8
Interpreting Crime Data and Statistics

Rachel Boba

The goal of this chapter is to provide knowledge of how to appropriately apply and interpret statistics relevant to crime analysis. This chapter includes a discussion of data, statistics, and application of statistics in crime analysis. The chapter does not cover all types of statistics but focuses on those that are most useful in crime analysis. Additionally, this is not a "how to" chapter in which formulas are included, but a discussion of data and statistics and their practical application. Throughout the chapter, there are additional resources recommended. One should view this as a first step to learning about applying and interpreting statistics for the purposes of crime analysis.

To understand crime and law enforcement statistics, one must first understand the data systems. A brief overview of systems in which police data are collected and housed is included. In addition, knowledge of what the data describe and what they may not describe is important in applying and interpreting statistics accurately and effectively. Even though there are many types of data used in crime analysis, for the sake of brevity, only three are discussed in this chapter, as they are the primary data sources for crime analysis. They are crime, arrests, and calls for service.

After a discussion of the data types and their effect on the interpretation of the statistics, the next section includes a specific discussion of selected statistics relevant to crime analysis work. Basic statistics, such as frequency and percent, are not discussed, as they will be covered in Chapter 9. Discussion of rates, measures of central tendency, inference, statistical significance, t-tests, and correlation are included.

Police Data Systems

Police departments typically have two data systems in which crime, arrests, and calls for service, among other data, are housed. Even though these systems are slightly different from agency to agency, the basic purposes and functions are the same and thus these systems are discussed generally here. The two types of systems are referred to as a computer aided dispatch system (CAD) and a records management system (RMS). Although many small agencies may not have these systems and others may have additional specialized databases housed in specific units, the CAD and RMS are key systems for data in policing and crime analysis.

A CAD or a computer-aided dispatch system is a highly specialized telecommunications and geographic display technology created to support police and public safety response operations. CADs are typically used for all emergency operations, which in addition to police include fire and ambulance services. In a police department, calls for service that are dispatched to or initiated by officers are collected and warehoused in a CAD system. All activity concerning these calls is recorded in the CAD, which includes the information about the call recorded by the call taker and dispatcher (e.g., officers dispatched to the scene, when they arrived, the disposition of the call) and correspondence between officers and dispatchers over the system (e.g., vehicle lookups, officers taking lunch). This information is typically called "unit history" as it records the activity of each police unit during a call.

It is important to note that a CAD system does not record all phone calls coming into the police agency, but those dispatched to or

initiated by officers. Because of the vast information contained in a CAD system, information is often purged after a certain amount of time. However, it is standard to retain a summary or snapshot of a call for service for a longer period of time for analysis purposes. This information is often downloaded into an RMS.

An RMS or records management system is a technology in which police records are stored as databases. The main purpose of an RMS is to store crime incidents and related data such as property data, evidence, vehicle data, and persons data (e.g., suspects, arrests, witnesses). Many departments also include other types of data in their RMSes, such as personnel data, investigation information (cases), traffic stop data, and calls for service information (summary data). An RMS is used in a police agency for quality control evaluation, to look up individual incidents, to track cases, and to run reports.

Historically and still today, it is a struggle for crime analysts to retrieve data from their agencies' CAD and RMS systems. In many cases, software vendors create functions for producing paper reports and retrieving one case at a time without creating functions for downloading large amounts of data needed for analysis purposes. When they do have access to the systems, crime analysts typically have access to the RMS system but not to the CAD system directly since the summary information is exported to the RMS. Less often, a police agency will not have an RMS at all, and the analyst only has access to the CAD system and, consequently, may access data directly through the CAD. However, in either case the data used is a summary of the calls for service instead of the complete CAD records. The primary exception is when data are used for staffing allocation, when all unit history and back up unit data is needed.

Key Crime Analysis Data Sources

To understand crime and law enforcement statistics, one must first understand the data on which the statistics are based. Knowledge of what the data do and do not describe is important in applying and interpreting statistics accurately and effectively. This section covers the primary types of data crime analysts use— crime, calls for service, and arrests. There are many considerations that must be taken into account with these data types as these issues can affect the selection and interpretation of subsequent statistical analyses.

Crime

Crime data is the primary type of data used in crime analysis. There are many issues about crime data that can lead to misinterpretation or misuse of statistics. The section covers some of the general issues with crime data and then discusses the various ways in which crime can be counted.

