**Interpreting R**

* *The R-squared value is a measure of how well the model explains the data.*
* Values between 0 and 0.3 (0 and -0.3) indicate a weak positive (negative) linear relationship via a shaky linear rule.
* Values between 0.3 and 0.7 (-0.3 and -0.7) indicate a moderate positive (negative) linear relationship via a fuzzy-firm linear rule.
* Values between 0.7 and 1.0 (-0.7 and -1.0) indicate a strong positive (negative) linear relationship via a firm linear rule.

**Interpreting P Value**

* *Does this model explain the data significantly better than would just looking at the average value of the dependent variable?*
* A p-value less than 0.05 (typically ≤ 0.05) is statistically significant. It indicates strong evidence against the null hypothesis, as there is less than a 5% probability the null is correct (and the results are random). Therefore, we reject the null hypothesis, and accept the alternative hypothesis.
* A p-value higher than 0.05 (> 0.05) is not statistically significant and indicates weak evidence against the null hypothesis. This means we fail to reject the null hypothesis and cannot accept the alternative hypothesis. You should note that you cannot accept the null hypothesis, but only find evidence against it.

*p*-values and *R-squared* values measure different things.  The *p*-value indicates if there is a significant relationship described by the model, and the *R-squared* measures the degree to which the data is explained by the model.  It is therefore possible to get a significant *p*-value with a low *R-squared* value.  This often happens when there is a lot of variability in the dependent variable, but there are enough data points for a significant relationship to be indicated.

\*The above definitions were copied from the following sources as a reference to guide our analysis. <https://www.simplypsychology.org/p-value.html>, <http://www.dmstat1.com/res/TheCorrelationCoefficientDefined.html>

**We hypothesize that...**

...counties with better high school attainment also have a lower poverty rate, higher income, and/or more residents with college degrees.

* + Poverty Rate = TRUE
  + Higher Income = TRUE
  + More bachelor’s degrees = TRUE

...counties that spend more money on education will also report better educational outcomes.

* FALSE

...counties with a higher median income and/or quality of life score will spend more money on education.

* FALSE

...Montgomery county will be in the top 10% in terms of income, educational outcomes and

* Highest income - TRUE
* HS attainment – FALSE
* BS degrees - TRUE

...Maryland will be in the top 25% of the country in terms of higher education degrees attained.

...regardless of educational spending, black and Hispanic students will have a lower rate of graduation than their white or Asian counterparts.

* FALSE

**High School Attainment %**

All our charts are arranged in order of p-value from lowest to highest from left to right. As such, items of moderate to high significance will be found on the left.

Several economic factors are correlated with higher high school graduation rates. These include Median Household Income, Percentage of Population in Poverty, Unemployment Rates, and Cost of living. Income and Cost of Living have a positive correlation – as those numbers increase, so does the percentage of graduates. The inverse is true for Poverty and Unemployment rates – as those numbers decrease, the percentage of graduates increases. Further there is a moderate positive correlation related to the education level of the population – the higher educated, the higher proportion of students who graduate high school.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dimension** | **P-Value** | **R2** | **Significance?** | **Direction** |
| Student-Teacher Ratio | 0.146112 | 0.093537 | None indicated | Positive |
| Expenditures per student | 0.233886 | 0.0637608 | None indicated | Negative |
| Population w/HS Degree % | 0.0002561 | 0.462498 | Moderate | Positive |
| Population w/ Bach Degree % | 0.0005056 | 0.429811 | Moderate | Positive |
| Cost of Living | 0.0003385 | 0.449303 | Moderate | Positive |
| Median HH Income | < 0.0001 | 0.65286 | **High** | Positive |
| Poverty % | < 0.0001 | 0.776032 | **High** | Negative |
| Unemployment % | 0.0003711 | 0.444887 | Moderate | Negative |

**High School Attainment % by Race**

Race did not have a strong correlation with HS graduation rates, except for the percentage of the population that identified as either Asian Only or Black only. Both these groups see a moderate positive correlation with increased HS graduation rates. The r-squared score for this was low, indicating that this is an area that could use further research to attain a more accurate model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dimension** | **P-Value** | **R2** | **Significance?** | **Direction** |
| Minority % | 0.169353 | 0.0840616 | None indicated | Positive |
| White Alone % | 0.170047 | 0.0838004 | None indicated | Positive |
| Hispanic or Latino Ethnicity % | 0.969145 | 6.957e-05 | None indicated | Negative |
| Asian Alone % | 0.0371789 | 0.182766 | Moderate | Positive |
| Native Hawaiian % | 0.484052 | 0.0235976 | None indicated | Positive |
| Two or More Races % | 0.0685503 | 0.142908 | None indicated | Positive |
| American Indian % | 0.4738 | 0.0235752 | None indicated | Negative |
| Black Alone % | 0.0311547 | 0.194161 | Moderate | Positive |
| Some Other Race % | 0.447715 | 0.0264461 | None indicated | Negative |

