

CARTOGRAPHIC APPROACHES TO THE VISUAL EXPLORATION OF VIOLENT  
CRIME PATTERNS IN SPACE AND TIME:

A USER PERFORMANCE BASED COMPARISON OF METHODS

by

Benjamin Green Anderson

---

A Thesis Presented to the  
FACULTY OF THE USC GRADUATE SCHOOL  
UNIVERSITY OF SOUTHERN CALIFORNIA  
In Partial Fulfillment of the  
Requirements for the Degree  
MASTER OF SCIENCE  
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

May 2015

## **ACKNOWLEDGMENTS**

I want to thank Dr. Robert Vos and Dr. Katsuhiko Oda for their support and advice as I worked on this project. I also want to thank the rest of the Spatial Sciences Institute faculty for their support as I completed the program. Finally, thank you to my family and friends, without whom I could not have made it this far.

## TABLE OF CONTENTS

|   |      |
|---|------|
| ACKNOWLEDGMENTS   | ii   |
| LIST OF TABLES  | vi   |
| LIST OF FIGURES   | vii  |
| LIST OF ABBREVIATIONS                                       | viii |
| ABSTRACT  | ix   |
| CHAPTER 1: INTRODUCTION                                     | 1    |
| 1.1 Motivation  | 2    |
| 1.2 Experiment Design Overview                              | 4    |
| 1.3 Research Questions                                      | 5    |
| CHAPTER 2: RELATED WORK                                     | 6    |
| 2.1 Visualization of Spatiotemporal Data with Static Maps   | 6    |
| 2.2 Visualization of Spatiotemporal Data with Animated Maps | 7    |
| 2.3 Interactivity of Spatiotemporal Maps                    | 9    |
| 2.4 Visualization of Crime Patterns                         | 10   |
| 2.5 Comparison between Static and Animated Maps             | 12   |
| 2.5.1 Previous Comparative Studies                          | 14   |
| 2.5.1.1 Socia's Comparative Study                           | 14   |
| 2.5.1.2 Adjustments to Socia's Methods                      | 15   |
| CHAPTER 3: METHODOLOGY                                      | 17   |
| 3.1 Mapped Data   | 18   |

|   |    |
|---|----|
| 3.1.1 Data Aggregation Methods                              | 19 |
| 3.2 Cartographic Design                                     | 20 |
| 3.3 User Performance Experiment Design                      | 22 |
| 3.3.1 Map Test Format and Performance Measurements          | 24 |
| 3.3.2 User-interface Design                                 | 25 |
| 3.3.2.1 Static Time-series Map Display Format               | 25 |
| 3.3.2.2 Animated Time-series Map Display Format             | 26 |
| 3.3.3 Recruitment of Participants                           | 27 |
| 3.3.4 Testing Procedures                                    | 28 |
| 3.4 Methodology for Analyzing the Results of the Experiment | 29 |
| 3.4.1 Statistical Analysis Methodology                      | 30 |
| CHAPTER 4: RESULTS AND ANALYSIS                             | 32 |
| 4.1 Experimental Population                                 | 32 |
| 4.2 Overall Performance Metrics                             | 34 |
| 4.3 Task Accuracy   | 36 |
| 4.4 Confidence  | 39 |
| 4.5 Completion Time   | 41 |
| 4.6 User-preferences  | 44 |
| 4.7 Correlations  | 46 |

|  |    |
|--|----|
| CHAPTER 5: DISCUSSION AND CONCLUSIONS                                    | 48 |
| 5.1 Task Accuracy  | 49 |
| 5.2 Confidence   | 50 |
| 5.3 Completion Time  | 50 |
| 5.4 User-preferences   | 51 |
| 5.4.1 Performance and User-preference Correlations                       | 52 |
| 5.5 Conclusions  | 52 |
| 5.5.1 Parallels with Socia’s Study                                       | 53 |
| 5.5.2 Study Strengths and Weaknesses and Suggestions for Future Research | 54 |
| 5.5.3 Implications for the Cartography Community                         | 56 |
| REFERENCES   | 58 |
| APPENDIX A: Chicago Homicide Hot Spot Maps, 2009-2013                    | 63 |
| APPENDIX B: Map Test Survey Interface Images                             | 70 |

## LIST OF TABLES

|  |    |
|--|----|
| Table 4.1 Age Distribution of Study Participants                       | 33 |
| Table 4.2 Education Level of Study Participants                        | 33 |
| Table 4.3: Accuracy, Confidence, Completion Time, and User-preferences | 35 |
| Table 4.4 Test Score Descriptive Statistics                            | 36 |
| Table 4.5: Confidence Score Descriptive Statistics                     | 39 |
| Table 4.6 Static and Animated Completion Time                          | 42 |
| Table 4.7 User-preference Descriptive Statistics                       | 44 |
| Table 4.8 Static Performance and Preference Correlations               | 46 |
| Table 4.9 Animated Performance and Preference Correlations             | 47 |

## LIST OF FIGURES

|   |    |
|---|----|
| Figure 3.1 Hot Spot Analysis Tool Settings                            | 21 |
| Figure 3.2 Chicago Homicide Hot Spots, 2009-2013, 12:00 AM to 3:59 AM | 22 |
| Figure 3.3 Animated Time-series Map Display Format                    | 27 |
| Figure 4.1 Static and Animated Accuracy Score Histograms              | 37 |
| Figure 4.2 Static and Animated Confidence Score Histograms            | 40 |
| Figure 4.3 Static and Animated Completion Time Box-plots              | 42 |
| Figure 4.4 Static and Animated Completion Time Histograms             | 43 |
| Figure 4.5 Static and Animated Preference Score Histograms            | 45 |

**LIST OF ABBREVIATIONS**

|       |  |
|-------|--|
| GIS   | Geographic Information System(s)         |
| NIBRS | National Incident-based Reporting System |
| SMMD  | Small-multiple Map-display               |



## ABSTRACT

This study provided an empirical comparison of static and animated cartographic representations of spatiotemporal phenomena in their application to basic choropleth map-based knowledge-extraction tasks to answer the following research questions: 1) Do animated maps provide heightened potential for accuracy in completing basic knowledge-extraction tasks over static time-series maps, or vice versa? 2) Do animated maps provide heightened potential for efficiency in completing basic knowledge-extraction tasks over static time-series maps, or vice versa? and 3) How do user preferences align or not align with measurements of accuracy and efficiency?

To this end, this study examined map readers' accuracy and efficiency in completing knowledge-extraction tasks through static and animated time-series maps about homicide patterns in the Chicago metropolitan area. Through an online user performance experiment, participants answered a series of questions about homicide hot spots and cold spots using both static and animated versions of the maps as the basis for their answers. They were also asked to indicate their level of confidence in the accuracy of their responses and to indicate which map type they preferred for completing the tasks. Task completion times were recorded for efficiency measurements. The results of independent samples t-tests indicate statistically significant differences between the static and animated maps in terms of task accuracy and completion time. Generally, users were able to complete the assigned tasks more accurately and much more efficiently using the static maps, as compared with their animated counterparts. Additionally, user-preferences were checked for correlations with task accuracy and completion time via Pearson's product-moment correlation coefficient calculations. The results indicate no significant correlations between performance measurements and user-preferences.

## CHAPTER 1: INTRODUCTION

Technological advances continue to stimulate new methods for the visualization of spatial and spatiotemporal information (Blok et al. 2001). While these methods can provide new and exciting ways to view and interpret spatial data, care must be taken to understand and prevent potential usability issues. With animated maps, for example, the images can sometimes be too fleeting to be perceived correctly (Betrancourt and Tversky 2002). Unfortunately, the excitement that often surrounds new technological developments in cartography, particularly temporally animated maps, interactive web-maps, and 3D spatiotemporal data visualizations, can sometimes distract cartographers and their audiences from any potential limitations that might accompany these new visualization tools. Furthermore, Andrienko et al. (2008) suggest that as animated maps become more and more common, it is increasingly important to understand the utility of temporal map animation for supporting basic map-based knowledge-extraction tasks.

This study endeavors to provide an empirical comparison of static and animated cartographic representations of spatiotemporal phenomena in their application to basic choropleth map-based knowledge-extraction tasks. These tasks simply prompt study participants to interpret basic information from the maps and use this information to answer a series of questions. To this end, this study examines map readers' performance and efficiency in completing choropleth map-based knowledge-extraction tasks, using static time-series maps and animated maps that depict homicide patterns in the Chicago metropolitan area as the basis for doing so. Through an in-depth user performance experiment, this study helps to provide insight as to the strengths and weaknesses of static time-series maps and animated maps as the basis for choropleth map-based knowledge-extraction tasks, to determine which of these visualization

methods map users prefer for carrying out these knowledge-extraction tasks, and to better understand which tool inspires more confidence in response accuracy.

## 1.1 Motivation

Historically, maps have been essential to mankind's understanding of geographic features and phenomena. The very nature of studying geographic phenomena requires a visual element, as without one there is no way for researchers to understand the spatial context of their data. As such, few would argue against the notion that cartographic visualization facilitates learning about geography. After all, a single map image can quickly and effectively communicate complex spatial processes, whereas it might be virtually impossible to effectively communicate the same volume of information without these visual tools. Furthermore, according to the picture superiority effect principle (Shepard 1967) complex concepts that are learned by viewing pictures are much more likely to be remembered than those learned by reading written words.

While there is little debate over the utility of maps for conveying spatial processes, there is still some contention over the practicality of modern visualization tools like animated maps and 3D data visualizations for supporting basic knowledge-extraction tasks. While static maps have been used for thousands of years, animated maps became popular much more recently, in the late 1990s (Harrower 2009). As such, cartographic design principles relating to the communicative effectiveness of static maps have been the subject of countless years of research while similar research regarding animated maps began very recently, by comparison (Fabrikant and Harrower 2007). Unfortunately, as suggested by Andrienko et al. (2008), cartographers still know rather little about the effectiveness of interactive graphical data depictions and visualization methods for knowledge-extraction, learning, and understanding dynamic processes.

A common theme underlying modern visualization research challenges is the lack of verified methods for identifying any positive or negative influences on people's map-based knowledge-extraction or decision-making through interactive visualization tools such as animated maps, interactive web-maps, 3D maps, and static time-series maps (MacEachren and Kraak 2001; Fabrikant 2005; Harrower 2007; Andrienko et al. 2008).

While maps are produced and used for many different purposes, they are most often used as tools for communication of complex spatial phenomena to an audience. The communication model, as defined by Board (2011), describes the map as a conduit for the transmission of a message from the mapmaker to the map user. As Board describes, "Cartographic communication emphasizes not only the medium but both the initiator and receiver of the information being communicated. It emphasizes a process rather than a product" (Board 2011, p. 37). As such, cartographers should take account of users' perceptual and cognitive limits, as well as their preferences, when designing maps.

Crime maps are most often produced and used by crime analysts at police departments and various other government agencies and non-governmental organizations. As such, many crime maps are too specialized or too advanced for the general public to easily understand. Efforts need to be made to provide the general public with user-friendly and informative visualization tools so that they too can take part in the conversation about crime in their cities.

Another significant issue in the realm of crime mapping is that the vast majority of crime maps are produced without any attention paid to temporal variances in the distribution of crime incidents, as if time played no part in these events. Particularly with violent crime, time plays an important role in the occurrence and spatial distribution of these events. Assaults, for example, are much more likely to be clustered near bars late at night, or near sports venues during game

time (Ratcliffe 2010). Clearly, space and time interact to create criminal opportunities. As such, efforts need to be made to understand these temporal variances in criminal events, and to communicate any patterns that might exist in the distribution of these events to the general public. Armed with this knowledge, communities may be better prepared to assist in combatting the proliferation of violent crime in their neighborhoods. Access to this information might also help concerned citizens to protect themselves and their families from becoming victimized.

## **1.2 Experiment Design Overview**

The primary aim of the experiment conducted for this study is to measure user performance in carrying out various map-based knowledge-extraction tasks. This experiment is modeled on the work of Kristie Socia, specifically on her 2011 thesis entitled *Small-multiples and Animation: Measuring User Performance with Wildfire Visualization*. Socia measured user performance via task accuracy and response time using static time-series maps (small-multiple map displays) and animated maps that depicted the progression of a wildfire outside of San Diego, California. She also conducted a survey of user preferences between static and animated maps, and of users' confidence in the accuracy of their responses. By comparing users' performance measurements including response accuracy and response time to their reported preferences and their confidence in the accuracy of their responses, Socia found that user preferences in her study did not coincide with the practical application of visualization tools for basic knowledge-extraction tasks.

Like Socia's study, performance measurements for this study are based on panel participants' response accuracy for each knowledge-extraction task and on the average amount of time it takes them to complete each task. By comparing participants' test scores and response times between the animated and static maps, this study helps to provide insight as to the

strengths and weaknesses of each visualization tool as the basis for various knowledge-extraction tasks.

### **1.3 Research Questions**

To reiterate, this study concentrated primarily on measuring user performance (accuracy and efficiency) in carrying out various map-based knowledge-extraction tasks using both static and animated time-series maps of Chicago crime incidents as the basis for doing so. This study also prompted participants to provide a subjective review of each map type based on their personal preferences and their confidence in the accuracy of their responses. Finally, user performance measurements were analyzed for correlations with user-preferences to find potential statistical associations that might suggest patterns and relationships among the different variables. This work was conducted to answer the following research questions:

1. Do animated maps provide heightened potential for accuracy in completing basic knowledge-extraction tasks over static time-series maps, or vice versa?
2. Do animated maps provide heightened potential for efficiency in completing basic knowledge-extraction tasks over static time-series maps, or vice versa?
3. How do user preferences align or not align with measurements of accuracy and efficiency?

## **CHAPTER 2: RELATED WORK**

As described in the previous chapter, this study endeavors to measure user performance, user preferences, and user's confidence in the accuracy of their responses in the contexts of static and animated crime map interpretation, as well as to reveal any relationships that might exist between these different variables. To provide the necessary background for this study, this chapter covers previous work that directly relates to this study in the fields of spatiotemporal data visualization, interactive maps, cartographic experimentation, and crime incident data visualization. This chapter is divided into five sections. Section 2.1 discusses the visualization of spatiotemporal data with static maps. Section 2.2 discusses the visualization of spatiotemporal data with animated maps. Section 2.3 discusses the interactivity of spatiotemporal maps. Section 2.4 discusses the visualization of crime patterns. Finally, Section 2.5 discusses the comparison of static and animated maps,

### **2.1 Visualization of Spatiotemporal Data with Static Maps**

Static maps can depict change over time with temporal snap-shots (Thrower 1959). The two most common types of static spatiotemporal maps are small-multiple map displays (SMMDs) and planimetric overlay map series (Baldwin 2014).

The term SMMD describes a series of small maps arranged next to each other that are used to portray change over time or to convey multiple thematic attributes for comparison to one another (Tufte 1995). As Tufte describes them, small multiple map displays are "illustrations of postage-stamp size [that] are indexed by category or a label, sequenced over time like the frames of a movie, or ordered by a quantitative variable not used in the single image itself" (Tufte 1995, p. 67). It is important to note that small-multiples do not necessarily have to be the size of a

postage stamp. In fact, the term has been used to describe map series of widely variable sizes. The key element of SMMDs is that they are series of static maps that can be used to depict change in the element of interest from one frame to the next.

Planimetric temporal overlay maps are similar to SMMDs in that several map layers representing different time periods or points in time are used together to communicate change over time. The key difference between planimetric overlay maps and SMMDs is that in planimetric overlay, the images are stacked one on top of another, like a layer cake, rather than side by side. Paper planimetric overlay maps are generally designed to be viewed from an oblique angle, which allows the user to view each layer individually, and thereby to differentiate between the different time periods or thematic elements (Baldwin 2014).

Boscoe et al. (1999) demonstrated the utility of static geographic visualizations as platforms for the exploration, analysis, synthesis, and presentation of georeferenced spatiotemporal information. Boscoe and his colleagues also explored the utility of static time-series maps for examining time sequences and displaying changes over time within the confines of a Geographic Information System. Through an empirical comparison of methods, similar to the experiment conducted for this study, Boscoe and his colleagues confirmed the utility of time-series maps in the presentation of spatiotemporal phenomena in multiple fields of research.

## **2.2 Visualization of Spatiotemporal Data with Animated Maps**

The animated map, as defined by Peterson (2014) “is a cartographic statement that occurs in time. Its interpretation is based on the human sensitivity to detect movement or change in a display” (Peterson 2014, p. 1). Change in this context, as defined by DiBiase et al. (1992), is divided into three distinct categories, each emphasizing a distinct type of change: change in



either position or an attribute, change in the location of some phenomenon, or change in the spatial distribution of an attribute. DiBiase et al. also note that animations can be subdivided into three distinct categories: time series (which depict chronological change), re-expressions (which depict attribute changes), and flybys (which depict spatial change). For the sake of brevity, re-expressions and flybys are not discussed in detail here as they are not relevant to this study. Time series, as Slocum et al. (2009) suggest, are by far the most common form of animated map. Time-series animations operate in much the same way as a movie clip. Map images (each frame depicting a specific moment in time or a specific timeframe) are sequenced chronologically and compiled into a video sequence.

As suggested by Slocum et al. (2009), the first animated time-series maps were developed in the 1930s, and by the late 1950s cartographers had acknowledged the potential utility of animated maps for conveying dynamic processes. Thrower (1959), one of the earliest proponents of animated maps, describes animation “by the use of animated cartography we are able to create the impression of continuous change and thereby approach the ideal in historical geography, where phenomena appear as dynamic rather than static entities” (Thrower 1959, p. 10). Despite their acceptance as useful tools for conveying dynamic processes, cartographic animations remained very rare until the early 1990s (Slocum 2009), due to the extremely high monetary costs associated with their production. Technological advancements in the early 1990s allowed the development of much more affordable hardware (Socia 2011). As production costs have come down over time, animated maps have grown increasingly common.

### **2.3 Interactivity of Spatiotemporal Maps**

Cartographic interaction, as defined by Roth (2013), is the dialog between a human and a map. This interaction is the basis for our ability to read maps and interpret their contents. The mode of interaction varies widely between different cartographic tools, particularly between static and animated maps. As Roth points out, the cartographic interaction dialog is often mediated through a computing device, though it applies to analog cartographic visualization as well, as the simple act of interpreting the information presented in a static map is one form of cartographic interaction. While he acknowledges that all maps are interactive to some extent, Roth suggests that digital map mediums typically provide a much wider array of interaction forms for manipulating cartographic representations, thereby allowing more flexible interaction. He also notes that maps with high interactivity are quickly growing in popularity. It can be expected then, as Roth suggests, that making design decisions that account for the different modes of cartographic interaction that are made possible by digital map media will only grow more fundamental to cartographic design as the dominant map prototype shifts from analog to digital.

It is important to note that interactivity has varying levels of intensity (Andrienko et al. 2008). Static maps have the lowest level of interactivity, but not zero. As Andrienko et al. (2008) suggest, static time-series maps, particularly SMMDs, afford mental interactivity in that people can control the viewing order of the static sequence, they can choose to go back to the beginning, and they can study the sequence in any order they choose at their own pace. While animated maps are a bit more externally interactive, in that they feature start, stop, and rewind buttons, they are actually less internally interactive, as the animation must be passively viewed in a pre-defined sequence (Fabrikant 2005; Andrienko et al. 2008).

Animations are transient by nature, often requiring viewers to keep track of multiple symbols and map elements that are changing simultaneously (Socia 2011). As Socia points out, when animations become too complex, it can become very difficult, if not impossible, to keep track of all of the different dynamic elements. Betrancourt and Tversky (2002) suggest that animations may be less effective than static representations because animations are often too complex or too fleeting to be perceived accurately. Harrower (2007) coined the term split attention to describe this effect. Split attention, according to Harrower, is a significant weakness unique to animated maps, particularly when temporal legends or other dynamic elements are employed.

While all animated maps are interactive in the sense that they provide play and stop buttons, not all animations provide the same level of interactivity. Without additional interactive elements such as time-sliders or other interface tools that allow the user to easily navigate the temporal extent of the animation, the user must attempt to remember and integrate changes between scenes, which may overload users' working (short-term) memory (Andrienko et al. 2008). Though static maps are not traditionally viewed as being interactive, SMMDs are interactive in the sense that the viewer can toggle between images at will, viewing the images in any order they choose or spending as much time on each image as they deem necessary (Fabrikant 2005; Andrienko et al. 2008).

