Sam Lauer
Professor Hwang
GEO 491
21 March 2019

Correlation Analysis of High School Graduation Rates and Per Pupil Current Spending in United States

Abstract

The research question being studied in this report is in response to rising concerns regarding how to best support student education and their future success. Higher education levels are often associated with higher income potential. Increased earnings benefit not only the future of individual students but also benefit the communities they live in and the greater overall economy. One research variable that is consistently explored in relation to pupil academic achievement is per pupil current spending. The research question explored in this research study examines whether there is any correlation between per pupil current spending and graduate rates at the state level within the United States. The null hypothesis of this study is that there is no correlation between the two variables. The alternate hypothesis is that there is a positive correlation between per pupil current spending and graduation rates. The alpha level or accepted level of error is a p-value of .05 or 5%. It uses the Spearman Rho correlation to test the level of correlation between the two variables and yields a correlation coefficient of .160 which demonstrates a positive, yet weak correlation. It also yields a significance level or p-value of .131 which is above the alpha level of .05 for this study. Based on this result, the null hypothesis cannot be rejected. It is therefore concluded that there is no correlation between per pupil current spending and graduation rates.

Background

Throughout the United States, school districts and states are consistently searching for ways in which they can improve student performance and also increase the percentage of high school graduation rates. However, the issue of how this can be best accomplished is hotly contested with no single agreed upon solution. Some studies have found that boosting per pupil spending can positively impact educational performance as well as increase the chances that children will eventually enroll in and graduate from college (Fleisher, 2017). In a study conducted by Northwestern Researcher Kirabo et al. (2018), for example, it concluded that states that cut funding following the Great Recession experienced lower test scores and lower high school graduation rates, especially amongst Hispanic students. This trend can be especially true of minority student and economically-disadvantaged students who often live in economically depressed communities (Barrett, 2018). However, other studies argue that funding itself does not contribute to improved performance. Instead, it is argued that schools should allocate their resources differently in order to boost performance (Lips & Watkins, 2008).

Regardless of the approach to improving academic performance, there seems to be a national census that improved educational performance translates to higher income potential and that increased individual earnings leads to real economic benefits. In 2017, the median income of an

individual with no high school diploma was just $520 on a weekly basis whereas those with a high school diploma earned $712 on average. This number jumps to $1,173 for individuals with a bachelor's degree (Torpey, 2018). These numbers demonstrate the economic benefits to an education and help prove that individuals who earn a high school diploma will make considerably more throughout their lifetime than those without a diploma. If these same graduates go on to earn a bachelor's degree, they will earn nearly $1 million more on average than high school graduates throughout their lifetime according to the U.S. Census Bureau (Longely, 2018). Much of the income generated by these individuals will be spent within their local economies demonstrating that when individuals make more money, their communities stand to benefit as well as the economy as a whole (Florida, 2016). These numbers help demonstrate why communities and states are so interested in boosting high school graduation rates and how increased academic performance can lead to increased economic prosperity.

Research Question

This study seeks to answer whether there is a statistically significant correlation between per pupil current spending and graduation rates in the United States. Data for per pupil current spending was obtained from the U.S. Census Bureau while graduation rate data was obtained from figures recorded by the U.S. Department of Education. Per pupil current spending is defined by the U.S. Census Bureau as including instruction, support services, and non-instructional functions such as teacher salaries, benefits, building maintenance, and transportation. High school graduation rates are based on students who receive a high school diploma. Both data sources originate from the year 2015 with the study area being the United States and the spatial unit of analysis being the state level. All 50 states are included in this analysis as well as the District of Columbia. The null hypothesis for this research question is that there is no correlation between per pupil current spending and graduation rates. The alternate hypothesis is that these two variables are positively correlated to a statistically significant level. The significance level is .05.

Methodology

Data sources for this research question were first input into Microsoft Excel and then uploaded into ArcMap. A shapefile containing US state boundaries was downloaded from the U.S. Census Bureau and also uploaded into ArcMap. These two tables were then joined using the state name to add the data for the two variables into the table demonstrating state boundaries. Choropleth maps were created to help visualize the variables for individual states. The results of the choropleth maps are outlined in the descriptive statistics analysis section. A cluster and outlier analysis was also conducted to determine whether any outliers existed for the two variables being tested. Hawaii and Alaska were both excluded in the cluster and outlier analysis because of their geographic location in relation to the 48 contiguous states. The cluster and outlier analysis found one high-low outlier state (Utah) for the per pupil current spending meaning Utah's per pupil current spending was higher compared with its neighbor states. There were four outliers for the graduation rates. Texas and Utah represented high-low outliers meaning they had higher graduation rates when compared with their neighbor states and Minnesota and New York represented low-high outliers meaning their graduation rates were lower when compared with neighboring states. All maps are represented within the appendix of this report.

