Reading Retrograde: An Analysis of Reading Capabilities Among American Youth

Background

In December 2019, the New York Times published an article describing the dismal state of reading ability among American students. Authors Erica Green and Dana Goldstein assert that "America's fourth and eighth graders are losing ground in their ability to read literature and academic texts." (Green & Goldstein 2019). They note that two out of three children did not meet the standards for reading proficiency set by the National Assessment of Educational Progress. The average eighth grade reading score declined in more than half of the states in 2019 as compared with 2017. Only 34% of eighth graders were proficient in reading in 2019 (Green and Goldstein).

Although, American students scored lower on average in 2019 as compared with 2017, is the trend in reading scores generally decreasing over the past two decades, or have the improvements to the American education system made a difference over the years? The National Assessment of Educational Progress (NAEP), often called the Nation's Report Card, is the largest assessment of American student's capability. Since 1969, the National Center for Education Statistics (NCES), a branch of the Department of Education, has administered the NAEP to asses the skill level of 4th, 8th and 12th grade students in math, science, and reading among other subjects ("NAEP Report Cards"). It takes approximately two years to develop the framework for the test, during which time the National Assessment Governing Board decides the standards of testing for each subject. The NCES then develops the assessment items in accordance with the framework and decides what questions to include on the survey portion of the assessment, which collects data on student, educators, and school demographics.

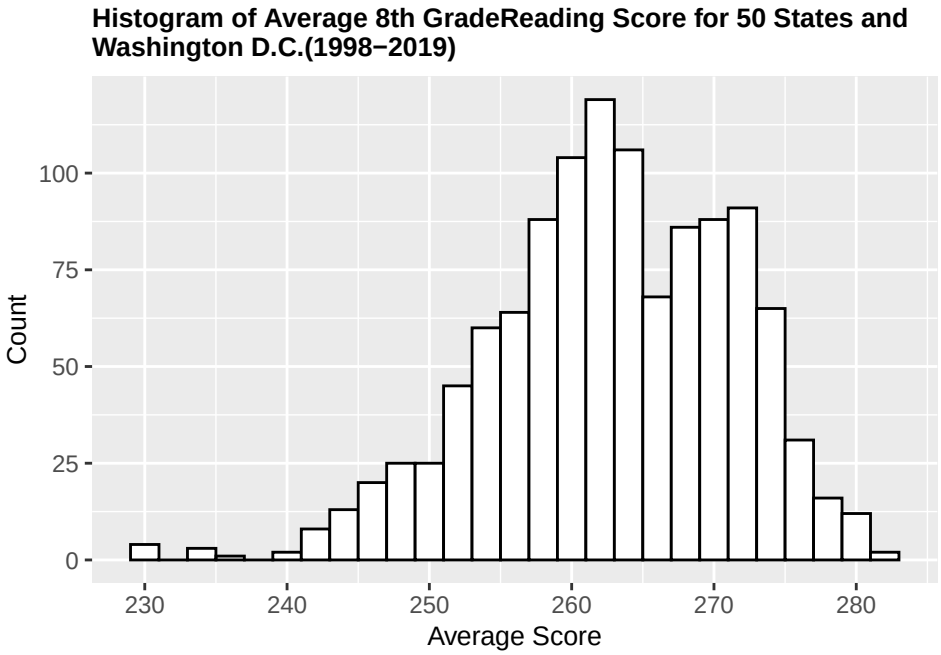
The NAEP is administered to a sample of about 600,000 students from a variety of school types throughout the nation to ensure that the results are representative of the population. The sampling method used is a multistage stratified random sample ("NAEP - Assessment Process"). The NCES first identifies the schools to test by dividing the country into its 3,000 counties designated by the U.S. Census Bureau. It then identifies both public and non-public schools within each county. Schools are classified by type of location (city, suburb, rural, etc.) and racial/ethnic composition. The schools are also categorized by student achievement level, generally the results of a state achievement test, to make sure schools of different levels of achievement are included. The NCES draws a separate sample of schools from each stratum with probability proportional to school size ("NAEP - Assessment Process"). Small schools, high minority schools, and private schools are sampled to ensure that they are adequately represented. Once the schools have been selected, the NCES randomly selects about 30 students for each test subject from grades 4, 8, and 12 to be tested.