## **2.4 Visualization of Crime Patterns**

GIS-based visualization and spatial analysis of crime are commonly used to reveal patterns in the distribution of crime incidents (Nakaya and Yano 2010; Chainey and Ratcliffe 2005). Compstat, for example, is a system that was developed by the NYPD. "Compstat is a goal-oriented,

strategic-management process that uses information technology [including Geographic Information Systems], operational strategy, and managerial accountability to guide police operations” (Vito and Walsh 2004, p. 51). According to Vito and Walsh, Compstat combines accurate and timely intelligence in the form of geocoded criminal incident data, rapid deployment, and effective tactics which allow police departments to react to crime outbreaks very quickly. This rapid response to crime outbreaks is made possible, in large part, by the efficient data collection, visualization, and spatial analysis tools that GIS provides. Upon implementing Compstat, New York City experienced a dramatic reduction in crime rates across the board (Vito and Walsh 2004). As Vito and Walsh describe, this success story led to the implementation of similar programs around the nation. Today, GIS-based crime analysis software packages are used, in some fashion, in virtually every police precinct in the United States.

Significant effort has been devoted to detecting geographic areas with particularly high crime density, commonly referred to as crime hotspots. Several methods have been utilized for visualizing these hotspots, including: pin maps, choropleth maps, shaded grid maps, risk terrain maps, kernel-density estimation maps, Getis-Ord hotspot maps, and inverse-distance weighted interpolation maps. However, the majority of this work has been done from an entirely spatial perspective. This is unfortunate because spatial analysis alone ignores the necessary interaction of space and time to create criminal opportunities (Grubestic and Mack 2008).

Previous crime studies suggest that the spatial distribution of crime incidents varies from one year to the next, between seasons of the year, between weekdays and weekends, and within the span of a single day (Bowers and Johnson 2004). Unfortunately, because the vast majority of crime visualization and analysis has been done from a wholly spatial perspective (Grubestic and

Mack 2008), many of the tools for crime data visualization are of limited use for comparing crime patterns between different time periods.

## **2.5 Comparison between Static and Animated Maps**

Socia (2011), drawing on previous work done by Larkin and Simon (1987) and Andrienko et al. (2008), indicates that in order to conduct a fair comparison between static time-series and animations for a specific purpose, the two visualizations must be informationally equivalent (Andrienko et al. 2008). As Andrienko et al. (2008) describe, informational equivalence (a term coined by Larkin and Simon 1987) describes the notion that any information inferable from one representation must also be inferable from the other, and vice versa, for any fair comparison to be made between them. Like Socia's study, this study was designed with the necessity of informational equivalence in mind. The static and animated maps that provided the basis for this experiment are identical, aside from the fact that one series is static and the other animated.

As suggested by Andrienko et al. (2008), several previous comparative cartographic experiments have been deemed inconclusive because these experiments attempted to determine which cartographic approaches were universally superior for representing dynamic processes. In opposition to these previous works, Andrienko et al. (2008) argue that the question of whether one cartographic method is comprehensively superior to another is not only an ill-conceived question, but an unanswerable one. They go on to suggest that visualization designers should, instead, be interested in determining how interactive visual displays work, determining when they are successful, and why. As such, this study was conducted only to understand the strengths and weaknesses of static time-series maps and animated maps in the specific context of this study, not to determine which tool is generally superior to the other.

Usability engineering, a method for evaluating a product or system's ease of use (Coltekin et al. 2009), can be used to measure the effectiveness of cartographic representations as tools for spatial knowledge-extraction. As Coltekin et al. (2009) explain, "Users are provided with a specific set of tasks based on a particular usage scenario, and in a specific context. Usability performance metrics such as satisfaction, efficiency, and effectiveness (SEE) are employed to assess how easy the product or system is to use. Satisfaction refers to the user's attitude or preferences about the system, efficiency refers to how quickly the tasks are completed, and effectiveness refers to whether or not a task is successfully completed" (Coltekin et al. 2009, p. 6). This study uses the usability engineering principles described above to evaluate the different visualization tools that were the focus of this comparison.

By focusing on the merits of different visualizations methods and cartographic design elements for supporting specific knowledge-extraction tasks in a specific context, it may be possible to gauge the strengths and weaknesses of each cartographic approach in its application to certain tasks. While one visualization technique might be better for a given task or type of tasks in the context of this study, that does not mean it is a universally superior visualization method. There are far too many potential applications for these tools to make a blanket claim of superiority. Acknowledging the importance of the intended application of these tools, this study sought only to investigate the suitability of each tool for supporting specific tasks pertaining to a specific set of maps. While this process did reveal patterns in tool usability and user performance, these patterns were not and should not be assumed to apply universally outside the specific context of this study.

### ***2.5.1 Previous Comparative Studies***

Previous comparative studies of visualization methods for spatiotemporal information (Brunsdon et al. 2006; Grubestic and Mack 2008) have compared fairly user-friendly visualization tools and discussed the importance of user-accessibility, but neglected to carry out the necessary panel review experiments to measure user performance on knowledge-extraction tasks with these tools. This study develops user-friendly static time-series maps and animations and emphasized the usability issue via the user performance experiment described in the next chapter.

Other previous studies (Bekele et al. 2009; Midtbo and Larsen 2005) compared static maps and animated maps via user performance experiments, seemingly to determine which methods were superior for demonstrating dynamic spatiotemporal processes. This study seeks only to understand the strengths and limitations of each tool in the specific context of this experiment. While this study hints at the strengths and weaknesses of each tool for demonstrating dynamic spatiotemporal processes, the patterns in user performance that are revealed by this study are not assumed to apply universally.

#### ***2.5.1.1 Socia's Comparative Study***

The experiment design for this study is based on a 2011 University of Michigan geography thesis by Kristie Marie Socia entitled *Small Multiples and Animation: Measuring User Performance with Wildfire Animation*. Socia measured user performance via task accuracy and response-time using static time-series SMMDs and animated maps that depicted the progression of a wildfire outside of San Diego, California, over time. Socia also conducted a survey of user preferences and of users' confidence in the accuracy of their responses.

In reviewing the results of her experiment, Socia found that small-multiples afforded study participants statistically significantly higher response accuracy scores (85.4% for small-multiples and 80.4% for animation). She also found that small-multiples provided users with a statistically significant advantage in terms of efficiency. Socia's study participants were able to complete the assigned tasks in an average time of 21.8 seconds using the small-multiple series, while it took them 26.1 seconds, on average, to complete each task using the animated maps. Socia also found that her study participants tended to be more confident in their responses when using the small-multiple series than they were when using the animated maps as the basis for their responses.

By comparing users' response accuracy and response time to the subjective feedback they provided on each of the visualization tools, Socia also found evidence to suggest that, at least in the context of her experiment, user preferences generally did not coincide with the practical application of the visualization tools for the basic knowledge-extraction tasks. Despite scoring better in both accuracy and efficiency using the static small-multiple series, the vast majority (72%) of Socia's study participants preferred the animated version.

### ***2.5.1.2 Adjustments to Socia's Methods***

In her concluding discussion, Socia (2011) notes several problems with the design of her study. Chief among these was her decision to include skip-to-time-stamp buttons as the primary navigation tool for her animated maps. Several of her study participants reported having difficulty navigating the animations as necessary to complete the tasks she assigned. Study participants attributed this difficulty to the skip-to-time-stamp navigation interface. This study endeavors to further Socia's work by using a time-slider (a scroll bar with which users are able to



seamlessly navigate the temporal extent of the animation) rather than skip-to-time-stamp buttons. This change both helps to simplify the user-interface and allows users to easily navigate the entire temporal extent of the animation, whereas Socia's participants were only permitted to skip back and forth between certain timestamps. This was particularly problematic in Socia's study because some of the test questions asked for information that was located in between two skip-to-time-stamp-markers. Several of Socia's study participants indicated having difficulty with the skip-to-time-stamp interface, particularly for the test questions that required them to mentally interpolate between scenes.

Further improvements are being made by ensuring that study participants are properly briefed on cartographic design and task format before beginning the test, and by ensuring that all tasks and/or questions are posed in very clear language. Like the decision to use a time-slider rather than skip-to-time-stamp buttons, these improvements are being made based on feedback from participants in Socia's study.

This study also furthers Socia's research by applying her general methodology to entirely different subject matter. By replicating and improving upon her methods and applying them to choropleth homicide incident visualization, rather than to raster-based wildfire visualization, this study helps to determine whether the results of Socia's study are applicable outside the specific context of her experiment.

## CHAPTER 3: METHODOLOGY

This study endeavored to provide an objective comparison of static and animated cartographic representations of spatiotemporal phenomena in their application to basic choropleth map-based knowledge-extraction tasks. To this end, this study compared static time-series maps with animated maps in their application to the visualization of homicide patterns in the Chicago metropolitan area. Through an in-depth user performance experiment, this study helped to provide insight as to the strengths and weaknesses of static time-series maps and animated maps as the basis for choropleth map-based knowledge-extraction tasks. It also helped to determine which of these cartographic tools map users prefer for carrying out these tasks, and to better understand which tools inspired the most confidence in response accuracy.

This chapter discusses the methodology for the user performance experiment itself and the maps developed for the experiment. The general experiment design that was used in this study was based loosely on the experiment design that Socia developed for her study, even though the maps that provided the basis for the two experiments were very different. As described in the previous chapter, this study aims to further Socia's work by attempting to adjust for some of the issues she encountered while conducting her experiment.

Chapter 3 is composed of four sections. Section 3.1 provides details on the data that provided the basis for the maps that were used for the experiment. Section 3.2 discusses the cartographic design for the static and animated map series. Section 3.3 discusses the user performance experiment design. Finally, Section 3.4 discusses the methodology for analyzing the results of the experiment.

### 3.1 Mapped Data

Crime hot spot maps provided the basis for the aforementioned comparison. A six-step time-series of static hot spot maps and an animated version of the same maps were created to visualize how the spatial distribution of homicide incidents varied from one time period to the next. While the visualization methods that provided the basis for this study could be applied to any point-incident-based datasets that provide specific point locations and specific times for each data point, this study focused on homicide incidents in Chicago, Illinois. Chicago was selected as the basis for this study for three reasons. First, Chicago has one of the highest homicide rates in the entire United States (Huffington Post 2013). Second, the city of Chicago provides a very detailed, accurate, and up-to-date crime dataset that is accessible via the City of Chicago Data Portal (<https://data.cityofchicago.org/>). Finally, among the several datasets for different cities studied (Atlanta, Chicago, Denver, San Diego, and Seattle), Chicago stood out because it contains records on the time of day at which each incident was reported, which have been thoroughly checked by the Chicago GIS team for consistency and accuracy (City of Chicago Department of Innovation and Technology 2014). Many of the other datasets that were considered contained incomplete records or limited metadata with which to verify the suitability of the data for this study.

All data used for this study are available through the City of Chicago Data Portal (<https://data.cityofchicago.org/>). This includes a very detailed National Incident-based Reporting System (NIBRS) crime dataset going back to 2001, containing point incident data for individual crime events in the Chicago metropolitan area, as well as police precinct and police beat shapefiles which helped to provide context for the crime data. The dataset chosen for this study is one of very few NIBRS datasets that has accurate time of day data (in addition to accurate

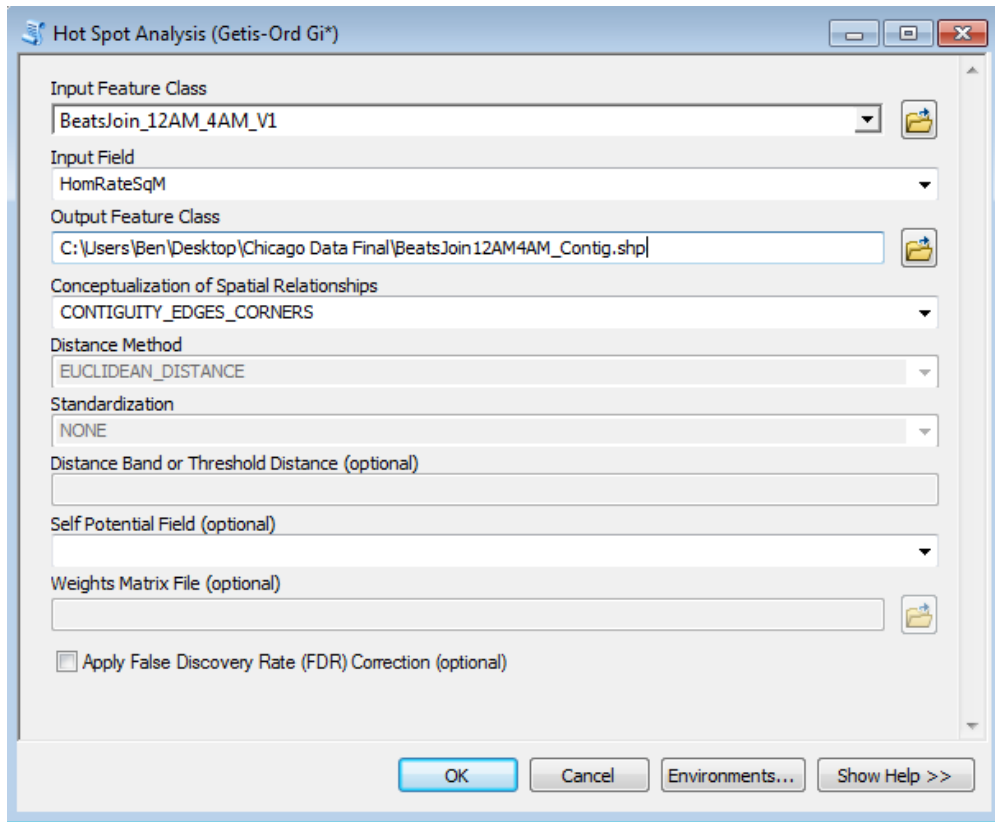
spatial data). Most local datasets have date information only. Importantly, time values in the dataset are based as closely as possible on the times at which the incidents actually occurred, rather than the time at which law enforcement officers arrived on scene (City of Chicago Department of Innovation and Technology 2014). The dataset also has detailed metadata that was used to verify its suitability for this study, something many other NIBRS datasets are lacking. This metadata file provides details on the method of data acquisition, expanded definitions of attribute data and data types, and estimations of spatial and temporal accuracy for the incident data.

### ***3.1.1 Data Aggregation Methods***

Since one day's worth of incident data does not provide a sufficient number of points for the effective visualization of spatiotemporal trends in these events, all homicide events that occurred over the course of several years (2009-2013) were visualized together as if they occurred in the same 24-hour period. The homicide incident points were divided into six separate datasets based on the time of day at which they took place. The six time periods were 12:00 AM to 3:59 AM, 4:00 AM to 7:59 AM, 8:00 AM to 11:59 AM, 12:00 PM to 3:59 PM, 4:00 PM to 7:59 PM, and 8:00 PM to 11:59 PM. Once the data points were divided by time period, the incident data for each time period were spatially joined with a polygon shapefile containing all of the police beats in the city of Chicago. Area measurements for the police beats were then utilized to produce a homicide rate (number of incidents per square mile) for each police beat and each time period. These homicide rate values for the police beat polygons provided the basis for the Getis-Ord  $G_i^*$  hotspot analysis calculations described in the next section.

### **3.2 Cartographic Design**

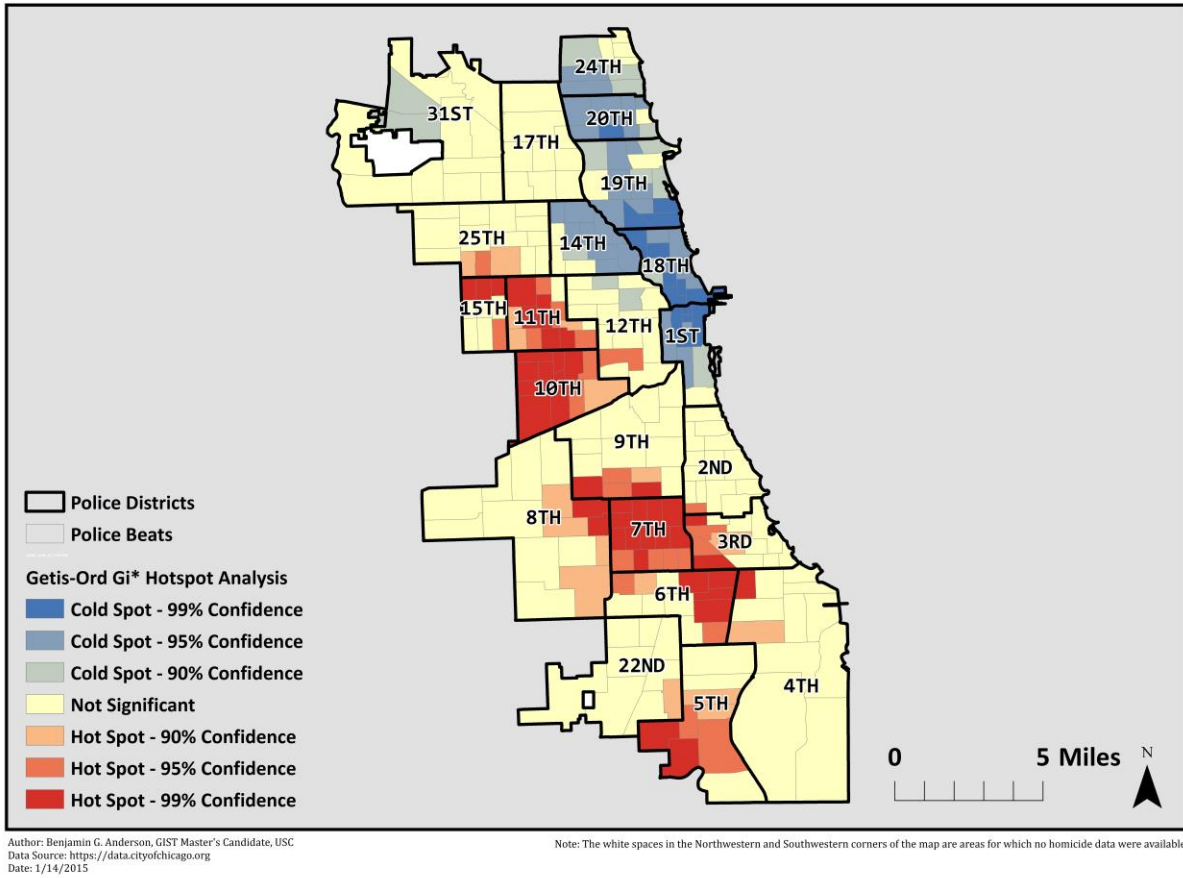
Getis-Ord  $G_i^*$  hot spot maps provided the basis for the comparison between static and animated visualizations in this study. To construct the hot spot maps, the Hot Spot Analysis (Getis-Ord  $G_i^*$ ) tool available in ArcGIS 10.2 (Figure 3.1) was used. Given a set of weighted features (the police beat polygons, weighted according to the homicide rate for each beat, in this case) the Getis-Ord  $G_i^*$  statistic identifies statistically significant clusters of hot spots and cold spots (clusters of police beats with particularly high or low homicide rates, in this case). In addition, the tool required the researcher to select an option for conceptualizing the spatial relationships between the features they wish to analyze. Since the incident data for this study were aggregated to police beat polygons and converted into homicide rates (number incidents per square mile calculated individually for each police beat), spatial relationships for this analysis were conceptualized using the Contiguity\_Edges\_Corners option. This setting allows researchers to include all polygons that share an edge or corner with the target polygon in the computations for that polygon. This setting was chosen because the police beats varied considerably in size. If one of the distance-based conceptualization options had been selected, the output might have been distorted due to the highly variable size of the police beat polygons.



**Figure 3.1: Hot Spot Analysis Tool Settings**

Once the Getis-Ord  $G_i^*$  statistic has identified the hot spots and cold spots (as well as areas that are neither hot or cold spots, referred to as not significant by the Getis-Ord  $G_i^*$  output), each police beat polygon is assigned a color, based on its status as a hot spot, cold spot, or neither. Hot spots and cold spots each have three confidence levels: 99%, 95%, and 90%. Each of these levels is assigned a different shade of either red or blue. The 99% confidence hot spots are marked in dark red, the 95% confidence hot spots in dark orange, and the 90% confidence hot spots in light orange. The 99% confidence cold spots are marked in dark blue; the 95% confidence cold spots are marked in light blue, and the 90% confidence cold spots are marked in light blue/green. Those areas determined to be insignificant, neither hot nor cold spots, are marked in pale yellow. Figure 3.2 is an example of one frame of a six map time-series.

**Chicago Homicide Hot Spots, 2009-2013**  
**12:00 AM to 3:59 AM**



**Figure 3.2: Chicago Homicide Hot Spots, 2009-2013, 12:00 AM to 3:59 AM**

### 3.3 User Performance Experiment Design

In order to measure the effectiveness of static time-series and animated maps as tools for the communication of complex spatiotemporal information to an audience, a panel of study participants was asked to complete a series of choropleth map-based knowledge-extraction tasks, using each of the aforementioned visualization tools in turn as the basis for doing so. Based on task accuracy and response time, it was possible to gain valuable insights as to which tools better facilitated successful and timely completion of the assigned tasks.