To test whether graduation rates and per pupil current spending are positively correlated to a statistically significant level, the data containing these variables was loaded into SPSS. Before determining the correlation test to be used, it first had to be determined whether the data met the assumptions of correlation analysis. These include linearity, numeric data, and normality. Is these assumptions are met, a parametric test such as the Pearson's test can be made. If these assumptions are not met, a nonparametric test such as Spearman Rho must be conducted. The Spearman Rho test was ultimately conducted to determine correlation based on the fact that the normality assumption was not met by this specific set of data. It was determined that normality was not met through the Shapiro-Wilk test. The results of the Shapiro-Wilk test were found through SPSS and will be discussed in the results section.

Descriptive Statistics Analysis

Descriptive Statistics were utilized in this study to help understand each variable independently. This information can be utilized to demonstrate how states compare with the descriptive statistics as evidenced by the choropleth maps as well as compare the descriptive statistics to the cluster and outlier analysis maps. Descriptive statistics can be found in Figure 1 of the appendix. Measurements of central tendency include mode, median, and mean. The US graduation rate mean is 83% and the per pupil current spending is $11,876.71. The median graduation rate is 85% and median per pupil current spending is $11,010. The mode for graduation rate and per pupil current spending was not included in the table. Measurements of dispersion include range and standard deviation. The range for graduation rates is 22% while per pupil current spending is $14.631. The standard deviation for graduation rates is 5.5 and for per pupil current spending it is 3232.90.

The skewness for graduation rates is -.851 indicating negatively skewed distribution. This also demonstrates that the mode and median are above the mean for graduation rates. The skewness of the per pupil current spending is .974 demonstrating positively skewed distribution with the mode and median below the mean. The kurtosis for the graduate rate is .025 with the positive value indicating leptokurtic or peaked distribution. The kurtosis for the per pupil current spending is .304 again indicating leptokurtic distribution.

Figures 2 and 3 (below) show a box plot of both graduation rates and per pupil current spending. These box plots are a visual representation of central tendency and dispersion data found within Figure 1. The box plot itself is a representation of the interquartile range (IQR). The top line is quartile 3, the middle line is the median, and the bottom line is quartile 1. The IQR for graduation rates is .08 while the IQR for per pupil current spending is 4875. The upper and lower whiskers extend to minimum and maximum data points within 1.5 box heights. There are no outliers in Figure 2 and one outlier in Figure 3 (Alaska).

Descriptives

			Statistic	Std. Error
Grad_15	Mean		.8298	.00773
	95% Confidence Interval for Mean	Lower Bound	.8143	
		Upper Bound	.8453	
	5% Trimmed Mean		.8333	
	Median		.8500	
	Variance		.003	
	Std. Deviation		.05523	
	Minimum		.69	
	Maximum		.91	
	Range		.22	
	Interquartile Range		.08	
	Skewness		-.851	.333
	Kurtosis		.025	.656
PPCS_15	Mean		11876.7059	494.70561
	95% Confidence Interval for Mean	Lower Bound	10883.0604	
		Upper Bound	12870.3513	
	5% Trimmed Mean		11678.6329	
	Median		11010.0000	
	Variance		12481415.81	
	Std. Deviation		3532.90473	
	Minimum		6575.00	
	Maximum		21206.00	
	Range		14631.00	
	Interquartile Range		4875.00	
	Skewness		.974	.333
	Kurtosis		.304	.656

Figure 1: Descriptive Statistics for graduation rates and per pupil current spending