The dataset used in this analysis compiles the results of the reading test for American 8th graders. The test is administered approximately every 2 years. Scores range from 0–500 and are based on statistical procedures called Item Response Theory, which are test analysis procedures that assume a mathematical model for the probability that an examinee will respond correctly to a specific test question ("NAEP - Assessment Process"). The national average for 8th grade reading in 2019 was 263. The column variables in the dataset include Year, Jurisdiction, Gender, Average Score, and Region. The variable Year specifies the year in which the test was administered which includes select years from 1998 to 2019. Jurisdiction specifies the state or city in which the test was administered and includes all 50 states and the city of Washington D.C. Gender specifies the biological sex of the examinees. The variable region describes the region to which the state specified belongs and is determined by the U.S. Census Bureau's categorization of regions. The variable average score is the mean score of the 8th grade reading test. There are 1420 rows of entries in the data and each row of the dataset gives a year, a jurisdiction, a gender, the average score for the specified variables, and the region to which the jurisdiction belongs. This study aims to determine if the general trend of 8th reading scores has been negative since 1992, and whether the trends vary for students with different characteristics like gender or the region in which they live.

Data Summaries

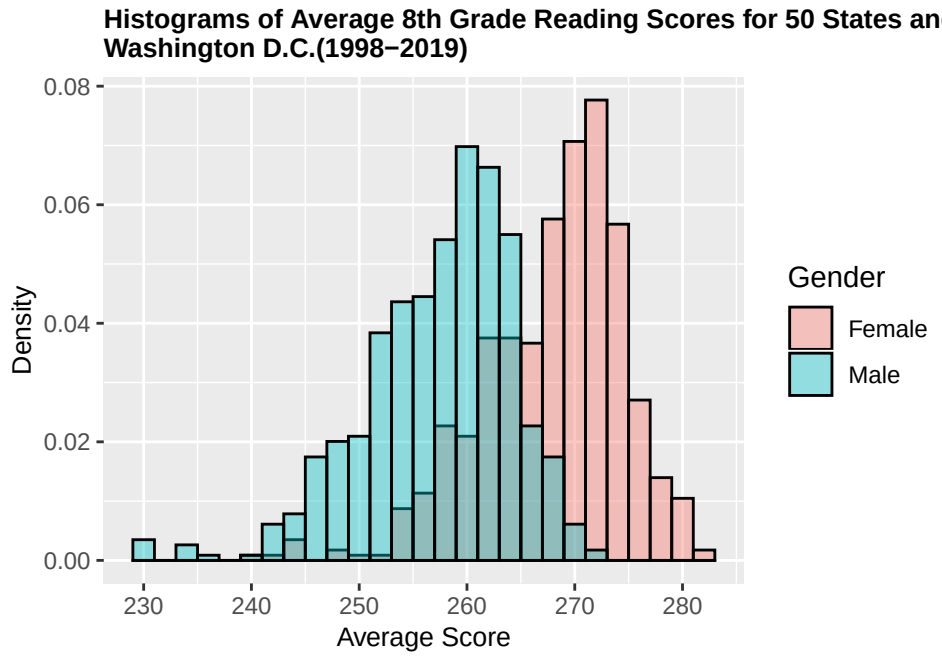
```
library(ggplot2)
ggplot(data,aes(x=Avg.score))+geom_histogram(color="black",
fill="white",binwidth = 2)+labs(title=
"Histogram of Average 8th GradeReading Score for 50 States and
Washington D.C.(1998-2019)",
x="Average Score",y="Count")+theme(plot.title=element_text(face="bold",size=10),
axis.title.y=element_text(size=10), axis.title.x=element_text(size=10))
```

```
## Warning: Removed 274 rows containing non-finite values (stat_bin).
```




```
ggplot(data,aes(x=Avg.score,fill=Gender))+geom_histogram(aes(y=..density..),
color="black",binwidth=2,position="identity", alpha=0.4)+labs(title=
"Histograms of Average 8th Grade Reading Scores for 50 States and
Washington D.C.(1998-2019)",
x="Average Score",y="Density")+theme(plot.title=element_text(face="bold",size=10),
axis.title.y=element_text(size=10),axis.title.x=element_text(size=10))
```