The user performance experiment outlined above provided the basis for the evaluation of each visualization technique as a tool for supporting users' spatial knowledge-extraction. The results of this evaluation were not intended to prove any cartographic methods to be universally superior or inferior in their application to the proposed tasks. This study was intended only to gain some general insights into the strengths and weaknesses of each technique for certain types of map-based knowledge-extraction tasks in the specific context of time-series crime maps.

The user performance experiment consisted of three distinct segments: a pre-test, map test, and post-test. The pre-test consisted of a brief tutorial on interpreting time-series hot spot maps, a description of what study participants could expect to encounter as they completed the exercise, and a short questionnaire that collected basic background information on study participants (age, sex, and level of education). The map interpretation tutorial gave participants an opportunity to learn how to interpret and interact with the maps before accuracy and efficiency measurements began. The map test measured study participants' accuracy, response time, and confidence in responses using each of the cartographic tools as the basis for their answers. The post-test consisted of a series of questions on user-preferences between the static and animated maps as well as a series of qualitative open-ended questions aimed at gaining insights on users' map-reading strategies and whatever other feedback they offered.

The user-preference assessment in the post-test was conducted to determine which of the proposed cartographic tools were best liked by review participants, which tools inspired the most confidence in responses, and to understand which tools participants felt to be most effective for the assigned tasks. These user-preference questions were structured using a forced-choice Likert-scale. For each of a series of statements, which were designed to gauge user-preferences between the static and animated maps, participants were asked to indicate their level of agreement



(strongly disagree, disagree, agree, or strongly agree). The neutral (neither agree nor disagree) option was omitted from the possible responses to prevent central tendency bias. Acquiescence bias was avoided by including an equal number of positive and negative statements in the qualitative review questions. Two statements favored the static maps and two statements favored animated maps. Participants were asked to indicate their level of agreement with each statement.

### ***3.3.1 Map Test Format and Performance Measurements***

The panel review process for this study focused primarily on user performance in carrying out map-based knowledge-extraction tasks through a series of static and animated choropleth maps. More specifically, this study was designed to test user's ability to visually discern temporal change in choropleth maps of homicide rates in Chicago analyzed at the scale of police precincts (while the maps are designed to show hotspots, the hotspots are displayed similarly to a choropleth map, so the interpretability of these maps is similar to choropleth map interpretation). The test questions asked participants to both discern the overall distribution of homicide hot spots and cold spots and to identify particular time periods where certain areas experienced particularly high or low homicide rates.

Performance measurements were based on each panel participant's accuracy scores for each set of map-based knowledge-extraction tasks and on the average amount of time it took them to complete each task. Each participant answered three questions using the static maps as the basis for their answers and three additional questions using the animated maps as the basis for their answers. Tight experimental control was maintained by ensuring that the questions for the static version of each cartographic technique corresponded very closely with the questions

for the animated version, without permitting the review participants to answer questions from memory based on previous exposure. This helped to ensure that neither map type was inadvertently put at a disadvantage. The map test consisted of six multiple-choice questions, three for the static map series and three for the animation. The testing sequence was randomized as much as possible while still ensuring that the testing sequence alternated between static and animated questions in turn. See Appendix B for a detailed account of the survey interface, including the test questions.

### ***3.3.2 User-Interface Design***

The user performance experiment was administered using Qualtrics, a web-based survey design and distribution platform. This platform was chosen because it provided an array of advanced tools for formatting questions, embedding images and videos at specific display sizes, and tracking response times. The user-interface displayed the static maps and animations along with test questions and radio buttons for the multiple choice responses that participants were asked to provide. Time measurements were recorded from the time at which participants submitted the previous page to the time they submitted the current page. Since each test question was on a separate page of the survey, this resulted in completion time values for each participant and each map test question.

#### ***3.3.2.1 Static Time-series Map Display Format***

The static time-series maps were presented full-size, the same size as the animated map video clip, approximately 6.5 inches tall by 8.5 inches wide on a single page. HTML was utilized to ensure that the images were displayed the exact same size on all screens, irrespective of potential

variances in browser and screen size choices among study participants. A scroll-bar was employed to allow study participants to easily scroll from one full-size map image/time-period to the next. It also allowed for the map images to be displayed at a much larger scale, which helped to avoid issues with label legibility.

### ***3.3.2.2 Animated Time-series Map Display Format***

The animated time-series were presented using a video clip embedded into the survey page, much in the same way as the static version. Like the static maps, the video image was approximately 6.5 inches tall by 8.5 inches wide. The imbedded video interface provided play, pause and start over buttons, as well as a time-slider with which users were able to seamlessly navigate the temporal extent of the series as necessary to complete the assigned tasks.

The time-slider helped to avoid problems with split-attention, one of the cognitive issues commonly associated with animated maps discussed in Chapter 2, by utilizing the temporal labeling of the animation itself as the temporal legend for navigating between time periods. Additionally, Adobe Premiere's cross-dissolve feature, which gradually transitions from one frame to the next, was employed to help prevent participants from being surprised by the transition between animation frames, which can negatively influence comprehension and memorization. See figure 3.4 on the next page for an image of the animated time-series map display interface.

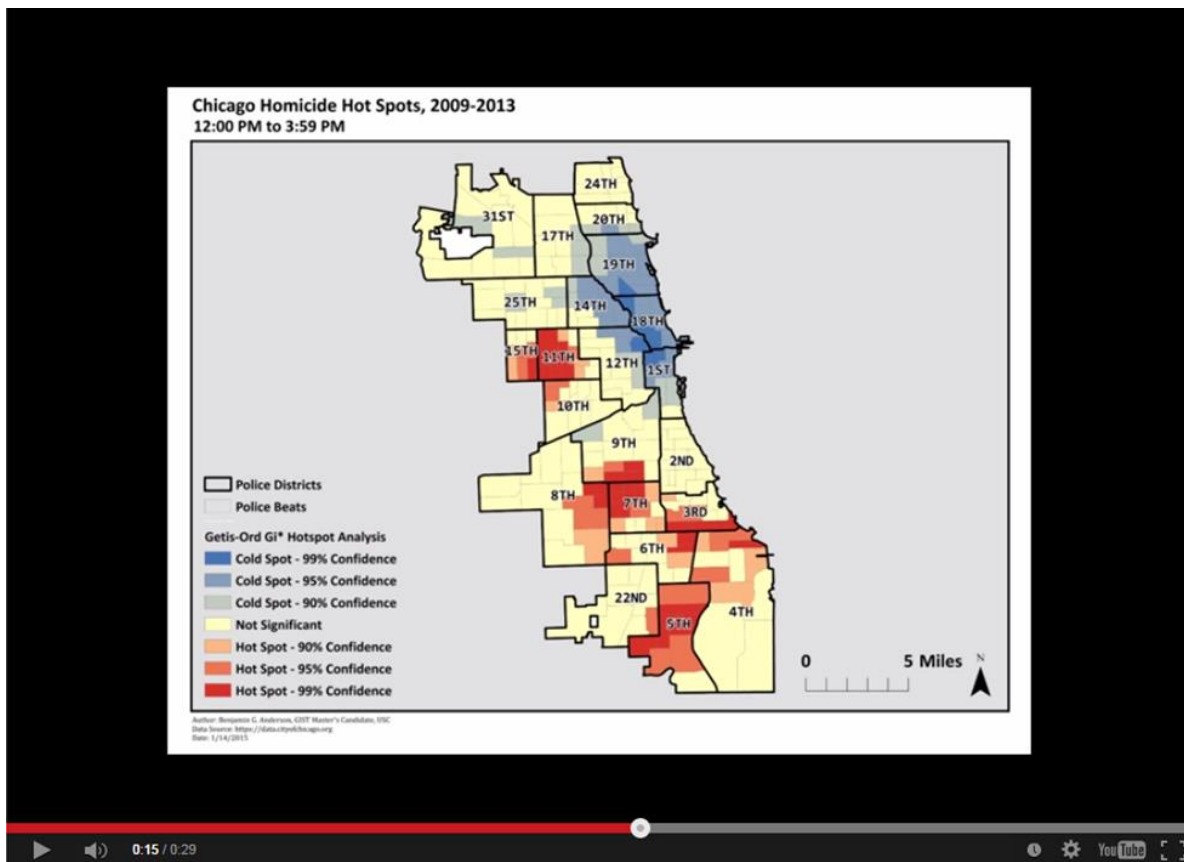


Figure 3.3 Animated Time-series Map Display Format. See Appendix B for larger image.

### 3.3.3 Recruitment of Participants

Participants were recruited via Qualtrics Panels. This service allows the researcher to define a set of criteria for selecting participants, as well as criteria for determining which participants and their responses should be included in the final analysis, based on completion time and various other measures of response quality. In total, 1,300 participants were recruited. To be included in the study, respondents simply had to be eighteen or older and have some level of college education (either in school or graduated). Due to screen-size requirements, participants were excluded from the final analysis for accessing the survey on tablets or smartphones. A quota system was used to ensure that respondents were 50 % male and 50% female. Attention questions were dispersed throughout the survey to ensure that participants were paying attention

and trying to answer the test questions correctly. For example, one question simply requested that respondents select the strongly disagree option. Any participants who failed to do so were disqualified. Additionally, any submissions completed in less than 30% of the mean completion time were excluded from the final analysis. Since participants were compensated for participation in the study, these measures helped to weed-out any participants who might have tried to game the system by rushing through the survey without actually trying to answer the questions correctly. Out of the 1,300 recruited participants, fifty met all of the conditions for inclusion in the final analysis. Once the fifty valid, in-quota, completes were collected, data collection ceased.

#### ***3.3.4 Testing Procedures***

Testing for this study was conducted via internet distribution, as described in the previous section. Upon clicking the link to begin the survey, study participants were asked to begin the pre-test by following the on-screen instructions.

The pre-test portion of the web-form provided participants with a brief tutorial on the cartographic techniques that provided the basis for the map exercise they were asked to complete. The pre-test also provided a brief description of the test format and collected some basic demographic information from participants before they began the map test. No informed-consent documentation was included in the pre-test, as this study was exempted from Institutional Review Board oversight.

Following the pre-test, participants were informed that the map test was about to begin, and that all map questions would be timed to measure task performance efficiency. Each page of the map test consisted of a static time-series or animation followed by a content question and a follow-up question regarding participants' confidence in the accuracy of their response. The

web-form automatically recorded the time it took for participants to move from one page to the next, as well as the timing of each click on each page.

Upon completion of the map test, participants were asked to answer a series of multiple-choice Likert-scale questions that were designed to gauge user-preferences. Participants were also given the opportunity to provide open-ended feedback or to describe their strategies for completing the exercises if they desired. This concluded post-test and the exercise.

### **3.4 Methodology for Analyzing the Results of the Experiment**

By comparing participants' aggregate test scores and response times between the animated and static map series, it was possible to examine the strengths and weaknesses of static and animated maps in their application to choropleth map-based knowledge-extraction tasks. Additionally, statistical analysis was employed to measure statistical significance and to better understand patterns and relationships in the test results and feedback provided by study participants. The first two research questions for this study are restated below as null and alternative hypotheses to help provide context for the subsequent statistical analysis discussion. Following these is another pair of null and alternative hypotheses for static and animated confidence scores. While the confidence data were not used directly to answer the research questions for this study, they were still included in the statistical analysis.

There is no significant difference between static and animated mean accuracy scores:

$$\mathbf{H_0: \mu_{static\ accuracy\ score} = \mu_{animated\ accuracy\ score}}$$

There is a significant difference between static and animated mean accuracy scores:

$$\mathbf{H_1: \mu_{static\ accuracy\ score} \neq \mu_{animated\ accuracy\ score}}$$

There is no significant difference between static and animated mean completion times:

**H<sub>0</sub>:**  $\mu_{\text{static completion time}} = \mu_{\text{animated completion time}}$

There is a significant difference between static and animated mean completion times:

**H<sub>1</sub>:**  $\mu_{\text{static completion time}} \neq \mu_{\text{animated completion time}}$

There is no significant difference between static and animated mean confidence scores:

**H<sub>0</sub>:**  $\mu_{\text{static confidence score}} = \mu_{\text{animated confidence score}}$

There is a significant difference between static and animated mean confidence scores:

**H<sub>1</sub>:**  $\mu_{\text{static confidence score}} \neq \mu_{\text{animated confidence score}}$

### ***3.4.1 Statistical Analysis Methodology***

The statistical analysis conducted for this study consisted of basic independent samples (two-tailed) t-tests and Pearson's product-moment correlation coefficient calculations. The independent samples t-test allows researchers to examine if the means of two different data sets are significantly different from each other. The t-tests were utilized to compare the mean test scores, mean confidence levels, and mean completion times between the static and animated map series, and to measure the statistical significance of these findings. These comparisons provided the information necessary to determine whether the static or animated condition better facilitated accurate and efficient retrieval of the information needed to complete the assigned tasks. They also helped to determine whether one visualization method inspired more confidence in response accuracy than the other. The Pearson's product-moment correlation coefficient allows the

calculation of correlation by dividing the covariance of the two variables by the product of their standard deviations. This results in a correlation coefficient between -1 and 1, with -1 indicating total negative correlation, 0 indicating no correlation, and 1 indicating total positive correlation. The Pearson's product-moment correlation coefficient calculations were conducted to determine whether or not user-preferences are correlated with accuracy or efficiency measurements, and to measure the statistical significance of these findings.



## CHAPTER 4: RESULTS AND ANALYSIS

This chapter summarizes the results of the user performance experiment and discusses the results of the statistical analyses that were conducted on the survey data. The findings of the experiment, which are the focus of this chapter, are divided into six sections. Section 4.1 discusses the experimental population. Section 4.2 discusses overall performance metrics. Section 4.3 discusses task accuracy. Section 4.4 discusses participants' confidence in their responses. Section 4.5 discusses completion-time. Section 4.6 discusses user-preferences. Finally, Section 4.7 discusses the relevant statistical correlations between these different variables.

### 4.1 Experimental Population

Approximately thirteen-hundred individuals agreed to take part in the study. The vast majority of these individuals were screened out because they didn't meet all of the requirements for inclusion in the final analysis. Many of the excluded responses were screened out for not meeting the college education requirement. Some were screened out for inadequate screen size or because their Flash Player version was out of date. Several more were screened out because they failed to enter the correct answers for the preliminary attention questions as well. Fifty-eight participants completed the entire survey. Of these, eight failed to pass the second series of attention-filters (questions designed to test whether participants are paying attention) and were disqualified from the final analysis, leaving fifty complete responses, twenty-five male and twenty-five female. All fifty respondents were college-educated and between the ages of eighteen and sixty-eight. See Tables 4.1 and 4.2, respectively, for more details on the age and level of education of study participants. Each participant answered three static map questions and three animated map questions. They also answered six confidence questions. These questions were asked after each

of the six static or animated map questions. They also answered a series of questions that was intended to measure user-preferences between the static and animated map series.

**Table 4.1 Age Distribution**

| Age        | Response | %    |
|------------|----------|------|
| 18-25      | 1        | 2%   |
| 26-34      | 9        | 18%  |
| 35-54      | 21       | 42%  |
| 55-64      | 14       | 28%  |
| 65 or over | 5        | 10%  |
| Under 18   | 0        | 0%   |
| Total      | 50       | 100% |

**Table 4.2 Education**

| Level of Education           | Response | %    |
|------------------------------|----------|------|
| Less than High School        | 0        | 0%   |
| High School / GED            | 0        | 0%   |
| Some College                 | 11       | 22%  |
| 2-year College Degree        | 4        | 8%   |
| 4-year College Degree        | 21       | 42%  |
| Master's Degree              | 9        | 18%  |
| Doctoral Degree              | 1        | 2%   |
| Professional Degree (JD, MD) | 4        | 8%   |
| Total                        | 50       | 100% |

## 4.2 Overall Performance Metrics

Table 4.3 shows the descriptive statistics for the accuracy scores, confidence scores, completion times and user-preferences. These scores and completion time values are grouped according to map type (static and animated). The accuracy scores represent how participants answered the static and animated map questions correctly. If a participant answered a question correctly, he or she received 1 point. Then, his or her total score was calculated by map type. The total score was averaged for further analysis. The confidence measurements were handled in essentially the same way. Confidence scores, however, were recorded out of three points possible. Each of the confidence questions has three scales of confidence: not confident, somewhat confident, and very confident. These scales were assigned the values of 1, 2, and 3 respectively to quantify users' confidence. Then, each user's scores were added together by map type and averaged. For the timing values, the number of seconds it took each participant to answer each test question was added together by map type. Then, the totaled timing values were averaged. Users' preferences between the static and animated maps were measured through the Likert-scale questions. For each of a series of statements, participants were asked to indicate their level of agreement. The six scales of agreement were very strongly disagree, strongly disagree, disagree, agree, strongly agree and very strongly agree. These scales were assigned values of 1 to 6, respectively. The three preference scores for each map type were then averaged together. This quantification enabled the author to calculate user-preference scores for each map type.

The static map series afforded users slightly higher accuracy scores, confidence scores, and preference scores while greatly reducing completion time. No substantial differences were discovered in test scores, confidence scores, or completion times between male and female

respondents. The age distribution of the sample population was not sufficiently diverse to warrant an analysis of differences in performance metrics between age groups.

The mean accuracy score for the static map series was .826 out of 1 (82.6%), while the mean accuracy score for the animated map series was .733 out of 1 (73.3%). The mean confidence score for the static map series was 2.59 out of 3, while the mean confidence score for the animated series was 2.48 out of 3. The mean completion time for the static map series was 50.64 seconds, while the mean completion time for the animated map series was 67.94 seconds. Finally, the mean preference score for the static map series was 4.05 out of 6, while the mean preference score for the animated series was 3.34 out of 6. The results of the descriptive and inferential statistics are discussed in greater detail in the following sections.

