Section 3: Week 7: Formulate a Strategic Plan

Nate Bachmeier

TIM-7101: Statistics for Technology Leaders

September 13, 2020

Northcentral University

# Formulate a Strategic Plan

NCU-CARES (NCU-C) is a politically neutral nonprofit organization that seeks to make the world a better place by lobbying for policy changes to social-economic challenges. The group begins each project by understanding the landscape of an issue through the lens of statical models. These models feed into every aspect of the decision process to maximize the impact while also minimizing resource expenses. After recently receiving a sizable donation, the institution can hire a dedicated staff to pursue one more initiative. Rarely does such an event occur, and NCU-C does not want to waste this opportunity. The organization chose to focus on police violence and devise a campaign to address its challenges.

# Executive Summary

In May, thousands of people joined in protest over the death of George Floyd, raising to the national debate stage several critical questions about police violence. This debate pointedly asks if systematic racism is disproportionally killing minorities, and what changes would be most effective at preventing these issues. It can be challenging to examine the situtation pragmatically because this is not a race issue, but a human issue, and full of emotional responses on both sides. A reoccurring theme in the conversation proposes defunding law enforcement budgets; instead, proposing those monies should flow to mental health and related civil services. Others suggest the opposite believe a solution requires more funding toward better training programs.

NCU-C wants to understand this highly-partisan environment so that it can introduce impactful changes at the crux of the problem. Accomplishing this goal requires a data collection, analysis, and inference of facts that answer these questions. Specifically, is police racially bias and would investing in mental health services resolve the scenario? Using the Washington Post Police shooting dataset, the organization concludes the answer to both is no. Police brutality is an onion, and it stinks. Ideally, no one should ever die, but that does not mean the officers are racially profiling victims. Nor can one entirely blame any mental illnesses in these situations. Peeling the next layer places the focus on weapons at the scene of the crime. According to simple aggregation counts, nearly 72% of all victims possessed a gun or knife during the fatal altercation. There needs to be additional research into this cluster as it presents the highest safety risk and greatest reward. Of the remaining people either unarmed or defending themselves with low-risk weapons (6-12%), these also warrant further thought. However, it might be impossible to eliminate the unarmed fatality group, due to only representing 0.00002% (119 of 53 million) of police interactions per year.

NCU-C should focus on strategies to reduce the interactions with police and armed civilians, specifically guns and knives. When

# Statement of Problem and Hypothesis

## Why this topic matters

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). On the one hand, an argument exists that defunding the police will force systematic change (BLM, 2020). After cutting the law enforcement budget, the state department could repurpose those monies into civil and medical services. On the other hand, are concerns that these changes would enable a “symbol of hate (Trump, 2020)” and reducing the safety of all parties. It is unlikely that either side is entirely right or wrong, and this situation requires an unbiased mediator to assess claims quantitively before the punitive rhetoric will abate (Smith, 2020). Working to restore public confidence and fill this gap represents a unique opportunity for the organization.

## Potential Research Areas

The central idea of the Black Lives Matter (BLM) movement is that police violence disproportionally victimizes people of color (Pierce, 2019). Assuming this statement is true, to what extent is this true? Data collectors are quick to cite that “black people represent 24% of all police killings, despite being only 13% of the population (KilledByPolice, 2020).” However, can these two data points be uniformly compared? Alternatively, does a demographically adjusted accounting provide greater insight into racial injustice hotspots? Processes that can uncover such disparity could lead to laser-focused policies versus broad debate on the national stage. These policies would not represent the final stage, but do offer a path for measurable short term improvements.

Another central theme is that shifting funding from police departments to civil services will change the risk calculus. Assuming this statement is true, to what extent? What portion of the population is going through a medical crisis during their time of demise? Until examining the data, it can be challenging to separate the norm from media machines selling advertising. Perhaps a more accurate perspective is that scenario-specific categories exist, and additional training programs can target those situations, reducing the mortality rates.

# Overview of Design

## Data Sourcing

There are numerous strategies for approaching this problem with varying levels of sophistication and planning. One standard solution is to perform statistical application analysis on the Washington Post’s police shootings data set (Nix, Campbell, Byers, & Alpert, 2017). This data source contains demographic, location, and contextual information on all publically known fatalities of police violence between 2015 to the present (Washington Post, 2020). While there are several limitations to this aggregate feed, it does provides an unbiased sampling of the broader population.

## Identifying Questions

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 efficiently 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 violent escalations that result in death.

## Collecting the Results

Figure 1: Victims by Race



NCU-C began the exploration process by collecting descriptive statistics about the victims in terms of *age*, *race*, and *sanity* (see Figure 1). Using a variate of measurements and pivots, then assessed the statistical effect of these features against one another. These suggest that people of color tend to have fatal altercations younger than White people (see Table 1).

