

# Big Data Analysis

*Examining the Meaning of Boston's 911 Call  
Data and Implications for Public Policy*

Dr. Jack McDevitt, Keller Sheppard, and Stephen Douglas

Northeastern University

School of Criminology and Criminal Justice

# Contents

- Contents ..... 2
- I. Introduction to Big Data Analysis ..... 3
  - A. Data Usage in Public Policy ..... 3
  - B. Guiding the Use of Big Data in Public Policy..... 4
  - C. The Involvement of Stakeholders in Big Data Analysis..... 5
  - D. Practical Concerns..... 6
  - E. Ethical Concerns ..... 7
  - F. Addressing the Limitations ..... 8
- II. Big Data Analysis Illustrated: An Introduction to 911 Data ..... 9
  - A. Incident to Operationalization: How Crimes Become Data ..... 11
    - 1. Barriers to Crime Reporting ..... 11
    - 2. Accuracy of Reporting ..... 13
    - 3. Summary of Potential Bias in Citizen Requests for Service ..... 14
  - B. Categorization of Incidents ..... 15
    - 1. Purposeful Categories ..... 15
    - 2. Accurate Incident Classification ..... 16
    - 3. Interpreting and Operationalizing the Data ..... 17
  - C. Methodological/Practical Decisions ..... 17
  - D. Ethical Considerations..... 19
- III. Proposed Approach ..... 20
  - A. Overview of Dataset..... 20
  - B. Categorizing Incidents ..... 21
  - C. Unit of Analysis: Levels of Aggregation ..... 24
  - D. Limitations, Technical Challenges, and Interpreting Results..... 28
  - E. Using the ‘Big Data’ Approach ..... 29
- IV. Political Implications ..... 30
- References..... 32
- Appendix – NIBRS Offense Grouping and Hierarchy ..... 37

# I. Introduction to Big Data Analysis

Big data is a term that is often used to describe the combination of information across various heterogeneous sources, primarily for the purpose of identifying greater patterns among populations (Davidson, 2017). Big data's prominence in data science has grown in recent years and is increasingly being used to inform decision making within a variety of policy fields. The perceived utility of big data in public policy is understandable, as it is capable of providing previously unobtainable levels of detailed information regarding patterns and trends among populations, groups, and individuals (Chan and Moses 2017). Policy makers and researchers who advocate for its use view this capability as an asset that facilitates the creation of more effective policy, with many going as far as to suggest there is no longer a need for theory, given that big data can provide 'all' of the necessary information (Chandler 2015). However, the increased usage of big data analysis should be accompanied by a healthy dosage of caution, given the complexity and potentially invasive nature of this type of analysis. This represents an issue that has received little attention within the existing literature. This paper will provide an overview of big data analysis, its current place in public policy, and an assessment of the general practical and ethical concerns related to its use. The Boston Police Department's (BPD) 911 call data will then be provided as an example of how big data may be treated and used effectively. Steps relating to the collection, operationalization and interpretation will be discussed in turn, with the overall aim of informing those involved in using this type of data of its inherent limitations and biases. Finally, a discussion of the potential political challenges related to the use of big data will assist those who seek to navigate the pitfalls associated with informing public policy.

## A. Data Usage in Public Policy

Overall, the evidence appears to indicate that big data analysis represents the next step in policy creation for many industries and government services (Barocas and Nissenbaum 2014). Its potential for public policy could be especially beneficial for social service agencies seeking to better serve their target populations. Data generated for the purpose of improving housing policy or traffic systems are generally collected and utilized by public entities (usually local, state, or federal government), which means that the use of the data is subject and accountable to established measures of oversight. Public perception of this phenomena, however, is unduly influenced by private industries' exploitation of consumer data. The use of big data by private commercial interests poses a different issue in that the purpose of the analysis is primarily profit driven, which has prompted ethical questions regarding these techniques. High-profile revelations about extensive data collection and analysis by social media and the telecommunications industry dominates the public's perception of big data analysis as a whole. The ability of private interests to monitor digital footprints and online consumer habits allows them to tailor their approach to commercial platforms; controversies surrounding the sharing of consumer data has brought big data analysis into the light and it has largely been met with scrutiny and criticism. As a result of these controversies, and public sentiments towards big data, policy makers' and academics' widespread participation in this trend must differentiate themselves from the motives of private industry. Even though public policy makers' use of big data is intended to improve government services, it can result in social harm and ethical quandaries if misused. Therefore, the ethical and effective use of big data analysis for policy decisions necessitates a careful consideration of issues surrounding data collection, storage, and analysis (Crawford et al., 2014). Its utility and appeal as a policy tool will only

expand as researchers, practitioners, and other parties develop innovative uses of current big data and organizations increase their data collection capacity.

Although big data represents a highly valuable resource for both researchers and practitioners, its use also carries a series of inherent concerns and challenges. Within the field of policy creation, these challenges primarily center upon the fact that data that was originally collected for a specific purpose is now increasingly being used to address other research questions and topics, a process typically referred to as secondary data analysis. These issues require extensive consideration of the data's quality and a critical assessment of the process by which raw data is transformed into actionable findings that can then be used to inform policy decisions (Davidson 2017). In order to fully understand the challenges presented by the use of big data, it is first vital to consider the nature of policy creation and the role of data in decision making. Despite the perception that the emergence of big data has coincided with the increased demand for informed decision making and evidence based policy, policy has been based on various forms of data for an extensive period of time (McElheran and Brynjolfsson 2016). Government bodies, politicians, and law enforcement officials have long relied upon record keeping, categorization, and other traditional forms of data collection and analysis in shaping policy decisions (Mattioli 2014). The crucial change that has occurred in the modern era has been the drastic increase in the amount of data that is available, as well as the improved capability of computers and statistical programs to collect, store, analyze, and draw meaning from such data (Mayer-Schonberger and Cukier 2013).

Aggregating information from multiple data sources capitalizes on the expanded capacity of data collection. Combining information in this way can often lead to confusing and unexpected findings, especially when one considers that most of these data sources were not intended to be merged in this manner. Resultantly, this type of analysis necessitates the use of specific analytic tools that can accurately identify these trends and translate the data into interpretable information capable of informing policy decisions. The potential of collecting and analyzing massive amounts of data in a rigorous and systematic manner is the major reason why big data analysis is so appealing to policy makers (Richards and King 2013). Big data's overwhelming size and apparent ability to produce quantifiable metrics of previously unmeasurable processes can be mistaken for unassailable method of constructing and assessing evidence based policy. Thus, it should come as no surprise that it is an increasingly common feature within policy creation and decision making processes. For example, policy focused upon generating affordable housing seized upon this movement. The ability to effectively identify the number of housing units created and assess the impact of such housing policy upon various outcomes (e.g. the number of people served by a given program or the number of jobs created by its implementation) has cemented big data's place in the decision making process (Chandler 2015). It is also prevalent within traffic management systems, being used to track traffic flow and calibrate traffic lights in Los Angeles (Jahanian 2015) and has proven to be an effective tool in fraud prevention, having been used to comprehensively identify fraudulent tax returns in Indiana, resulting in a saving of approximately \$85 million dollars (Davidson 2017). These examples provide evidence of big data's ability to inform policy programs and guide the generation of new policy initiatives.

## **B. Guiding the Use of Big Data in Public Policy**

The potential of big data analysis is especially enticing for public agencies that are under significant pressure to produce effective policy change within a limited time frame. The pressure cooker

environment of public policy, which features limited resources, serious social and crime problems, and increased scrutiny from the public and relevant stakeholders, demands immediate implementation of quantifiably effective programs and policies without the luxury of methodical evaluations (Chan and Moses 2017). For many, big data represents an efficient method for identifying complicated patterns among populations that may conceivably allow policy makers to achieve their goal of optimal and demonstrable performance. Agencies can point to constructed metrics arising from big data analysis and claim they represent the efficacy of their policies and programming. Reliance upon such measures of performance could ultimately serve to expedite the entire decision making process. However, although there is value in quantifying the effects of public policy, there is a significant concern that policy makers will proceed to build policy solely upon the results of big data analysis. Justifications for policy decisions that rely on big data analysis may appear to be concrete on the surface, but the complexity of this process threatens the infallibility of this approach. Instead, these processes should be complemented with a knowledge of the caveats and limitations that will assist in guarding against the common pitfalls of big data.

### **C. The Involvement of Stakeholders in Big Data Analysis**

Any attempt at harnessing the power of Big Data in the public sector must be preceded by the development of a process for identifying limitations in the raw data sets, controlling for these limitations, and responsibly interpreting any results. Prior to developing a specific analytic strategy, it is important for those using the data and those informing its use to have a clear collective vision of what the research is intended to achieve. This can be regarded as the first stage where relevant stakeholders can play an important role. Stakeholders can take the form of any group or body that hold an interest in the research being conducted, and could conceivably affect or be affected by its results (Miles 2017). A collaboration between public policy makers, researchers, and relevant external bodies—e.g. community groups, other social services agencies, and the general public—would permit a more nuanced consideration of the various facets potentially affecting the quality and interpretability of any data source being considered for analysis. For example, involving multiple stakeholders has become increasingly common in the field of law enforcement. Although this cooperative approach has led to many practical and technological advances, the mere presence of stakeholders is no guarantee of success as the quality of relationships and extent collaboration results in varying degrees of success (Engel and Whalen 2010). Recognizing the often strained relationship between law enforcement and researchers, Rojek et al. (2012) discuss the importance of abiding by a ‘governance framework’ that aims to implement a guideline for cooperation that can result in meaningful and positive change. Requiring mutual respect, this framework will result in the cultivation of partnerships between law enforcement, community organizations, academic institutions and political figures, with the goal of creating and ensuring the correct implementation of policy change. This type of cooperative framework can and should be applied to efforts by other agencies and public policy makers in their pursuits to improve their services. Big data analysis represents the next advancement in public policy assessment and this type of cooperative framework should be a central feature of its use in public policy.

In regards to the use of big data in business economics, Gupta et al. (2018) discuss the importance of the ‘supply chain’, a figurative term for partnerships where information is shared freely among all partners with the collective aim of providing benefits for all. Crucially, they note that trust between those involved in the analysis of big data is essential for this type of partnership to be effective (Gupta et al.,

2018). Considering this point, the process of shaping the research question should be a collaborative process, with amendments made on all sides of the 'supply chain' (Gupta et al. 2018). Forming the research question in this way will allow for a focused and narrow analyses that will gather and assess the necessary information in a systematic manner. Additionally, abiding by Rojek et al.'s (2012) conception of a governance framework, this cooperation can also ensure that the expectations for any resultant findings are managed appropriately. Stakeholders who form cooperative relationships with those directly involved in the collection and storage of the data source are more likely to be informed of the inherent limitations of the data, and be aware of their impact upon any interpretation of the findings. This type of cooperative framework is regarded as being essential if actionable results are desired from the analysis of big data.

Overall, these points highlight the importance of trust and cooperation among those conducting and informing the analysis of big data, both for generating focused research questions and managing the interpretation of findings. However, these features are also important in ensuring the effective implementation of policy changes that may result from big data analysis. Relating to this point, Davidson (2017) discusses the necessity for a feedback loop when implementing this kind of policy change. To illustrate this point, it is important to consider the nature of big data, which inherently includes crossovers between different policy fields and jurisdictions. For example, the Boston 911 call data is nested within the very real context of Boston, with calls for service placed within specific addresses and streets, which reside within neighborhoods, school districts, police jurisdictions, college campuses, and other geographical boundaries. This crossover, combined with the transitory nature of big data, makes it vital that stakeholders are able to evaluate and monitor the implementation as it occurs. Further, policy implementation is likely to change in line with developments in the 'research event'. Law (2004) describes the research event as a fluid phenomenon, which is particularly salient when analyzing 911 calls, where patterns and trends can develop or change within the space of hours. Crawford and Hutchinson (2015) add to this point by highlighting how projects of security (such as law enforcement policy) seek to provide concrete assurances about the future and generate expectations people can depend upon, as the occurrence of crime and public safety in general represents a vital concern for entire communities. These insights stress the need for effective communication between stakeholders and those using the data at all stages of big data analysis, but particularly underline the importance of continuing such evaluation once a new policy direction has been decided upon.

