8th Grade NAEP Reading Exam Scores as a Factor of Family Income, Divorce, and State Educational Expenditure

Lewis Collum (0621539)

Updated: May 1, 2020

Contents

1 Introduction

We are studying the National Assessment of Education Progress (NAEP) reading exam performance of 8th graders, across the states. The test is out of 500 points. NAEP administers tests across subjects at the 4th, 8th, and 12 grade levels, with randomly selected students. We would like to see if the exam performance has nationally increased since 1990. We would also like to see if we can find a significant relationship between reading performance and average family income, divorce rates, or educational expenditure, across the states.

2 Methods

2.1 Is There an Increase in NAEP Reading Exam Scores Since 1990?

We use a simple linear regression with the model,

$$\texttt{scôre} = \beta_0 + \beta_1 \cdot \texttt{year}. \tag{1}$$

Our hypotheses are

$$H_0 : \beta_1 = 0;$$
$$H_A : \beta_1 \neq 0.$$

2.2 Does State Educational Expenditure have a Significant Affect on Scores?

We use a simple linear regression with the model,

$$\texttt{scôre} = \beta_0 + \beta_1 \cdot \texttt{expenditure} \tag{2}$$

Our hypotheses are

$$H_0 : \beta_1 = 0;$$
$$H_A : \beta_1 \neq 0.$$

2.3 Does State Divorce Rate have a Significant Affect on Scores?

We use a simple linear regression with the model,

$$\texttt{scôre} = \beta_0 + \beta_1 \cdot \texttt{divorce\_rate} \tag{3}$$

Our hypotheses are

$$H_0 : \beta_1 = 0;$$
$$H_A : \beta_1 \neq 0.$$

2.4 Does State Median Family Income have a Significant Affect on Scores?

We use a non-linear regression with the model,

$$\texttt{scôre} = \beta_0 + \beta_1 \cdot \texttt{income} + \beta_2 \cdot \texttt{income}^2 \tag{4}$$

Our hypotheses are

$$H_0 : \beta_1 = 0;$$
$$H_A : \beta_1 \neq 0.$$

3 Data

All hypotheses in section 2 use the `U.S. Education Datasets: Unification Project` dataset (section 6.1) which is designed to bring together multiple facets of U.S. education data into one dataset. We use the `total expenditure` and `average reading score` for eigth grade, from this dataset.

To find if state median family income has a significant affect on reading scores (section 2.4) we use an American Community Survey (section 6.2) for median family income by state from 2014 to 2018.

To find if state divorse rate has a significant affect on reading scores (section 2.3) we use a CDC study (section 6.3) for divorse rates in 2015.

4 Analysis

4.1 Is There an Increase in NAEP Reading Exam Scores Since 1990?

We compare the NAEP reading scores from 1990 to 2018, to see if there is a significant change. We find that 85% of the variance in scores is explained by year (i.e. R^2). We also find that the p-value is much less than the significance level of 0.05. As such we reject the null hypothesis (eq. 1) that scores have not changed since 1990. We see that scores have increased since 1990 and the concrete model is

$$\texttt{scôre} = -1110 + 0.69 \cdot \texttt{year}. \tag{5}$$

The intercept is not interable since we do not have scores for students for year 0, since the NAEP study started in 1990. For every increase in year of 0.69, the national reading score increases by 1 (out of 500 total points).

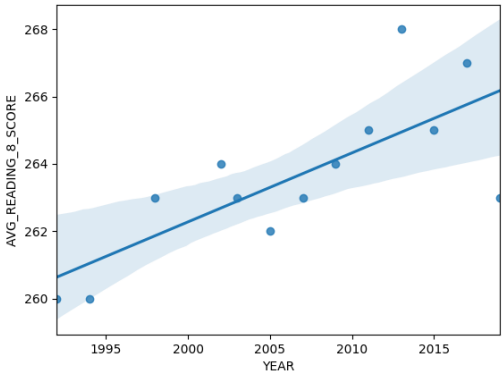


Figure 1: Shows a positive trend between reading scores and year. This indicates reading scores have increased since 1990.

4.2 Does State Educational Expenditure have a Significant Affect on Scores?

We compare the NAEP reading scores from 2015 to total educational expenditure, per state, to see if there is a significant affect. Since each state has a different enrollment count (due to population), we divide the total expenditure by the number of students enrolled. This gives us the expenditure per student approximately. We find that the p-value is 0.49, which is much greater than the significance level of 0.05. As such we fail to reject the null hypothesis (eq. **??**). We see no significant effect from expenditure on scores.

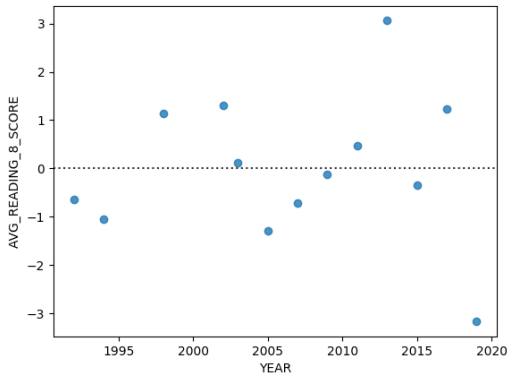


Figure 2: Reading Scores per Year Residuals. The residuals have no obvious bias and are homoscedastic.