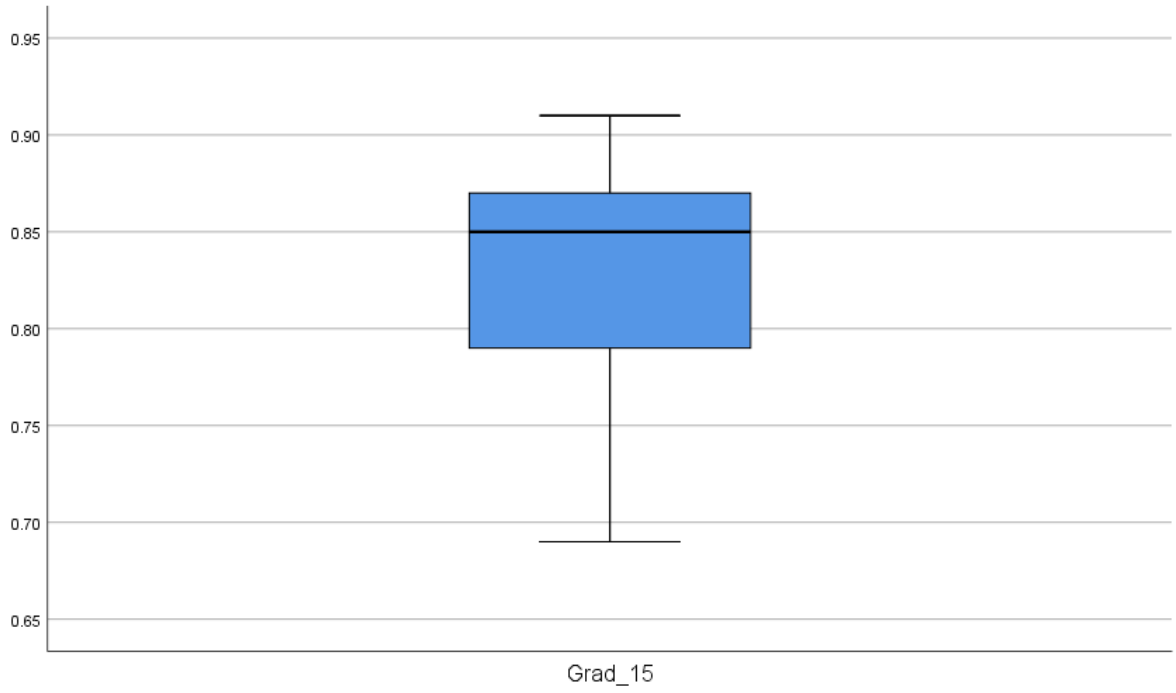


Figure 2: Box Plot of Graduation Rate

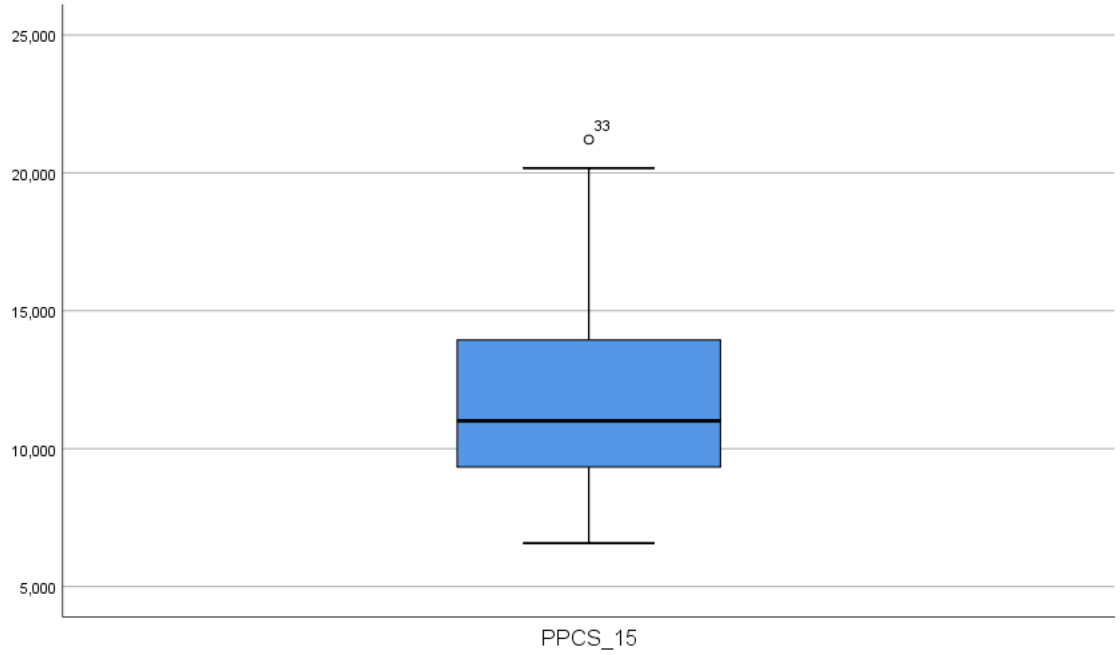


Figure 3: Box Plot of Per Pupil Current Spending

Results

As stated previously, the Spearman Rho test was conducted to determine the correlation between the two variables being tested in this research analysis. It was determined that the normality assumption was not met through Shapiro-Wilk test. The results of the test can be found in Figure 4 below.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Grad_15	.178	51	.000	.913	51	.001
PPCS_15	.159	51	.003	.914	51	.001

a. Lilliefors Significance Correction

Figure 4: Shapiro-Wilk Test

In the Shapiro-Wilk test, normality is determined through the significance level of the variables being tested (abbreviated above on the right-hand side as Sig.). The null hypothesis for the Shapiro-Wilk test is that data is normally distributed. If the p-value is less than the alpha level (.05), the null