## **D. Practical Concerns**

Regarding inherent limitations of big data analysis, the first major concern relates to the issue of data quality, specifically centering upon what the data is actually intended to represent and its suitability in doing so. Data collection is often filtered through unmeasured social phenomena and further biased by categorization into a predetermined instrument. These features of data collection complicate the interpretation of the data itself, which proves to be especially true regarding big data. 911 calls exemplify the complexity of this process and vulnerability to unintended errors. In order for an incident to be reported, an individual has to perceive a reason to call the police. The way in which the information is then passed along to different elements in the chain of communication (from caller to dispatch to officer) will ultimately shape the final categorization of the incident, which may differ significantly from the actual event. The possibility of inaccuracies occurring in this manner can be related to numerous factors, including inherent bias, where Mattioli (2014) states that collected data is

often influenced by the subjective judgements of those collecting it. Understanding the political and organizational influences on the development and implementation of any dataset is key to any effective analysis of that data source.

Further, the presence of outliers within data of this kind represents a significant concern, especially for public policy. It is important to consider the nature of outliers in this process, as social harms may arise if outliers are permitted to dictate a policy response. For example, fire departments may build particular strategies around the possibility of a large fire or other emergency, which can ultimately be viewed as an acceptable occasion for an 'outlier' to define policy, as it is unlikely to result in unintentional social harm. If police agencies were to amend policy based upon the occurrence of outliers, such as an increase in the number of calls for assistance during a major storm, it is possible that it will result in greater public harm. Police officers who are reallocated as a consequence of rare events will then be active in areas where there was no demand for an increased police presence, which may lead to further issues surrounding interactions between the police and particular communities. On the other hand, social harm may also occur if significant events are dismissed as outliers when they actually hold considerable implications for policy. These points illustrate the importance of placing possible outliers firmly within the spatial and temporal context of their occurrence.

Beyond the technical and practical concerns of the public Big Data sources, it is possible for public opinion to both influence the use and interpretation of big data. For example, public perceptions of advancements in technology may hold considerable implications for the general use of big data in public policy. The recent controversy surrounding big data analysis is key to this discussion, considering that it may have been a symptom of wider technological suspicion. Regarding this point, Orlikowski and Gash (1994) discuss technological frames as a mechanism for how members of social groups conceive and respond to technological change. They believe that the success of a technological change can mostly be explained by the congruence or incongruence of technological frames between the architects and users of new technology. An appropriate example of this may be provided by the recent focus in American law enforcement on implementing Body Worn Cameras (BWCs), which has faced significant criticism from law enforcement as well as the wider public (Crow et al. 2017; White et al. 2018). Perceptions of this type of data may hinder or adjust its development and use. Michael and Lupton (2016) also add to this point, stating that understanding and trust from the public towards those using and creating this data is essential for its success. Ultimately, any organization attempting to analyze big data needs to be aware of the wider public perception. This is especially important to those involved in creating law enforcement policy, where certainty and assurances in predicting trends is a key concern (Crawford and Hutchinson 2015).

## **E. Ethical Concerns**

Before any specific practical steps can be taken to interpret or implement findings from big data, it is important for those involved in its collection and analysis to fully understand the potential harms that can result from its use. At the moment, Institutional Review Boards (IRBs) hold the responsibility of assessing the potential harms to research subjects that can result from university research, whatever the specific form of the research may be. However, on the whole, there has been long standing tension regarding the general role of ethical frameworks in data science and the social sciences, given that these frameworks have been adopted from clinical medical research standards that advocate for the rigorous

assessment and evaluation of research (Metcalf and Crawford 2016). Big data has added an extra dimension to this debate, given that big data sources are already in existence and are often freely available to the public. Indeed, this is the argument often utilized by those who seek to defend big data analytics, where the public nature of such data sources can be portrayed as evidence that it poses little threat to the privacy of individuals (Metcalf and Crawford 2016). However, it has recently been posited that big data analysis requires a different form of ethical assessment. Despite the widespread opinion that big data sources that have been anonymized and aggregated do not pose a threat, advanced tracking and matching techniques can make this innocuous data harmful by merging it with other data sources, which creates the possibility for inferences to be made about specific individuals. Crawford and Schulz (2014) describe this potential identification as a predictive privacy harm, where existing data is combined and repurposed in such a way that makes the identification of those within the data a possibility. As the use of big data analysis develops, it increasingly appears that data sources are able to be merged and matched in potentially innumerable ways that will enable even identifying information that is supposedly anonymized be re-identified. As a result, big data challenges the existing conception of data research ethics, as current frameworks still rely upon the assumption that data is temporally and contextually contained by technical infrastructures and financial cost (Metcalf and Crawford 2016).

Big data essentially removes the researcher from the process of data collection, and allows for the collection of data without interacting with the research subjects. Even with the increased distance afforded by secondary data analysis, it is vital for those using the data understand that big data analysis is still technically a form of human subjects research, and one that can cause harm to notions of privacy and discrimination (Crawford and Schulz 2014). Current ethical frameworks assume that data that is publicly available cannot cause any further harm to an individual, which is factually incorrect when considering big data sources. Metcalf and Crawford (2016) highlight that although these data sources already exist, their potential harm stems from their capability to be merged with other data sources, which effectively creates new knowledge. As a result, ethical assessments need to focus less on how or where the data was collected, and more on the purpose of the research and the exact ways in which the data will be manipulated. Despite this concern, big data analysis is on course to fall into a hypothetical grey area, where strict regulations are not enforced as it is believed that existing data sets can cause no further harm. If the existing framework for assessing ethical concerns in big data analysis doesn't change, relationships between the public and researchers may suffer as their information is increasingly manipulated and repurposed without their explicit consent. Ethical regulations should be considered as the bedrock of trust between researchers and participants, and big data analysis should be evaluated on a case by case basis to determine the potential harms that can result. As the usage of big data analysis develops, IRBs will play an important role in the establishment of such regulations.

## **F. Addressing the Limitations**

To begin addressing these concerns, there are several short term steps that can be taken. First, big data analysis must be reviewed with an appropriate amount of skepticism, and every action must be taken to ensure that the interpretation of the data is narrow and informed by the various limitations of the data source. Researchers and practitioners need to be aware of the personal and greater social harms that can accompany the quantification of individuals, and be ready to amend policies based on big data should they prove to infringe upon these concerns. Some of this risk can be minimized by ensuring that the data is anonymized and aggregated to the greatest possible extent, a step that will be discussed



further in regards to the Boston 911 call data within the next sections of this paper. In the long term, there requires to be a shift in how big data is perceived as a whole. Davidson (2017) discusses this in terms of outputs and outcomes. Traditionally, policy has been based upon the interpretation of outputs, the immediate, short term consequences resulting from policy change. Outputs usually took the form of raw numbers, where the number of rental units leased, police officers employed, crimes reported (to name but a few) would be the primary indicators of success for a particular policy implementation. This short sighted focus was both due to a lack of analytical capability and a general unawareness on behalf of policy makers of the potential wide reaching consequences of policy change. As a result, these outputs rarely took into account the context surrounding policy change. However, the advent of big data allows for a greater ability to understand and analyze the wider contextual impact. Davidson (2017) stresses that outcomes focus upon longer term impacts, and recommends that those evaluating policy should be careful to consider policy consequences for health, education, and the general well-being of those individuals subjected to policy change. Understanding the greater contextual impacts of policy will assist in better evaluating the efficacy of public policy, and it is important for big data analysis to be used for this purpose. However, when assessing the contextual impact with multiple big data sources, investigators must critically assess the validity of each one.

Having discussed the general implications of big data analysis for public policy, this paper will now illustrate the process of big data analysis with a specific example, Boston's 911 call data. This data represents a case study of the opportunities and challenges offered by various big data sets in addressing a variety of research questions. The vast majority of police agencies across the United States use a 911 call system for identifying public safety problems and allocating responses to those problems. Specifically, when a community member identifies a public safety problem that requires an immediate response from the police or other emergency service, they call 911 and request assistance. Information regarding the incident is then obtained from the community members request and the resulting public agency response. This data is recorded 24 hours a day, 365 days a year, and a number of researchers see this source of big data as a currently untapped method of furthering the understanding communities' social realities and experiences with crime and disorder.

## II. Big Data Analysis Illustrated: An Introduction to 911 Data

The Boston Police Department (BPD) provides a publicly available summary of 911 calls for service and officer initiated crime incident reports through the City of Boston's open data hub Analyze Boston. This incident report summary provides information on the preliminary details of requests for assistance from BPD officers, covering a period of June 2015 to the present. BPD restricts the publicly available data to a reduced set of variables intended to capture the time, location, and type of incidents. This system was designed with the express purpose of informing the police response to public safety problems by identifying the specific time, place and, nature of police interactions with the public. Incidents captured within this data source come from two sources, citizen requests for service and officer-initiated contacts. The former most often originate from the utilization of the 911 emergency system, which represents the only 24 hour 7 day a week method for community members to request an immediate response to serious public safety problems (Hagan et al. 2018). Since its implementation, it has become the primary means of contact for citizens who require an immediate police response and has been used

to provide a general indicator of crime across time and space (O’Looney 1997; Klinger and Bridges 1997). While calls for police service from community members represent the source of most incidents, an unknown minority of cases present in the BPD’s dataset come from officer-initiated contacts. These contacts include instances when an officer identifies a crime in progress or suspicious activity in the normal course of their patrol activities. We generally recommend against the use of datasets that mix calls for service with officer incidents, however, this dataset is still suitable for the purposes of this discussion and will focus on highlighting of the issues surrounding the validity of citizen requests for assistance recorded by this 911 system.

Researchers and practitioners have long recognized the value of this data source and have used it to address a variety of questions about crime, police-public relationships, and many other social processes. Prior research utilizing 911 calls for service has attempted to measure the prevalence and types of crime across communities. This type of question can consider differential crime patterns at numerous geographic levels and explain variation between crime and disorder at specific addresses (O’Brien 2017; Sherman 1989), between neighborhoods (Warner and Pierce 1993), and across cities (Bursik and Grasmick 1993). Additional research has sought to measure citizen’s willingness to report to the police (Desmond et al. 2016) and believe this data can provide insight into the public’s perception of the police (Kirk and Papachristos 2011). The ability to track spatial and temporal patterns of calls for service is this type of data’s main appeal to practitioners and researchers. The inclusion of location and time information enables investigators to explore how certain areas or businesses—bars, nightclubs, casinos, etc.—might be associated with crime. Others could explore how underlying socioeconomic conditions or cultural conditions of neighborhoods may lead to differential experiences with crime and utilization of police services. Groups not strictly interested in crime, public-police interaction, and these more traditional uses of 911 data could still benefit from it. Interest groups working with specific populations might explore this data to identify areas where their clients (e.g. homeless individuals, victims of domestic violence, etc.) may be interacting with the police and thus inform where they might best allocate their resources. The potential applications for 911 call data ensure a variety of parties could benefit from this rich data source.

As tempting as it might be to tap into this invaluable data, it would be inappropriate without first considering its limitations. Instead of telling the definitive story about crime and disorder in the City of Boston, the 911 reports tell a story that is filtered by police-public relationships, the allocation of police resources, and the reporting patterns of certain crimes in particular neighborhoods. It is important to recognize that the propensity to contact the police differs across individuals, places, and situations. These concerns can have a profound impact on how crime in a specific community is represented in these types of datasets. Even when crimes are reported to the police, features of the BPD and other agency’s data collection process further threaten the data’s validity. Categorization of incidents within system responsible for storing calls for service data is fraught with challenges. Failing to account for these potential sources of bias would inhibit a researcher’s ability to reliably answer their research question. These two sources of potential bias will be explored in depth and used to illustrate how unreliable these dataset can be. Such a painstaking tour of the biases should reinforce the need for caution investigators should possess when dealing with call for service data. Finally the process of operationalizing (defining how a phenomenon or concept will be measured) the raw data will show how these sources of bias will complicate investigators’ attempts to draw conclusions from the data.