**Table 4.3: Accuracy, Confidence, Completion Time, and User-preferences**

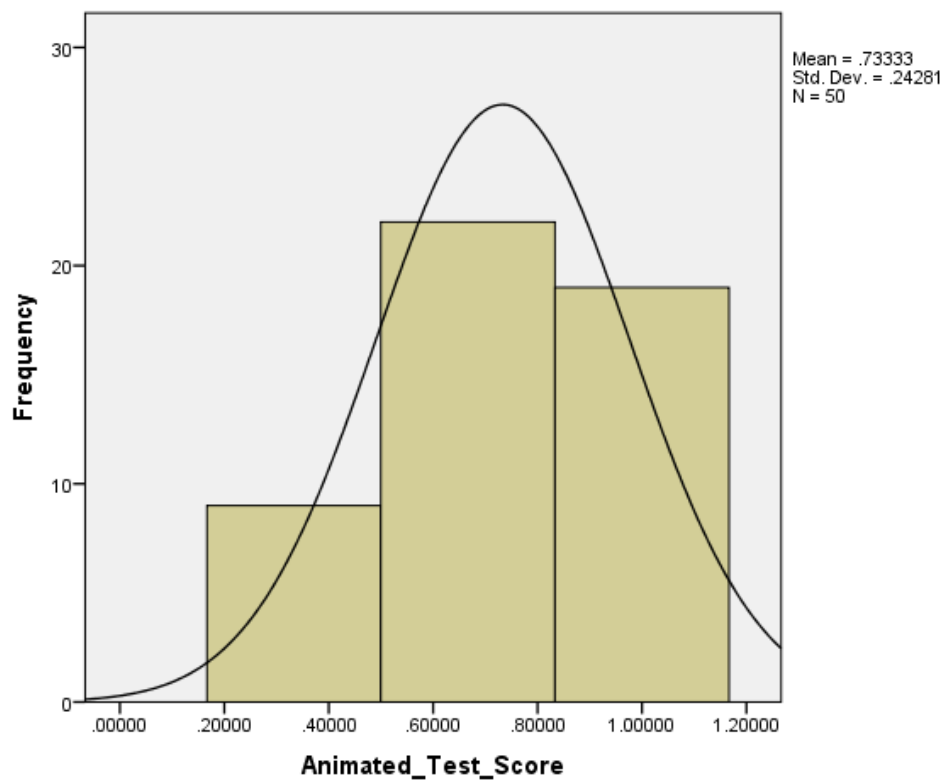
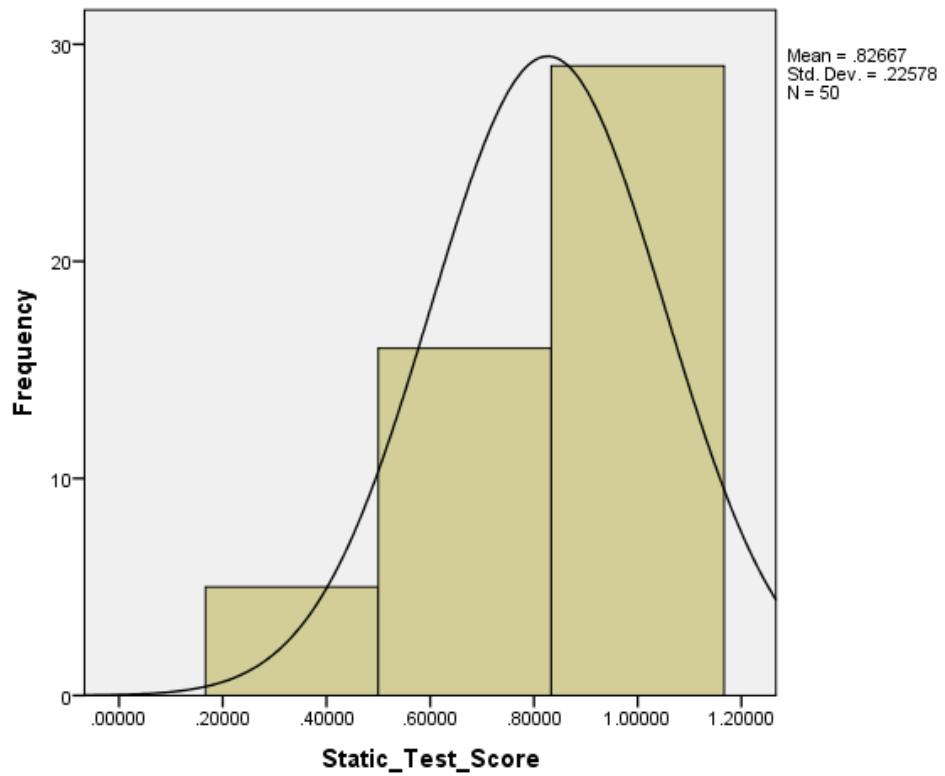
|                           | N  | Minimum | Maximum | Mean   | Std. Deviation |
|---------------------------|----|---------|---------|--------|----------------|
| Static Accuracy Score     | 50 | 0.333   | 1.000   | 0.827  | .226           |
| Animated Accuracy Score   | 50 | 0.333   | 1.000   | 0.733  | .243           |
| Static Confidence Score   | 50 | 2.000   | 3.000   | 2.593  | .352           |
| Animated Confidence Score | 50 | 1.666   | 3.000   | 2.480  | .331           |
| Static Completion Time    | 50 | 21.049  | 94.559  | 50.636 | 16.818         |
| Animated Completion Time  | 50 | 16.666  | 158.018 | 67.944 | 30.339         |
| Static Preference Score   | 50 | 1.333   | 6.000   | 4.053  | 2.550          |
| Animated Preference Score | 50 | 1.000   | 5.667   | 3.347  | 1.017          |

### 4.3 Task Accuracy

The mean accuracy score for the static maps was 82.6% while the mean accuracy score for the animated maps was 73.3%. The static scores are somewhat more skewed than the animated scores, while the animated scores are somewhat more kurtotic. Table 4.4 summarizes the relevant descriptive statistics for the static and animated map test scores. Figure 4.1 contains histograms of the static and animated map test scores.

**Table 4.4 Test Score Descriptive Statistics**

|                        | Static Accuracy Score | Animated Accuracy Score |
|------------------------|-----------------------|-------------------------|
| Mean                   | 0.827                 | 0.733                   |
| Median                 | 1.000                 | 0.667                   |
| Mode                   | 1.000                 | 0.667                   |
| Std. Deviation         | .226                  | .243                    |
| Skewness               | -.951                 | -.330                   |
| Std. Error of Skewness | .337                  | .337                    |
| Kurtosis               | -.238                 | -1.023                  |
| Std. Error of Kurtosis | .662                  | .662                    |
| Minimum                | 0.333                 | 0.333                   |
| Maximum                | 1.000                 | 1.000                   |



**Figure 4.1 Static and Animated Accuracy Score Histograms**

To test the hypothesis that the static and animated maps afforded users statistically significantly different mean accuracy scores, an independent samples (two-tailed) t-test was performed. As can be seen in Table 4.4 the animated and static map series test scores are sufficiently normally distributed for conducting a t-test (i.e., skew < |2.0| and kurtosis < |9.0|; Bayer, Buhner, Danay, Schmider, and Ziegler 2010). Additionally, the assumption of homogeneity of variances was tested and satisfied via Levene's F-test  $F(98) = .005, p = .945$ . The t-test was associated with a statistically significant effect,  $t(98) = 1.990, p = .049$ . Thus, study participants achieved higher test scores with the static map series than with the animated map series, with a mean difference of approximately 9.3%.

#### 4.4 Confidence

The mean confidence score for the static maps was 2.59 out of 3 while the mean confidence score for the animated maps was 2.48 out of 3. The static confidence scores are somewhat less skewed and more kurtotic than the animated confidence scores. Table 4.5 summarizes the relevant descriptive statistics for the static and animated confidence scores. Figure 4.2 contains histograms of the static and animated confidence scores.

**Table 4.5: Confidence Score Descriptive Statistics**

|                        | Static Confidence | Animated Confidence |
|------------------------|-------------------|---------------------|
| Mean                   | 2.593             | 2.480               |
| Median                 | 2.667             | 2.667               |
| Mode                   | 3.000             | 2.667               |
| Std. Deviation         | .352              | .331                |
| Skewness               | -.297             | -.478               |
| Std. Error of Skewness | .337              | .337                |
| Kurtosis               | -1.137            | .005                |
| Std. Error of Kurtosis | .662              | .662                |



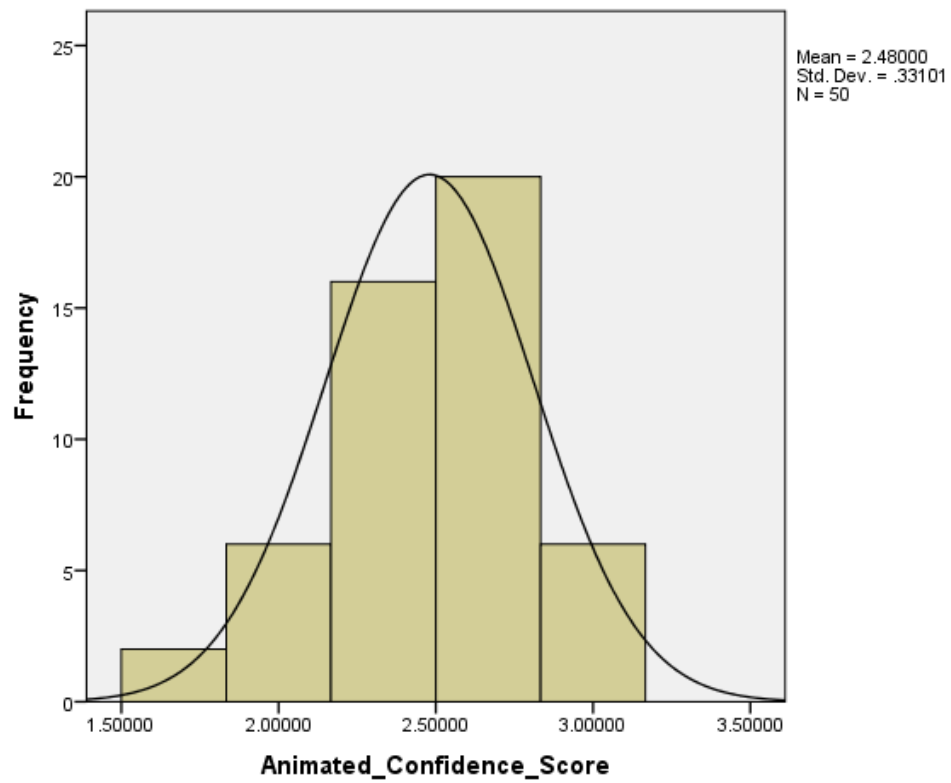
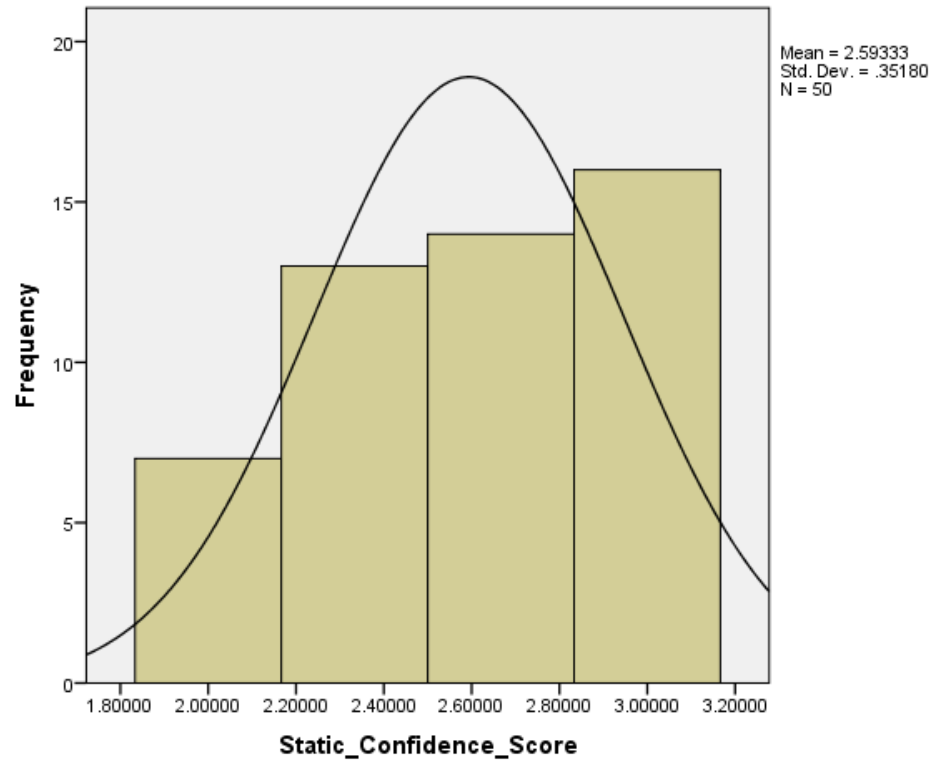


Figure 4.2 Static and Animated Confidence Score Histograms

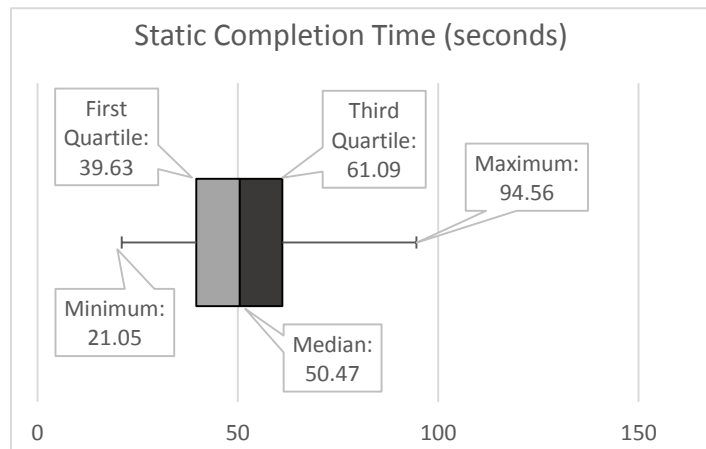
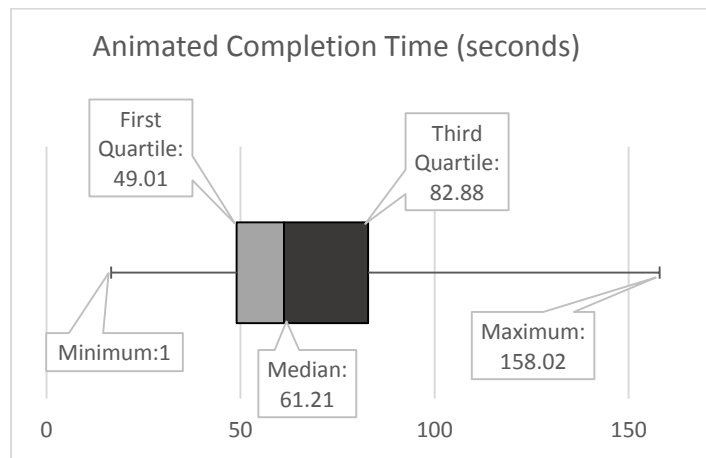
To test the hypothesis that the static and animated maps afforded users statistically significantly different mean confidence scores, an independent samples (two-tailed) t-test was performed. As can be seen in Table 4.5, the animated and static map series confidence scores are sufficiently normal for conducting a t-test (i.e. skew < |2.0| and kurtosis < |9.0|). Additionally, the assumption of homogeneity of variances was tested via Levene's F-test and was not confirmed, thus equal variances were not assumed,  $F(98) = .578, p = .049$ . The t-test was not associated with a statistically significant effect  $t(98) = 1.659, p = .100$ . Thus, the confidence scores for the static map series were not significantly different from the mean confidence scores for the animated map series.

#### **4.5 Completion Time**

The static maps were associated with a mean completion time of 50.64 seconds. By comparison, the animated maps were associated with a numerically higher mean completion time of 67.94 seconds. The mean static completion times were slightly less skewed and somewhat less kurtotic than the mean animated completion times. Table 4.6 summarizes the relevant descriptive statistics for the static and animated completion times. Figure 4.3 contains box-plots of static and animated completion times, respectively. Figure 4.4 contains histograms of static and animated completion times, respectively.

**Table 4.6 Static and Animated Completion Time**

|                        | Static Completion Time | Animated Completion Time |
|------------------------|------------------------|--------------------------|
| Mean                   | 50.636                 | 67.945                   |
| Median                 | 50.467                 | 61.212                   |
| Mode                   | 21.050                 | 16.667                   |
| Std. Deviation         | 16.818                 | 30.339                   |
| Skewness               | .517                   | .831                     |
| Std. Error of Skewness | .337                   | .337                     |
| Kurtosis               | -.048                  | .532                     |
| Std. Error of Kurtosis | .662                   | .662                     |
| Minimum                | 21.050                 | 16.667                   |
| Maximum                | 94.559                 | 158.018                  |



**Figure 4. 3 Static and Animated Completion Time Box-plots**

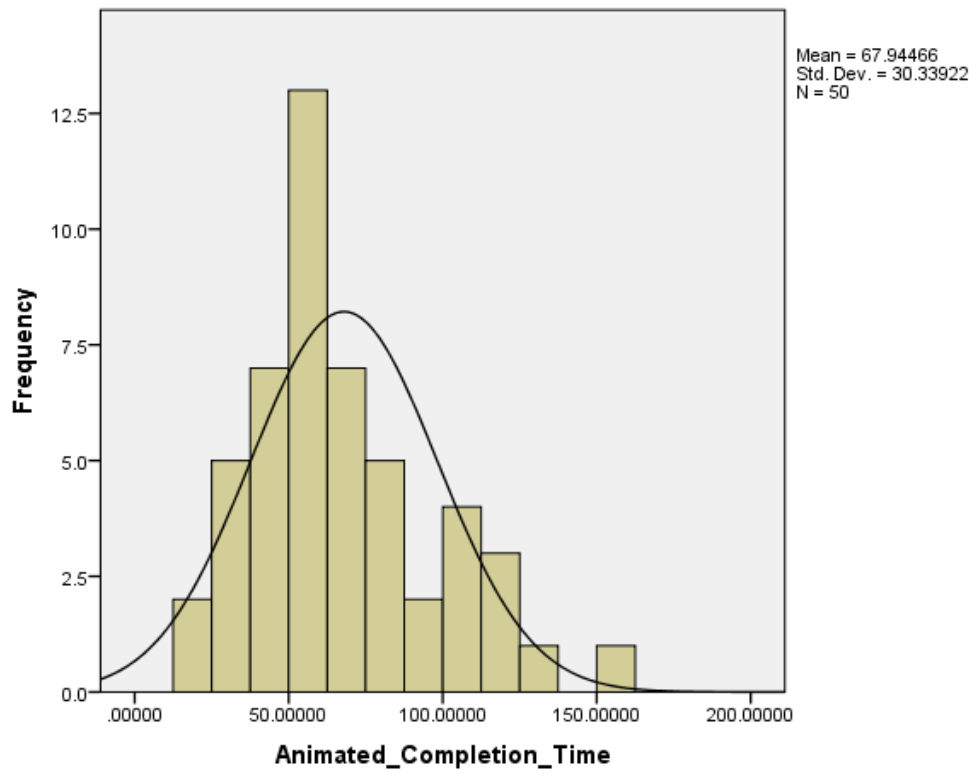
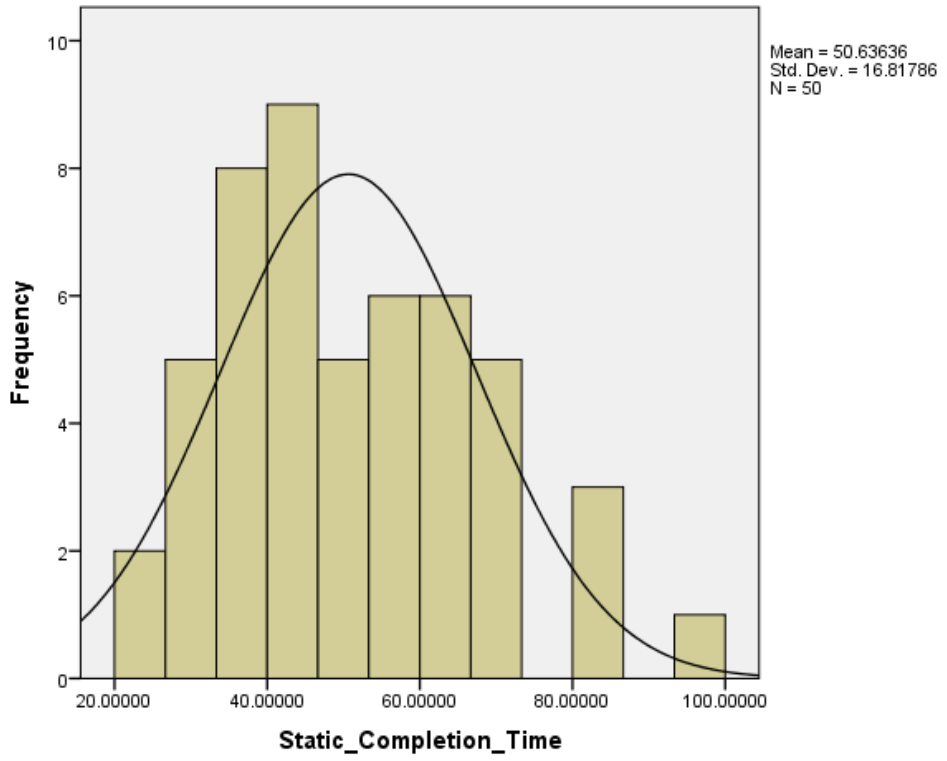


Figure 4.4 Static and Animated Completion Time Histograms (in seconds)

To test the hypothesis that the static and animated maps afforded users statistically significantly different mean completion times, an independent samples (two-tailed) t-test was performed. As can be seen in Table 4.6 the animated and static map series completion times are sufficiently normally distributed for conducting a t-test (i.e. skew  $< |2.0|$  and kurtosis  $< |9.0|$ ). Additionally, the assumption of homogeneity of variances was tested via Levene's F-test and was not confirmed, thus equal variances were not assumed,  $F(98) = 11.79$ ,  $p = .001$ . The t-test was associated with a statistically significant effect  $t(98) = -3.528$ ,  $p = .001$ . Thus, study participants were able to complete the static map tasks in a significantly shorter mean time than the animated map tasks, with a mean difference of 17.31 seconds.

#### 4.6 User-Preferences

The mean static preference score was 4.05 out of 6, while the mean animated preference score was 3.35 out of 6, resulting in a mean difference of .70 in favor of the static map series. Static preference scores were slightly more skewed and substantially more kurtotic than animated preference scores (though both were within acceptable regions for statistical analysis). Table 4.7 provides details on the descriptive statistics of user-preferences. Figure 4.5 contains static and animated preference score histograms.