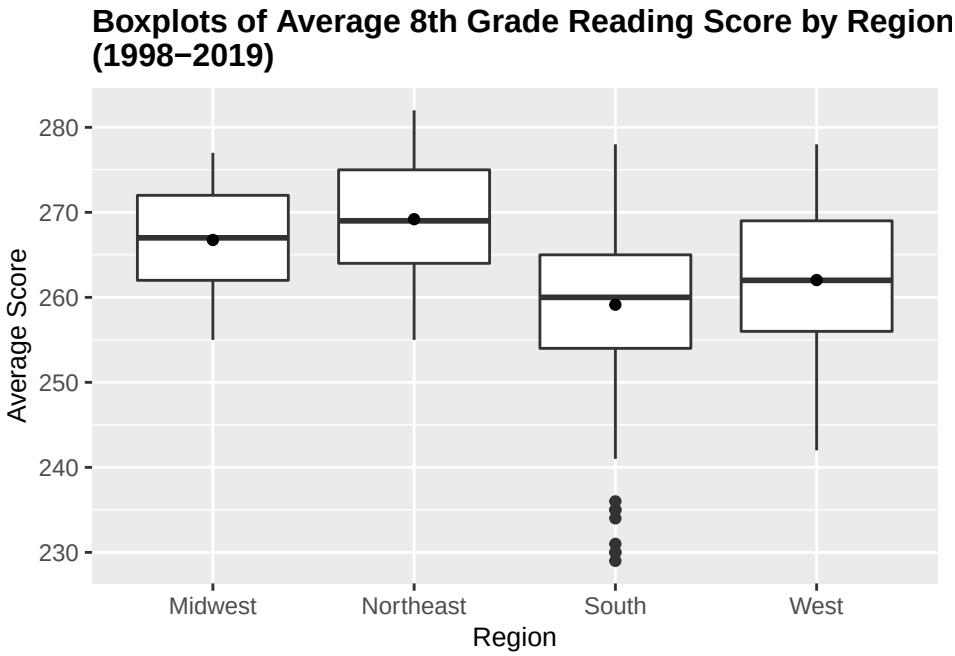
```
## Warning: Removed 274 rows containing non-finite values (stat_bin).
```