hypothesis is rejected, and it can be determined that the variable is not normally distributed. Conversely, if the p-value is greater than the alpha level, the null hypothesis is accepted, and it is determined that the variable is normally distributed. As the significance level for both graduation rates and per pupil current spending are below the alpha level (.001<.05), the

null hypothesis can be rejected in both cases and determined that the variables are not normally distributed. A scatterplot of the two variables also demonstrates that there does not appear to be any linear relationship between the two variables. This can be seen below in Figure 5.

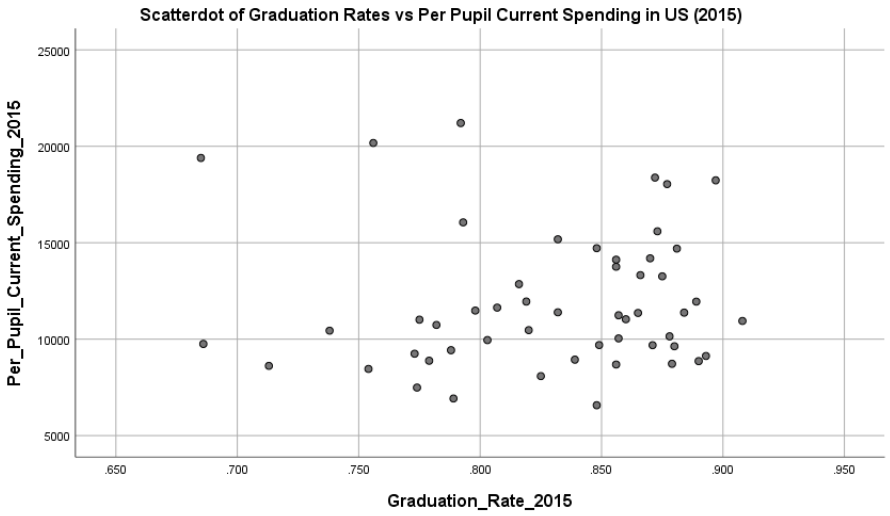


Figure 5: Scatterplot of research variables

Additionally, a histogram was created for each individual variable demonstrating that neither variable conforms to the symmetric bell curve pattern of the Gaussian function. The results are shown in Figure 6 and Figure 7.