**Table 4.7 User-preference Descriptive Statistics**

|                        | Static Preference<br>Score | Animated<br>Preference Score |
|------------------------|----------------------------|------------------------------|
| Mean                   | 4.053                      | 3.346                        |
| Median                 | 4.000                      | 3.333                        |
| Mode                   | 3.667                      | 3.00                         |
| Std. Deviation         | 0.850                      | 1.017                        |
| Skewness               | -.263                      | .010                         |
| Std. Error of Skewness | .337                       | .337                         |
| Kurtosis               | 1.358                      | .108                         |
| Std. Error of Kurtosis | .662                       | .662                         |

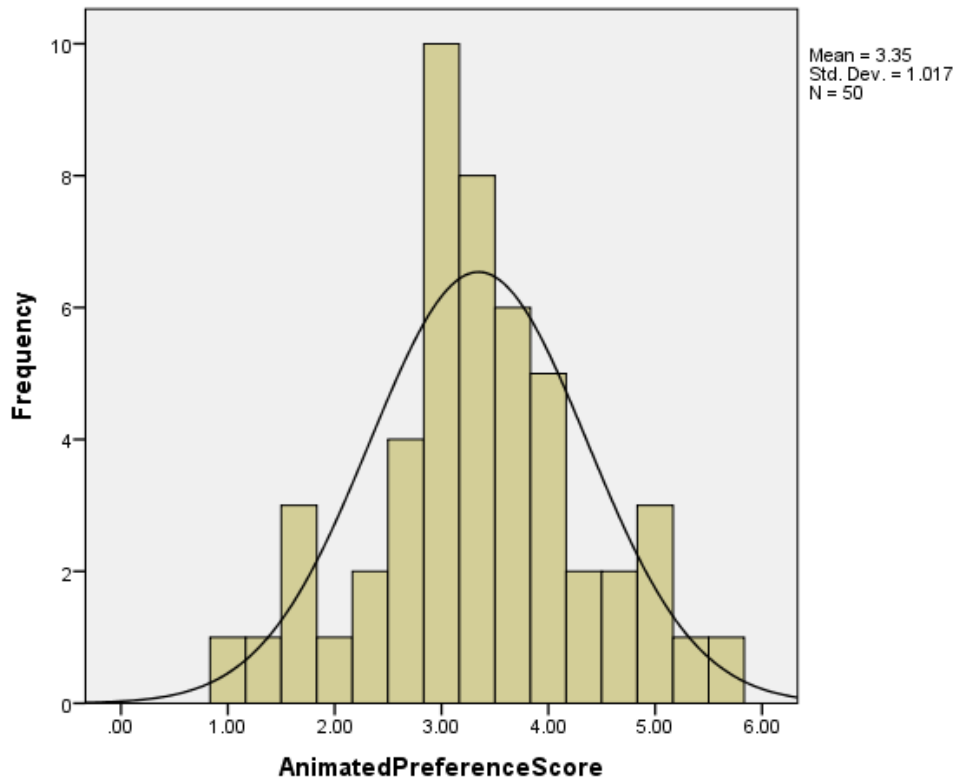
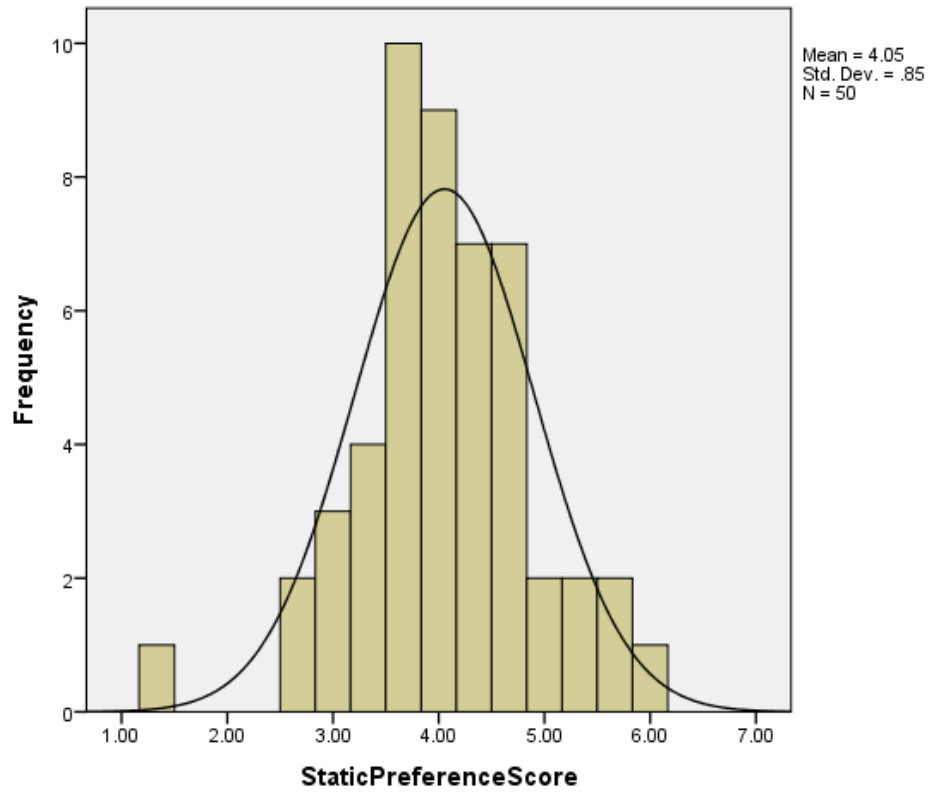


Figure 4.5 Static and Animated Preference Score Histograms

## 4.7 Correlations

Pearson's product-moment correlation coefficient calculations were used to determine whether test scores or completion times were positively or negatively correlated with user-preference scores. Animated completion time and animated preference score were the most strongly correlated of the pairs of variables, with a correlation coefficient of -0.165. Static accuracy score and static preference score were the least strongly correlated of the different pairs of variables, with a correlation coefficient of -0.001. As can be seen in Tables 4.8 and 4.9, however, no correlations with significance values below .05 were found between these variables for the static map test or for the animated map test. Thus, no statistically significant correlations were found.

**Table 4.8 Static Performance and Preference Correlations**

|                         |                     | Static Accuracy<br>Score | Static Completion<br>Time | Static Preference<br>Score |
|-------------------------|---------------------|--------------------------|---------------------------|----------------------------|
| Static Accuracy Score   | Pearson Correlation | 1                        | -.001                     | -.057                      |
|                         | Sig. (2-tailed)     |                          | .993                      | .693                       |
|                         | N                   | 50                       | 50                        | 50                         |
| Static Completion Time  | Pearson Correlation | -.001                    | 1                         | .126                       |
|                         | Sig. (2-tailed)     | .993                     |                           | .383                       |
|                         | N                   | 50                       | 50                        | 50                         |
| Static Preference Score | Pearson Correlation | -.057                    | .126                      | 1                          |
|                         | Sig. (2-tailed)     | .693                     | .383                      |                            |
|                         | N                   | 50                       | 50                        | 50                         |

**Table 4.9 Animated Performance and Preference Correlations**

|                           |                     | Animated Accuracy Score | Animated Completion Time | Animated Preference Score |
|---------------------------|---------------------|-------------------------|--------------------------|---------------------------|
| Animated Accuracy Score   | Pearson Correlation | 1                       | .001                     | .033                      |
|                           | Sig. (2-tailed)     |                         | .997                     | .820                      |
|                           | N                   | 50                      | 50                       | 50                        |
| Animated Completion Time  | Pearson Correlation | .001                    | 1                        | -.165                     |
|                           | Sig. (2-tailed)     | .997                    |                          | .252                      |
|                           | N                   | 50                      | 50                       | 50                        |
| Animated Preference Score | Pearson Correlation | .033                    | -.165                    | 1                         |
|                           | Sig. (2-tailed)     | .820                    | .252                     |                           |
|                           | N                   | 50                      | 50                       | 50                        |



## CHAPTER 5: DISCUSSION AND CONCLUSIONS

This study endeavored to provide an empirical comparison of static and animated cartographic representations of spatiotemporal phenomena in their application to basic choropleth map-based knowledge-extraction tasks. To this end, this study examined map readers' performance and efficiency in completing choropleth map-based knowledge-extraction tasks, using static time-series maps and animated maps that depict homicide patterns in the Chicago metropolitan area as the basis for doing so. This study helped to provide valuable insight as to the strengths and weaknesses of static time-series maps and animated maps as the basis for choropleth map-based knowledge-extraction tasks. It also helped to determine which of these visualization methods map users prefer for carrying out these knowledge-extraction tasks.

The results of the user performance experiment clearly indicate that the choice between a static or animated display interface can greatly influence map-readers' ability to read hot spot maps accurately and efficiently. Generally, users were able to complete the assigned tasks more accurately and much more efficiently using the static maps, as compared with their animated counterparts. Interestingly, while the user-preference metrics, in aggregate, indicate that study participants preferred the static maps over their animated counterparts, no significant correlations were found between individual test scores and individual user-preferences. This suggests the possibility that user-preferences are not based entirely on the practical application of the tools to the assigned tasks.

Chapter five is devoted to further discussion of these findings. This discussion is divided into five sections. Section 5.1 discusses task accuracy. Section 5.2 discusses users' confidence in the accuracy of their responses. Section 5.3 discusses completion time. Section 5.4 discusses user preferences and finally, Section 5.5 discusses the conclusions drawn from this study, as well as

its limitations, its parallels with Socia's study, its implications for the cartography community, and the author's suggestions for future research.

### **5.1 Task Accuracy**

As discussed in Chapter 4, the static maps were associated with a mean test score of 83%, while the animated maps were associated with a mean test score of 73%. While the difference in mean test scores between the static and animated maps was fairly small, this difference was found to be statistically significant. These findings indicate that, at least in the context of low temporal resolution choropleth time series visualization, static maps are likely to better facilitate the retrieval of accurate information than their animated counterparts.

These findings are bolstered by some of the feedback participants provided in the open-ended questions at the end of the map test survey. Two participants indicated that they had to rewind the animated version and replay it at least once in order to glean the information necessary to complete the assigned tasks, whereas they were able to quickly and easily jump back and forth between time-frames using the static map series. Based on the disparity between accuracy scores for the static and animated maps, in conjunction with this user feedback, it seems likely that some study participants answered the animated test questions incorrectly because they failed to recall the information from previous frames in the video sequence correctly. While this evidence is far from conclusive, it does lend credence to previous claims (Betancourt and Tversky 2002; Andrienko et al. 2008) that animations may be too fleeting or have too many moving parts to be perceived accurately.

## **5.2 Confidence**

The static maps were associated with a mean confidence score of 2.59 out of 3 while the animated maps were associated with a mean confidence score of 2.48 out of 3. Interestingly, these confidence scores correspond fairly closely with accuracy and efficiency measurements. Accuracy scores and confidence scores were both higher for the static maps than for their animated counterparts. Likewise completion times were shorter for the static maps than for their animated counterparts. These findings, while not conclusive, suggest two possibilities; that study participants were more likely to be confident in their responses when they had answered the question correctly, or that the static maps inspired more confidence in response accuracy than the animated maps, independently from correctness or completion time. The findings of the correlation calculations support the latter conclusion. However, given that the correlation calculations produced no statistically significant results, it is difficult to say for certain.

## **5.3 Completion Time**

The static maps were associated with a mean completion time of 50.64 seconds. By comparison, the animated maps were associated with a mean completion time of 67.94 seconds. It took study participants 17.3 seconds longer, on average, to complete the animated map test questions than it took them to complete the static map test questions. As detailed in Chapter 4, this disparity in completion times was found to be statistically significant. In fact, the difference between static and animated mean completion times is quite substantial. It took study participants approximately 35% longer to answer the test questions using the animated maps, as compared with their static counterparts. Based on these findings, it seems prudent to conclude that, at least in the context of low temporal resolution homicide hot spot time-series maps, static maps are

likely to be more efficient than their animated counterparts. However, this may vary substantially depending on the temporal resolution of the mapped data, as animation may begin to outperform static time-series with increased temporal granularity.

There could be many potential explanations for the significant disparity in completion times between the static and animated map series. Perhaps, as previous researchers have suggested (Betancourt and Tversky 2002; Andrienko et al. 2008), animations can be too complex or too fleeting to be perceived accurately. This would explain why some participants seemingly had to stop the sequence and replay it, or use the time-slider to navigate back to a previous time period to retrieve the information necessary to complete the assigned task, thereby increasing the completion time for that task. It could also be due to the complexity of the user-interface for the animated maps. The buttons to play, pause and rewind the video sequence, as well as the time-slider, added a degree of complexity to the user-interface. This might have interfered with users' efficient interaction with the animated maps.

#### **5.4 User-Preferences**

As detailed in Chapter 4, the mean static user-preference score was 4.05 out of 6, while the mean animated preference score was 3.35 out of 6, resulting in a mean difference of .70 favoring the static maps. It is not entirely clear why users tended to prefer the static maps over their animated counterparts. One study participant indicated that they experienced difficulty navigating back and forth between time periods as necessary to complete the assigned tasks. Most study participants, however, reported no such difficulty. Two participants indicated that they had to stop and rewind the video sequence to complete the assigned animated map interpretation task, whereas they were able to complete the corresponding static map interpretation task more easily.

In these cases, it seems that preferences were closely associated with the practical application of the static and animated map products to the tasks at hand. Based on the findings of the correlation coefficient calculations between users' map task performance metrics and quantified preferences (which are the focus of the next section), however, it seems inappropriate to generalize these findings to the entire sample population.

#### ***5.4.1 Performance and User-preference Correlations***

As discussed in Chapter 4, no statistically significant correlations were found between accuracy and user-preferences or between efficiency and user-preferences. In this instance, however, the lack of any strong correlations between performance metrics and user-preferences is an interesting finding in itself. One might expect to find a positive correlation between static test scores and static preference scores, for example, based on the assumption that users are likely to prefer the method that allows them to complete the task at hand most accurately and efficiently. Surprisingly, however, no such correlations were found for the static maps or for their animated counterparts. Two study participants indicated that they were displeased with the efficiency of the animated map interface, but given that two participants only constitutes 4% of the sample population, this sentiment was far from typical. As such, it seems that user-preferences, at least in the context of this experiment, were largely independent from user performance metrics.

### **5.5 Conclusions**

Overall, this study produced some very interesting results. Substantial differences were found between static and animated performance metrics, largely favoring the static map series. This study also found evidence to suggest that user-preferences are not strongly correlated with the

practical application of cartographic tools for specific knowledge-extraction tasks. This is arguably the most interesting result of this analysis. Since maps are generally used as tools, rather than for entertainment purposes, one might expect that map-readers would prefer the cartographic tool that enables them to most effectively and efficiently complete the task at hand. While this study does not contradict this notion directly, the fact that no statistically significant correlations were found whatsoever between performance metrics and user preferences certainly lends credence to previous claims that user-preferences, particularly in relation to animated maps, might not be based entirely on the practical application of these tools for learning about geographic features or spatial processes. Perhaps the growing popularity of animated maps really is due to the ‘fun-factor’ that is commonly associated with innovative new cartographic products.

#### ***5.5.1 Parallels with Socia’s Study***

Overall, the results of this study tend to reinforce Socia’s findings. In both instances, study participants were able to complete the assigned tasks more accurately and more efficiently using the static maps than they were able to using the animated maps as the basis for their answers. Interestingly, neither study found a statistically significant difference between confidence levels for the static maps and for the animated maps, despite the significant differences in performance metrics.

The time-slider was introduced for this study on the basis that it might alleviate users’ operational difficulty and split-attention. The decision to include a time-slider, rather than skip-to-time-stamp buttons, doesn’t seem to have had a substantial effect on task accuracy or completion time. Accuracy scores were comparable between the two studies despite the

differences in animated map display interface. As such, it seems that the time-slider did not have a substantial effect on user performance

One area where the two studies differed is in their findings on user-preferences. Socia's participants, despite lower test scores and slower response times, expressed a strong preference for the animated maps over the static maps that provided the basis for her study. Participants in this study, on the other hand, tended to prefer the static map series. User-preferences in this study aligned better with user performance metrics than they did in Socia's study. Participants in this study tended to prefer the cartographic product that allowed them to complete the assigned tasks most accurately and efficiently, whereas the same cannot be said for Socia's study participants. However, given that no significant correlations were found between performance metrics and user preferences at the level of the individual participant, one would be ill-advised to conclude that there is a causal relationship between accuracy or efficiency and user-preferences.

### ***5.5.2 Study Strengths and Weaknesses and Suggestions for Future Research***

Overall, this study was quite successful in addressing the research questions it set out to answer. Statistically significant differences were revealed between static and animated accuracy scores and between static and animated completion times. Overall, these findings suggest that static maps are more effective and efficient for communicating complex spatiotemporal phenomena like homicide patterns to an audience. While this study did not find any statistically significant correlations between accuracy or completion time and user-preferences, the lack of correlations are interesting findings in themselves. As such, this study unearthed evidence to support both of its proposed hypotheses.

There were, however, a couple of areas that could be improved upon for future research. One way to improve upon the methods for this study would be to include a comparison between different levels of temporal granularity. For example, one map sequence could cover the 24 hours of the day in six four-hour segments, while another could cover the same 24-hour period in twenty-four 1-hour segments. In this context, perhaps the animated maps would begin to outperform the static maps as granularity is increased. To facilitate this comparison, two separate sample populations would need to be tested. One for the low temporal granularity series and one for the high temporal granularity series. Otherwise participants may become fatigued by the length of the test. Another area where improvements could be made to this study is in the open-ended qualitative user-preference questions. Responding to these questions was optional for this study. As such, many study participants chose not to answer them. If the open-ended questions had been required, it may have been possible to gain additional insights into the reasons behind users preferences between the static and animated maps. At the outset of this study, the researcher assumed that the correlation coefficient calculations between performance metrics and user-preference metrics would be sufficient to reveal any potential relationships between these different variables. In practice, however, these data did not reveal any significant correlations between performance and user-preferences. Future research in this area should emphasize detailed, open-ended personal accounts of the reasoning behind each participant's preferences between static and animated maps. It may also be interesting to inform participants of their scores and completion times before asking them to indicate their preferences between the two types of maps.



### ***5.5.3 Implications for the Cartography Community***

The primary implication of this research for the cartographic community is quite simple. Animated choropleth maps just might not be practical for specific knowledge-extraction tasks, as compared with their static counterparts. While animated maps can be very visually appealing, this study demonstrated how animation can hinder the effective and efficient retrieval of specific geographic information. In light of these findings, cartographers should consider the potential costs and benefits associated with the choice between static and animated display interfaces very carefully. It seems that animation may be best-suited to providing a very general overview a geographic phenomenon, whereas static map series are much better suited to more specific knowledge-extraction tasks. Cartographers should keep this in mind when deciding between static and animated display interfaces.

Another interesting implication of this research is that user-preferences between static and animated choropleth maps do not seem to be directly correlated with the practical application of these tools for routine spatial knowledge-extraction. While this could be attributed to the likelihood that some users just find interactive maps to be more entertaining, it also suggests the possibility that many map-users do not fully understand the impact that animation has on their ability to accurately perceive the information contained in a cartographic animation. The disparity between static and animated confidence metrics in this study lends credence to this notion. Participants indicated similar confidence levels for the static and animated tasks, despite scoring substantially higher on the static map test. Perhaps cartographers should work against the tide of popular opinion and consider the likely disconnect between performance and user-preferences when designing maps. In some situations, users may prefer a certain type of

visualization, even when the cartographic design choices for that visualization have a markedly negative impact on their ability to accurately and efficiently interpret its contents.

## REFERENCES

- Andrienko, G., N. Andrienko, S. Fabrikant, D. Montello, and S. Rebich-Hespanha. 2008. "Novel Method to Measure Inference Affordance in Static Small-multiple Map Displays Representing Dynamic Processes". *The Cartographic Journal*. 45 (3): 201-215
- Baldwin, S. 2014. *Institutional Inscription in the Minorcan Quarter of Saint Augustine, Florida*. University of Southern California.
- Bayer, L., M. Buhner, E. Danay, E. Schmider, and M. Ziegler. 2010. "Is it Really Robust? Reinvestigating the Robustness of ANOVA against Violations of the Normal Distribution Assumption". *European Journal of Research Methods for the Behavioral and Social Sciences*. 147-151.
- Bekele, T., J. Cinnamon, M. Chipman, M. Cusimano, R. Glazier, T. Hernandez, M. Marshall, C. Rinner. 2009. "Evaluating Web-based Static, Animated and Interactive maps for Injury Prevention". *Geospatial Health*. 4 (1): 24-46.
- Betrancourt, M., B. Tversky. 2002. "Animation: Can it Facilitate?" *International Journal of Human-Computer Studies*. 57: 247-262.