## **A. Incident to Operationalization: How Crimes Become Data**

### **1. Barriers to Crime Reporting**

Not every instance of crime or disorder will come to the attention of the police. Crime reporting is contingent upon an individual being aware that a crime was committed, recognizing a criminal action as serious or harmful enough to merit a response, and deciding to call the police instead of pursuing some other course of action. Individuals may not deem some offenses worthy of the police's time, fear police involvement or reprisals from criminal suspects, do not wish to see a loved one arrested, or decide against notifying the police for numerous other reasons (Felson et al. 1999; Baumer 2002; Fleury et al. 1998). Reasons for non-reporting can be personal, situational, or structural, but there is a general consensus that all crime datasets reliant on citizen reporting will underreport the true extent of crime and do so in a manner that alters the integrity of the data (MacDonald 2009). This practice of underreporting criminal offenses injects bias into official crime statistics, such as the BPD's 911 database, as they only represent crimes known to the police. The Dark Figure of Crime, which symbolically represents the number of unreported crimes, encompasses many offenses (e.g. drug offenses, sexual assault, white collar crime), which while present in the BPD's system, are subject to significant underreporting that may misrepresent their quantity and distribution across the city (Fisher et al. 2003; Braithwaite 1985; Beckett et al. 2006). In addition to victims of crime deciding not to report, a host of 'victimless crimes' will be sorely misrepresented in these types of datasets (Meier and Geis 1997). As an example it is unclear who might report a drug transaction since neither the seller nor the buyer could be categorized as a victim. A detailed discussion of these barriers to crime reporting and resulting biases will justify the repeated calls for caution.

Recognizing a particular situation as both criminal and serious enough to merit a response represents the first step of reporting to the police. Labeling a behavior or situation as going beyond a mere nuisance or tolerable disorder and instead recognizing it as a criminal act differs among individuals. Some people may be compelled to report the unauthorized sale of bottled water by a child (BBC 2018), while others may hesitate to alert the police to the presence of an open drug use on their street corner (Marcelo 2017). Crime reporting is further complicated by the fact that individual's assessment of crime seriousness differs across types of crimes. While homicide is universally regarded as a serious criminal behavior, the criminality of recreational marijuana usage tends to be a source of contention. Disparities in the seriousness of crimes will lead to systematic reporting differences between different types of crime. More serious crimes are more likely to be reported than their more trivial counterparts (Skogan 1984). 911 data will reflect these reporting behaviors and the accuracy of the data will vary between types of crimes.

The individual calculus involved in determining a crime's seriousness is crime-type specific and highly dependent on both individual and contextual factors (Warr 1989). Differential reporting behaviors introduce bias into the 911 data when the individual reporting propensities vary systematically across the population and cannot be effectively measured. This kind of variation is subject to these systematic biases as differential assessment of crime seriousness has not only been observed to be an individual process, but also a social one. Whole communities can collectively exhibit different degrees of tolerance as exhibited in studies on reporting behaviors. Tolerance of certain forms of deviance such as underage drinking, marijuana usage, and fist fights differs based on the characteristics—e.g. concentrated disadvantage, collective efficacy, racial composition—of the neighborhoods in which they occurred

(Sampson and Barusch 1998). Identical crimes types could be occurring in two neighborhoods at identical levels, but different reporting behaviors will ensure calls for service data show a nonexistent difference in quantity of these offense. Potential suppression and inflation of criminal offenses will occur in the 911 database as it is dependent on individual's decisions to report to police.

While some individuals may opt against reporting due to offense seriousness, other motivations have also been reported as barriers to reporting criminal acts. Victims and witnesses regularly cite the fear of reprisal from the criminal offender(s) or their associates, be it an intimate partner fearing their abuser or a resident avoiding the persistent gang presence in their neighborhood (Felson et al. 2002; Rosenbaum et al. 2002). Beyond the fear of future victimization, interpersonal connections between victims, witnesses, and criminal offenders may inhibit individuals from reporting the criminal behavior of family, friends, or acquaintances. Intimate partner violence advocates draw attention to victims' reluctance to report as an example of this barrier (Logan and Valente 2015). Some victims continuing affection toward their abuser or their reliance upon them for financial support could serve as an impediment to reporting. Certain types of crimes are more likely to involve the social dynamics described above that could result in different underreporting by type of crime. A business owner victimized by a shoplifter is far less likely to have the personal relationship with the criminal than a victim of intimate partner violence. This again introduces variation in the validity between crime times. Such motivations can be place dependent and neighborhoods exhibiting different degrees of social cohesion could strength of relationships between residents. Pattillo's (1998) ethnography of a south Chicago neighborhood demonstrates how dense social networks that result in connections between gang members, drug dealers, and law abiding peers might dissuade individuals from calling the police. These close social ties meant that residents in this neighborhood were reluctant to call the police to report crime committed by these individuals, despite the fact that they personally objected to this kind of behavior.

The circumstances precipitating a criminal act may also deter reporting. Victims may feel a degree of culpability which can suppress their willingness to report their experiences to the police. Skogan (1984) provides a succinct description of this motive to not report in stating, "due either to shame, embarrassment, or a concern about their own labeling by the police, victims may have reasons for not wanting to get involved with the authorities" (Skogan 1984 p. 124). If individuals perceive any sort of social stigma attached to types of victimization, they will not readily report their experience. Victims of sexual victimization have often reported feelings of shame as factoring into their reluctance to contact the police (Weiss 2010). Others may fear that their role in an incident, such as a fight, may subject them to potential blame or arrest (Bowles et al. 2009). Reservations stemming from potential culpability may prevent these crimes from being reported. Importantly for the validity of the incident summary dataset, this type of non-reporting is particular to certain types of crime, meaning some crime types are more likely than others to go unreported.

An individual's distrust of the police and their ability to resolve problems may also dissuade them from reporting crime. Kirk and Papachristos (2011) used the concept of legal cynicism to illustrate this process. Legal cynicism is "a cultural orientation in which the law and the agents of its enforcement, such as the police and courts, are viewed as illegitimate, unresponsive, and ill equipped to ensure public safety (Kirk and Papachristos 2011 p. 1191)." Individuals whom distrust the police or question their effectiveness will not call the police in situations where others eagerly seek police assistance. In the wake of social movements such as Black Lives Matter, variation in perceptions of the police among different racial groups and communities is as evident as any point in the history of American policing.

Given the extent of racial residential segregation in many American cities, it is easy to imagine how racial disparities in perceptions of the police could result in systematic errors in reporting behaviors (Peterson and Krivo 2010). 911 calls in African American neighborhoods in the city of Milwaukee, Wisconsin experienced a marked decline following a high profile incident involving officer use of force against an unarmed black man (Desmond et al. 2016). The researchers attributed the subsequent suppression of 911 utilization in predominately African American neighborhoods to reservations of the police's efficacy in resolving disputes in these communities. Even if an individual possesses a generally positive view of the police, they may have alternative means of resolving criminal situations. Anderson's (2000) classic ethnography of intercity Philadelphia depicted a street culture that emphasized self-reliance in conflict resolution. Victims of violent or property crime would not call the police, but rather attempt to seek revenge through the perpetration of a retaliatory action as both a form of justice and protection from future victimization. Reporting behaviors can be susceptible to demographic and structural conditions, which are not uniform across any community.

## **2. Accuracy of Reporting**

While unreported crime and inconsistent reporting patterns represent the most pressing limitation of this type of data, even those crimes that are reported to the police may not accurately reflect the incidents from which they originate. Sherman (1989) sought to use 911 data to identify places where significant amounts of criminal activity were concentrated in small geographic areas, called crime hot spots. Initially, the results came as a surprise, with concentrations of crime appearing in unexpected places. He determined that certain locations, namely hospitals, convenience stores, and other public locations, tended to host the first point of communication between crime victims/witnesses and the police. In many cases, these locations were erroneously reported as the incident location in the 911 dataset, when in fact the criminal incident occurred at another location entirely. Reporting of secondary locations as incident locations misrepresents the distribution of crime across a city. The frequency of incidents at a hub of police citizen contact, such as hospitals, will then appear to be greater than it should in these databases. Other features of the 911 system were found to similarly challenge the accuracy of 911 datasets in terms of reporting the actual patterns of criminal incidents. Public criminal events may garner significant attention and result in multiple calls for service. Flooding 911 dispatchers with multiple calls for a single incident occasionally resulted in the creation of duplicate incident reports for a single incident (Sherman 1989). Duplicate incident reports can also occur when dispatchers or the police fail to consolidate separate components or events that are related to a single incident. Complex incidents, such as the transportation of a victim to the hospital following an assault or the interviewing of witnesses to a shoplifting in a secondary location, are especially vulnerable to these inaccuracies. In this way, incident complexity risks erroneous classification of single cases with multiple components as separate incidents in the 911 system. As particular types of crimes involve varying degrees of complexity, this could skew the 911 system's reflection of certain crime types. Inaccuracies, stemming from misclassifications and duplications, further divorce the representation of crime and disorder in the 911 data from their patterns and trends they purportedly represent.

Although the BPD's dataset includes time stamps of each incident, the accuracy of this variable should be especially scrutinized. Frequently there is a substantial lag between the time an incident occurred and the time the incident was reported. The timeliness of reporting, and therefore accuracy of the time reported in the 911 call systems, likely differs across types of crimes. For example, crimes such as

burglary may be reported hours or days after the actual criminal event occurred. A victim returning from work or vacation to find their home ransacked will contact the police a significant period of time after the burglary's occurrence. Other crimes, such as an armed robbery, are more likely to be reported immediately as the victims or witnesses were present during the crime.

Misidentification of incidents at the time of reporting poses yet another threat to the validity of the offense information present in the database. For citizen-generated incidents in the report, the initial categorization of a criminal offense is dependent on the citizen's description. Citizens may only have a vague understanding of the nature of the criminal offenses and what information is important for classification purposes (Reiss 1971). Their ability to accurately relay the description of an incident which 911 dispatchers would in turn assign to a particular crime category is hindered by the potential errors in these descriptions. Considering an average citizen's unfamiliarity of the law, much less proper offense categorization, one would expect a significant degree of misidentification (Williams and Hall 1972). Citizens' lack of knowledge prevents descriptions that will allow call dispatchers and officers to properly distinguish between types of crimes. Researchers would be wise to question the veracity of this aspect of the data, as proper classifications is dependent on an unreliable party.

Beyond instances of these unintentional errors in the form of citizen misidentification, misuse and abuse of the 911 system further challenges the validity of this dataset. Fraudulent calls which fabricate or exaggerate a situation still require police attention and may be recorded in the data as it was initially reported (Sampson 2002). Fake bomb threats are unfortunately a routine problem for many police departments in the United States and even when any threat is discredited these types of incident still appear in 911 records (Bowman 2004). Vengeful neighbors, disgruntled employees, or a group of teens can generate false records in any police department's incident management system. In addition to the fraudulent calls, some incidents may contain inaccurate information due to individuals providing false information to the police or 911 dispatchers (Sherman 1989). Police scrutiny may dissuade some individuals or business owners from providing accurate information regarding the location or details of an incident. However, changing such details may result in inaccurate entries in the incident records system. Although the errors stemming from fraudulent reports or misleading information likely accounts for a small proportion of citizen-generated records in the BPD system, it proves impossible to completely remove erroneous records from any analysis.

### ***3. Summary of Potential Bias in Citizen Requests for Service***

In a paper using 911 data to approximate crime, the authors downplayed the validity concerns of this type of data, stating, "to the extent that these data present a biased picture of crime, they are biased only by citizens' willingness to report crimes (Warner and Pierce 1993 p. 512)." As is evident from the discussion above, individuals routinely decide against reporting personal victimizations or observed criminal behaviors to the police. Unreported crimes that escape the attention of the police will not be captured in the BPD incident summary dataset. Even those incidents that are reported do not necessarily result in a classification that accurately reflects the reality of situation. These sources of error, while still undesirable, could be tolerated by researchers if they each had a similar and predictable effect on the data quality. If all crime types were reported at similar rates in all communities, researchers could accept that while the 911 database would not present a perfectly accurate picture of crime, its bias would vary consistently across types of crimes and places. Such concerns would not prevent many research questions as patterns between crime types and their distribution across time

and place could still be determined. However, this assumption that all crime types are subject to the same degree of underreporting across all places is disputed by the literature presented above.