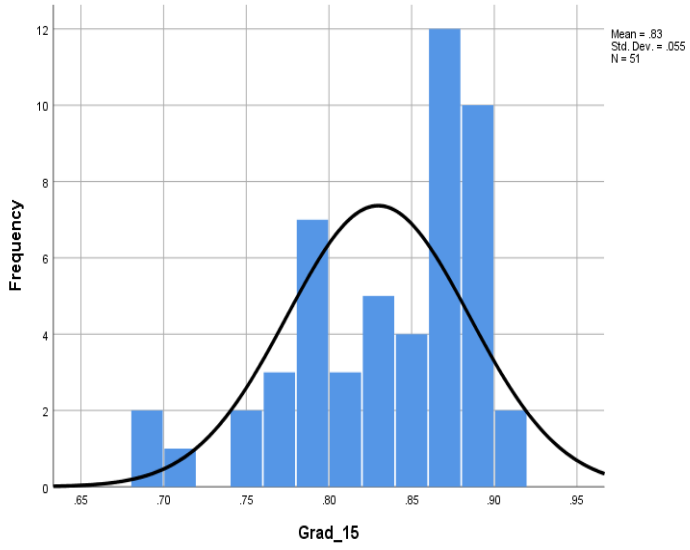


Figure 6: Histogram of Graduation Rates

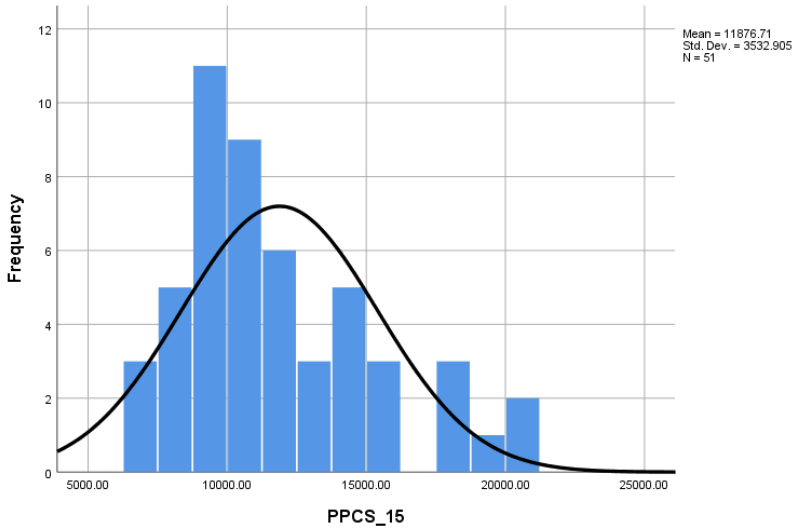


Figure 7: Histogram of Per Pupil Current Spending

Based on a culmination of these results, the Spearman Rho test was conducted to test correlation. It was conducted as a one-tailed (directional) test through SPSS. The results can be seen below in Figure 8:

Correlations

			Grad_15	PPCS_15
Spearman's rho	Grad_15	Correlation Coefficient	1.000	.160
		Sig. (1-tailed)	.	.131
		N	51	51
	PPCS_15	Correlation Coefficient	.160	1.000
		Sig. (1-tailed)	.131	.
		N	51	51

Figure 8: Spearman Rho test determining correlation

The Spearman Rho test demonstrates a correlation coefficient of .160 which is a positive, yet weak correlation between the two variables being tested. It also demonstrates a significance level or p-value of .131.

Conclusion

The p-value of .131 found within the Spearman Rho test demonstrates that there is a 13.1% chance of making a type I error. This is greater than the allowed alpha level of 5% level of error and therefore the null hypothesis cannot be rejected. In other words, the null hypothesis is accepted for this research question, and it is determined that there is no correlation between graduation rates and per pupil current spending.

Limitations & Future Research Opportunities

Data for this research question was taken from the state level which yields several limitations. The data set of 51 entries is quite limited in that each individual set of data point carries a larger degree of weight in determining the final outcome. Any outliers or data points not demonstrating any correlation are able to impact the final result to a greater extent than they would in a much larger set of data points. Therefore, larger and more robust data sets such as Chicago Public Schools which has over 600 schools might yield different results as each individual data point carries less weight in how it impacts the final results. Also, examining the issue of funding per pupil and graduation rates from the state level spatial unit assumes that schools and pupils within these states are homogenous entities when this is most certainly not the case. Pupils who experience language or learning barriers, for example, face an entirely different set of challenges than students who do not face the same circumstances. Socioeconomic factors can also impact a child's learning environment. A child facing adverse conditions outside of the classroom such as an abusive household or homelessness will experience difficulties that are likely not experienced by a child living within a stable household within a more affluent environment. A study that accounts for different sets of circumstances impacting the education of student populations may yield different results. Another factor to be considered is the per pupil current spending definition utilized by the US Census Bureau includes monies that go to non-instructional spending such as building maintenance and student transportation that may have less of an impact in student achievement than classroom resources. Therefore, another opportunity to research the impact of per pupil current spending could include solely those funds being directly spent on the instruction of pupils. Exploring any combination of the aforementioned factors would result in a more comprehensive and robust study that takes into account the complexity of measuring how to best ensure the success of students in their educational endeavors.

