year_2013$G08_AM_F + year_2013$G08_AM_M
AS_8 = year_2013$G08_AS_F + year_2013$G08_AS_M
HI_8 = year_2013$G08_HI_F + year_2013$G08_HI_M
BL_8 = year_2013$G08_BL_F + year_2013$G08_BL_M
WH_8 = year_2013$G08_WH_F + year_2013$G08_WH_M
HP_8 = year_2013$G08_HP_F + year_2013$G08_HP_M
TR_8 = year_2013$G08_TR_F + year_2013$G08_TR_M
FEM_8 = year_2013$G08_AM_F + year_2013$G08_AS_F +
  year_2013$G08_HI_F + year_2013$G08_BL_F +
  year_2013$G08_WH_F + year_2013$G08_HP_F +
  year_2013$G08_TR_F
MAL_8 = year_2013$G08_AM_M + year_2013$G08_AS_M +
  year_2013$G08_HI_M + year_2013$G08_BL_M +
  year_2013$G08_WH_M + year_2013$G08_HP_M +
  year_2013$G08_TR_M

testing_2013 = year_2013[grep('..._._._MATHEMATICS$|..._.._._MATHEMATICS$|..._._._READING$|..._.._._REA


testing_2013 = cbind(year_2013$YEAR, year_2013$STATE, AM_4,
AS_4, HI_4, BL_4, WH_4, HP_4, TR_4, FEM_4, MAL_4, AM_8 , AS_8,
HI_8, BL_8, WH_8, HP_8, TR_8, FEM_8, MAL_8, testing_2013)
```

Let's ignore states for now, and just go ahead and sum the data into one row.

```
test_2013_row = testing_2013[,-c(1:2)]
test_2013_row = colSums(test_2013_row)
AM = test_2013_row[1] + test_2013_row[10]
AS = test_2013_row[2] + test_2013_row[11]
HI = test_2013_row[3] + test_2013_row[12]
BL = test_2013_row[4] + test_2013_row[13]
WH = test_2013_row[5] + test_2013_row[14]
HP = test_2013_row[6] + test_2013_row[15]
TR = test_2013_row[7] + test_2013_row[16]
FEM = test_2013_row[8] + test_2013_row[17]
MAL = test_2013_row[9] + test_2013_row[18]
fem_reading = test_2013_row[23] + test_2013_row[43]
```

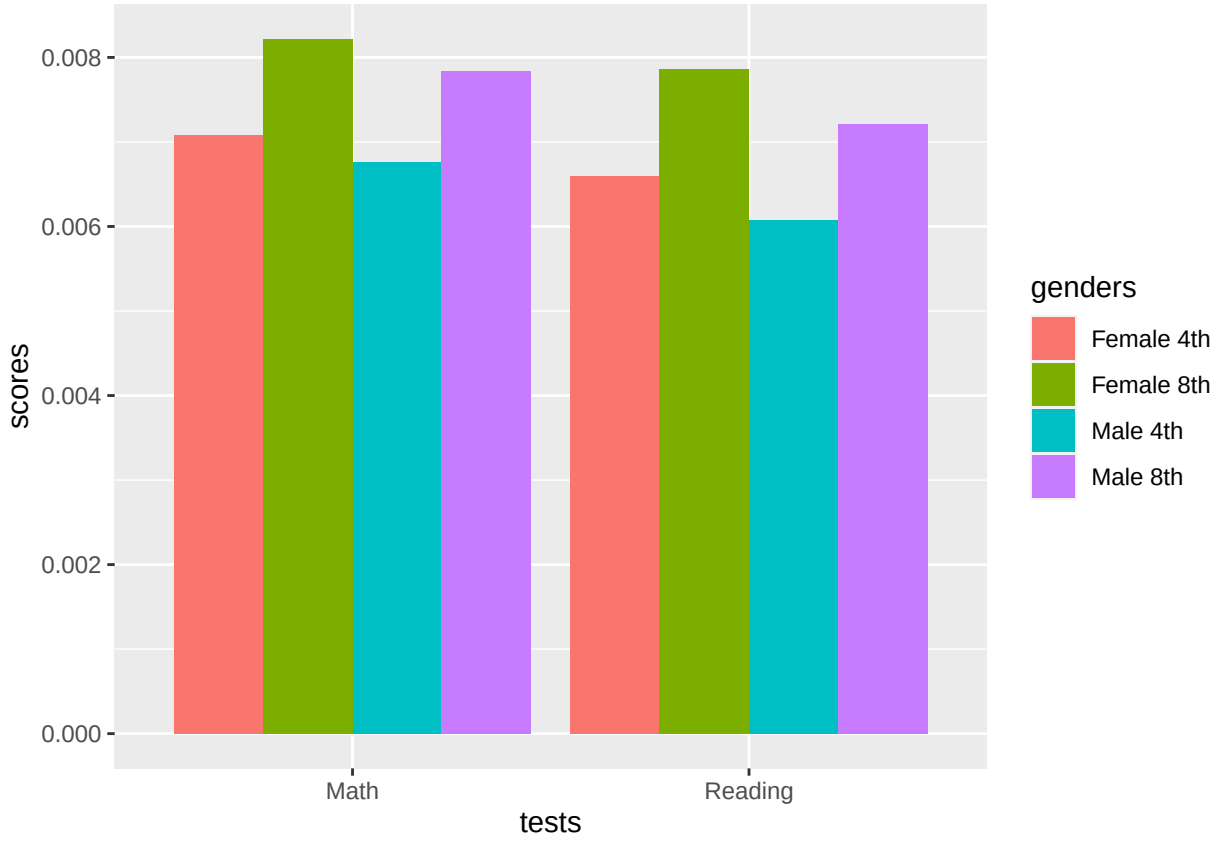
```
mal_reading = test_2013_row[21] + test_2013_row[41]
fem_math = test_2013_row[24] + test_2013_row[44]
mal_math = test_2013_row[22] + test_2013_row[42]