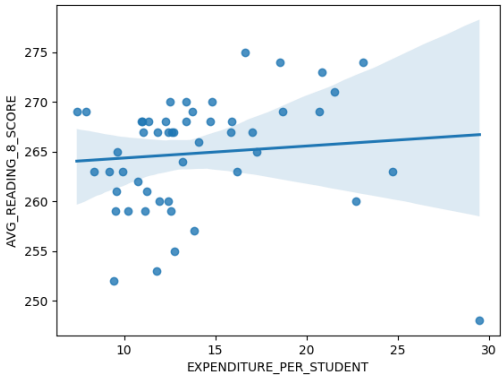


Figure 3:

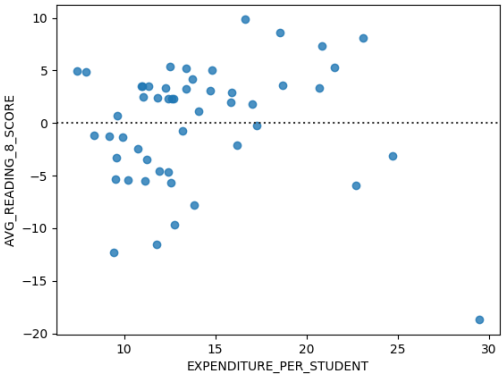
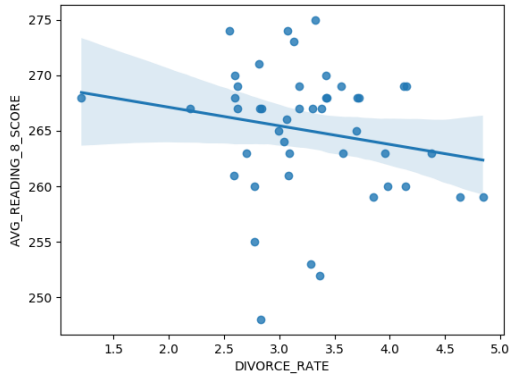


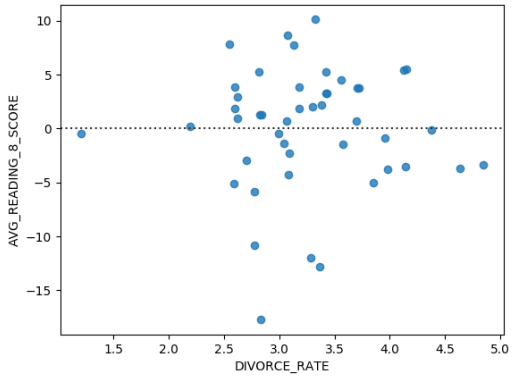
Figure 4:

4.3 Does State Divorce Rate have a Significant Affect on Scores?

We compare the NAEP reading scores from 2015 to divorse rate, per state, to see if there is a significant affect. We find that the p-value is 0.92, which is much greater than the significance level of 0.05. In fact, there is almost no relation between the two variables. As such we fail to reject the null hypothesis (eq. **??**). We see no significant effect from divorse

rate on scores.





4.4 Does State Median Family Income have a Significant Affect on Scores?

4.4.1 Non-Linear Model

We compare the NAEP reading scores from 2015 to median family income, per state, to see if there is a significant affect. We find that 30% of the variance in scores is explained by year (i.e. R^2). We also find that the p-value is much less than the significance level of 0.05. As such we reject the null hypothesis (eq. **??**). We see that scores non-linearly increase as income increases. With high error towards the higher incomes, this model is likely only useful for lower to mid-range incomes. The concrete model is

$$\texttt{scôre} = 156.1 + 0.0034 \cdot \texttt{income} + -2.58 \cdot 10^{-8} \cdot \texttt{income}^2 \tag{6}$$

The intercept is not interable since we do not have scores for students for year 0, since the NAEP study started in 1990. For every increase in year of 0.69, the national reading score increases by 1 (out of 500 total points).

5 Conclusion

Ultimately, we found that NAEP reading scores at the eigth grade level have increased since 1990. We then looked at the reading scores only in 2015, and attempted to find a factor that significantly affects the scores. Such factors were, median family income, divorse rate, and educational expediture. Of these factors, median family income was the only to significantly affect reading scores.

Furthermore, it may be useful to note, that since expenditure did not have a significant effect on scores, we cannot conclude that increasing money spent on education per state will have an impact on score performance.

Also, this study used divorse rate per state, however it may be beneficial to replace this factor with the number of single-family homes with children in school. This may provide a factor that is more directly related to students in school, since divorse rate includes all couples, even those without children in school.

In the future, it would be appropriate to find more factors that affect reading scores across states. It is unrealistic to expect that a univariate analysis could accurately model a complex subject such as student reading performance across states.

6 Datasets

6.1 Education Unification Project

`https://www.kaggle.com/noriuk/us-education-datasets-unification-project#states_all_extended.csv`

6.2 2014-2018 Income American Community Survey 5-Year Estimates

`https://www.census.gov/search-results.html?q=Median+Household+income&page=1&stateGeo=none&searchtype=web&cssp=SERP&_charset_=UTF-8`

6.3 Divorce Rates

`https://www.cdc.gov/nchs/data/dvs/state-divorce-rates-90-95-99-18.xlsx`