- Blok, C., S. Fuhrman, N. Hedley, B. Jiang, A. Koussoulakou, D. Montello, T. Slocum. 2001. "Cognitive and Usability Issues in Geo-visualization". *Cartography and Geographic Information Science*. 28 (1):, 61-75.
- Board, C. 2011. "Cartographic Communication". *The Map Reader: Theories of Mapping Practice and Cartographic Representation*. 1: 37-47.
- Boscoe, F. P., D. Haug, A. MacEachren, L. Pickle. 1999. *Geographic Visualization: Designing Manipulable Maps for Exploring Temporally Varying Georeferenced Statistics*. Pennsylvania State University.
- Bowers, K. J., S. Johnson. 2004. "The stability of space-time clusters of burglary". *British Journal of Criminology*.
- Chainey, S. J. Ratcliffe. 2005. *GIS and Crime Mapping*. Wiley Publications.
- City of Chicago. 2014. Crime Incidents 2001 to Present. Taken from: [www.data.cityofchicago.org](http://www.data.cityofchicago.org) accessed on 7-1-2014.
- Coltekin, A., S. Fabrikant, S. Garlandini, B. Heil. 2009. "Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study with Eye-Movement Analysis". *Cartography and Geographic Information Science*. 36 (1): 5-17.

- DiBiase, D., J. Krygier, A. MachEachren, C. Reeves. 1992. "Animation and the Role of Map Design in Scientific Visualization". *Cartography and Geographic Information Systems*. 19 (4): 201-214, 265-266.
- Fabrikant, S. 2005. "Towards an Understanding of Geovisualization with Dynamic Displays: Issues and Prospects". Proceedings, American Association for Artificial Intelligence 2005 Spring Symposium Series: Reasoning with Mental and External Diagrams. Stanford University, Mar. 21-23, 2005.
- Fabrikant, S., M. Harrower. 2008. *The Role of Map Animation for Geographic Visualization. Geographic Visualization: Concepts, Tools, and Applications*. John Wiley and Sons, Ltd.
- Grubestic, T. H., E. Mack. 2008. "Spatiotemporal Interaction of Urban Crime". *Quant Criminol.* 24: 285-306.
- Harrower, M. 2007. "The Cognitive Limits of Animated Maps". *Cartographica*. 42 (4): 349-357.
- Huffington Post. 2013. *Fourth Of July Violence in Chicago: Homicide Rate Passes 200 after Long Holiday Weekend of Shootings*. Taken from:  
[http://www.huffingtonpost.com/2013/07/07/fourth-of-july-violence-c\\_n\\_3558463.html](http://www.huffingtonpost.com/2013/07/07/fourth-of-july-violence-c_n_3558463.html)  
[Accessed on 6/1/2014.](#)

- Kessler, F. C., H. Howard, B. McMaster, T. Slocum. 2009. *Thematic Cartography and Visualization*. Pearson, Prentice Hall.
- Kraak, M., A. MacEachren. 2001. "Research Challenges in Visualization". *Cartography and Geographic Information Science*. 28 (1): 3-12.
- Larkin, J., H. Simon. 1987. "Why a Diagram is (Sometimes) Worth a Thousand Words". *Cognitive Science*. 11: 65-69.
- Nakaya, T., K. Yano. 2010. "Visualizing Crime Clusters in a Space-Time Cube: An Exploratory Data-Analysis Approach Using space-time Kernel Density Estimation and Scan Statistics". *Transactions in GIS*. 14 (3): 223-239.
- Ratcliffe, J. 2010. "Crime Mapping: Spatial and Temporal Challenges". *Handbook of Quantitative Criminology*. 5-23.
- Shepard R N, 1967. "Recognition memory for words, sentences and pictures." *Journal of Verbal Learning and Verbal Behavior*. 6: 156-163.
- Socia, K. M. 2011. *Small-Multiples and Animation: Measuring User Performance with Wildfire Visualization*. Michigan State University.

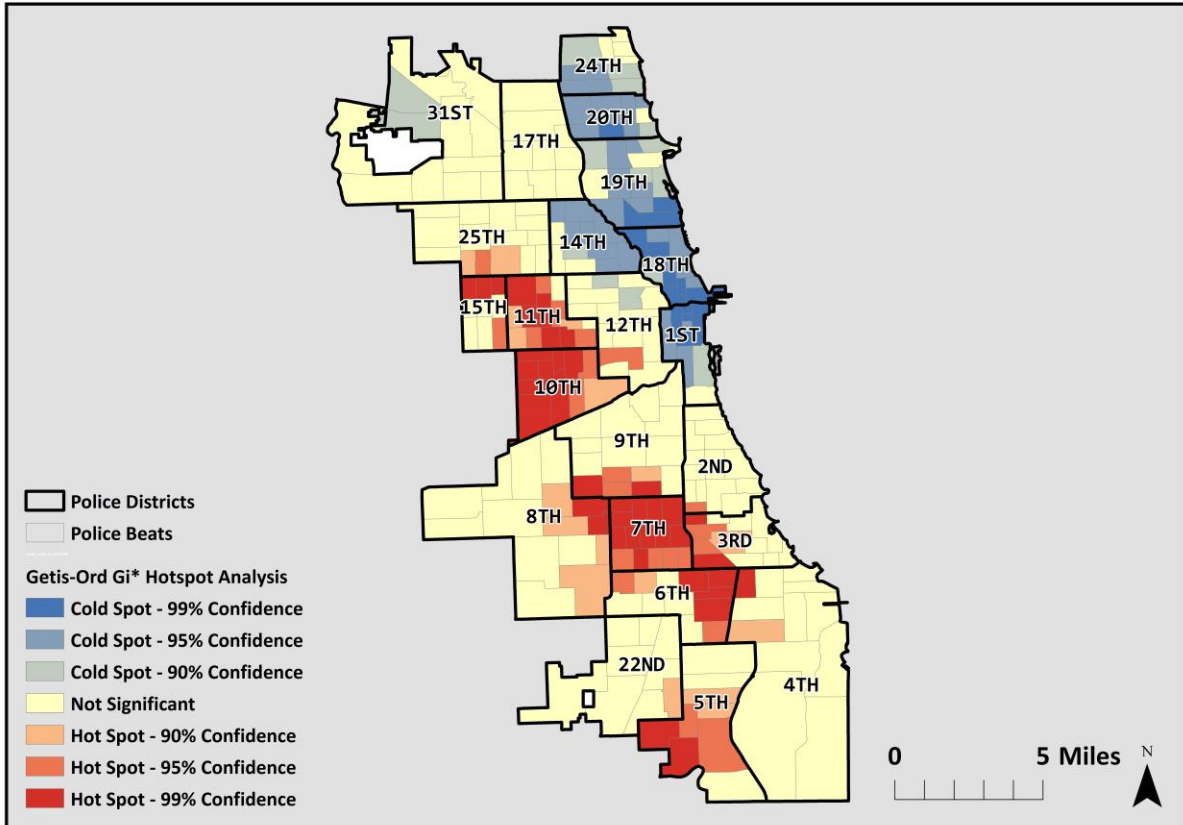
Thrower, N. 1959. "Animated Cartography." *The Professional Geographer*. 11 (6)

Tufte, E. 1995. *Envisioning Information*. Graphics Press.

Vito, G., W. Walsh. 2004. "The meaning of Compstat Analysis and Response". *Journal of Contemporary Criminal Justice*, 20 (1): 51-69. Sage Publications.

## APPENDIX A: CHICAGO HOMICIDE HOT SPOT MAPS, 2009-2013

**Chicago Homicide Hot Spots, 2009-2013**  
 12:00 AM to 3:59 AM

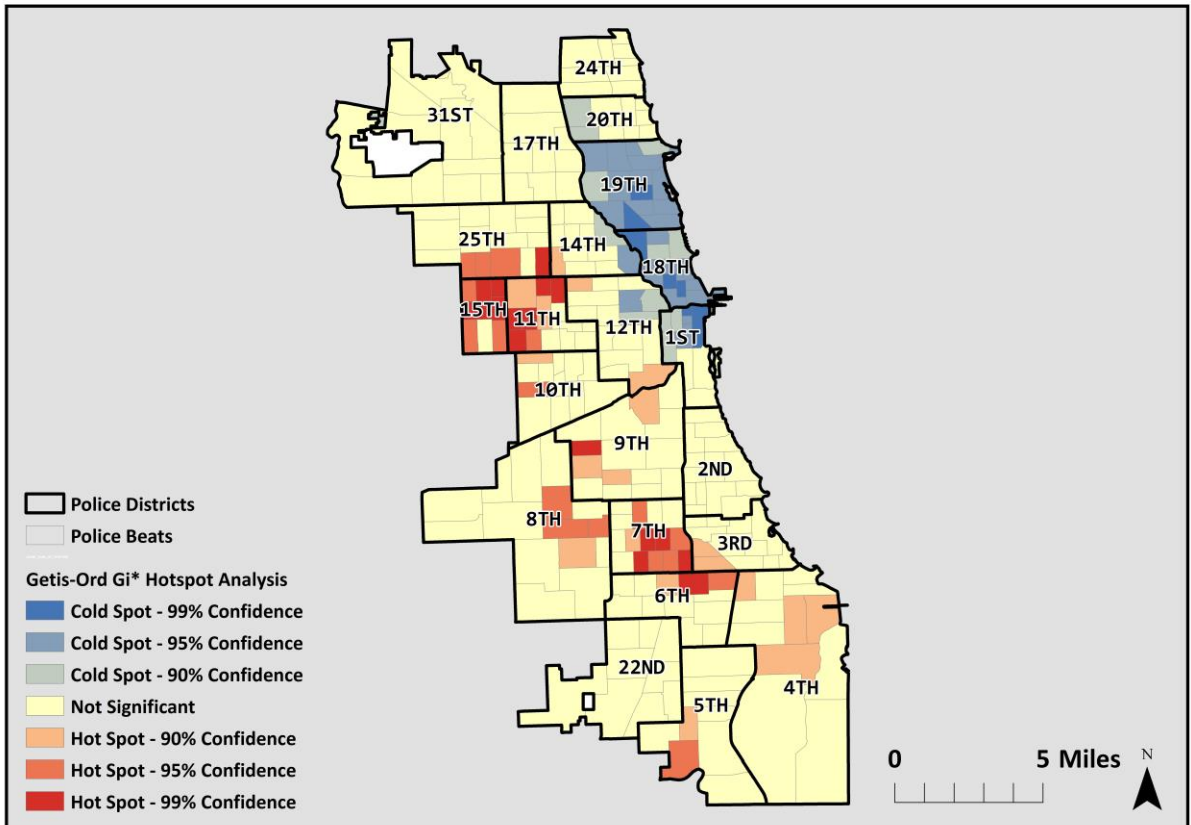


Author: Benjamin G. Anderson, GIST Master's Candidate, USC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

Note: The white spaces in the Northwestern and Southwestern corners of the map are areas for which no homicide data were available.



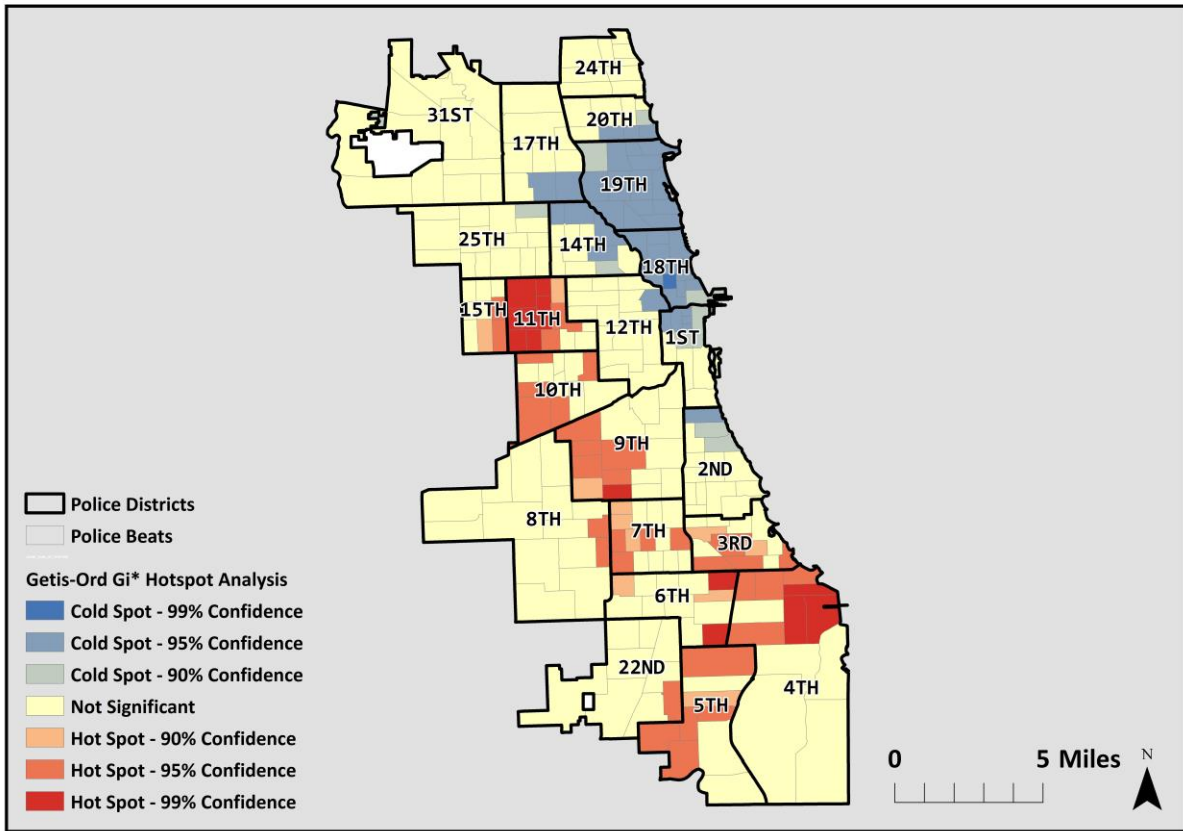
**Chicago Homicide Hot Spots, 2009-2013**  
**4:00 AM to 7:59 AM**



Author: Benjamin G. Anderson, GIST Master's Candidate, USC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

Note: The white spaces in the Northwestern and Southwestern corners of the map are areas for which no homicide data were available.

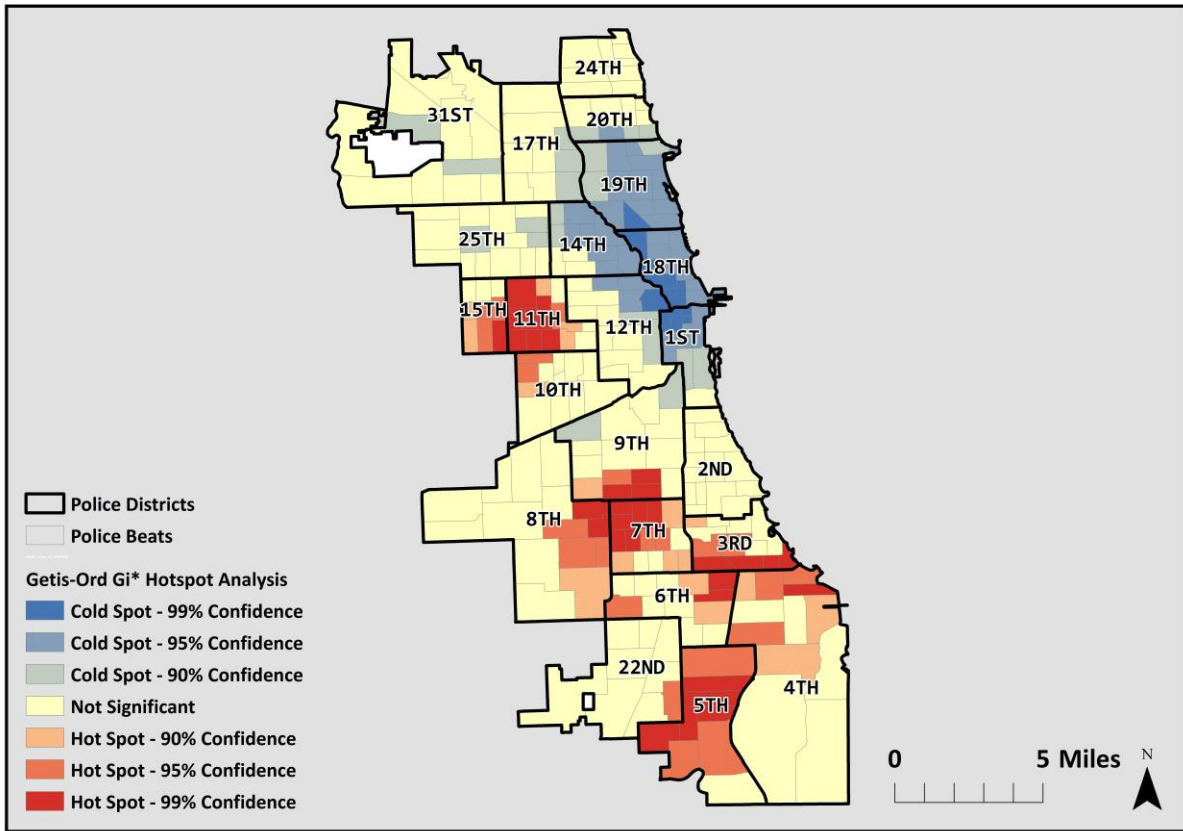
**Chicago Homicide Hot Spots, 2009-2013**  
**8:00 AM to 11:59 AM**



Author: Benjamin G. Anderson, GIST Master's Candidate, USC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

Note: The white spaces in the Northwestern and Southwestern corners of the map are areas for which no homicide data were available.

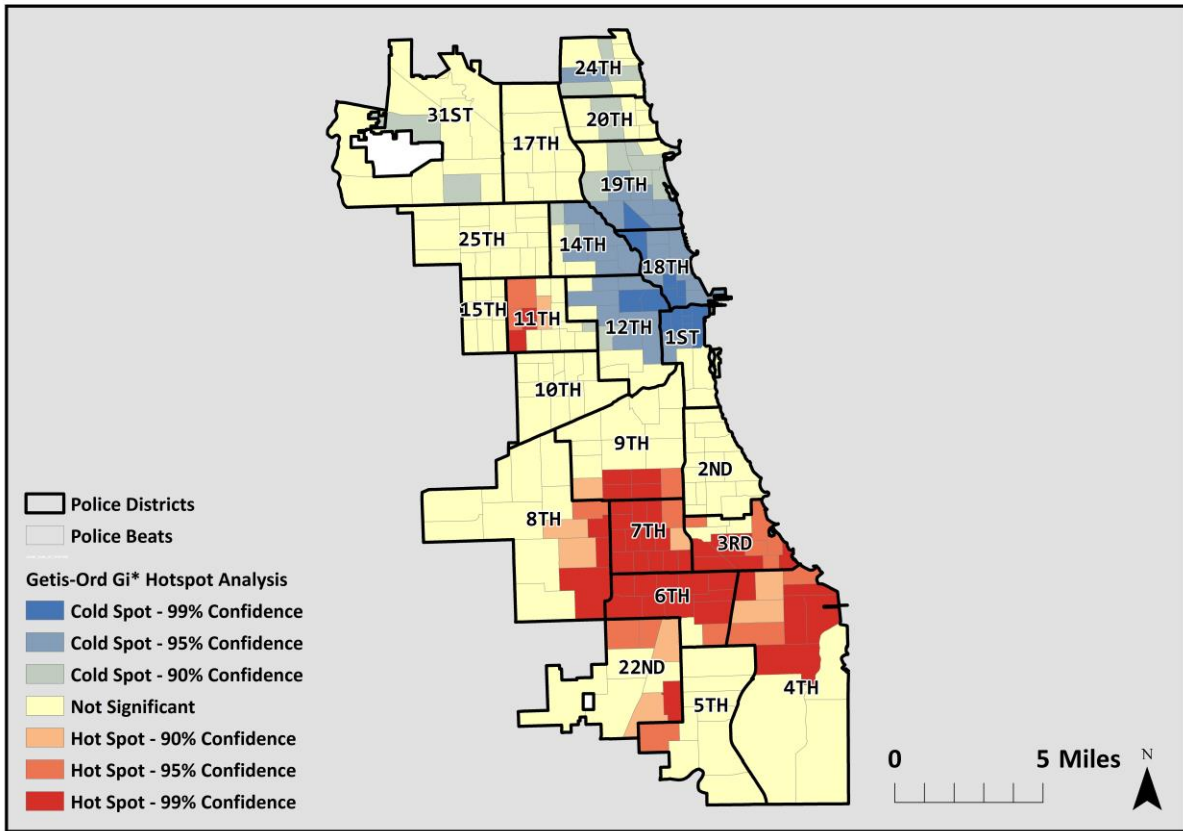
**Chicago Homicide Hot Spots, 2009-2013**  
**12:00 PM to 3:59 PM**



Author: Benjamin G. Anderson, GIST Master's Candidate, USC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

Note: The white spaces in the Northwestern and Southwestern corners of the map are areas for which no homicide data were available.