Instead, biases resulting from differential crime reporting and accuracy that varies by types of crimes and between neighborhoods is an undeniable challenge for researchers using this data. It is necessary to develop strategies to address these issues as they introduce systematic bias according to type and location of crimes, ultimately disguising their actual quantities and distributions. Although the potential impact of these forms of bias will be dependent on particular research question, researchers using this type of data must find a way to control them. At the very least, researchers should acknowledge their presence and include a discussion of their potential limitations. Failing to do so can result in misleading findings and lead to misguided policy decisions. We will later propose a process for addressing these concerns, however, we must first discuss the second major source of bias in 911 data, the categorization and storage of incidents.

## **B. Categorization of Incidents**

Researchers and practitioners using a secondary dataset must ask two questions: why was the data collected and how was the data stored. Police agencies, in an effort to more effectively serve their jurisdiction, seek to track the patterns and trends of citizen requests for service. This information guides their response to a host of public safety concerns. However, as the preceding discussion on citizen-generated data hopefully demonstrates, this particular measurement may not necessarily portray the reality of crime and disorder throughout a community due to systematic underreporting. When considering the second question posed, how the data was stored, it will become clear that validity concerns emerge from complications of categorization and digital storage of data. When a citizen or officer describes a particular offense, police dispatchers are tasked with assigning the incidents into distinct categories. A complex incident, which could be described at length, will be sorted into one of BPD's 241 offense types in their 911 system. Condensing the details of an incident affords interpretability and convenience of data analysis at the expense of detail and this process is fraught with threats to the validity of the data. Specifying the original purpose of any data collection effort can help reveal the motivations behind the decision making in constructing a data storage process. As will be demonstrated, how the data is stored will have an impact on how an investigator should assess its quality and account for its potential flaws.

### ***1. Purposeful Categories***

Incident descriptors are created by their respective agencies and as such, distinctions between the types of incidents are tailored toward the needs and orientation of a police department. In this case, the Boston Police Department likely constructed their crime definitions to differentiate between different types of crime which had implications for their operations. Most descriptors appear to be drawn broadly from legal definitions of crimes. Larceny offenses exemplify this as they include legally relevant dollar amounts of stolen property to differentiate between sub-categories of larceny offenses for example "Larceny under \$100". Others draw attention to the circumstances surrounding the incident such as Burglary descriptions that include the type of property where the crime occurred (i.e. Commercial, Residential, or Other), use of force, and whether the offense was completed. Still others offer vague descriptions of the service aspects of law enforcement such as the descriptors of "Animal Incidents" and "Sick/Injured/Medical". Though these definitions lack some detail that may be useful to

police practitioners (namely the resources required to handle a particular incident) they do offer some basic information that speak to this purpose. Definitions indicate the seriousness of crimes and certain crime types may prompt certain police responses.

For example, larceny, generally defined, is theft of personal property without force. However, criminal actions which broadly fall under this umbrella term can be drastically different in terms of their societal harm, demand on police resources, and best response by a police agency. Distinguishing between two larceny categories included in the BPD dataset, pick-pocketing and shoplifting, illustrates this point. Police administrators' response to pick-pocketing and shoplifting would likely differ as one, pick-pocketing, would likely place greater demands on a responding officer. Victims of pick-pocketing likely have less experience interacting with an officer when compared to a shop or business owner who is regularly victimized. Filing reports and conducting an investigation require different amounts of investigative resources and the victims of these offenses may have different expectations of a police response. Shoplifting victims may simply want a police report for the purposes of documenting the incident while a person who lost a wallet may expect a more thorough investigation. BPD's prevention strategies for these offenses could likely differ. The installation of Closed Circuit Televisions (CCTV) in pick-pocketing hotspots has been advanced as one strategy, although it is likely that many businesses already have CCTV systems in place (Troelsen and Barr 2012). Police departments may be better suited to address the shoplifting problem by launching an informational campaign with regularly victimized business owner on shoplifting prevention techniques (Clarke and Petrossian 2002). This type of distinction between the types of larceny and other crimes benefit the BPD.

Different types of organizations—academic, medical, social work, etc.—would almost assuredly create a unique set of crime categories to suit their particular needs. Distinctions between call types and definitions which have meaning for the organizations creating them. For the Boston Police Department, their crime categories decisions and distinctions likely came from a desire to inform their resource allocation and public safety strategies. It is difficult to estimate the potential impact BPD's process of category creation may have on how the incidents are categorized, but it should be understood and acknowledged. In dealing with secondary data sources, researchers should be encouraged to obtain detailed notes on all aspects of the data collection process, which includes codebooks or instructions informing how the categories are differentiated in the case of categorical variables such as offense type (Cheng and Phillips 2014). BPD's process for creating offense categories was likely an organic process and justifications for definitions distinguishing between crimes are likely nonexistent. However, researchers should still consider how such decisions may impact the data quality as it pertains to their research question.

## ***2. Accurate Incident Classification***

As previously mentioned, both citizens and officers are prone to misidentifying the true nature of an incident at the time that they report to an officer or 911 dispatcher. However, the process of classifying incidents can be further obstructed by the complexity of many criminal incidents that do not readily lend themselves to easy classification. Some incidents are comprised of multiple offenses, which would be labeled as different crime types if each component were evaluated individually. Without a well-defined and consistently utilized process for designating such incidents, the classification of the incident will be an arbitrary decision. For example, an individual arrested for homicide is found to be in possession of illegal narcotics. This incident could be labeled as either homicide or possession of drugs as both crimes



are present. Many agencies utilize a hierarchy rule to resolve this dilemma when reporting their official crime statistics to the Federal Bureau of Investigation as part of the Uniform Crime Reporting and National Incident-Based Reporting System. If an incident contains multiple offenses, agencies applying this rule will report the incident based on the most serious offense that occurred (see James 2008 for discussion of hierarchy rule). In the previous example, the hierarchy rule would result in the categorization of homicide for this incident as it represents a more serious offense than drug possession. The BPD does not have a publically available statement of the process categorizing multi-offense incidents in their system, however, they likely use something akin to the hierarchy rule.

In addition to the confusion and potential for misclassification presented by multi-offense incidents, some crime are more difficult to define. Especially complex cases such as human trafficking illustrates this issue. The incident summary report includes two offense descriptions related to human trafficking, “Human Trafficking-Commercial Sex Acts” and Human Trafficking-Involuntary Servitude”. Law enforcement’s inability to first accurately identify criminal behaviors that are consistent with human trafficking and then properly labeled an incident as such is stymied by numerous factors (Farrell and Reichert 2017). Commercial sex acts for example can often be mistaken for prostitution or kidnapping and involuntary servitude could be mistaken as a civil matter concerning an employee-employer dispute. Although incidents that could constitute a Human Trafficking offense in many cases are more likely to be labeled as “Prostitution” or a catch-all term of “Investigate Person”. Human trafficking estimates will therefore be drastically underestimated and although the dataset includes them, it will not be capable of providing a valid measure of its prevalence or distribution. These types of problems with incident categorizations are extremely common and should remind investigators about the folly of trusting the classification of incidents.

### ***3. Interpreting and Operationalizing the Data***

Whether a 911 or police incident summary dataset is used as the sole dataset or combined with other datasets, investigators should undergo a similar process to ensure their use of the data and any conclusions based on it are appropriate. Consequences for forgoing this process could be disastrous. Big Data, such as this incident summary report, tantalizes practitioners and individuals shaping public policy with promises of quantifiable metrics upon which to identify problems and evaluate interventions. Failing to account for the bias in data risks ineffective and potentially harmful policies. Responsibly interpreting the data from this and similar datasets requires investigators to abide by methodological and ethical considerations. The litany of data quality concerns presented above, while significant, do not render this dataset unsuitable for exploring Bostonian’s experiences with crime, disorder, and the police. However, it does limit the suitability for some research questions and mandates that certain precautions be employed. Investigators using the incident summaries should be prepared to make a set of methodological decisions and ethical considerations that emerge from both the practical limitations of the data and the purpose of the research question(s).

## **C. Methodological/Practical Decisions**

Formulating a research question or set of questions is one of the first steps in social science research. Research questions represent the point of orientation and specific areas of focus for all research investigations from which researchers tailor their data selection, project design, and methods (Bryman 2007). These questions serve as the foundation to methodologically rigorous exploration and should

dictate process of preparing and analyzing the data. As previously discussed, research questions concerned with crime, disorder, and their prevalence or spatial and temporal patterns often consult 911 datasets. Other researchers have used calls for service as a proxy of perceptions of police legitimacy. All these studies sought to test their specific research questions with the best available data and after selecting 911 data made a series of methodological decisions to with the data's strengths and weaknesses in mind. Calls for service were theoretically associated with some social processes and a quantifiable manifestation of latent phenomena relevant to their research question. Preparing the data to properly accommodate a research question requires several important methodological decisions. We draw attention to two decisions researcher working with this type of data must make to make the data suitable for their research question. Researcher should make meaningful offense groups and aggregate to an appropriate geographic unit of analysis.

BPD's dataset contained 241 distinct offense codes. Researchers seeking to utilize this will need to make important decisions regarding which types of offenses they utilize and if they need to group theoretically similar offenses together. Given the validity concerns stemming from the incident categorization process, researchers regularly aggregate individual crime types into broader crime or offense groups or categories. For example, many studies opt to use broadly defined categories such as violent, property, and disorder crimes which encompass dozens of crime types. While researchers should be mindful of many factors influencing these grouping decisions, their researcher question should serve as the primary guide. Both the concepts these groupings should represent and the specific offenses that comprise them will be derived from the researcher's research questions. Following the selection of meaningful groups, investigators must determine the geographic unit of analysis that best corresponds with their research questions.

Geographic units range from micro-spatial places such as addresses or street segments to larger areas such as cities, counties, states, or counties. Rengert and Lockwood (2009) note "the appropriate unit of analysis to be used depends both on the research question we wish to address and the availability of data ... [and] if the data are available, research generally begins with the smallest level of aggregation possible (Rengert and Lockwood 2009 p. 110)." Distribution of data across potential geographic units of aggregation is a central consideration as it has implications for the statistical viability of particular research question. Fortunately for investigators utilizing crime data from major metropolitan areas, micro-geographies such as the address level is suitable for many research questions and has been frequently utilized (Weisburd et al. 2016). This type of data is suitable for higher levels of analysis as Bursik and Grasmick (1993) advocated for the use of calls for service as an additional indicator of crime trends at the city or county level. However, some levels of geographic analysis may be more appropriate than others depending on the type of research question. O'Brien et al.'s (2017) analysis of Boston residents' call for service and non-emergency requests demonstrates how results from this type of data are sensitive to the geographic unit of analysis. In their study, requests for service were found to cluster at the address, street segment, and tract level, however, the two higher-order geographic levels better accounted for the persistence of crime from year to year. Researchers must be both theoretically driven and statistically cognizant when selecting their geographic unit of analysis as this decision has direct implications for the viability of a project and the validity of the results.

## **D. Ethical Considerations**

Researcher's use of big data sources is complicated by the potential for subject identification. 911 datasets, and especial those featuring detailed location and offense information, are at risk of violating a basic tenant in human subject research. Researchers should take certain precautions to ensure they avoid any ethical missteps in their use of this sensitive data source. Protections for human-research subjects have been codified into law since the mid-1960s and the U.S. Department of Health and Human Services is currently tasked with regulating human subject research (Schrag 2010). Most academic research institutions feature an Institutional Review Board (IRB) which must approve of research proposals involving human subjects. First and foremost, researchers should submit their research questions and project to an IRB if using any data source that could identify individuals. This will almost assuredly be the case for any researcher utilizing 911 databases. Although this type of data may be exempt from IRB regulation in some cases, the presence of potentially identifying information means that investigators must make every effort to protect the individuals who could be identified in the data. IRB's experience with de-identification and other human subject safeguards could assist investigators in their obligation to ethically protect their data's subjects. In addition to contacting an IRB, abiding by the following recommendations can address some of the most pressing ethical concern related to this type of data, identification of subjects.