genders = rep(c("Female", "Male"), 2)
tests = c(rep("Reading", 2), rep("Math", 2))
scores = c(fem_reading/FEM, mal_reading/MAL, fem_math/FEM, mal_math/MAL)

scores_by_gender = data.frame(genders, tests, scores)
rownames(scores_by_gender) = c()

library(ggplot2)
ggplot(scores_by_gender, aes(fill=genders, y=scores, x=tests)) +
    geom_bar(position="dodge", stat="identity")
```



Looks like females, on average, perform better on both math and reading, with the reading difference being greater than the math difference. I wonder if these performance gaps between the genders differ from grade 4 to grade 8.

```
fem_reading_04 = test_2013_row[23]
fem_reading_08 = test_2013_row[43]
mal_reading_04 = test_2013_row[21]
mal_reading_08 = test_2013_row[41]
fem_math_04 = test_2013_row[24]
fem_math_08 = test_2013_row[44]
mal_math_04 = test_2013_row[22]
```

```
mal_math_08 = test_2013_row[42]

fem_04_pop = test_2013_row[8]
fem_08_pop = test_2013_row[17]
mal_04_pop = test_2013_row[9]
mal_08_pop = test_2013_row[18]

genders = rep(c("Female 4th", "Male 4th", "Female 8th", "Male 8th"), 2)
tests = c(rep("Reading", 4), rep("Math", 4))
scores = c(fem_reading_04/fem_04_pop,
           mal_reading_04/mal_04_pop,
           fem_reading_08/fem_08_pop,
           mal_reading_08/mal_08_pop,
           fem_math_04/fem_04_pop,
           mal_math_04/mal_04_pop,
           fem_math_08/fem_08_pop,
           mal_math_08/mal_08_pop)

scores_by_gender = data.frame(genders, tests, scores)
rownames(scores_by_gender) = c()