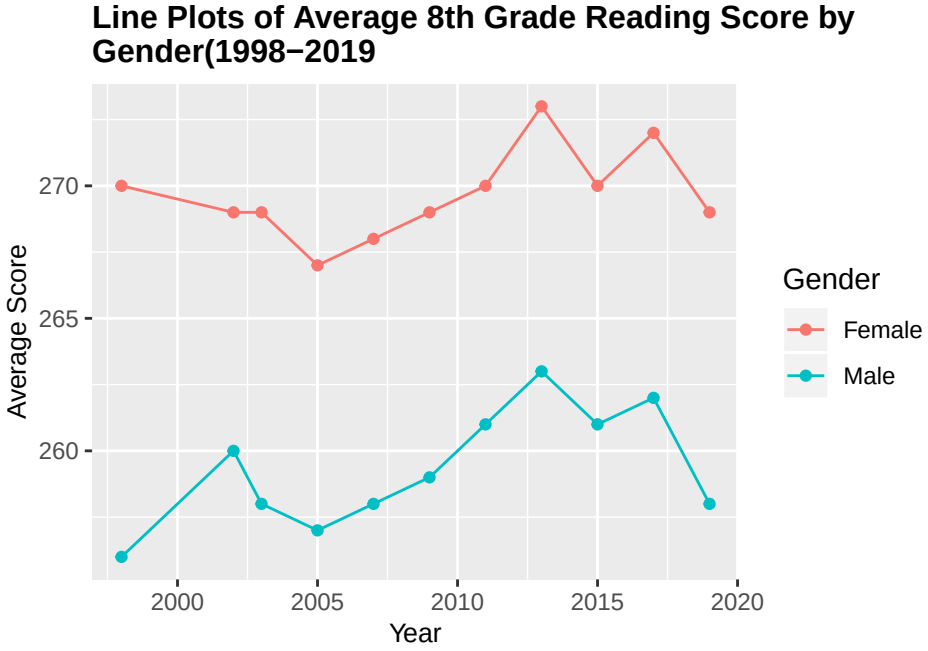

```
ggplot(data,aes(x=Region,y=Avg.score))+geom_boxplot()+stat_summary(fun.y=
"mean",geom="point")+labs(title=
"Boxplots of Average 8th Grade Reading Score by Region
(1998-2019)",x="Region",
y="Average Score")+theme(plot.title=element_text(face="bold",size=12),
axis.title.y=element_text(size=10),axis.title.x=element_text(size=10))
```

```
## Warning: Removed 274 rows containing non-finite values (stat_boxplot).
```

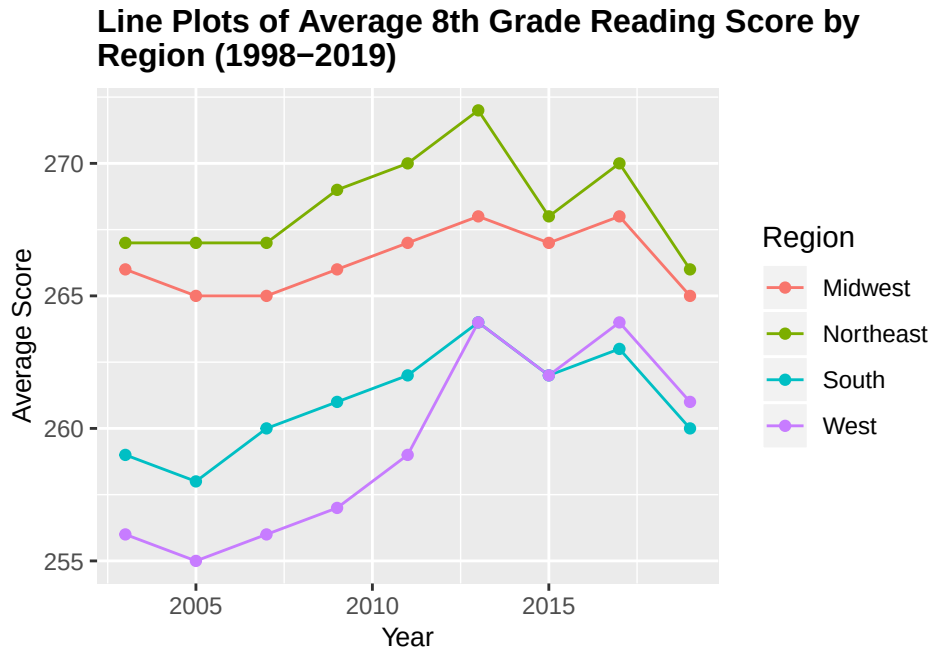
```
## Warning: Removed 274 rows containing non-finite values (stat_summary).
```




```
ggplot(data.,aes(x=Year,y=Avg.score, color=Gender))+
geom_line()+geom_point()+labs(title=
"Line Plots of Average 8th Grade Reading Score by
Gender(1998-2019",x="Year",
y="Average Score")+theme(plot.title=element_text(face="bold",size=12),
axis.title.y=element_text(size=10),axis.title.x=element_text(size=10))
```



```
ggplot(data1,aes(x=Year,y=Avg.score,color=Region))+geom_line()+geom_point()+
labs(title="Line Plots of Average 8th Grade Reading Score by
Region (1998-2019)",
x="Year",y="Average Score")+theme(plot.title=element_text(face="bold",size=12),
        axis.title.y=element_text(size=10),
        axis.title.x=element_text(size=10))
```



