A Statistical Graphic Exploration of Crime Rate for the States of USA

Ruben Thoplan
Lecturer in Statistics, Department of Economics and Statistics, Faculty of Social Studies and Humanities, University of Mauritius, Réduit, Mauritius

ABSTRACT

In this paper, an in-depth graphical exploration of different types of crime for the states of USA is proposed. The trellis plot, Chernoff face, star plot and geo-visualization are graphical methods adopted in this paper for the exploration of crime rates. The explorations of the different types of crime show that there is a sustainable effort of USA to decrease crime rate. However, the rate of decrease for some types of crime is relatively slow. Overall, there is not much change in terms of geographical concentration for each type of crime since 2008 to 2012. Findings also show that there is a higher tendency for most type of crimes to be located in the southern states.

Keywords: trellis plot; Chernoff face; star plot; geo-visualization; crime

1. Introduction

According to the Federal Bureau of Investigation (2009), crimes also referred to as offences can be classified into two categories; Part I and Part II. Part I offenses are considered as serious offenses which occur on a frequent basis and Part II offenses are all other offenses not reported in Part I. Part I offenses include criminal homicide, forcible rape, robbery, aggravated assault, burglary, larceny-theft (excluding motor vehicle theft), motor vehicle theft and arson. In this paper, only Part I offenses are considered except manslaughter with negligence which is part of criminal homicide and arson. Part I offenses are generally easier to measure as they are more likely to be reported to the police compared to transnational crimes where there is a strong likelihood of under-reporting due to fear of consequences as a result of reporting.

It is important to study the evolution of crime over time and space to understand the dynamics that will allow policy makers to target their policies in view of fighting crime. To understand better the nature and causes of crime, it is crucial to take into account the significant variation of crime in terms of region (Harries, 1974). Evans & Herbert (1989) demonstrates that geographic location and environmental influences have a role to play in crime and the fear of crime.

Apart from the geographical influences, there are many other factors that have an impact on crime in a country. For example, unemployment is one of the factors known to have some effect on certain types of crime like homicide, rape, aggravated assault, robbery, burglary, larceny-theft and motor vehicle theft (Cantor & Land, 1985). In addition, poverty and ethnicity are two other factors which are linked to violent crimes (Short Jr, 1997). According to the Federal Bureau of Investigation (2009), violent crimes include murder and non-negligent manslaughter, forcible rape, robbery and aggravated assault. In the United States of America (USA), violent crime has been decreasing since 2008 to 2011 but increased in 2012.

One of the consequences of crime is fear which itself hinders the well-being of an individual. In that sense, it is essential to prevent crime in view of having a future without fear (UNODC, 2014). The fear of crime can impact on an individual’s life in terms of avoidance behaviors which involve limitations applied on one’s activity. It can also affect someone in terms of defensive behaviors which involve actions taken by the person to relieve the feeling of fear like carrying weapon for self-defense (Addington, 2010).
In this study, emphasis is laid upon the identification of patterns in the crime rates of the states of USA (excluding Alaska and Hawaii). Crime rates across the states are explored graphically (i) by states to see if there is any geographical link with specific types of crime, (ii) over time to understand the trend in different types of crime and (iii) over both time and states to see any decline or increase in specific types of crimes by state. Section 2 describes the different methods of graphical tool employed herein to identify patterns in crime rates. Section 3 presents the findings and section 4 attempts recommendations and conclusions based on the findings.

2. The Statistical Graphics

Developing an appropriate graphical display to portray the maximum information from a data structure takes time just as with written language (Wilkinson, 2005). There has been much debates over centuries on the idea of what is a good and a bad statistical graphic. An excellent statistical graphic is one which portrays the maximum amount of information in the shortest time with the least ink in the smallest space (Tufte, 2001).

Crime Statistics from the Federal Bureau of Investigation (Federal Bureau of Investigation, 2014) are gathered from 1993 to 2012 and by states (excluding Alaska, Hawaii and Puerto Rico) with respect to different types of crime including violent crime as a whole category. In this section, the graphical tools employed in this paper which are deemed appropriate to portray the maximum information in the shortest time with the least ink are outlined.

2.1 Trellis Plot

The trellis graphic enables the acquisition and interpretation of multivariate information within a single space and in a very short time frame. The credit for the design and implementation of trellis graphics are mainly attributed to Rick Becker, Bill Cleveland, and Ming Shyu from AT&T Research and Bell Labs. The name trellis graph refers to the garden trellis as it is a graphical method which contains rectangular array of plots. A trellis graph can be applied for time series plot, scatter plots and many others.

A trellis plot displays a variable or the relationships between variable, conditioned on one or more other variables (Kabacoff, 2014). The main idea for drawing a trellis graph is the comparison of variables in a multivariate data. While drawing a trellis graph, careful attention to the aspect ratio is vital in order to highlight information in the data that cannot be seen otherwise (Becker, 1996). Moreover, it is important that the scales are similar for each panel in the trellis plot to enable comparison.

