**Problem:** Healthcare costs are rising worldwide. The United States is one of the biggest spenders of total GDP on healthcare, but with middling results. The American Heart Association projects that by 2035, cardiovascular disease costs will exceed $1 Trillion. By looking for influential factors, we might be able to slow the upward trajectory of these costs.

**Client:** US Policyholders and Healthcare systems

**Dataset:** A combination of the US Census Poverty dataset and the University of Washington’s Institute for Health Metrics and Evaluation Cardiovascular Mortality Rate dataset.

**Other Potential Datasets:** Due to our initial findings below, we see that our dataset may require additional supplementation. We could use datasets that include income, education, and occupation. These datasets would probably be available from the US Census Website.

**Initial Findings:** Our initial suspicion was that poverty rates would be good indicator of cardiovascular mortality. It does not seem to be the case. At least, there does not seem to be a strong correlation between the two. The only variable in our initial dataset that correlates strongly with the Mortality Rates is the year variable.

We were able to create a new variable that would assign each state to a geographical location. Grouping the data by location did not affect the statistical significance of the Year variable.

It is clear that there are unaccounted factors driving down the mortality rates over time. The question to left to answer is, can we use the Poverty Rate as a statistically significant predictive variable?