**Grade Crime Correlation Analysis**

Date Completed: June 1st, 2020

By Brendon Carroll

**Introduction**

The goal of this analysis is to find if there is a correlation between highschool GPA and the amount of crime in a given state.

**Data**

The dataset used for this study is a combination of two datasets. A crime report for all 50 states between 1960 and 2012, along with a school grade report. The crime dataset contains total counts of crime along with rates, it also includes the type of crime, e.g., total violent robbery. The grade data set contains information about GPA within a subject, SAT scores by subject, and salary info for students families.

**Process** / **Explanation**

In order to combine datasets they must be sorted by a common variable. The variable chosen to sort by is the state. After combining datasets I run a proc corr using the pearson correlation method on the new dataset. I correlate the average GPA within ELA subjects and the average GPA within math subjects with total counts of all property crime and all violent crime along with total counts of all violent robbery. In one instance for my study of Illinois, I found a negative correlation coefficient between average ELA GPA and total property crimes of -0.90212 and a P value of 0.0001. A correlation coefficient ranges from -1 to 1. -1 being a strong negative correlation, 0 being no correlation, and 1 being a strong positive correlation. The P value measures the odds of my findings being random (lower is better). For instance, a P value of .0354 means there is a 3.54% chance my findings are random. Along with the proc corr I printed scatter plots between the variables I correlated to help the viewer see the connection. Below these I printed a proc sgplot (scatter plot) showing the year on x plane along with all violent crime and ELA average GPA on the y plane. I repeated this process for both Illinois and North Carolina.

**Conclusion**

I have concluded that there is a strong correlation between GPA and Crime rates in a particular state. All results contained very strong negative correlation coefficients along with miniscule P values. As grades go up, crime goes down. One interesting find was ELA scores appear to have a stronger correlation with crime than math scores. Of course, correlation doesn't necessarily mean causation.

<https://corgis-edu.github.io/corgis/csv/state_crime/>

<https://corgis-edu.github.io/corgis/csv/school_scores/>