Section 1: Week 1: Business Case for Statistical Applications

Nate Bachmeier

TIM-7101: Statistics for Technology Leaders

August 2, 2020

North Central University

# Business Case for Statistical Applications

## Organization Background

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. NCU-C’s portfolio contains several high profile efforts such as reducing climate, improving access to clean water, and providing medical resources to underserved nations. After receiving a sizable donation, the institution can hire a dedicated staff and pursue one more initiative. Rarely does such an event occur, and NCU-C does not want to waste this opportunity.

## Selection Criteria

Inadequate planning, scope creep, and poor communication are three common reasons that projects fail (Jain, 2018). NCU-C addresses these concerns with acceptance criteria that new efforts must adhere too. For instance, the effort needs to be (SMART) Specific, Measurable, Attainable, Relevant, and Time-based (Indeed, 2020). When the goals are unrealistic or prohibitively expensive, then it leads to waste. Completing a social-economic development requires years of involvement, and this necessitates the ability to measure progress along the way. One of the challenges with addressing macro-trends is the costs that arise from data collection. Instead, there need to be aggregate data sources or random sampling procedures that can feed into statical models and produce generalizable insights about the population. Outside of these operational requirements, the project must align with the core mission of making the world a better place.

# Section I: Problem Statement

## Background

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. On the one hand, an argument exists that defunding the police will force 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. The 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 NCU-C.

## Research Questions

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 represent a path for measurable improvements in the short term.

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, improving the mortality rates.

# Section II: Measurements

## What are statistical applications

The purpose of statistical applications is to create a model that describes the interactions of a system (Denis, 2015). A model attempts to map input parameters to output results while taking into account the likelihood of those decisions. For instance, assume that a fair coin flip lands on heads fifty percent of the time. This distribution can feed into a model that derives the chances of getting heads three times in a row is 12.5% (= 0.53). George Box is famous for stating that “all models are wrong, but some are useful (Denis, 2015, p. 3).” The quote infers that it can be challenging to include all aspects of the environment, and these external factors can create a delta between expectations and reality.

## Forming the statistical models

Answering the research questions for this topic requires demographic information about victims of the police, traffic stops metrics and county census data. These sources can feed into a model that assesses the risk of mortality present to each class of citizens. That assessment needs to express descriptive statistics across both horizontals (e.g., age and income) and verticals (e.g., nationality and race) groupings. These groups can bubble up latent feature dependencies, such as poverty-stricken people might be more influential than race. Uncovering these intricate details requires further investigation into the relative log-likelihood between grouping.

Another critical piece of the puzzle is determining the validity in defunding the police and pivoting toward civil service investments. According to a cursory investigation, most efforts on this front have been symbolic at best (Smith, 2020). Instead, a model needs to exist for better categorizing risk levels during police stops. For instance, how many fatalities involve mental illness, drug abuse, or are unarmed? These features might unlock additional latent features that enable the officer to operate differently while still ensuring personal safety.

## Hypothesis Testing

The purpose of a hypothesis is to predict the outcome of an experiment using empirical evidence (Denis, 2015). This prediction often accompanies a null-hypothesis, which makes the opposite claim. Depending on the scenario proving the null-hypothesis might be more straightforward than the underlying hypothesis. For instance, an initial assumption might state that civil services are necessary at *every* police stop. There is virtually an unlimited number of scenarios that result in police stops, and its instead easier to find an instance that does not gain value.

One of the challenges with writing an adequate hypothesis test is that it must be precise and avoid ambiguity. For example, another assumption could state that large scale civil and mental services are not useful because people are generally sane. However, what defines ‘sane’ in this context? While many people fit the clinical definition, a misunderstanding during the stressful confrontation with patrols could trigger a ‘fight-or-flight’ response. Are these same people still sane during this altercation? If an altercation can trigger temporary insanity, what prevents the officer, another human, from having a similar episode? It would be highly complex to include every possible feature within a research study. Instead, efforts must convey the scope of the study and which aspects lack consideration.

## Moments and Expectations

When an analyst can quantify the likelihood of a value appearing in the data, it enables them to

A moment refers to statistical properties about the data distribution, such as mean and variance (Denis, 2015).