Concerns surrounding the identification of subjects are more pronounced at smaller levels of geographic units. Excluding an individual's name from a dataset does not effectively anonymize the data, as addresses can be combined with other publically available data to re-identify individuals (Ohm 2010). Reports utilizing this data should omit any references to specific addresses or details of high profile incidents that could easily be linked with other information to identify the person(s) involved. Just as the research question should inform the level of geographic units employed in a study for methodological reasons, so too should investigators consider the ethical mandates when determining an appropriate unit. Investigators should strive to use the highest level of geographic aggregation that their research question and dataset permit. This recommendation should minimize the risks of re-identification while preserving the functionality of the dataset. Finally, combining the raw BPD data with other sources of data can offer researchers insight into compelling social processes, but further risks pairing crime data with potential identifiers. Following any data mergers, investigators should remove all information that is not necessary for their analysis. This coupled with basic measures of protecting raw and merged datasets—password protection, limiting physical access, encryption, etc.—limits the potential exposure of identifying and private information. IRBs are valuable resources for researcher unfamiliar with these techniques and can provide recommendations tailored to the specifics of a researcher's own project and data.

### III. Proposed Approach

#### A. Overview of Dataset

Prior to discussing specific strategies for working with this dataset, a brief overview of the raw dataset’s contents and structure should assist in comprehending the more technical recommendations. The Boston Police Department (BPD) summary of 911 calls for service and officer initiated crime incident reports was retrieved from Analyze Boston on 8/24/2018. After downloading the files as a comma-separated values (CSV) file, we recommend importing this file into the investigator’s preferred statistical program. Raw data files contain 18 variables which are summarized in Figure 1. Offense Descriptors provide the most detail regarding the type of offense and should be starting point for investigators looking to group similar offense types. Several variables indicate the location of the incidents, but the geographic coordinates represent the most versatile spatial information. From this information, researchers working with this data can aggregate information up to any geographic unit of analysis, including address, street segment, census block, census tract, etc. Researchers considering the temporal patterns of crime can use the time and date variable which indicates the time the incident was recorded by the BPD.

If the incident is considered the unit of analysis, the publically available file is stored in the long format with incidents containing multiple offenses that are reported as multiple observations for the same incident. Incidents contain 10 digit string values which correspond with the BPD’s 911 system’s internal report number. However, these incident values do not uniquely identify each row in the raw data file. The sample dataset’s 315,958 observations represent 279,711 unique incidents and nearly 10% of the incidents feature multiple offenses, as illustrated in Table 1. Failing to account for the multiple offense incidents would violate the independence assumption demanded by any analysis utilizing regression and must be addressed early in the cleaning process.

*Table 1*

<b>Number of Offenses per Incident</b>	<b>N</b>	<b>%</b>
1	252,507	90.25
2	21,193	7.58
3	4,153	1.48
4	1,292	0.46
5	427	0.15
6	132	0.05
7	45	0.02
8	11	<0.01
9	6	<0.01
10	3	<0.01

11	1	<0.01
13	1	<0.01
Total	279,711	100

Figure 2 displays a few examples of multi-offense incidents that would need to be reduced to a single case. Although strategies for this process will be dependent upon the research question, researchers must sacrifice detail afforded by multi-offense incidents to ensure the data accurately depicts the distribution of incidences across both time and space. Projects concerned with specific offenses or groups of offenses will use their research question to dictate their collapsing strategy.

## B. Categorizing Incidents

Systematic biases in underreporting requires research questions utilizing 911 data to reflect general crime patterns and trends across space or time should refrain from using all types of crime. Other analyses may be interested in differentiating between types of crime that would require investigators to group similar offense types. Categorizing incidents into meaningful groupings is the first step for most research projects. Only projects interested in the total volume of incidents would not have to categorize incidents, although even these would benefit from looking at differential patterns by type of call. The recommended process of grouping incident types is guided by two principles: offense validity differs by crime type and offense seriousness begets measurement validity.

To ensure some degree of uniformity with other studies on crime, assigning offense types to their respective categories in the National Incident-Based Reporting System (NIBRS) is highly recommended. Most of BPD’s 241 offense descriptions correspond with one of NIBRS 49 offense categories and one of the 18 NIBRS offense groupings. Figure 3 provides a visual representation of this grouping process. The few BPD offenses that do readily lend themselves to a particular NIBRS offense category can be assigned to the ‘all other offenses’ category. These types of offense descriptions consist primarily of officer initiated incidents such as serving an arrest warrant. Utilizing a nationally recognized set of offense categories and offense grouping allows for researchers to more easily compare their 911 data set to the broader literature. It also assigns offenses to groups with similar types of offense that often have comparable levels of seriousness. Generally, the seriousness of the offense corresponds with the validity of a this data’s ability to measure general crime patterns as individuals are more likely to report serious crimes, so ensuring offense grouping exclude offenses with drastically different likelihoods of reporting is ill advised (Gove et al 1985; Skogan 1984; Felson et al. 2002). Appendix 1 shows the NIBRS offenses categories and their groupings ranked by their seriousness. Ranking seriousness of offenses is a subjective process, so researchers should consider the best ranking scheme for the purposes of their specific research question.

Figure 1

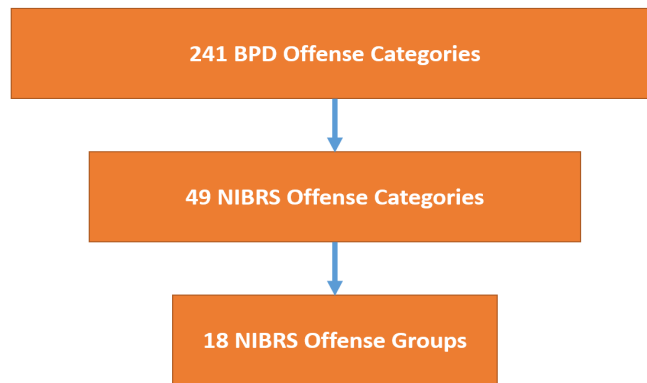
Variable Name	Description	Example
Internal BPD Report Number	Incident number assigned by BPD	I172034759; I152104844; I182067682

Numerical Offense Code	Unique offense code number assigned to each unique offense	3115; 619; 1402
Offense Code Group Name	Offense Group descriptor	Larceny; Simple Assault; Other
Offense Descriptor	Detailed Offense Description beyond Offense Group	Threats to do Bodily Harm; Sudden Death; Trespassing
Police District	BPD Police District	C11; B3; D4
Reporting Area	BPD designated Geographic Unit Smaller than Patrol District	257; 450; 203
Shooting Dummy Variable	Indicates the report of a gunshot associated with the incident	(Blank); Y
Earliest Possible Date and Time	Date and time the incident initially logged	9/13/2016 22:57:00; 10/1/2015 6:00; 7/5/2015 6:59
Year	Year of the incident	2017; 2016; 2015
Month	Month of the incident	1; 2; 3
Day of the Week	Day of the incident	Monday; Tuesday; Wednesday
Universal Crime Reporting Part Number	BPD designated crime reporting part number	Part One; Part Two; Part Three
Street Name	Street associated with the incident	Tremont St; River St; Summer St
Latitude and Longitude	XY coordinates indicating location of the incident	42.00000, -71.074746

We arrived at this particular ranking it corresponded with the Uniform Crime Reports hierarchy reporting process. While the details will vary, this process of making theoretically meaning groupings and determining a hierarchy of offense seriousness is a critical step. We created two new variables for the NIBRS offense category and its seriousness ranking and attached these to each record in the dataset. Multi-offense incidents will have multiple records in the dataset, which will each assigned to their respective offense category and be assigned a seriousness ranking. The next step in the process will allow researchers to condense these multiple records to a single one that contains the most meaning assignment for the purposes of the research question.

Multi-offense incidents must be consolidated to single records prior to analysis, and this process will be dictated by the investigator’s research question. Approximately 10% of the records in the BPD dataset represent multi-offense incidents. The final offense categorization should retain the most serious of all offenses in a particular incident, which is consistent with the FBI’s hierarchy rule. Figure 2’s first incident will be used to demonstrate this process. For such a research question, investigators should strive to retain only the most serious, and therefore most consistently reported, offense for their analysis. Skogan’s (1984) observation that the “elements of seriousness are by far the strongest predictors of

Figure 2: Offense Category and Grouping Process



reporting [crimes to the police]” would prompt an investigator to retain the most serious offense category in consolidating multi-offense incidents. Two offenses, Disturbing the Peace and Larceny Shoplifting, are both attached to the same incident. Following their assignments to an NIBRS offense category they would be assigned to Misdemeanors/Non-Criminal Acts and Larceny, respectively. According to the FBI hierarchy rules and an evaluation of offense seriousness, the Larceny category designation would be preferred over the Misdemeanors/Non-Criminal Acts. This multi-offense incident would then be recoded to a single offense Larceny Incident.

Figure 3

Incident ID (Offense #)	Offense Code	Offense Code Group	Offense Description
I152104844	(Offense 1) 2403	Disorderly Conduct	Disturbing the Peace
	(Offense 2) 613	Larceny	Larceny Shoplifting
I172034759	(Offense 1) 3125	Warrant Arrests	Warrant Arrest
	(Offense 2) 1874	Drug Violation	Drugs-Other
	(Offense 3) 1810	Drug Violation	Drugs- Sale/Manufacturing

Another investigator may be less interested in approximating crime generally and is instead interested in measuring the neighborhoods' differential tolerance of drug use. Sampson and Barusch (1998) explored this concept of neighborhood differences in tolerance of several deviant behaviors including marijuana usage, albeit with a different measure in the form of community surveys. Investigators may be interested the association between neighborhood factors such as concentrated disadvantage or racial composition and calls for service to the police regarding drug activities. If the investigator seeks to measure intolerance of drug use by the number of drug related incidents, they would prioritize categorizing multi-offense incidents involving drugs as drug incidents instead of adhering to the offense seriousness strategy described above. The second incident in Figure 2 is comprised of three offenses, two of which are related to drug offenses and the third being a Warrant Arrest offense. These would be assigned to the NIBRS categories of Drug Offenses and All Other Offense or Incident Types. Using drug offenses as a measure of citizen's concern with drug use would dictate the retention of the Drug Offense category when consolidating the multiple offenses. Comparing these consolidation strategies demonstrates how the research question can dictate the decision making process at every stage, including this first crucial step of preparing the data for analysis by eliminating multi-offense incidents.

Offenses and offense groups selected for a project will ultimately depend on the specific research question, however, we discourage investigators from blindly trusting the accuracy of BPD's offense categorizations. Instead, researchers should attempt group offense types into broad groupings whenever possible. Grouping data does not fully remedy the validity concerns of these types of datasets, however, it does prevent the more pronounced biases present in specific offenses to be offset by their less biased counterparts. Investigators should be cognizant of the many barriers to reporting and the accurate categorization of incidents. No matter which groupings an investigator selects, the process will likely have some degree of bias introduced by the differential use of 911 and factors influencing officer behaviors. An acknowledgement of these limitations should accompany any results generated from this data. Analyses intending to inform policy makers must qualify any conclusions resulting from this dataset and seek corroboration in other datasets prior to acting upon these analyses.

### **C. Unit of Analysis: Levels of Aggregation**

The importance of geography cannot be overstated when it comes to the theoretical and methodological implications for research on crime. Weisburd et al. (2016) extensive discussion of "the appropriate scale of analysis of criminological enquiry (Weisburd et al. 2016 p. 16)" demonstrates not only the variety of methodological approaches that could be used to analyze the spatial properties crime, but also their theoretical import. A comprehensive discussion of these decisions and their consequences would go beyond the scope of this paper, however, we bring it up to emphasize the need to carefully consider the spatial units through which investigators view patterns and trends of 911 and incident summary data. As with the incident categorization decisions, the decisions regarding geographic units of analysis should be dictated by an investigator's research question(s). Understanding the different levels available and their respective strengths and weaknesses should help investigators begin the process of selecting the most appropriate spatial scale for their purposes.