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 | -- |

DeCarlo (2018) states that quasi-experiments are particularly useful in social welfare policy research (see chapter 12.2). Under a quasi-experiment, the researcher team does not use random assignment and instead looks at different populations. This method could be highly effective for examining the impact of both *race* and *sanity* variables. Using this method to perform all-versus-one comparisons concludes that *race* and *age* are not the dominant traits. Similarly, *sanity,* asreported in the signs of mental illness column, does not a robust statistical prediction of these records (see Table 2).

Table 2: Effect of Mental Health

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mental Illness | Mean (Age) | Standard Deviation | Effect vs. All | Effect vs. Sane |
| Yes | 39.74 | 13.87 | 0.20 | 0.25 |
| No | 36.35 | 12.80 | -0.06 | -- |

## Alternative Variable

The data set includes the ‘armed’ column that contains free-form text describing any weapons on the victim. One of the challenges with analyzing this field comes from the various subtle differences in its values (e.g., *baseball bat* versus *baseball bat and bottle*). Enhancements of each record include categorical-features that bucket the weapons by genre. These buckets are named projectiles, sharp/blunt instruments, tool/small objects, explosive, unspecified, vehicles, and unarmed. When the suspect has multiple weapons, such as both gun and knife, the higher risk object dictates the category. According to these categorical-aggregations, roughly 58% of victims had a firearm, plus another 18% had a sharp/blunt instrument (see Figure 2).

Figure 2: Victim Weapon Category



# Research Limitations and Challenges

The four major threats to research projects are internal, external, statistical conclusion, and construct validity (Parker, 1993). While efforts to minimize these risks do exist, the time and resources of this project are relatively finite.

## Internal Threats

An internal threat exists when the researcher does not accurately represent the results. This scenario could exist from biases during the categorization and groupings of the victims. For instance, the Washington Post data set uses a free-form text field to record the weapon and threat level. The analysis uses a few general buckets versus other researchers propose using more fine-grained options (Nix, Campbell, Byers, & Alpert, 2017). Minor forms of selection bias might also exist due to the filtration rules of the data set. Expressly, the analysis only includes records that provide the racial demographic and location of properties.

## External Threats

An external threat comes from a variable that is outside of the researcher’s control. The government does not require law enforcement agencies to report incidents that result in police brutality. Since official sources do not exist, researchers must rely on open-source data sets like the ones provided by the Washington Post. The Post uses news and social media reports, which could be both erroneous and lossy. There are also risks that the manual entry process could have inaccurate values for a record in the table (e.g., wrong *race*). Another class of risks comes from the data set being immature and starting in 2015. Ideally, having more longitudinal data to understand trends or alternative sources for cross-validation would improve the validity of results.

## Statistical Conclusion

Invalid statistical conclusions arise from not having sufficient samples or encountering too many uncontrolled parameters. The filtered data set only contains 4937 records, which does not provide sufficient evidence for some pivots (see Figure 5). For example, assessing Asian and Native American victims independently likely result in overfitting. Due to time constraints, the evaluation of only a subset of features took place, and this could have missed an important or confounding variable (e.g., age group).

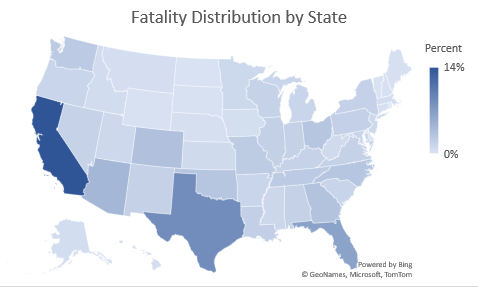
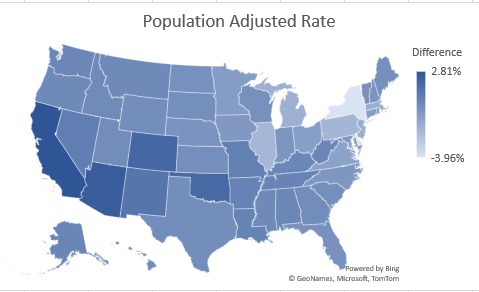
## Construct Validity

Threats to the construction of the experiment occur when controls do not protect against information leaks between tests or controls between variables. These risks might exist due to the analysis evolving with the research project. While a general outline and strategy exist, the budget to procure sufficient causes a more relaxed set of requirements than during the onset. It is also possible that initial observations encouraged exploration of specific portions of the dataset because it more easily aligns with the topic. Instead, a more thorough effort could exist that examines other pivots and asks other questions of the information.