Crime represented by police crime data does not represent all crime occurring in society. Thus, we typically see reports that name this type of data, "reported crime" or "crimes known to the police." This is an important distinction for anyone interpreting or trying to understand crime problems based on police data: that is, that we may not be aware of the entire problem and the police data may portray a biased picture. For example, we know that a very low number of rapes are reported to the police. Thus, if we report rape statistics rising or falling we must be very cautious and mention the fact that even though the police are seeing an increase or decrease, the actual number may be changing in a different way since we do not know the actual number of rapes being committed. This is particularly relevant for certain types of crimes (domestic violence, drug crimes, white collar crimes) and not as much for others (motor vehicle theft, arson, murder).

Crime data captured by police agencies is dynamic, not static. In other words, information about crime incidents is constantly being updated. For example, a person may report a crime that occurred two days before to an officer who then takes a day or two to complete the report. The case is assigned to a detective who begins to investigate the crime by identifying suspects a few days, weeks, or years later. The victim may call the police department with new information about the crime or correct erroneous information from the original report. The officer may arrest a suspect and clear the case. As you can see, the data for this incident are constantly being changed and updated, and a report including this case could change based on when the crime analyst downloaded the data for statistics.

That the data surrounding crime incidents are constantly changing creates an issue for analysis. One of the issues is "real-time" data. Many agencies emphasize performing statistics on the most recent data—as recent as an hour ago. Yet the likelihood of that data changing soon after the initial report is recorded is fairly high because of additional information from investigation, identification of suspects, errors in the original report, and so on. Theoretically, however, after a certain amount of time, the likelihood of the data being changed substantially is significantly lower. Unfortunately, there is no research in this area and analysts must take an educated guess about when the likelihood for further changes to the data is low. In many cases, crime data will be downloaded around the fifteenth of the month following when it is reported and then updated in six month or one year intervals to capture arrests and clearances. In any case, the use of "real-time" data should be done with extreme caution.

All crime data recorded by the police have two sets of dates. The first is the date the report of the crime was written. (In some cases, the time the report was written is also recorded, but this information is not essential, except for certain operations studies.) The second is the date(s) and time(s) the crime actually occurred. This is included in crime data because the date a crime was reported is not always the same as when it occurred.

In many cases, the exact date and time that the crime occurred are known (e.g., robbery, assault); however, in other types of crime, the exact date and time of the crime are not known because they were not witnessed by anyone (e.g., burglary, auto theft, larceny). In these crimes, victims report their best estimate of when the crime "could" have occurred. These are called "first possible" and "last possible" date and time or "from" and "to." The difference between these two sets of dates (when the crime was reported vs. when the crime occurred) is very important as they serve different purposes depending on the nature of the analysis.

All crime reported to the police is counted based on the date it is reported, because it would be impossible to count it by when it occurred, since this is not always known. As noted above, crime data is dynamic and crimes can be reported days, weeks, months, even years after they occur. Counts would need to be constantly updated based on date of occurrence and the ranges would cause further issues in counting. Date of report is constant, it does not change, and thus counts of crime are based on this variable. However, it can affect the interpretation of crime statistics. For example, if a large number of burglaries occur over the Christmas/New Year's holiday, victims may not report them until after the new year when they return home. Thus, a large number of crimes would be counted in the following month and even the following year when they may have occurred the previous month and year. Although there is nothing an analyst or police can do to correct this problem, it should be considered during the

analysis process. Finally, the date of occurrence, though not adequate for counting crime, is important for analysis and is used for identifying patterns and series—when the crime occurred is much more important that when it was reported.

The various ways in which crime is measured affect statistics created from them. Below are the four major ways in which crime is measured and considerations for each.

Crime Reports

"Crime reports" refer to individual crime incidents that are reported to the police and are contained within the agency's RMS. It is important to note that even though all crime data collected by police agencies is generally the same, how and when incidents are recorded can vary by region, state, and even by police agency. For example, one agency may call tools taken from the back of a truck a theft, whereas another may call it a burglary from vehicle, two very different crime classifications even though it is the same crime. Thus, how an agency records its data will affect its statistics. Because each agency has its own policies, procedures, and RMS (there are no national standards for recording crime incidents), it is problematic comparing statistics based on crime report data across agencies. This is why crime statistics of different agencies are compared using federal standards (e.g., UCR or NIBRS).