Researchers interested in studying crime have recently begun focusing on the concept of crime in micro-spaces (Weisburg et al. 2008). Findings indicating that crime concentration occurs at geographic units as small as addresses, block faces, and street segments and their relative stability have validated this



approach. This unit of analysis permits the exploration of “intra-neighborhood variance in crime that is lost when neighborhoods are examined as homogenous unit” (Groff et al. 2010 p. 8). Real estate’s maxim to purchase the worst house in the best neighborhood emerges from the recognition that individual variation exists within larger geographic units. If each house is considered as the unit of analysis, property values can fluctuate within a neighborhood. However, once the unit of analysis is raised to the neighborhood level and average property values are considered, this variation between property values is no longer discernable. In this way, problem households (those responsible for significant amounts of crime) cannot be properly studied if the unit of analysis is set at a higher level of aggregation, such as the neighborhood or county. Research questions focused on associations between incidents and factors that vary across small geographic spaces should consider the use of micro-spatial units of analysis.

Figures 4 and 5 provide a visualization of how the level of aggregation can affect the ability to detect variation of crime based on the unit of analysis. Figure 4 displays a density map for all crimes aggregated to the census block level while Figure 5 displays the same aggregation to the census tract level. Census blocks are nested within census tracts, so Figure 4 is capable of presenting the intra-tract variance of crimes reported to the police. This information is lost at the census tract level as it becomes impossible to distinguish between patterns within tracts. However, as the discussion of micro-spaces above indicates, even the rich detail of variation provided by block level aggregation in Figure 4 cannot describe how reporting differs at smaller levels of aggregation such as the address, street block, or street segment. Just as the human eye is dependent upon the level of aggregation to detect important variations, so too are statistical programs dependent on these levels to perform analysis looking at different spatial levels.

Although this level of aggregation permits a detailed and geographically specific analysis of 911 calls and incidents, difficulties emerge when seeking to pair it with other data sources or information. Publicly available information on specific addresses, for example, are exceedingly rare. For the purposes of illustration, consider a research question interested in racial disparities in reporting domestic violence to the police. While an investigator may be able to assign each domestic violence incident in the BPD’s dataset to a specific address, they will not be able to determine the race of the residents at this address. Without racial information on addresses, testing this research question will prove impossible at this level of aggregation.

Figure 4

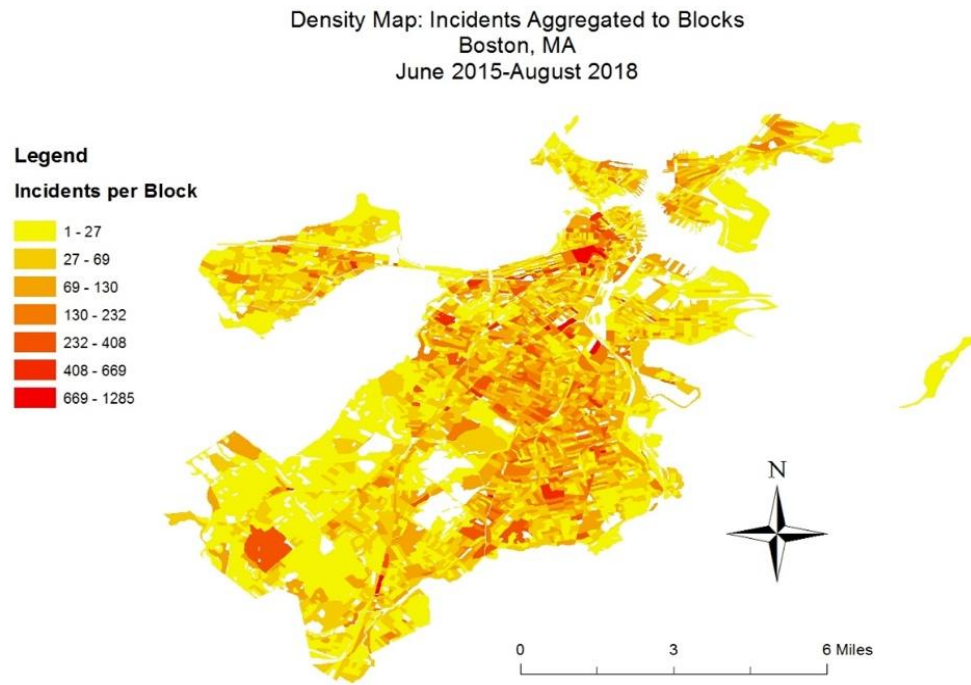
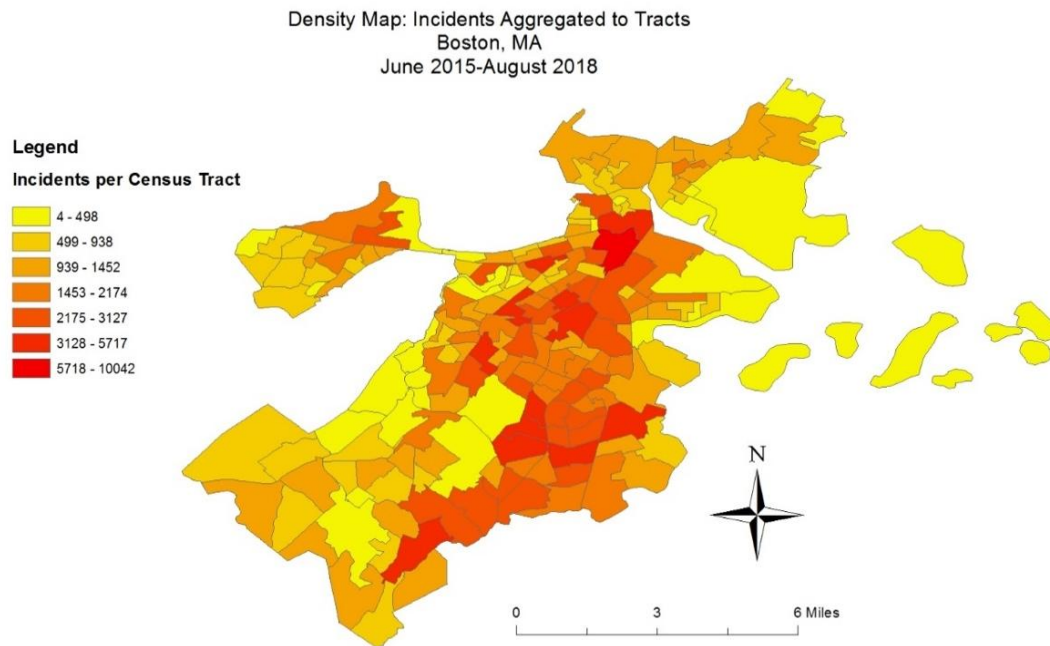


Figure 5



However, if the investigator were instead to consider domestic violence incidents at a higher level of aggregation, such as the block level, they could pair Census information which provides an estimate of each block's racial composition. While this would alter their research question slightly, its ability to include racial characteristics of a place makes questions related racial difference in reporting this crime to the police feasible.

Aggregating to geographic units associated with the U.S. census has been a cornerstone of the criminological literature on place and crime. Commonly used sources of socioeconomic and structural data can be found in the U.S. decennial census and the ongoing American Communities Survey (ACS) (Citation for CENSUS summary; ACS site for possible citation). These data sources report aggregate information on a variety of community characteristics of importance to investigators considering potential associations with 911 and incident information. Racial composition, structural disadvantage, residential mobility, and other measures are reported in aggregate form at various levels. From smallest to largest, census blocks, block groups, census tracts, counties, and state level geographic units contain rich information regarding the general characteristics of these places. For example, a researcher interested in a possible link between levels of unemployment and 911 calls for service may consider using this type of unit. Unemployment data from the ACS at the census tract level would prompt investigators with this research question to aggregate their 911 usage data up to this geographic unit.

Census geographies, while boasting significant contextual information, suffer from a lack of culturally meaningful boundaries. Indeed, these boundaries are strictly based on population sizes and do not necessarily correspond with meaningful distinctions of the individuals living within them. Coulton et al.'s (2001) study on residents' perceptions of neighborhood boundaries demonstrated that these census geographies, while capable of approximation, are incongruent with individuals' understanding of a neighborhood. Researchers interested in explaining the differences between Boston neighborhoods in

larceny shoplifting will need to develop a strategy for defining the neighborhood boundaries. While this task is certainly feasible, it requires a more nuanced understanding of neighborhoods with cultural significance and the acknowledgement of any proxy measures used. Only research questions interested in culturally defined neighborhood differences would need to consider these issues. Aggregating to these units of analysis also fails to capture the intra-unit variation that exists at the previously discussed smaller units of analysis. Research questions that attempt to explore variation within the census geographies should refrain from aggregating to this level, as they may not correspond perfectly with them.

Practitioners may be interested in looking at crime in administratively relevant geographic units. For example, metropolitan police department jurisdictions are often divided into districts and some are further divided into reporting or patrol areas. Investigators seeking to examine how the types of crimes differ across police districts may be interested in selecting these type of geographic unit. For example, Klinger's (1997) theory that the differential workload between patrol districts affects police behavior could examine the 911 and incident summaries aggregated to this level. Other studies could measure calls for service to determine the crime deterrence effects particular types of patrol strategies may have. A randomized controlled trial of the Philadelphia Police Department's foot patrol initiative was interested in measuring crime and deviance changes at the patrol area level, which is a subunit of the patrol district (Ratcliffe et al. 2011). These types of units, however, become less viable when a research question requires detailed community information or seek to explain processes occurring within these places.

BPD's dataset lists XY coordinates for each incident, which are capable of being aggregated to any of the geographic levels mentioned above. Investigator's research questions and data availability should dictate the selection of this level. If a particular research question does not require lower levels of geographic aggregation, such as the address or block face, detailed location information should be promptly removed from the investigator's dataset. This is especially important if the call and incident data is combined with any other form of data that contains potentially identifying information or information that could be used for to re-identify subjects. ArcMap 10.4 and other mapping programs allow users to join locations of incidents (reported in the BPD dataset as XY coordinates) to the higher levels of geography previously discussed. Geographic files of the desired geographic level are the only prerequisite to this process, which are readily available online. Aggregating incidents to units should occur early in the process so the detailed location information, unless needed for the analysis, can be removed from the investigator's working dataset.

## **D. Limitations, Technical Challenges, and Interpreting Results**

Researchers that seek to describe the data contained within the BPD's 911 and incident summary report can easily provide broad descriptive statistics about the quantity, location, and trends of the data. General conclusions can be derived from this type of approach, however, the conclusions capable of being drawn from this are restricted. Advanced spatial and statistical analysis afford researchers greater insight into the underlying associations and potential causal relationships between crime data and other available information. Engaging in these types of analyses requires a technical proficiency in and an appreciation for the limitations of spatial and statistical methods. The problem of spatial autocorrelation provides an example of the many technical challenges facing investigators interested in

going beyond broad descriptions of the data. Basic regressions models, which are commonly used to assess the relationship among variables, is dependent upon the assumption of independence. This means data is not connected to or influenced by factors not included in our statistical model. Researchers concerned with the spatial causes of crime have observed that criminal behavior occurring in one place is influenced by crime in adjoining places, which has been termed spatial autocorrelation (Mencken and Barnett 1999; Rice and Smith 2002). Because criminal behavior is affected by proximate criminal behavior, its assumption of independence is violated; similar independence violations almost assuredly affect 911 and incident data and necessitate the use of analytical techniques that account for this problem. Involving practitioners and academics familiar with these approaches can help investigators less familiar with spatial and statistical analysis in properly answering their research question. The processes described above will permit researchers to begin testing almost any research question.