Conclusion

The data analysis process began by considering whether American students have lost reading and comprehension skills over time. To answer this question, I obtained the average scores of the 8th grade reading assessment for years 1998-2019 from the NAEP database. I included variables describing the gender and location of the students in order to asses whether students who fall into different categories have different trends in their scores over time. The histogram of all students displays the count and distribution of all scores from 1998-2019. It is bimodal and slightly left skewed with one peak around the score 262 and another around 272. The histograms that divide the students by gender shows that the two peaks are the peaks of the distributions of the genders. The mean among male students appears to be 262 and the mean among female students appears to be 272. The spread of the distributions appears to be similar, however the

scores for female students occupy a slightly larger range. The boxplots of average score by region shows that the southern states have the lowest median and mean average scores, both around 260, and the northeastern states have the highest median and mean average scores, both around 270. The scores of the Southern states are the most variable, and the scores of the Midwestern states are the least variable. The New York Times article was correct in stating that 8th grade reading scores fell from 2017 to 2019; however, the article doesn't mention that between 2003 and 2011 reading scores improved for every year of testing. The line plots describing the trends for gender show that average scores for both genders increased from 1998 to 2013, peaked in years 2013 and 2017 and fell again in 2019. The plots show that the average scores for female students have declined slightly, but the average score for male students has increased overall from 1998-2019. The line plots describing the trends for region show a similar pattern. If the score declined, like for the Midwest and the Northeast, they only declined slightly, but if the score improved, like for the South and the West it improved fairly dramatically from 1998-2019. The scores of the 8th grade reading exam have declined over the short term, but since 1998 they have remained fairly stable overall, with average scores remaining in the 255-270 range regardless of the isolated characteristic of the student. Although things haven't become much worse, the policies that have been implemented over the last 20 years may not be achieving the goals set out by the Department of Education.There are many more factors that determine a student's performance in school. The dataset attempts to account for differences in resources by isolating scores based on region. However, the regions are large so the influence that the level of resources a school and therefore a student has may not be clearly captured within the data. The data also doesn't distinguish between the different types of schools. It may also be that students attending the different types of schools have extensively different backgrounds which affects their scores. Another issue with using the National Report Card data is that the test is optional, and students may lack motivation to take it seriously if they know that it is not part of their grade. Asking already sleep deprived students to take a test they don't feel matters to them may not be the best way to assess their ability. Although the test has been administered since 1969, the data available online spans from 1992-2019, and much of the data from 1992-1998 is unavailable. The ability to examine a longer period of time might help to asses how changes in the education system have affected students.

References

1. Abadi, Mark. "Even the US Government Can't Agree on How to Divide up the States into Regions." Business Insider, Business Insider, 10 May 2018, <www.businessinsider.com/regions-of-united-states-2018-5#the-us-census-bureau-divides-the-united-states-into-four-regions-theres-the-northeast–1.>

2. "Different Ways to Set Figure Size in RMarkdown." Sebastian Sauer Stats Blog, 2 Nov. 2016, <sebastiansauer.github.io/figure_sizing_knitr/.>

3. Green, Erica L., and Dana Goldstein. "Reading Scores on National Exam Decline in Half the States." The New York Times, The New York Times, 30 Oct. 2019, <www.nytimes.com/2019/10/30/us/reading-scores-national-exam.html.>

4. "NAEP - Parents." National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education, <nces.ed.gov/nationsreportcard/assessment_process/.>

5. "NAEP Report Cards - Home." The Nation's Report Card, <www.nationsreportcard.gov/.>