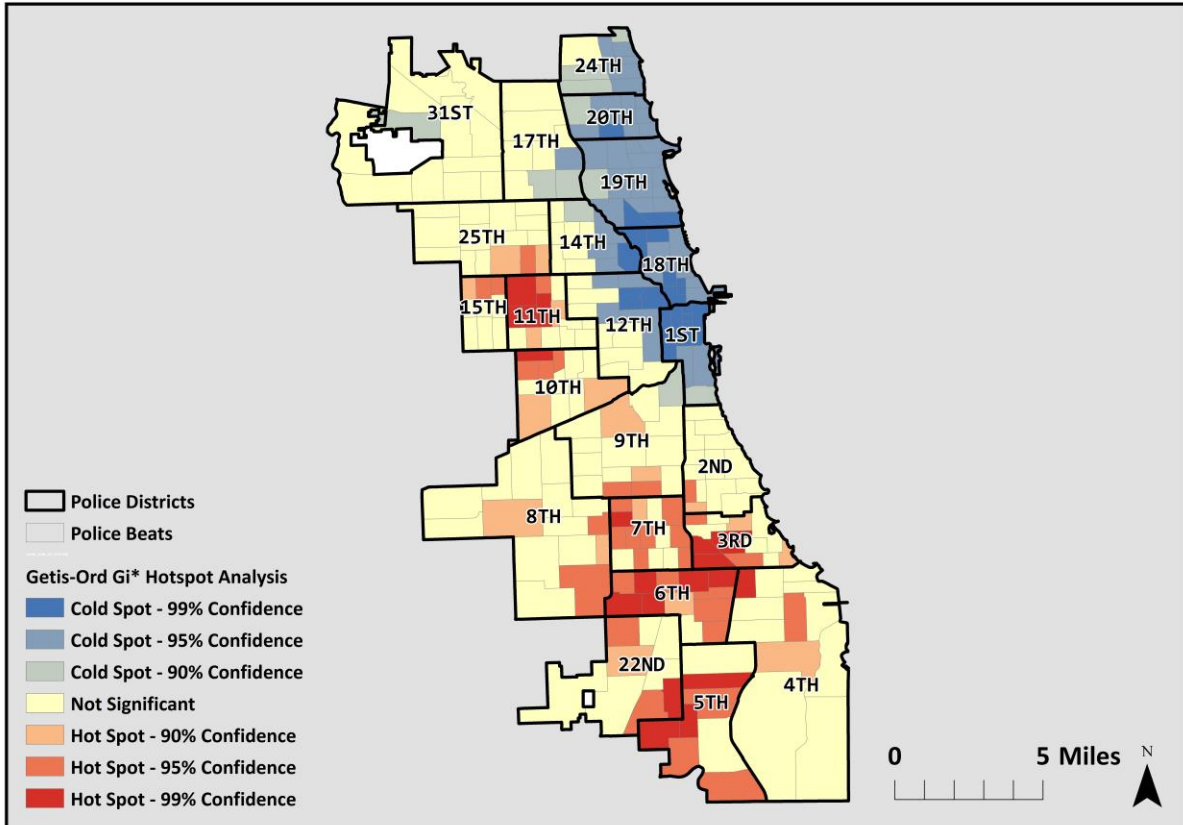
**Chicago Homicide Hot Spots, 2009-2013**  
**4:00 PM to 7:59 PM**



Author: Benjamin G. Anderson, GIST Master's Candidate, USC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

Note: The white spaces in the Northwestern and Southwestern corners of the map are areas for which no homicide data were available.

**Chicago Homicide Hot Spots, 2009-2013**  
**8:00 PM to 11:59 PM**



Author: Benjamin G. Anderson, GIST Master's Candidate, USC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

Note: The white spaces in the Northwestern and Southwestern corners of the map are areas for which no homicide data were available.

## APPENDIX B: IMAGES OF MAP TEST SURVEY INTERFACE

This study aims to measure user-performance (accuracy and efficiency) in carrying out various map-based knowledge-extraction tasks (interpreting maps to answer basic questions about their contents) using both static and animated time-series maps of Chicago crime incidents as the basis for doing so. As a participant, you will be asked answer a series of questions about each map series and to provide a subjective review of each tool based on your personal preferences and your confidence in the accuracy of your responses. Through this user-performance experiment, this study will help to provide insight as to the strengths and weaknesses of static time-series maps and animated maps as the basis for various map-based knowledge-extraction tasks, to determine which of these visualization methods map users prefer for carrying out these tasks, and to better understand which tools inspire the most confidence in response accuracy.

The survey that will provide the basis for this experiment consists of three sections: a pre-test, a map test, and a post-test. The pre-test will collect some basic information about you and provide a short tutorial on interpreting spatiotemporal maps. The map test will ask you to answer a series of questions about the distribution of homicide hot spots in Chicago, using both static and animated maps as the basis for your answers, and to indicate your level of confidence in the accuracy of your responses. Finally, the post-test will prompt you to provide feedback on the different tools. The map test will be timed to measure differences in efficiency between the two tools. Once you begin the map test, please finish it in one sitting. The entire survey should take no longer than 20 minutes to complete.

---

#### Browser Meta Info

*This question will not be displayed to the recipient.*

Browser: **Chrome**

Version: **41.0.2272.89**

Operating System: **Windows NT 6.1**

Screen Resolution: **1920x1080**

Flash Version: **17.0.0**

Java Support: **1**

User Agent: **Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/41.0.2272.89 Safari/537.36**

### How old are you?

- Under 18
- 18-25
- 26-34
- 35-54
- 55-64
- 65 or over

0%  100%

Save & Continue

Survey Powered By [Qualtrics](#)



## USC Dornsife

Dana and David Dornsife  
College of Letters, Arts and Sciences

### 1. What is your gender?

Male

Female

### What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

0%  100%

Save & Continue

Survey Powered By [Qualtrics](#)

# USC Dornsife

Dana and David Dornsife  
College of Letters, Arts and Sciences

What type of device are you using to complete this survey?

- Desktop PC
- Laptop PC
- Tablet
- Smart Phone

0%  100%

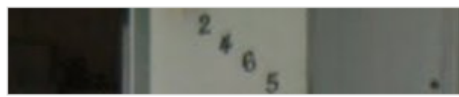
Save & Continue

Survey Powered By [Qualtrics](#)

# USC Dornsife

Dana and David Dornsife  
College of Letters, Arts and Sciences

Please type text shown in the image below in the space provided.



Type the text

reCAPTCHA™ [Privacy & Terms](#)

0%  100%

Save & Continue

Survey Powered By [Qualtrics](#)



## Map Interpretation Tutorial:

The maps that will serve as the basis for this experiment were designed to demonstrate how homicide hot spots (and cold spots) in Chicago move throughout the course of the day. For each map series, homicide incidents that occurred between January 1st, 2009 and December 31st, 2013 were divided into six groups based on the time of day at which they took place (12:00 AM to 3:59 AM, 4:00 AM to 7:59 AM, 8:00 AM to 11:59 AM, 12:00 PM to 3:59 AM, 4:00 PM to 7:59 PM, and 8:00 PM to 11:59 PM). For each time period, the homicide incidents were used to create a hot spot map. These maps, when used together, demonstrate how the spatial clustering of homicide incidents varies throughout the course of the day.

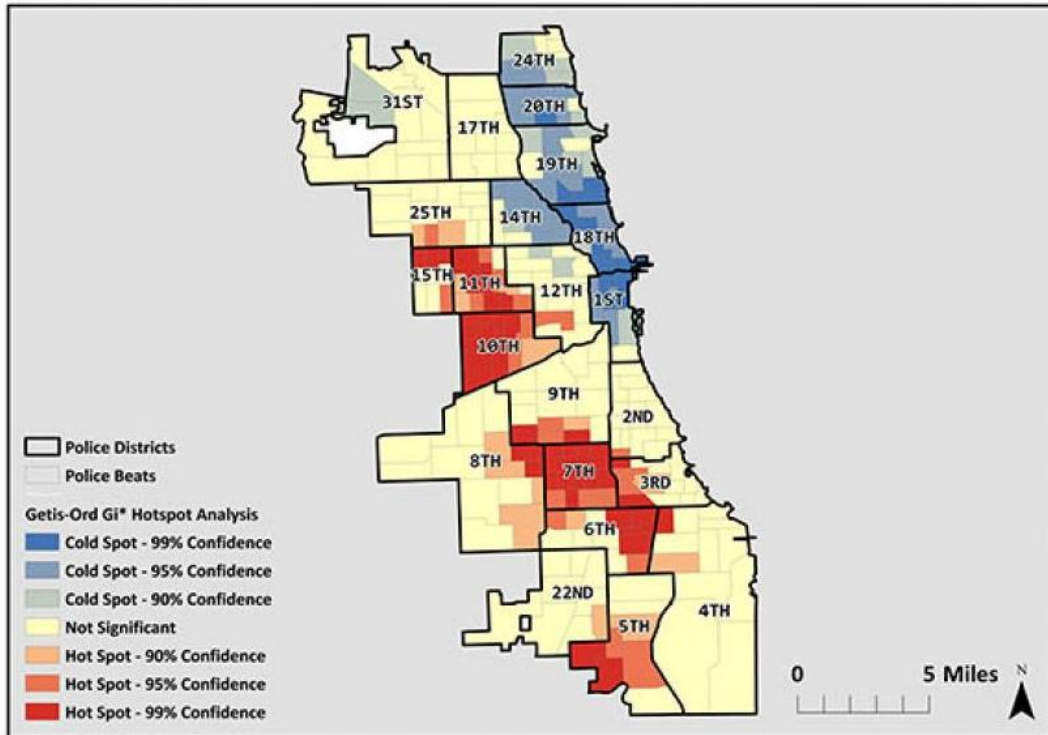
The method of hot spot analysis that was used to create the maps that serve as the basis for this experiment identifies statistically significant spatial clusters of high values (hot spots) and low values (cold spots). The dark red coloration indicates a 99% confidence hot spot (an area with a significantly higher than average homicide rate). The dark blue coloration indicates a 99% confidence cold spot (an area with a significantly lower than average homicide rate). The intermediate values indicate areas that fall somewhere in between these extremes. All of the shades of red indicate hot spots of varying statistical significance, while all of the shades of blue indicate cold spots of varying statistical significance. The intermediate yellow hue indicates areas that are not significant, meaning that they are neither hot or cold spots.

These hot spot maps are presented in two formats: static and animated.

The static maps are stacked one on top of another, with each successive map portraying the next time period. See the map series below for an example.

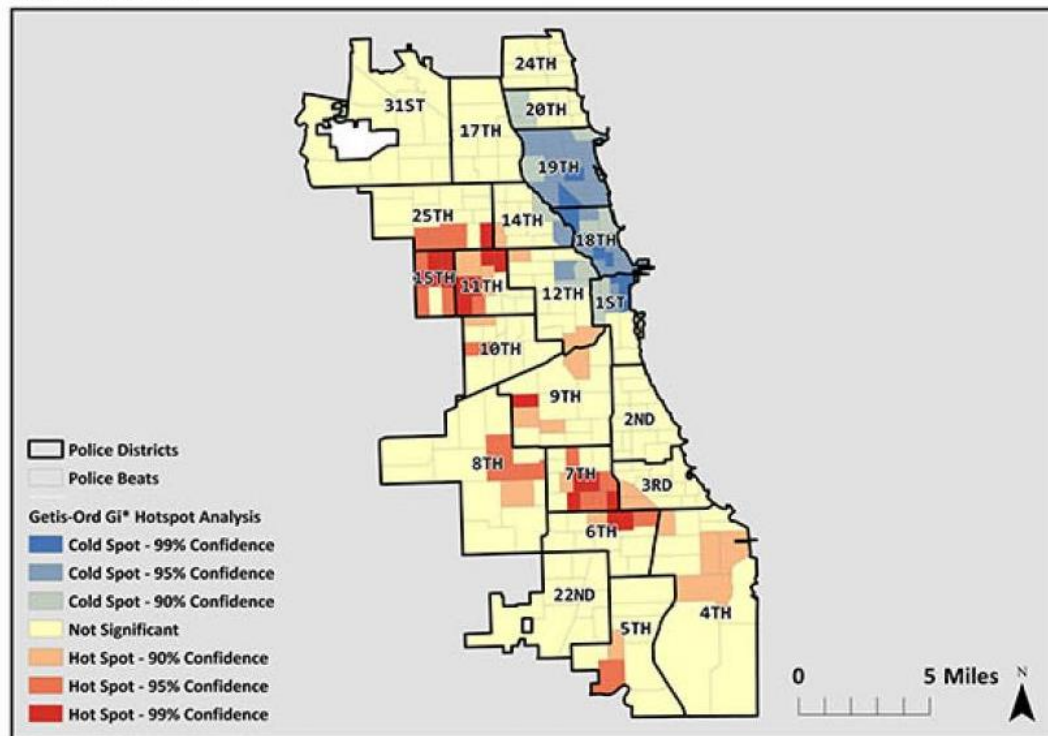
---

**Chicago Homicide Hot Spots, 2009-2013**  
**12:00 AM to 3:59 AM**



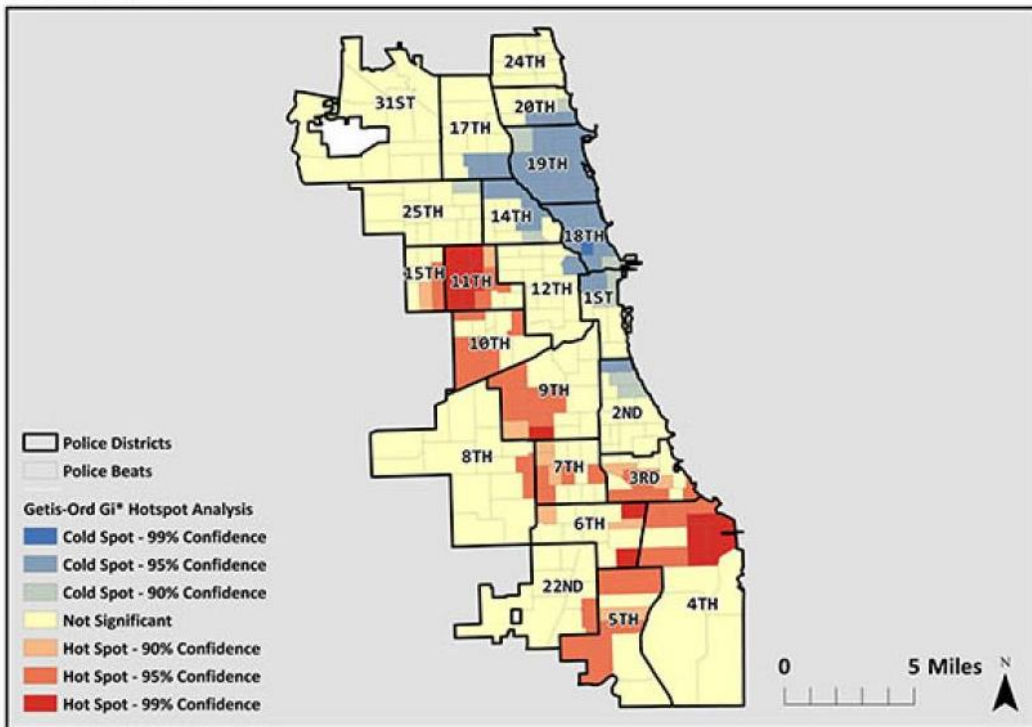
Author: Benjamin G. Anderson, GIS Master's Candidate, USK  
 Data Source: <http://data.cityofchicago.org>  
 Date: 1/14/2015

**Chicago Homicide Hot Spots, 2009-2013**  
**4:00 AM to 7:59 AM**



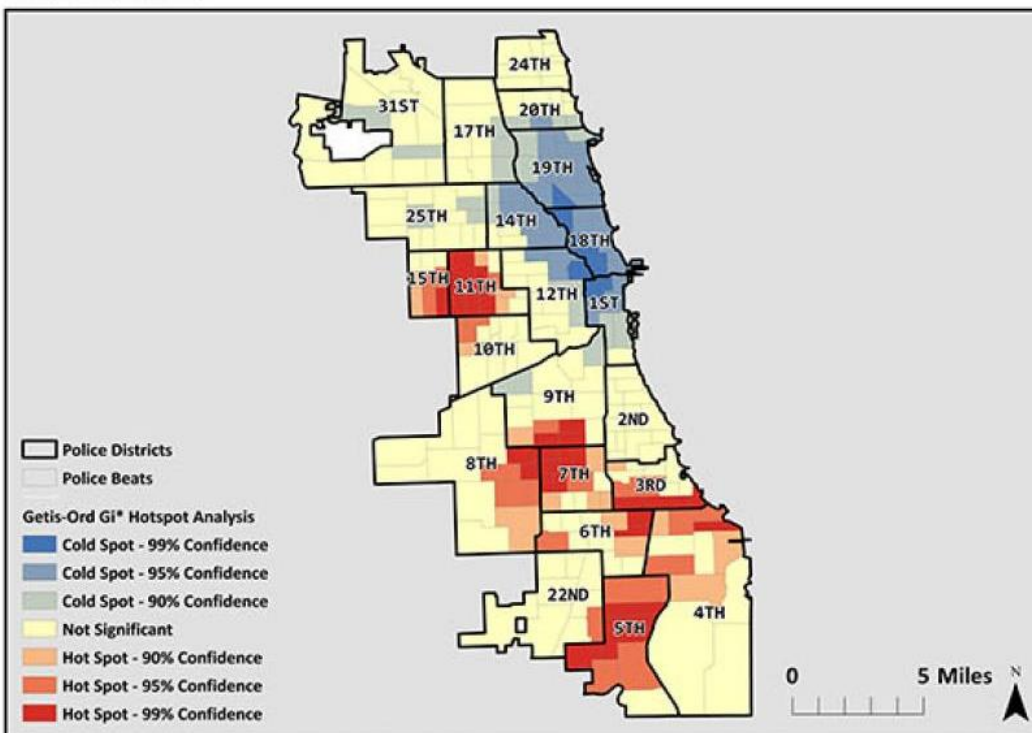
Author: Benjamin G. Anderson, GIS Master's Candidate, USK  
 Data Source: <http://data.cityofchicago.org>  
 Date: 1/14/2015

**Chicago Homicide Hot Spots, 2009-2013**  
8:00 AM to 11:59 AM



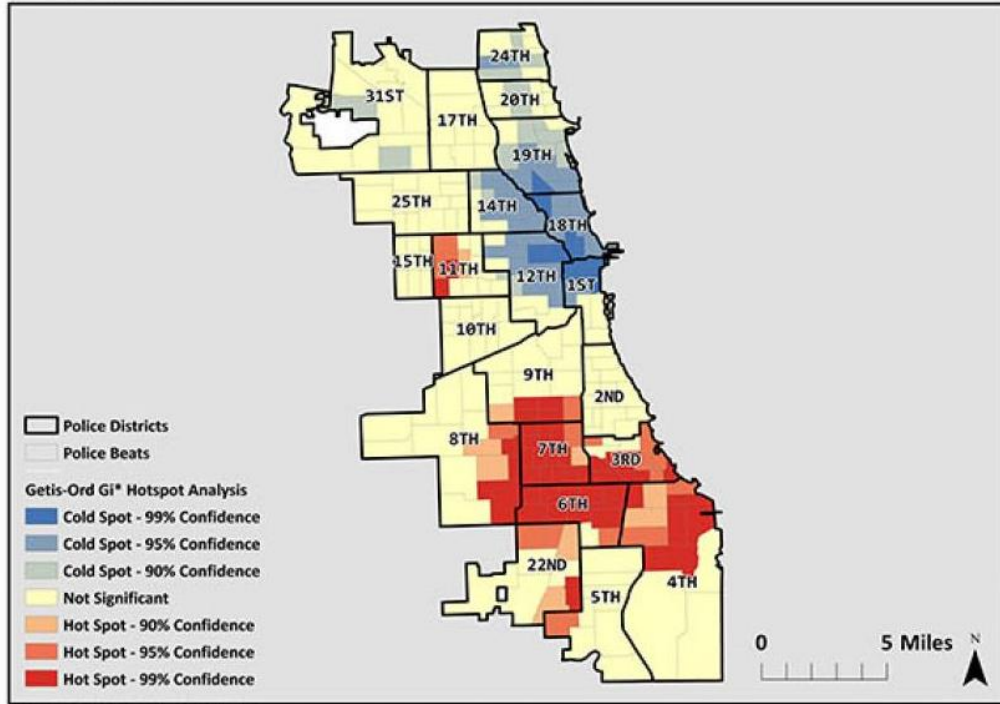
Author: Benjamin G. Anderson, GIS Master's Candidate, USC  
Data Source: <http://data.cityofchicago.org>  
Date: 1/24/2015