Whether they are generated by advanced spatial analysis or simple descriptive statistics, the results of any research question need to be scrutinized and contextualized. Stakeholder involvement is especially valuable in process, considering their intimate knowledge of the social and organizational processes that may influence the data. For example, a researcher interested in examining alcohol related disorder will understandably view the spatial and temporal clustering of this type of incident as a phenomenon that requires police intervention. However, relevant stakeholders involved in the analysis may inform the researcher that these incidents coincided with a victory parade for a local sports team that is not used to winning very often. What could have been misinterpreted as a rash of alcohol incidents attributed to other spatial and/or temporal factors is actually the result of a single event. Police practitioners will then be included to look at this cluster of incidents as an outlier that does not necessitate a policy response. The stakeholder interpretation of the relatively straightforward problem in the preceding example is good for illustrative purposes, but does not fairly represent the complexity of factors leading to crime and disorder and reporting behaviors. Stakeholder involvement at all stages of the research process, but especially at this final stage, allows researchers to leverage the institutional and local knowledge that could provide alternative explanations of results, provide further context, and suggest possible policy implications.

## **E. Using the 'Big Data' Approach**

The Big Data movement not only seeks to capitalize on the volume of new datasets, but also their ability to be combined with other datasets. BPD's detailed information on the location and time of incidents facilitates its compatibility with other data sources. Work by the Boston Area Research Initiative illustrates some of the capabilities of this Big Data approach (BARI WEBSITE CITATION). Their use of 911 calls in combination with other Big Data sources such as city administrative data from 311 calls, building permits, and emergency medical services exemplifies this approach and effectively demonstrates the capability of one dataset to complement another. As the digital data capacity of public and private institutions continues to expand, so too will the possible uses of this type of dataset. Respecting the privacy of individuals generating this data will continue to be an imperative for those using the 911 and incident data alone or in combination with other sources. Further, the limitations of this specific data set are in no way alleviated by their pairing with other data. Instead the biases engrained in the data will permeate into any Big Data project and potentially impact the interpretability of the data. Making

strategic offense groupings and geographic unit decisions should consider the other sources of data and inform this process.

## IV. Political Implications

This paper has attempted to highlight the capabilities of Big Data in the public policy arena and through the example of 911 call data demonstrate the thoughtful approach that must be brought to big data analysis. Public policy in a variety of fields can benefit from the creation of effective, focused strategies based on this type of analysis. These strategies offer the promise of quantifying the impact of policy and programing in easily interpretable ways. However, this paper has also stressed the need for caution in building policy solely upon the analysis of Big Data, given the inherent limitations that are present within these data sources. These issues, stemming from the complexity of Big Data, present unique challenges to public policy makers who are embedded within a political environment. City officials, police agencies, and other public institutions who seek to make use of big data analysis must be cognizant of the potential dangers. As has been discussed, policy that relies solely upon quantifiable metrics, such as those often derived from big data sources, is at risk of becoming overly dependent on these measurements at the expense of other potentially important considerations. Scrutiny from political adversaries or media outlets can rally behind flawed big data results to torpedo an otherwise effective policy. Overzealous policy makers can implement ineffective policy on the basis of flawed big data results.

In addition to the possible criticism that may result from the implementation of ineffective policy, Big Data analysis may also lead to problems concerning the evaluation of policy effectiveness. Ineffective policy results from situations in which policy and program evaluation becomes dependent on quantifiable metrics based on faulty big data analysis. This paper has already outlined how policy makers tend to focus on short term outputs, and judge the effectiveness of policy on metrics such as the number of individuals within a certain program or the number of arrests made. This tendency may mean that policy built upon the findings of big data may continue to be judged on metrics that ignore the wider societal harms that may result from such policy change. Relevant to this issue is the analysis of data from the Drug Abuse Resistance Education (DARE) program by Gorman and Huber (2009). In this paper, the authors found that the methods of data analysis used to evaluate the effectiveness of drug prevention programs such as DARE were inadequate and based on limited measurements, resulting in a number of ineffective programs being categorized as proven and evidence-based. The authors reported that those spearheading these drug prevention programs were happy with this categorization, and did not question the quality of the analysis as a result (Gorman and Huber 2009). This example highlights the danger in accepting favorable results of data analysis without question, and should inform how politicians and policy makers review the findings of big data analysis.

Further, the issue of resource allocation based on the findings of big data analysis may pose political implications. The notion of opportunity loss demonstrates the potential harm from implementation of ineffective policy or programing. If a policy resulting from a flawed analysis or misleading data is implemented, the potential societal benefits from a potentially effective program are squandered (Benson et al. 2001). Police and social service agencies with limited resources are then dependent on quality data and proper interpretation to ensure they achieve their optimal effectiveness. Echoing this papers earlier discussion on policy outliers, rare events should be placed firmly within the context of its

occurrence, in order to avoid the possibility of policy being created or changed to deal with an issue that does not actually exist. Socially complex problems may manifest in or be tangentially related to criminal or disorderly incidents. Such incidents could be misinterpreted after being stripped of important details and stored in a 911 data. This is a particularly salient point when one considers the tendency of police agencies to treat wider social problems as a 'police' problem, which is often the case regarding homelessness (Amster 2003) and persons with mental illness (Hirschfield et al. 2006). These problems are indistinguishable from other types of crime in many 911 datasets, however, the most effective responses would not necessarily be implemented. Cooperation between public agencies and community representatives can serve as a safeguard against these situations and reaffirms the need to involve stakeholders in big data analysis. Ultimately, big data analysis has the potential to improve how policy is formulated, and assist in its evaluation, but requires to be used in an informed manner by individuals who are fully aware of its limitations and considerate of the wider implications that may result.

## References

- Permit Patty: Woman 'Calls Police' On Eight-Year-Old for Selling Water. (2018). Retrieved 10/01/2018, 2018
- Amster, R. (2003). Patterns of Exclusion: Sanitizing Space, Criminalizing Homelessness. *Social Justice*, 30(1 (91)), 195-221.
- Anderson, E. (2000). *Code of the Street: Decency, Violence, And The Moral Life Of The Inner City*: W. W. Norton & Company.
- Barocas, S., & Nissenbaum, H. (2014). Big Data's End Run around Anonymity and Consent. *Privacy, Big Data, and the Public Good: Frameworks for Engagement*, 1, 44-75.
- Beckett, K., Nyrop, K., & Pfingst, L. (2006). Race, Drugs, And Policing: Understanding Disparities In Drug Delivery Arrests\*. *Criminology*, 44(1), 105-137. Doi: Doi:10.1111/J.1745-9125.2006.00044.X
- Benson, B. L., Leburn, I. S., & Rasmussen, D. W. (2001). The Impact of Drug Enforcement on Crime: An Investigation of the Opportunity Cost of Police Resources. *Journal of Drug Issues*, 31(4), 989-1006. Doi: 10.1177/002204260103100410
- Bowles, R., Garcia Reyes, M., & Garoupa, N. (2009). Crime Reporting Decisions and the Costs of Crime. *European Journal on Criminal Policy and Research*, 15(4), 365. Doi: 10.1007/S10610-009-9109-8
- Bowman, D. H. (2004). Bomb Threats Taking Financial Toll. *Education Week*, 24(1), 1.
- Braithwaite, J. (1985). White Collar Crime. *Annual Review of Sociology*, 11(1), 1-25. Doi: 10.1146/Annurev.Soc.11.080185.000245
- Bryman, A. (2007). The Research Question In Social Research: What Is Its Role? *International Journal of Social Research Methodology*, 10(1), 5-20. Doi: 10.1080/13645570600655282
- Bursik Jr, R. J., & Grasmick, H. G. (1993). The Use of Multiple Indicators to Estimate Crime Trends In American Cities. *Journal of Criminal Justice*, 21(5), 509-516.
- Chan, J., & Bennett Moses, L. (2017). Making Sense of Big Data for Security. *The British Journal of Criminology*, 57(2), 299-319.
- Chandler, D. (2015). A World without Causation: Big Data and the Coming of Age of Posthumanism. *Millennium*, 43(3), 833-851.
- Cheng, H. G., & Phillips, M. R. (2014). Secondary Analysis of Existing Data: Opportunities and Implementation. *Shanghai Archives Of Psychiatry*, 26(6), 371-375. Doi: 10.11919/J.Issn.1002-0829.214171
- Clarke, R. P., Gohar. (2013). *Shoplifting Problem-Oriented Guides For Police Problem-Specific* (Vol. 11). Washington, DC: Office Of Community Oriented Policing Services.
- Coulton, C. J., Korbin, J., Chan, T., & Su, M. (2001). Mapping Residents' Perceptions of Neighborhood Boundaries: A Methodological Note. *American Journal of Community Psychology*, 29(2), 371-383. Doi: 10.1023/A:1010303419034
- Crawford, A., & Hutchinson, S. (2015). Mapping the Contours of 'Everyday Security': Time, Space and Emotion. *British Journal of Criminology*, 56(6), 1184-1202.
- Crawford, K., Gray, M. L., & Miltner, K. (2014). Big Data | Critiquing Big Data: Politics, Ethics, Epistemology | Special Section Introduction. *International Journal of Communication*, 8, 10.



- Crawford, K., & Schultz, J. (2014). Big Data And Due Process: Toward A Framework To Redress Predictive Privacy Harms. *BCL Rev.*, 55, 93.
- Crow, M. S., Snyder, J. A., Crichlow, V. J., & Smykla, J. O. (2017). Community Perceptions of Police Body-Worn Cameras: The Impact of Views on Fairness, Fear, Performance, and Privacy. *Criminal Justice and Behavior*, 44(4), 589-610.
- Davidson, N. M. (2017). Affordable Housing Law and Policy in an Era of Big Data. *Fordham Urb. LJ*, 44, 277.
- Desmond, M., Papachristos, A. V., & Kirk, D. S. (2016). Police Violence and Citizen Crime Reporting In the Black Community. *American Sociological Review*, 81(5), 857-876.
- Engel, R. S., & Whalen, J. L. (2010). Police–Academic Partnerships: Ending the Dialogue of the Deaf, the Cincinnati Experience. *Police Practice and Research: An International Journal*, 11(2), 105-116.
- Farrell, A., & Reichert, J. (2017). Using U.S. Law-Enforcement Data: Promise and Limits in Measuring Human Trafficking. *Journal of Human Trafficking*, 3(1), 39-60. Doi: 10.1080/23322705.2017.1280324
- Felson, R. B., Messner, S. F., & Hoskin, A. (1999). The Victim-Offender Relationship And Calling The Police In Assaults\*. *Criminology*, 37(4), 931-948. Doi: Doi:10.1111/J.1745-9125.1999.Tb00510.X
- Felson, R. B., Messner, S. F., Hoskin, A. W., & Deane, G. (2002). Reasons for Reporting and Not Reporting Domestic Violence to the Police\*. *Criminology*, 40(3), 617-648. Doi: Doi:10.1111/J.1745-9125.2002.Tb00968.X
- Fisher, B. S., Daigle, L. E., Cullen, F. T., & Turner, M. G. (2003). Reporting Sexual Victimization to the Police and Others: Results from a National-Level Study Of College Women. *Criminal Justice and Behavior*, 30(1), 6-38. Doi: 10.1177/0093854802239161
- Fleury, R. E., Sullivan, C. M., Bybee, D. I., & Davidson, W. S. (1998). "Why Don't They Just Call The Cops?": Reasons For Differential Police Contact Among Women With Abusive Partners. *Violence and Victims*, 13(4), 333-346.
- Gorman, D. M., & Huber, J. C. (2009). The Social Construction Of "Evidence-Based" Drug Prevention Programs: A Reanalysis Of Data From The Drug Abuse Resistance Education (DARE) Program. *Evaluation Review*, 33(4), 396-414. Doi: 10.1177/0193841x09334711
- Gove, W. R., Hughes, M., & Geerken, M. (1985). Are Uniform Crime Reports A Valid Indicator Of The Index Crimes? An Affirmative Answer with Minor Qualifications\*. *Criminology*, 23(3), 451-502. Doi: Doi:10.1111/J.1745-9125.1985.Tb00350.X
- Groff, E. R., Weisburd, D., & Yang, S.-M. (2010). Is It Important To Examine Crime Trends At A Local "Micro" Level?: A Longitudinal Analysis Of Street To Street Variability In Crime Trajectories. *Journal Of Quantitative Criminology*, 26(1), 7-32. Doi: 10.1007/S10940-009-9081-Y
- Gupta, S., Chen, H., Hazen, B. T., Kaur, S., & Gonzalez, E. D. S. (2018). Circular Economy and Big Data Analytics: A Stakeholder Perspective. *Technological Forecasting and Social Change*.
- Hagan, J., Mccarthy, B., Herda, D., & Chandrasekher, A. C. (2018). Dual-Process Theory of Racial Isolation, Legal Cynicism, and Reported Crime. *Proceedings of the National Academy of Sciences*, 115(28), 7190-7199.
- Hirschfield, P., Maschi, T., White, H. R., Traub, L. G., & Loeber, R. (2006). Mental Health And Juvenile Arrests: Criminality, Criminalization, Or Compassion?\*. *Criminology*, 44(3), 593-630. Doi: Doi:10.1111/J.1745-9125.2006.00058.X