References

2016 Public Elementary-Secondary Education Finance Data. (2016). *United States Census Bureau.* Retrieved from https://www.census.gov/data/tables/2016/econ/school-finances/secondary-education-finance.html

Barrett, K. (2018). The Evidence is Clear: More Money for Schools Mean Better Student Outcomes. *NeaToday.* Retrieved from http://neatoday.org/2018/08/01/money-matters-in-education/

Fleisher, C. (2017). A Long-Term Investment in Education. *American Economic Association.* Retrieved from https://www.aeaweb.org/research/school-spending-increase-impact-student-performance

Florida, R. (2016). The Economic Impact of College Graduates on Local Economies. *CityLab*. Retrieved from https://www.citylab.com/life/2016/02/the-economic-impact-of-college-graduates-on-local-economies/435414/

High School Graduation Rates by State. *Governing.* Retrieved from http://www.governing.com/gov-data/high-school-graduation-rates-by-state.html

Jackson, C., Wigger, C., & Xiong, H. (2018). Do School Spending Cuts Matter? Evidence from the Great Recession. *Northwestern University*. Retrieved from http://works.bepress.com/c_kirabo_jackson/35/

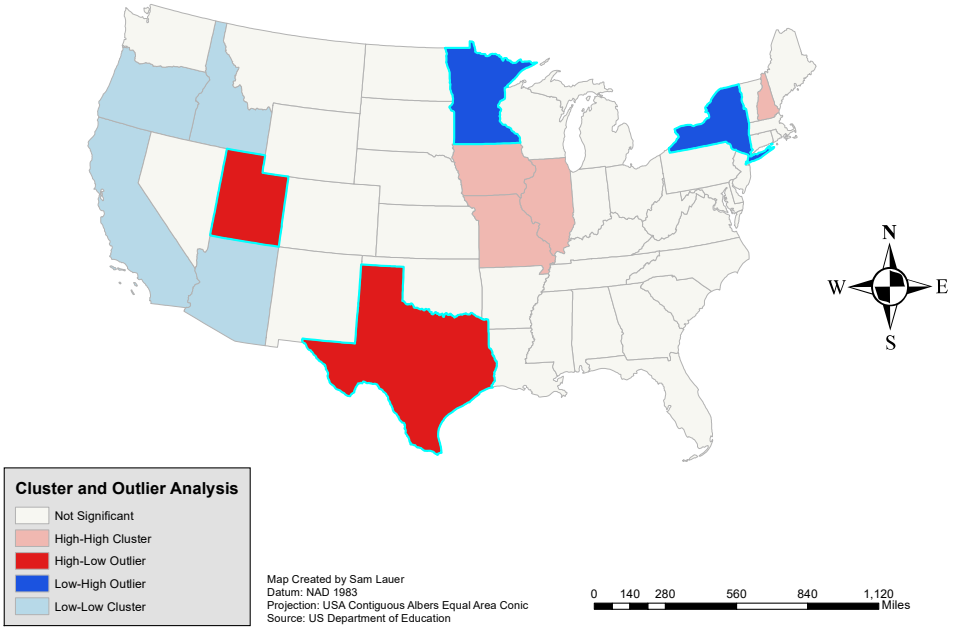
Lips, D. & Watkins, S. (2008). Does Spending More on Education Improve Academic Achievement. *The Heritage Foundation.* Retrieved from https://www.heritage.org/education/ report/does-spending-more-education-improve-academic-achievement

Longely, R. (2018). Lifetime Earnings Soar with Education. *ThoughtCo.* Retrieved from. https://www.thoughtco.com/lifetime-earnings-soar-with-education-3321730