Uniform Crime Report (UCR)

The UCR is a national standard of classifying crime uniformly that disregards state laws (see www.fbi.gov/ucr for further details). It was created to provide a way to consistently count crime across the United States. The UCR does not collect information from agencies about all crime, only selected types (e.g., homicide, rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft, and arson) and has a

hierarchical coding system (i.e., if a rape and a murder are committed to the same person, only the murder is counted). However, it does collect information on all arrests no matter the type of crime. Compliance is voluntary, so not all police agencies report their statistics (though over 95 percent of the U.S. population is represented in the UCR statistics). The numbers are typically reported on a monthly basis to a state agency that then reports it to the FBI. These numbers are not updated (static). The most important differences between crime reports and UCR for a crime analyst are that UCR are only aggregate numbers that can not be broken down once reported, they only include certain types of crimes counted hierarchically, and the numbers are static. These are key reasons why UCR crime statistics and crime analysis statistics for the same jurisdiction will rarely match.

National Incident Based Reporting System (NIBRS)

The NIBRS system is an attempt to improve the UCR but is similar in nature.[1] For example, example, it also requires a standard of classifying crime uniformly that disregards state categories. It is an improvement over UCR in that additional primary crime types are tracked; that individual cases (not aggregate counts) are submitted voluntarily to the FBI; and that the system allows agencies to update incident submissions as new data becomes available. Often, analysts use UCR or NIBRS categories for analysis of incident data. Many records management systems recode the incidents accordingly. There are no national NIBRS statistics because a minority of agencies report their data in this way.

National Crime Victimization Survey (NCVS)

The national survey of crime victimization is another way in which criminal activity is

---

[1] See www.fbi.gov/ucr for further details.

measured, but it focuses on the experiences of victims. The purpose of the survey is to avoid counting only crimes reported to the police and to have statistics independent of police activity. It began in 1973 and samples 100,000 residents age 12 or older, from nearly 50,000 households.[1] Respondents are interviewed every six months over a 3-year period. The survey gathers information such as location of the crime, month, time of day, physical setting, characteristics of the victims (e.g., sex, race, income, extent of injury, self-protective actions and results of those actions, and relationship with the offender) and provides additional information on crimes that rarely are reported to the police (e.g., rape and domestic violence). It does not collect information on homicides, arson, societal crimes, or commercial crimes; and victims determine whether the incident was a crime, not the law.

The NCVS is most helpful to social scientists and provides national trends of victimization, but local level police agencies are less apt to use this information in their jurisdictions. The primary reason is because a sample is taken from the nation, and there are not enough surveys conducted at smaller levels to generalize to a local population. If we ask: for a city of 150,000, how many would be surveyed in the NCVS? The following estimate can be generated:

NCVS	City
100,000 surveyed=	X surveyed
200,000,000	150,000
Estimate of U.S.	City
population over 12	Population
X = 100,000*150,000/ 200,000,000	
X = 75 would be surveyed in the city	

Even though 75 surveys may be enough to conduct statistics, it would not inform about the prevalence of crime victimization in that

[1] See www.icpsr.umich.edu/NACJD/NCVS/ for further details.

jurisdiction. In addition, the survey does not provide information about crime incidents but experiences of victims. While the national survey may not provide sufficient cases for local findings, many police agencies conduct their own citizen victimization surveys to help inform their crime prevention and control activities as well as the satisfaction with their activities.

Crime Data Summary

All of these types of data serve a purpose in providing a picture of crime. However, depending on which data are used the picture may be very different. A major issue crime analysts confront is why their statistics do not match the "official" statistics reported to the FBI. From the discussion above, we can see how this happens; any of the following could be true:

- The crime analyst downloaded the data at a different time than the UCR or NIBRS data were downloaded
- The crime analyst uses state crime codes, not UCR or NIBRS codes
- The crime analyst uses only those crimes determined by the police (not the victim) to be a crime.
- The crime analyst counts by number of reports, not number of victims or by a hierarchy of crimes.

Arrests

Arrests occur when there is sufficient evidence (probable cause) that a person has committed a crime. In crime analysis, the arrest is usually the final outcome of a crime incident, and we are not concerned with whether the person was actually found (or even is) guilty or sentenced for the crime. Police agencies are often measured by how many arrests occurred for a certain number of crimes (clearance rate). This will be discussed in the next section. This