# Overview of Analysis

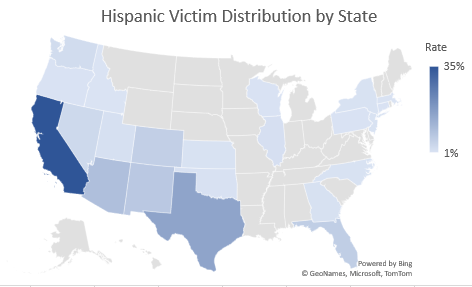
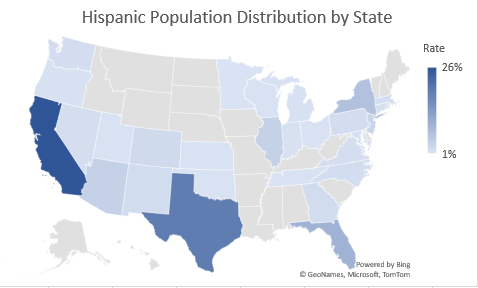
## Demographically Adjusted Results

Figure 3: Fatalities by State (2015-2019)

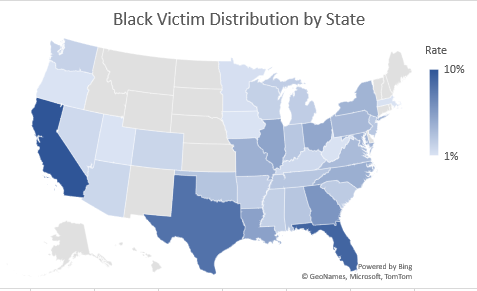
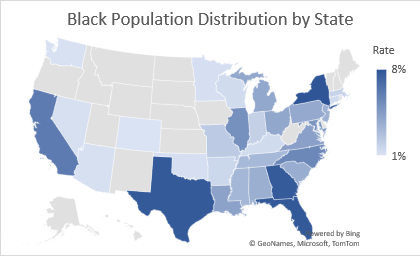
NCU-C’s investigations show that the method of sub-grouping and classification has significantly more impact on the results than any arbitrary feature. For instance, of the 4937 records, California and Texas own 14% and 9% of incidents, respectively (see Figure 1). There might be specific challenges within those states; however, they are also home to 12% and 9% of the national population (Census Bureau, 2019). Assuming all other things are equal, the actual number of fatalities less expected fatalities based on population shows the rates are relatively stable across the country.

Figure 4: Hispanic Victims by State

Next, an analysis of the total number of Hispanic victims in each state shows a strong correlation with the group’s natural population distribution (see Figure 2). This perspective highlights some of the challenges associated with treating the country as a homogenous cluster. Instead, decomposing America into regions allows for a more accurate assessment of the relevant population counts to include. For example, since at least 2015, no police officer has killed a Hispanic person in Montana (Washington Post, 2020). However, this group also only represents 38,000 (less than 4%) of that state’s inhabitants (Census Bureau, 2019). Then consider that Texas has over 11 million (39.4%) residents of Hispanic and Latino descent. Police violence in this state has killed 432 people during this same period, of which 143 (33%) were Hispanic.

Figure 5: Black Victims by State

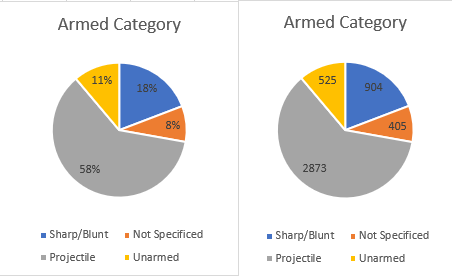
 

An extrapolation of comparable ratios exists in many other locations and ethnic groups. For example, approximately 39 million Black citizens live across the United States, of which 1277 have been killed by police violence since 2015. By plotting the distribution of these groups in terms of population per state, results in similar charts (see Figure 3). These pivots suggest that the deaths are not racially motivated, and instead, a function of the locale-specific population make-up. If that was not the case, then more pronounced outliers should exist with substantially higher victim rates relative to the group’s population. This outcome also hints that some other motivator outside of the *race* is likely to exist.

## Potential Motivators

There are fifty-three million people who have an interaction with the police each year (BJS, 2015). Annually, 167 (+/- 6) officers and 66 (+/- 16) unarmed civilians die during these altercations (NLEOMF, 2019). These situations have to introduce a wide range of emotions and hostility, which creates challenges while balancing the safety of officers and civilians. NCU-C should invest additional resources into exploring solutions to reduce this friction. For instance, of the incidents that result in death, 76% (3853 of 4937) involve the suspect having a weapon (see Figure 4). Training and procedures could exist to address these scenarios in a manner that improves the probability of a peaceful resolution. While it does not address all of the concerns, this is a significant source of police brutality. Specifically, the unarmed and unspecified groups need further consideration to reduce the loss of life for these subgroups. However, it might be impossible to eliminate due to this group representing 0.00002% (119 of 53 million) interactions per year.

Figure 6: Suspect Armed Category 2015-2019 Total



# Strategic Plan of Action

## Components of Plan

## RACI Actors

# Conclusions