The trellis display is very powerful in identifying how a response variable depends on explanatory variables in a study (Becker, et al., 2012). Since its inception, the trellis graph has been used in many instances like quality control (Gunter & Brideau, 2003), environment (Andersen & Kay, 1999), chemistry (Belle, et al., 2011), simulation studies (Sanchez & Lucas, 2002) and many others.

2.2 Chernoff Face and Star Plot

The Chernoff face is a representation of a multivariate data by faces where features of the faces are represented by variables. “It is effective in revealing rather complex relations not always visible from simple correlations based on two-dimensional linear theories” (Chernoff, 1971). People can easily detect very small changes in the facial expression of an individual and the use of the Chernoff face facilitates the identification of similarities and differences in observations by means of many variables in a relatively quick and easy way.
The Chernoff face has been used in many instances like in laboratory analysis (Lott & Durbridge, 1990), computer traffic system (Ren, et al., 2006), cartography (Neslon, 2007), behavioral accounting research (Waller, 1988), music (Hiraga, et al., 2002) and others. The number of face parameters depends a lot on the software being applied. For example, in the use of the “aplpack” package in R (Wolf & Bielefeld, 2013), the following parameters are considered:

1. Height of face
2. Width of face
3. Shape of face
4. Height of mouth
5. Width of mouth
6. Curve of smile
7. Height of eyes
8. Width of eyes
9. Height of hair
10. Width of hair
11. Styling of hair
12. Height of nose
13. Width of nose
14. Width of ears
15. Height of ears.

Additionally, Wolf & Bielefeld (2013) allow the inclusion of different colors in the face by averaging the main features of the face; eyes, iris, lips, ears, nose, hair and face. The use of the colors enhances the faces by showing similarities between faces and colors. An advantage of drawing the Chernoff face is its ability in revealing clusters in a multivariate data.

Apart from the Chernoff face, the star plot is another useful graphic which helps in revealing information from a multivariate data (Chambers, et al., 1983). The star plot has been applied in different fields like in the analysis of power systems (Mahadev & Christie, 1993), genome studies (Vinogradov, 2006), toxicology (Kim, et al., 2010), analysis of trace element (Tamás & Abony, 2002), analysis of drugs (Wu, et al., 1998) and other fields.

In the same vein as a Chernoff face, the star plot also known at the segment diagram or spider web is a way of visualizing multiple variables for each observation in a data set. The area of each segment in a star plot relates to the size of a specific variable. The advantage that the star plot has over the Chernoff face is that it is possible to visualize specific variables behavior more easily compared to that of the Chernoff face. Outliers and clusters can also be identified quite easily in a star plot just like in the case of the Chernoff face.

2.3 Geo-visualization

“Geo-visualization integrates approaches from visualization in scientific computing (ViSC), cartography, image analysis, information visualization, exploratory data analysis (EDA), and geographic information systems (GISystems) to provide theory, methods, and tools for visual exploration, analysis, synthesis, and presentation of geospatial data (with data having geospatial referencing)” (MacEachren & Kraak, 2001). In simple terms, a geospatial data has a geographical component. Geo-visualization allows one to recognize any pattern which may arise in a data related to geography.

In the case of a spatio-temporal data, it is still possible to apply geo-visualization by taking into account the order of time. A spatio-temporal data can be represented in three different ways; long format, time-wide format and space-wide format. For the purpose of data exploration here, the long
format is used. This implies that the data is structured such that column 1 contains information related to the states, column 2 contains information related to time and the other columns relate to the crime rate for the different types of crime.

Geo-visualization has been quite used as a decision support in many fields (MacEachren, et al., 2004). It has been applied in water harvesting position (Patel, et al., 2011), population distribution (Su, et al., 2005), human activity patterns (Kwan, 2004), natural disaster planning (Yeletaysi, et al., 2009), environmental processes (Brewer, et al., 2000) and other fields.

3. Graphical Exploration of Crime Rates

The first exploration of crime starts by analyzing the evolution of rate of decrease in crime since 1993 to 2010. For this purpose, a trellis plot of the log of crime rates since 1993 to 2010 for each type of crime is depicted as per figure 1. The reason why a semi-logarithmic graph has been considered is to be able to track down the rate of change for each type of crime. As observed robbery, larceny-theft and motor-vehicle theft show a relative strong rate of decrease in crime rates as compared to the other types of crime. However, the rate of decrease for aggravated assault, burglary and forcible rape is very slow or nearly constant. The advantage of this plot is that it allows a comparison of the rate of decrease for each type crime. It is clear that much effort needs to be applied in prevention strategies for those types of crimes which have been recording a slow rate of decrease over time.