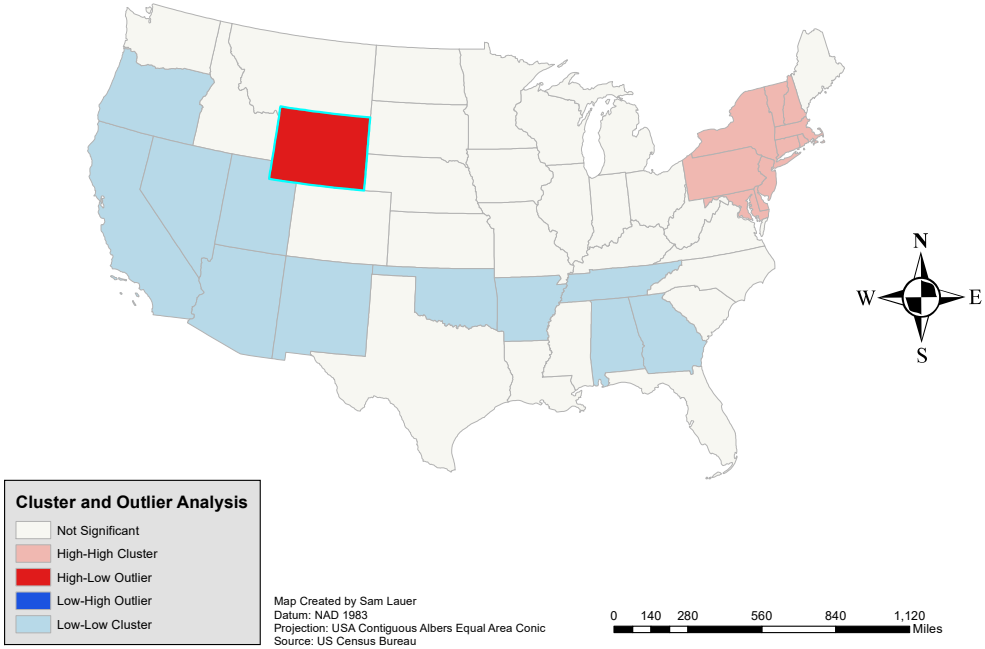
Torpey, E. (2018) Measuring the value of education. *Bureau of Labor Statistics.* Retrieved from https://www.bls.gov/careeroutlook/2018/data-on-display/education-pays.htm

Appendix

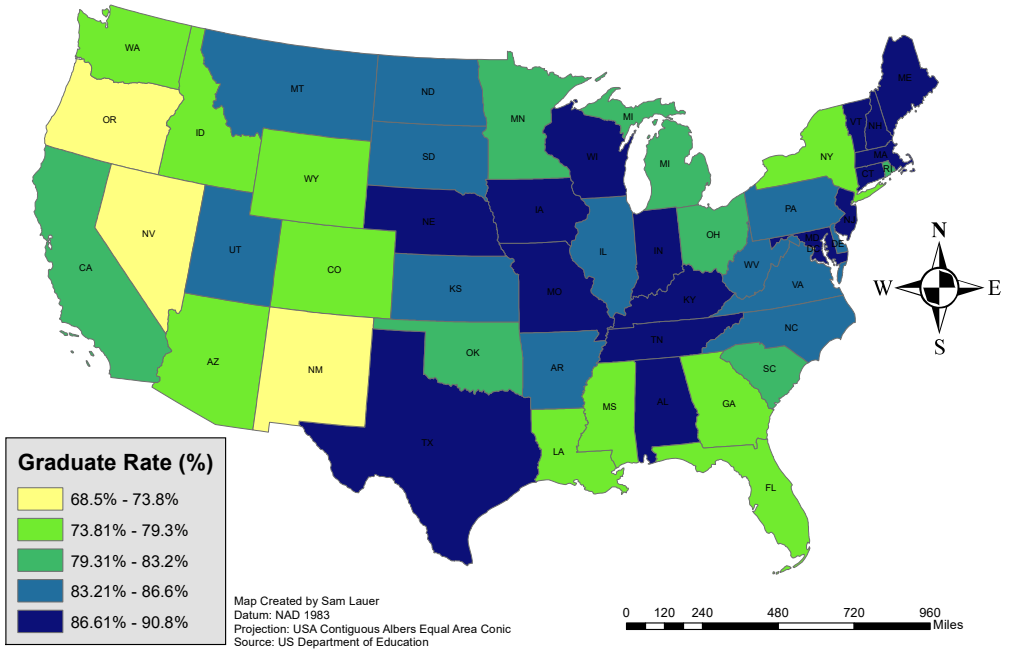
High School Graduate Rates in United States (2015)



Per Pupil Current Spending in United States (2015)



High School Graduation Rates in United States (2015)



Per Pupil Current Spending in United States (2015)