- Initiative, B. A. R. (2018). Boston Area Research Initiative. Retrieved 10/01/2018, 2018
- Jahanian, F. (2014). The Policy Infrastructure for Big Data: From Data to Knowledge to Action. *ISJLP*, 10, 865.
- James, N. (2008). How Crime In The United States Is Measured. Washington DC: Library of Congress.
- Kirk, D. S., & Papachristos, A. V. (2011). Cultural Mechanisms and the Persistence of Neighborhood Violence. *American Journal of Sociology*, 116(4), 1190-1233. Doi: 10.1086/655754
- Klinger, D. A. (1997). Negotiating Order In Patrol Work: An Ecological Theory Of Police Response To Deviance\*. *Criminology*, 35(2), 277-306. Doi: Doi:10.1111/J.1745-9125.1997.Tb00877.X
- Klinger, D. A., & Bridges, G. S. (1997). Measurement Error in Calls-For-Service as an Indicator of Crime. *Criminology*, 35(4), 705-726.
- Law, J. (2004). *After Method: Mess In Social Science Research*: Routledge.
- Macdonald, Z. (2002). Official Crime Statistics: Their Use and Interpretation. *The Economic Journal*, 112(477), F85-F106. Doi: Doi:10.1111/1468-0297.00685
- Marcelo, P. (2017). In Boston, 'Methadone Mile', an Open-Air Drug Market. Retrieved 10/01/2018, 2018.
- Mattioli, M. (2014). Disclosing Big Data. *Minn. L. Rev.*, 99, 535.
- Mayer-Schönberger, V., & Cukier, K. (2014). Big Data: A Revolution That Will Transform How We Live, Work, And Think: Citeseer.
- Mcelheran, K., & Brynjolfsson, E. (2016). The Rise Of Data-Driven Decision-Making Is Real But Uneven. *IEEE Engineering Management Review*, 45(4), 103-105.
- Meier, R. F. G., Gilbert (1997). *Victimless Crime? Prostitution, Drugs, Homosexuality, Abortion*. Los Angeles, CA: Roxbury Publishing Co.
- Mencken, F. C., & Barnett, C. (1999). Murder, Nonnegligent Manslaughter, and Spatial Autocorrelation in Mid-South Counties. *Journal of Quantitative Criminology*, 15(4), 407-422. Doi: 10.1023/A:1007592124641
- Metcalfe, J., & Crawford, K. (2016). Where Are Human Subjects In Big Data Research? The Emerging Ethics Divide. *Big Data & Society*, 3(1), 2053951716650211.
- Michael, M., & Lupton, D. (2016). Toward A Manifesto for the 'Public Understanding of Big Data'. *Public Understanding of Science*, 25(1), 104-116.
- Miles, S. (2017). Stakeholder Theory Classification: A Theoretical and Empirical Evaluation of Definitions. *Journal of Business Ethics*, 142(3), 437-459.
- O'Brien, D. T., & Winship, C. (2017). The Gains of Greater Granularity: The Presence and Persistence of Problem Properties in Urban Neighborhoods. *Journal of Quantitative Criminology*, 33(3), 649-674.
- Ohm, P. (2010). Broken Promises of Privacy: Responding To the Surprising Failure of Anonymization. *UCLA L. Rev.*, 57.
- O'Looney, J. (1997). *Emergency 911 Services: A Guide For Georgia Local Governments*: Carl Vinson Institute of Government, University of Georgia.
- Orlikowski, W. J., & Gash, D. C. (1994). Technological Frames: Making Sense of Information Technology in Organizations. *ACM Transactions on Information Systems (TOIS)*, 12(2), 174-207.

- Pattillo, M. E. (1998). Sweet Mothers And Gangbangers: Managing Crime In A Black Middle-Class Neighborhood\*. *Social Forces*, 76(3), 747-774. Doi: 10.1093/Sf/76.3.747
- Peterson, R. D., & Krivo, L. J. (2010). *Divergent Social Worlds: Neighborhood Crime and the Racial-Spatial Divide*: Russell Sage Foundation.
- Ratcliffe, J. H., Taniguchi, T., Groff, E. R., & Wood, J. D. (2011). The Philadelphia Foot Patrol Experiment: A Randomized Controlled Trial of Police Patrol Effectiveness in Violent Crime Hotspots\*. *Criminology*, 49(3), 795-831. Doi: Doi:10.1111/J.1745-9125.2011.00240.X
- Reiss, A. J. (1973). *The Police and the Public*. New Haven: Yale University Press.
- Rengert, G. F., & Lockwood, B. (2009). Geographical Units of Analysis and the Analysis of Crime. In D. Weisburd, W. Bernasco, & G. J. N. Bruinsma (Eds.), *Putting Crime in Its Place: Units of Analysis in Geographic Criminology* (Pp. 109-122). New York, NY: Springer New York.
- Rice, K. J., & Csmith, W. R. (2002). Socioecological Models of Automotive Theft: Integrating Routine Activity and Social Disorganization Approaches. *Journal of Research in Crime and Delinquency*, 39(3), 304-336. Doi: 10.1177/002242780203900303
- Richards, N. M., & King, J. H. (2013). Three Paradoxes of Big Data. *Stan. L. Rev. Online*, 66, 41.
- Rojek, J., Smith, H. P., & Alpert, G. P. (2012). The Prevalence and Characteristics of Police Practitioner–Researcher Partnerships. *Police Quarterly*, 15(3), 241-261.
- Rosenbaum, J. E., Reynolds, L., & Deluca, S. (2002). How Do Places Matter? The Geography of Opportunity, Self-Efficacy and a Look inside the Black Box of Residential Mobility. *Housing Studies*, 17(1), 71-82. Doi: 10.1080/02673030120105901
- Sampson, R. (2002). *Misuse and Abuse of 911*. Washington, DC: Office Of Community Oriented Policing Services.
- Sampson, R. J., & Bartusch, D. J. (1998). Legal Cynicism and (Subcultural?) Tolerance of Deviance: The Neighborhood Context of Racial Differences. *Law & Society Review*, 32(4), 777-804. Doi: 10.2307/827739
- Schrag, Z. M. (2010). *Ethical Imperialism: Institutional Review Boards and the Social Sciences, 1965–2009*: JHU Press.
- Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place. *Criminology*, 27(1), 27-56.
- Skogan, W. G. (1984). Reporting Crimes to the Police: The Status of World Research. *Journal of Research in Crime and Delinquency*, 21(2), 113-137. Doi: 10.1177/0022427884021002003
- Troelsen, L. H. B., Lindsey. (2012). Combating Pickpocketing in Public Transport. *Public Transport International*, 61(1).
- Valente, T. L. R. (2015). *Who Will Help Me? Domestic Violence Survivors Speak Out About Law Enforcement Responses*. Washington, DC: National Domestic Violence Hotline.
- Warner, B. D., & Pierce, G. L. (1993). Reexamining Social Disorganization Theory Using Calls To The Police As A Measure Of Crime. *Criminology*, 31(4), 493-517.
- Warr, M. (1989). What Is The Perceived Seriousness Of Crimes? *Criminology*, 27(4), 795-822. Doi: Doi:10.1111/J.1745-9125.1989.Tb01055.X

- Weisburd, D., Bruinsma, G. J. N., & Bernasco, W. (2009). Units of Analysis in Geographic Criminology: Historical Development, Critical Issues, and Open Questions. In D. Weisburd, W. Bernasco, & G. J. N. Bruinsma (Eds.), *Putting Crime in Its Place: Units of Analysis in Geographic Criminology* (Pp. 3-31). New York, NY: Springer New York.
- Weisburd, D., Eck, J. E., Braga, A. A., Telep, C. W., Cave, B., Bowers, K., . . . Yang, S.-M. (2016). *Place Matters: Criminology for the Twenty-First Century*. Cambridge: Cambridge University Press.
- Weiss, K. G. (2010). Male Sexual Victimization: Examining Men's Experiences of Rape and Sexual Assault. *Men and Masculinities*, 12(3), 275-298. Doi: 10.1177/1097184x08322632
- White, M. D., Todak, N., & Gaub, J. E. (2018). Examining Body-Worn Camera Integration and Acceptance among Police Officers, Citizens, and External Stakeholders. *Criminology & Public Policy*, 17(3), 649-677.
- Williams, M., & Hall, J. (1972). Knowledge of the Law in Texas: Socioeconomic and Ethnic Differences. *Law & Society Review*, 7(1), 99-118. Doi: 10.2307/3052831

## Appendix – NIBRS Offense Grouping and Hierarchy

Rank	Offense Group	Description	NIBRS offense code
1	Homicide	Homicide (murder and non-negligent manslaughter)	9
2	Sexual Offences	Rape, attempted rape	11A
3	Robbery	Robbery	120
4	Aggravated Assault	Aggravated Assault	13A
5	Kidnapping/Abduction	Kidnapping/Abduction	100
6	Burglary	Burglary/Breaking & Entering	220
7	Larceny (person)	Pocket-picking Purse-snatching All Other Larceny	23A 23B 23H
	Larceny (business)	Shoplifting Theft From Building Theft From Coin-Operated Machine or Device Stolen Property Offenses	23C 23D 23E 280
8	Motor Vehicle Theft	Motor Vehicle Theft Theft From Motor Vehicle Theft of Motor Vehicle Parts or Accessories	240 23F 23G
9	Arson	Arson	200
10	Major offenses to society	Animal Cruelty Peeping Tom/HARASSMENT  Prostitution Sex Offense Other Weapon Law Violations	720 90H  40A  520
11	Drug Offenses	Drug Violations Drug Equipment Violations	35A 35B

		Drugs	
12	Less serious assaults	Simple Assault	13B
		Intimidation	13C
13	Vandalism	Destruction/Damage/Vandalism of Property	290
14	Fraud/Crimes of Dishonesty	Counterfeiting/Forgery	250
		Bribery	510
		Embezzlement	270
		Fraud	26
		Gambling	39
		EXTORTION OR BLACKMAIL	
15	Pornography Offense	Pornography/Obscene Material	370
16	Drunkenness/Liquor offense	DrivingUndertheInfluence	90D
		Drunkenness	90E
		Liquor Law Violations	90G
17	Misdemeanor/Non-criminal acts	Bad Checks	90A
		Trespass of Real Property	90J
		Disorderly Conduct	90C
18	All Other Offenses or Incident Types	All Other Offenses	90Z
		Traffic/Motor Vehicle	
		Youth/Status Offenses	
		City Ordinance/Civil Enforcement	
		Lost/Missing/Damaged Property	
		Missing Persons Investigations	
		Social Services (Non Law Enforcement)	
		Investigate Person/Property	
		HUMAN TRAFFICKING - COMMERCIAL SEX ACTS	
		HUMAN TRAFFICKING - INVOLUNTARY SERVITUDE	
		Offenses Against Child / Family	
		Restraining Order Violations	
		Manslaughter: KILLING OF FELON BY POLICE	
		Warrants	

