Section 2: Week 4: Analyze Statistical Output

Nate Bachmeier

TIM-7101: Statistics for Technology Leaders

August 23rd, 2020

North Central University

# Analyze Statistical Output

NCU-Cares (NCU-C) is a politically neutral non-profit that seeks to improve the work through targeted lobbying efforts. The death of George Floyd has risen the debate of police violence and reform to the national stage (Crary & Morrison, 2020). While the topic rests on American’s hearts and minds, it has also become highly partisan with many efforts to undermind the conversation (McCaskill, 2020). It can be challenging to find complete records as federal regulators do not mandate data collection policies. The Washington Post attempts to fill this gap with its dataset based on news events, exposing the location, context, and several demographic markers for each fatality (Washington Post, 2020). Researchers can use these results to help determine if the solution to police violence comes from racial issues or another aspect better explains the national issues.

# Section I: Describe the Problem and Hypothesis

Annually nearly one thousand citizens die from police violence, which raises the central question, why? The collective perspective of the Black Lives Matter movement is that police exert disproportionate force against people of color (BLM, 2020). This perspective often comes with the quote that “Black civilians were more than twice as likely as White civilians to be unarmed (Nix, Campbell, Byers, & Alpert, 2017)” during the fatality. Nevertheless, others argue the brutality victims are experiencing a mental health crisis, and this is the actual reason (Lamb, Weinberger, & DeCuir, 2014). While these perspectives effectively drive media headlines, are they both missing the forest among the trees? Does another factor more accurately explain the challenges that are occurring? Instead, NCU-C hypothesizes that neither sanity nor race is the driving cause of police violence. Alternatively, provocation might better explain the need for forceful escalations that result in death.

# Section II: Describe the Data Set

The Washington Post provides demographic and contextual information about victims from January of 2015 to the present day. Each entry captures the threat level, flee status, any weapons, age, gender, race, and city. NCU-C enhanced these 5489 records to include several nominal features, such as ‘has a projectile,’ to simplify analysis on the free form weapons column. The inclusion of an ‘age group’ property also exists for smoothing visualization charts by partitioning into five-year windows. Aside from these transformations, no alterings of the original data set are present.

## Exploring by Race

America’s racial make-up is approximately 63% white, 15% Hispanic, 13% black, and 9% other (Census Bureau, 2019). If all things are equal, then looking at the raw victim statistics should convey a similar breakdown. These initial expectations are comparable though slightly skewed in Washington Posts’ data set when grouping by *race* (see Figure 1). After adding a second level of grouping by *year*, it also raises an observation that the number of victims is relatively stable across time. From January 2015 to December 2019, the mean death rate is 905, with a standard deviation of 35. While the situation is not getting any better, it is also not becoming worse.

Figure 1: Victims by Race



## Exploring by Age

A normal distribution exists for the victim’s age around the mean of 37 with a standard deviation of 13 years. After grouping by race, the data shows that minorities encounter deadly confrontations with the police roughly seven years younger. From these initial range values, it is possible to calculate the statistical effect of a person’s *race* and *age* relative to a similar group. An effect size is a measurement in z-scores with values typically between zero to one. This two-level comparison conveys that a medium-level effect exists for Whites, and a minimal difference exists between Blacks and Hispanics (see Table 1). These results roughly align with the exploration of race, which suggests that a skew exists in the data, but its not the smoking gun.

Table 1: Influence of Age

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Race | Mean | Standard Deviation | Effect vs. White | Effect vs. Black | Effect vs. Hispanic |
| All | 37.12 | 13.12 | -0.11 | 0.38 | 0.30 |
| White | 39.95 | 13.37 | -- | 0.61 | 0.54 |
| Black | 32.47 | 11.33 | -0.38 | -- | -0.10 |
| Hispanic | 33.54 | 10.87 | -0.30 | 0.10 | -- |

## Exploring by Sanity

An argument exists that the solution to police violence is defunding the police, and using those resources for drug rehabilitation and civil service programs (BLM, 2020). Assuming those changes went into effect, would it make a difference?

# Section III: Limitations and Findings

# Conclusion