**Public School Spending**

It is difficult to tell what is driving these numbers. Most data points have no correlation with spending. A modest correlation was found with unemployment rate and the percentage of the population that identified as White Alone or our calculated percentage of those who did not. Unemployment rates have a positive correlation with public school spending, indicating that more funds are allocated in areas of need. There are clearly other factors at play, perhaps tax rates, gambling income, or federal allocations, and this is an area for further research.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dimension** | **P-Value** | **R2** | **Significance?** | **Direction** |
| Cost of Living | 0.580122 | 0.0141293 | None indicated | Negative |
| Median HH Income | 0.344896 | 0.0406322 | None indicated | Negative |
| White % | 0.0786337 | 0.13392 | None indicated | Negative |
| 2018 HS Attainment % | 0.233886 | 0.0637608 | None indicated | Negative |
| 2018 HS Enrollment | 0.436352 | 0.0277772 | None indicated | Positive |
| Poverty % | 0.0940195 | 0.122219 | None indicated | Positive |
| Population w/ Bach Degree % | 0.786263 | 0.0034133 | None indicated | Positive |
| Unemployment % | 0.0355238 | 0.185708 | None indicated | Positive |
| Minority % | 0.0787755 | 0.133802 | None indicated | Negative |

**Class Sizes (Student/Teacher Ratio)**

Class size appears to be driven predominantly by economic factors. The only data points that showed a correlation to class size were the cost of living index and median household income. As those rates increase, so does the Student-to-teacher ratio. This could indicate that students in wealthier counties have perhaps more enrichment/services available to them, leading to fewer behavioral problems, thus requiring less teachers to manage. This is purely speculation and is an area for further research.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dimension** | **P-Value** | **R2** | **Significance?** | **Direction** |
| Cost of Living | 0.0394886 | 0.178865 | Moderate | Positive |
| Median HH Income | 0.0519998 | 0.160971 | Moderate | Positive |
| White % | 0.614531 | 0.011724 | None indicated | Negative |
| 2018 HS Attainment % | 0.146112 | 0.093537 | None indicated | Positive |
| 2018 HS Enrollment | 0.147566 | 0.0928984 | None indicated | Positive |
| Poverty % | 0.231539 | 0.0643832 | None indicated | Negative |
| Population w/ Bach Degree % | 0.369035 | 0.0368228 | None indicated | Positive |
| Unemployment % | 0.395725 | 0.03456 | None indicated | Negative |
| Minority % | 0.614132 | 0.0117503 | None indicated | Negative |

**Quality of Life**

As expected, the data points that specifically factor into a quality of life calculation had high statistical significance. This included Median household income, percent of families in poverty and unemployment. Cost of living increases as the median household decreases and the percentage of families in poverty and unemployment rate decreases.

Interestingly, race seemed to show up as moderately significant, though it may be more telling of that race’s predilection toward higher income levels. On the left, are populations that identified as Asian Alone, Hispanic/Latino (as an ethnicity, not race), Biracial, or other. Each of those showed a moderate to highly significant positive correlation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dimension** | **P-Value** | **R2** | **Significance?** | **Direction** |
| Median HH Income | < 0.0001 | 0.912834 | **High** | Positive |
| Poverty % | < 0.0001 | 0.52556 | **High** | Negative |
| Unemployment % | 0.0086164 | 0.274341 | Moderate | Negative |
| Hispanic or Latino Ethnicity % | 0.0036072 | 0.325394 | **High** | Positive |
| White Alone % | 0.137717 | 0.0973588 | Not Indicated | Negative |
| Minority % | 0.138014 | 0.0972197 | Not Indicated | Negative |
| Asian Alone % | 0.0003536 | 0.447216 | **High** | Positive |
| Native Hawaiian % | 0.33195 | 0.0428208 | Not Indicated | Positive |
| Two or More Races % | 0.0012316 | 0.384366 | **High** | Positive |
| American Indian % | 0.30247 | 0.0482343 | Not Indicated | Positive |
| Black Alone % | 0.657609 | 0.0090926 | Not Indicated | Positive |
| Some Other Race % | 0.0361968 | 0.184496 | Moderate | Positive |