![Figure 1: Trellis Plot of Evolution of Rate of Decrease in Crime](image)

In a second instance, to identify if there is any relationship between any two types of crime, a scatter plot matrix is produced as per figure 2. There are some positive linear relationships between (1) burglary and property crime and (2) property crime and larceny-theft. The other types of crime do not show any apparent strong relationships. Some outliers can also be identified for some types of crime like murder and non-negligent manslaughter and forcible rape.

Figure 3 shows the Chernoff face for each state. The parameters of the face are as follows: Height of face – Violent crime, Width of face – Murder and non-negligent manslaughter, Structure of face – Forcible rape, Height of mouth – Robbery, Width of mouth – Aggravated Assault, Smiling – Property Crime, Height of eyes – Burglary, Width of eyes – Larceny-theft, Height of hair – Motor vehicle theft, Width of hair – Violent crime, Style of hair – Murder and non-negligent manslaughter, Height of nose – Forcible rape, Width of nose – Robbery, Width of ear – Aggravated Assault, Height of ear – Property crime. The bigger the face, the higher is the crime is as whole. Louisiana and South Carolina
stand out with higher crime rates as a whole. On another note, Idaho, Iowa, Maine, Minnesota, New Hampshire, Utah, Vermont and Wyoming seems to share similar crime characteristics with low crime rate as a whole.

![Scatter Plot Matrix for Different Types of Crime in Year 2012](image1)

**Figure 2:** Scatter Plot Matrix for Different Types of Crime in Year 2012

![Chernoff Face of Crime Rate for US States in Year 2012](image2)

**Figure 3:** Chernoff Face of Crime Rate for US States in Year 2012

A star plot is also depicted as per figure 4. From the star plot, same observations as the Chernoff face can be concluded in terms of the overall crime rate aspect. However, some key features stand out in the star plot. For example, South Dakota stands out in terms of forcible rape and Louisiana stands out for murder and non-negligent manslaughter. This sheds light on the outliers observed in figure 2 above. From the star plot, it seems that robbery is a main concern for many states as the blue segment relating to robbery stands out in many states and even in the states with the lowest crime rates overall. Since the data contains a geographical component, it is important to consider the geo-visualisation of the crime rate as well. Figure 5 is a map of the USA showing the intensity of the different types of crime in year 2012. From figure 5, property crime is more alarming in the state of North Dakota. Another remarkable feature from the map is the concentration of violent crime in the central eastern
part of USA which relates to Kansas, Missouri, Indiana and Michigan. Some investigations are required to understand this phenomenon in that region.

Figure 4: Star Plot of Crime Rates for US States in Year 2012

Figure 5: Intensity of Crime Rates according to Type for Year 2012

Instead of viewing the crime rates at only one specific point in time, a temporal component can be considered for each type of crime to understand the evolution of the crime across states. Figures 6a, 6b, 6c, 6d and 6e show the evolution of the states in terms of the different types of crime since 2008 to 2012. Based on figure 6a, violent crime rate is relatively low in the northern states as compared to the southern states. The evolution has not much changed for violent crime in terms of states. Overall, nearly all the states show a slight decrease over time in terms of violent crime. Regarding property crime, there have not been many changes in relation to violent crime over the years. Property crime is overall higher in the south. Washington is the only northern state with a high level of property crime. Same comments apply to burglary. Regarding larceny-theft, there is no specific geographical correlate. A general decrease in crime rate over time can still be observed. For the case of motor vehicle theft, the crime rate is relatively low for all states except for that of California, Nevada and Washington. Nevada, Michigan, Georgia and South Carolina have noticeably shown a lot of effort to decrease the number of motor vehicle theft over time.
Figure 6a: Evolution of Violent Crime since 2008 to 2012.

Figure 6b: Evolution of Property Crime since 2008 to 2012.

Figure 6c: Evolution of Burglary since 2008 to 2012
4. Conclusions

The patterns of different types of crime in terms of state and over time have been therefore explored here. There is a clear determination by the USA to reduce its crime rate as a whole. But the rate of decrease for some types of crime is not noticeable. In this regard, these specific types of crime like aggravated assault, burglary and forcible rape require some more attention by the authorities.

It is also noted that some types of crime seem to be more concentrated in the southern parts of the USA. Further investigation is required to understand this phenomenon. Research has been made to understand the lower crime rates in the northern states to be able to curtail the high crime rate in the southern parts of the USA. However, the pattern for each type of crime has not much changed by state since 2008 to 2012.

Using the results of the exploration of crime rates from this paper, targeted policies may be devised to lower the crime rate faster. If each state focuses in fighting the main types of crime it faces, the overall crime rate will decrease at a faster rate. This will also reduce the economic costs borne by the Americans in terms of medical care, locale spending on policing, prosecuting and incarcerating (Shapiro & Hassett, 2012). But more importantly, a fast decrease in the rate of crimes will reduce the fear of crime.
References


