**Crime by City Data Analysis**

**MIKE ANDERSON // MARK ATHANANSON // HEATHER BELL // KELLY BROWN // THOMAS STOKES**

**INTRODUCTION:** Hopefully, you have had some downtime in your life between the year 2000 and now to have watched the hilarious movie *Miss Congeniality*. There is an infamous scene where the host of the Miss United States pageant asks Sandra Bullocks’ character (who was an undercover cop posing as a beauty contestant), “What is the most important thing our society needs?” Her answer: “… harsher punishment for parole violators …. And world peace!” This scene was a (semi) perfect portrayal of what many people desire: world peace, no more violence! It is the most important thing our society needs.

In this study, we gathered data on five fairly high-populous cities to assess whether the cities’ violent crime rates had any impact on the productivity of these societies.

**HYPOTHESIS:** If a city’s violent crime rate increases, there will be an inevitable drastic and economically devastating crash in graduation rate and median income and a heart-wrenching increase in the city’s overall poverty level, widening the chasm that keeps us all from experiencing the utopian world peace that is so desperately desired. However, if the violent crime rates plummet, then surely there is hope for us all! We will compare all results for each factor against violent crime rates with the national averages.

**NULL HYPOTHESIS:** It will be utterly sad to discover that the rate of violent crime has absolutely no impact on societal growth (graduation rate, median income, and poverty level), thus causing us all to reconsider our vicarious vote for Bullock’s character Gracie.

**ANALYSIS:** Using Jupyter Notebook, we organized, compiled, and analyzed our data for New York City, NY, Los Angeles, CA, Chicago, IL, Tampa, FL, and Washington, D.C with the national average being the comparative standard. *Please note that in the year 2014, Chicago changed its classification laws concerning violent crime, which resulted in obsolete violent crime data for Chicago between 2006-2013.* All statistics and line regressions were retrieved with ANOVA.

Violent Crime Rate Comparison

Violent crime can be defined as a crime in which a perpetrator uses force upon his or her victim(s). The arrests were categorized based on the cities’ specific classifications of “force upon his or her victim”. Each city displayed its own trend, though there seems to be a commonality of decrease in violent crime rate between 2006 and 2009. The national average continued to steadily decline through 2017 while the Five Cities experienced several fluctuations for the remaining eight years.

|  |
| --- |
| ANOVA  F = 11.3855  P = 0.0000118 |

The pvalues for Tampa and the U.S. show to be of significance whereas the other cities’ pvalues disprove the hypothesis.

|  |  |  |
| --- | --- | --- |
| **City** | **rvalue**  Linear Regression | **pvalue**  Hypothesis Test |
| Los Angeles | -0.186 | 0.563 |
| New York City | NaN | NaN |
| Chicago | NaN | NaN |
| Tampa | -0.897 | 0.0000757 |
| Washington, D.C. | -0.449 | 0.143 |
| U.S. Average | -0.871 | 0.000222 |

Graduation Rate Comparison

High school graduation rates in the Five Cities proved to increase between 2006-2017, with Washington, D.C. reaching far beyond the national average and Los Angeles lagging significantly behind.

The pvalue is very small, indicating that it supports the idea of increasing graduation rate throughout the years of the data collection.

|  |  |  |
| --- | --- | --- |
| **City** | **rvalue**  Linear Regression | **pvalue**  Hypothesis Test |
| Los Angeles | 0.994 | 0.00000500 |
| New York City | 0.992 | 0.0000119 |
| Chicago | 0.993 | 0.00000848 |
| Tampa | 0.975 | 0.00019 |
| Washington, D.C. | 0.994 | 0.0000059 |
| U.S. Average | 1.0 | 1.9206 x 10-50 |

|  |
| --- |
| ANOVA  F = 96.20  P = 8.48 x 10-20 |

Poverty Rate Comparison

Poverty rate was calculated by the percentage of households earning less than $25,000 annually. Impressively, all cities and the national average experienced a relatively steady decline in percentage of poverty over the eight-year period assessed in this study, with Washington, D.C. being the only city with a smaller percentage of people experiencing poverty within its borders than the national average.

The pvalue is very small, indicating that it finds significance in relation to the hypothesis and the decrease in poverty rate over the years.

|  |  |  |
| --- | --- | --- |
| **City** | **rvalue**  Linear Regression | **pvalue**  Hypothesis Test |
| Los Angeles | -0.893 | 0.00286 |
| New York City | -0.958 | 0.000178 |
| Chicago | -0.957 | 0.000187 |
| Tampa | -0.912 | 0.00158 |
| Washington, D.C. | -0.95 | 0.000295 |
| U.S. Average | -0.908 | 0.00177 |

|  |
| --- |
| ANOVA  F = 216.92  P = 7.49 x 10-29 |

Median Income Comparison

All Five Cities experienced a steady rise in the median income earned by working people.

The pvalue is very small, indicating that it finds significance in relation to the hypothesis and the increase median income over the years.

|  |  |  |
| --- | --- | --- |
| **City** | **rvalue**  Linear Regression | **pvalue**  Hypothesis Test |
| Los Angeles | 0.989 | 0.166 |
| New York City | 0.989 | 2.47 x 10-5 |
| Chicago | 0.971 | 0.000272 |
| Tampa | 0.963 | 0.000482 |
| Washington, D.C. | 0.988 | 3.1 x 10-5 |
| U.S. Average | 0.828 | 0.0215 |

|  |
| --- |
| ANOVA  F = 176.65  P = 2.82 x 10-24 |

To gather the impact that violent crime had on its cities between 2006 and 2017, scatter plots were developed yielding the following information:

Crime Rate vs. Graduation Rate

Though high school graduation rates proved to have increased across the board, violent crime rate remained inconsistent and without pattern.

Crime Rate vs. Poverty Rate

Though the level of poverty decreased rather significantly within the Five Cities, violent crime rate remained inconsistent and without pattern.

Crime Rate vs. Median Income

Though the working class earned greater income, the violent crime rate remained inconsistent and without pattern.

Graduation Rate vs. Median Income

It is crystal clear that as the city’s high school graduation rate increased, so did the city’s median income.

With a positive rvalue for all points, we can generalize that as graduation rates increased, median income also increased.

The pvalue is indicative of significance between the two variables.

Regression for All Points

R = 0.705

P = 1.82 x 10 -7

**CONCLUSION:** The conclusion of this report is accompanied with intense grief and sunken hearts. According to the data, there is no generalized relationship between violent crime rate and the societal factors examined. The city variable has the strongest relationship of the other independent variables, according to the ANCOVA. Crime rate largely depends on the city and can not be compared from one city to the next. Unfortunately, graduation rate, median income, and poverty rate do not play large factors in the violent crime rate.

So, what does this mean for us world peace seekers? Perhaps we should hop on the next rocket ship to Mars to begin our Peace Colony there. Ready for take off!