library(ggplot2)
ggplot(scores_by_gender, aes(fill=genders, y=scores, x=tests)) +
    geom_bar(position="dodge", stat="identity")
```

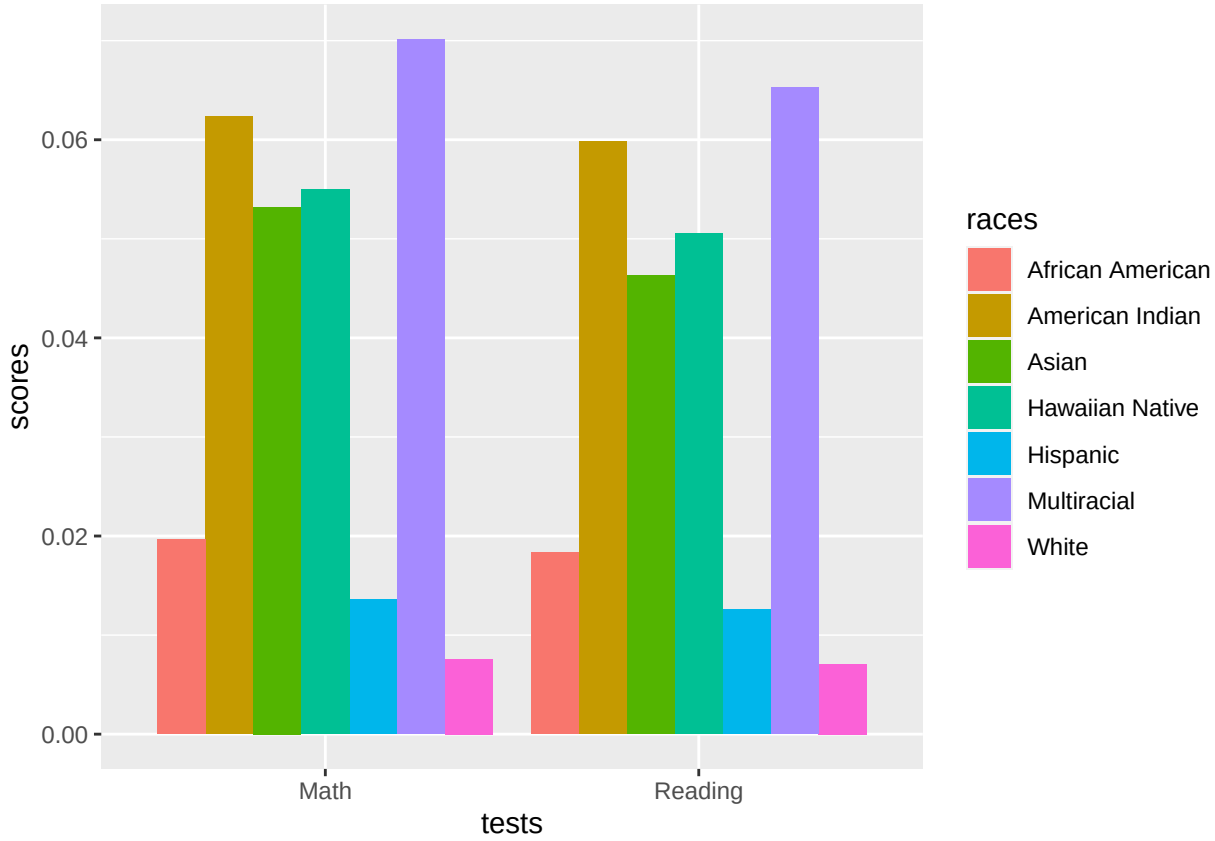


Looks like the gap between females and males increases by 26% for reading and by 22% for math between 4th

and 8th grade. I wonder what educational factors are key drivers in both the initial creation and widening of this gap over time, and what policies or educational strategies can be deployed to improve upon discrepancies.

Now that we have got a good understanding of the statistics around standardized test performance by gender, let us examine the differences by race.

```
AM = test_2013_row[1] + test_2013_row[10]
AS = test_2013_row[2] + test_2013_row[11]
HI = test_2013_row[3] + test_2013_row[12]
BL = test_2013_row[4] + test_2013_row[13]
WH = test_2013_row[5] + test_2013_row[14]
HP = test_2013_row[6] + test_2013_row[15]
TR = test_2013_row[7] + test_2013_row[16]

AM_reading = test_2013_row[33] + test_2013_row[53]
AS_reading = test_2013_row[31] + test_2013_row[51]
HI_reading = test_2013_row[29] + test_2013_row[49]
BL_reading = test_2013_row[27] + test_2013_row[47]
WH_reading = test_2013_row[25] + test_2013_row[45]
HP_reading = test_2013_row[35] + test_2013_row[55]
TR_reading = test_2013_row[37] + test_2013_row[57]

AM_math = test_2013_row[34] + test_2013_row[54]
AS_math = test_2013_row[32] + test_2013_row[52]
HI_math = test_2013_row[30] + test_2013_row[50]
BL_math = test_2013_row[28] + test_2013_row[48]
WH_math = test_2013_row[26] + test_2013_row[46]
HP_math = test_2013_row[36] + test_2013_row[56]
TR_math = test_2013_row[38] + test_2013_row[58]


races = rep(c("American Indian", "Asian", "Hispanic",
              "African American", "White", "Hawaiian Native", "Multiracial"), 2)
tests = c(rep("Reading", 7), rep("Math", 7))
scores = c(AM_reading/AM, AS_reading/AS, HI_reading/HI,
           BL_reading/BL, WH_reading/WH, HP_reading/HP,
           TR_reading/TR, AM_math/AM, AS_math/AS, HI_math/HI,
           BL_math/BL, WH_math/WH, HP_math/HP, TR_math/TR)

scores_by_race = data.frame(races, tests, scores)
rownames(scores_by_race) = c()

library(ggplot2)
ggplot(scores_by_race, aes(fill=races, y=scores, x=tests)) +
    geom_bar(position="dodge", stat="identity")
```



The bar graphs show clear distinctions between different racial categories and their performances on reading and math standarized tests. However, I would put down most of the discrepancy to data quality issues, since for example, multiracial had the most missing enrollment data (maybe inflating their testing numbers slightly), while categories like "White", "Hispanic", and "African American" had missing test score data but nearly all of their enrollment data filled in, which could possibly deflate their testing numbers.

Overall, I am concerned about the quality of this data even though it comes from reputable sources like the U.S. census and standardized testing boards. I wonder if some of the data collection issues are in themselves caused by racial or economic factors that make it harder/easier to collect data. It seems a worthy investment to focus on education data collection so that policymakers have more accurate information so they can craft better legislation and improve U.S. schools.