**Chicago Homicide Hot Spots, 2009-2013**  
12:00 PM to 3:59 PM



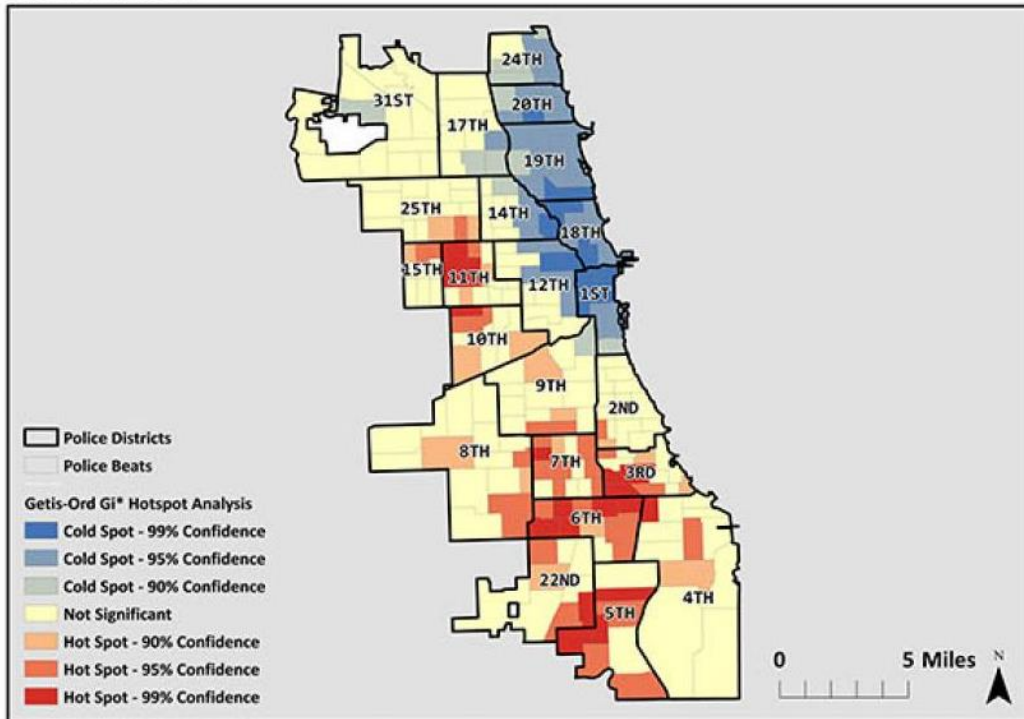
Author: Benjamin G. Anderson, GIS Master's Candidate, USC  
Data Source: <http://data.cityofchicago.org>  
Date: 1/24/2015

**Chicago Homicide Hot Spots, 2009-2013**  
**4:00 PM to 7:59 PM**



Author: Benjamin G. Anderson, GIS Master's Candidate, ISRC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

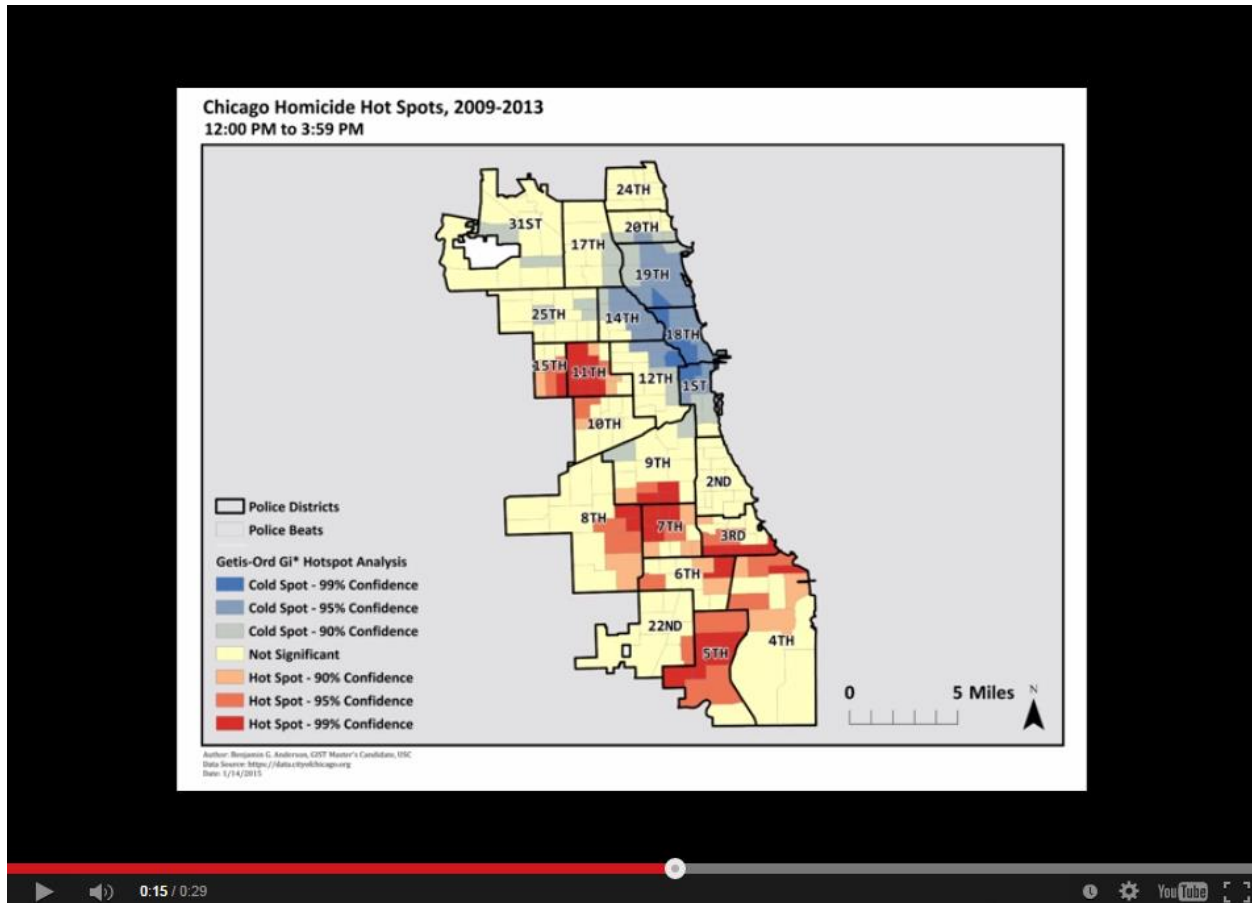
**Chicago Homicide Hot Spots, 2009-2013**  
**8:00 PM to 11:59 PM**



Author: Benjamin G. Anderson, GIS Master's Candidate, ISRC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

The animated maps are sequenced in the same way as the static version, with each frame portraying a different time period. However, the frames have been compiled into a video clip, rather than being stacked on top of one another. The video starts with the earliest time period and progresses through each successive time period as the video clip plays through.

You can click play and simply watch the video sequence play in its entirety, or you pause the video and use the time-slider at the bottom to navigate between time periods. See the video clip below for an example.



\*This image was shrunken down to fit on this page. The display size was adjusted to account for the empty space around the margins so that it was displayed the same size as the static maps in the survey interface.

The map test that is the emphasis of this experiment will ask you to use the static and animated map series in turn to answer some basic questions about how the spatial distribution of homicide hot spots varies between time periods. Please answer these questions to the best of your ability in as little time as possible and in one sitting. Accuracy, efficiency, and confidence measurements will be used to compare the strengths and weaknesses of each cartographic tool for supporting the assigned tasks.

0%  100%

Save & Continue

Survey Powered By [Qualtrics](#)



**USC Dornsife**

Dana and David Dornsife  
College of Letters, Arts and Sciences

Please use the map series below to answer question one. Afterward, please indicate your level of confidence in the accuracy of your response. You will need to scroll down to see all of the maps. Once finished, please click the save and continue button at the bottom of the page.

Which of the following statements best characterizes the overall distribution of homicide hot spots in Chicago across all time periods?

The hot spots tend to be clustered in

the Northwestern and central  
Southern districts.



The hot spots tend to be clustered in  
the Northeast.



There are no significant clusters of hot  
spots.



How confident are you in the accuracy of your response to question one?

Not confident



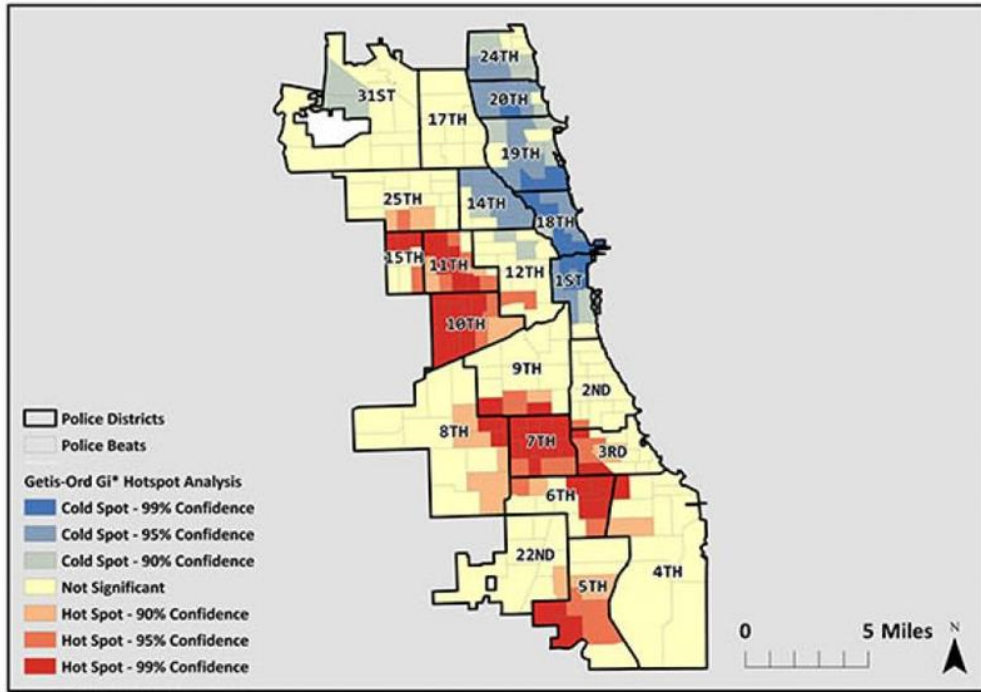
Somewhat confident



Very Confident

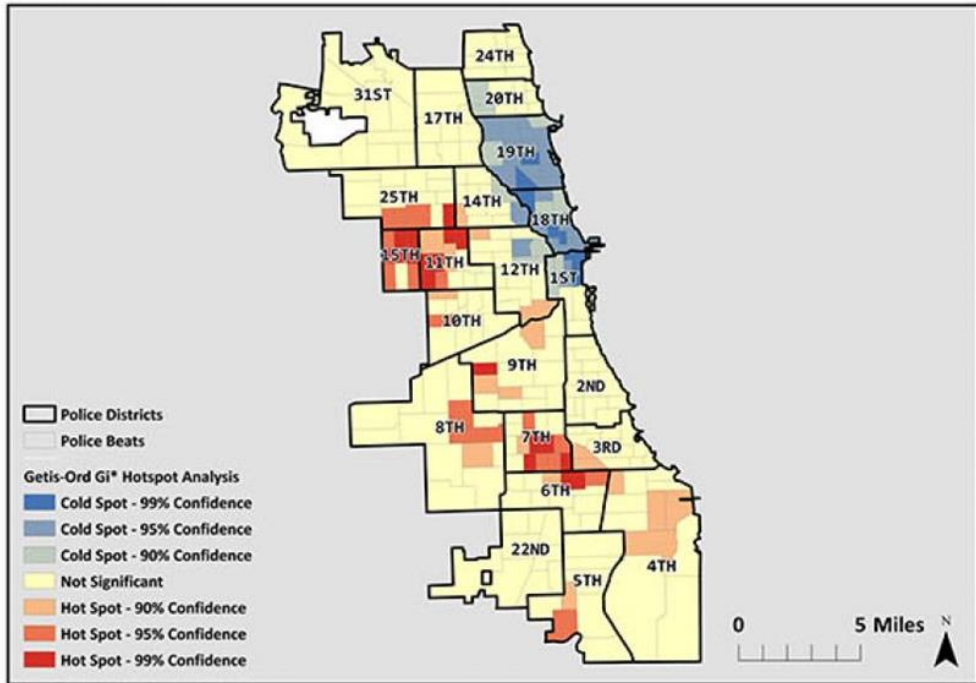


**Chicago Homicide Hot Spots, 2009-2013**  
**12:00 AM to 3:59 AM**



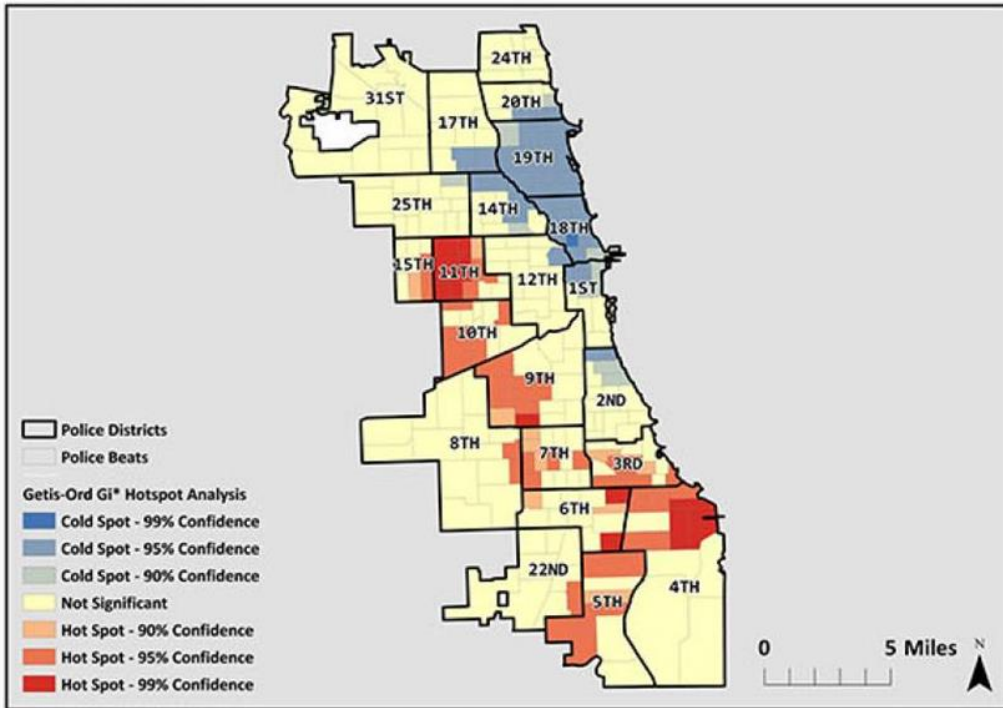
Author: Benjamin G. Anderson, GIS Master's Candidate, UIUC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

**Chicago Homicide Hot Spots, 2009-2013**  
**4:00 AM to 7:59 AM**



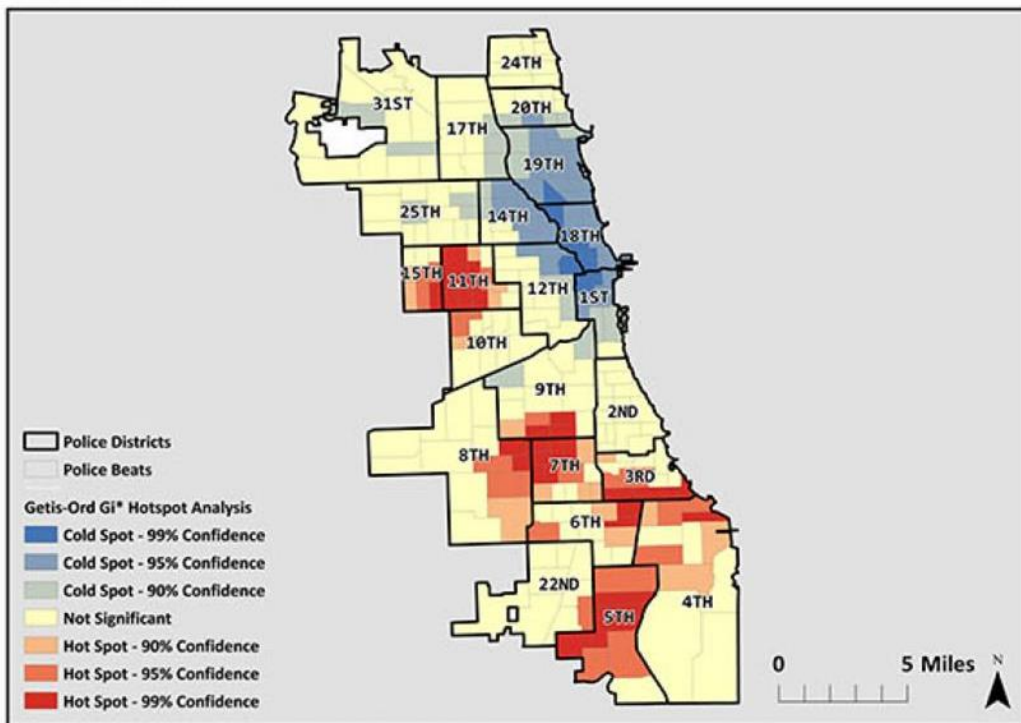
Author: Benjamin G. Anderson, GIS Master's Candidate, UIUC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

**Chicago Homicide Hot Spots, 2009-2013**  
**8:00 AM to 11:59 AM**



Author: Benjamin G. Anderson, GST Master's Candidate, UIC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

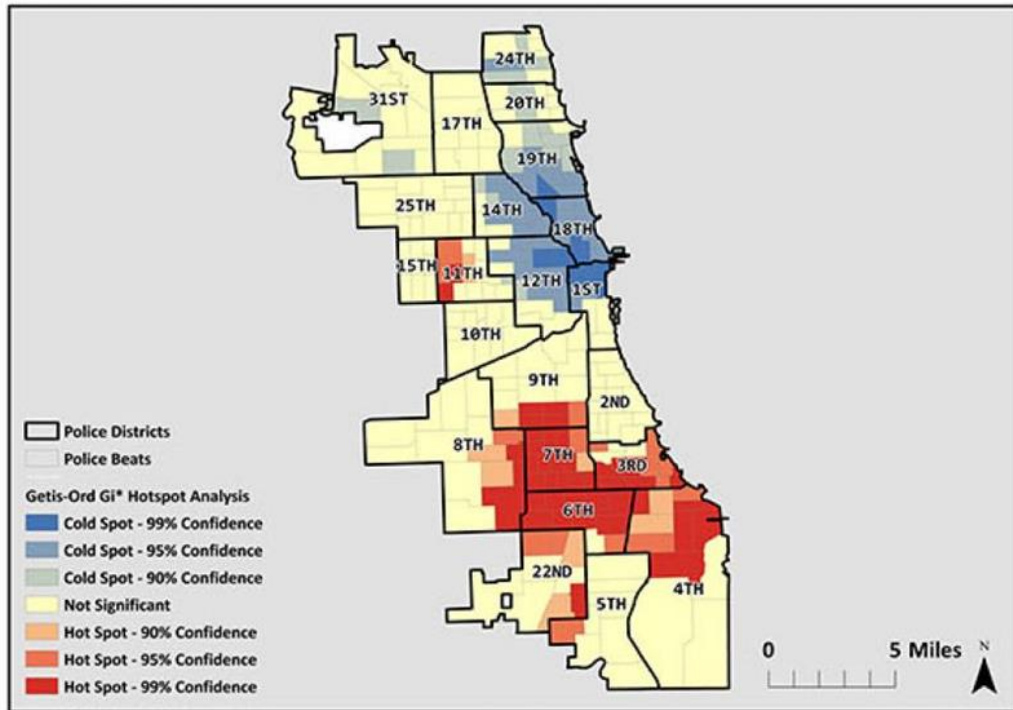
**Chicago Homicide Hot Spots, 2009-2013**  
**12:00 PM to 3:59 PM**



Author: Benjamin G. Anderson, GST Master's Candidate, UIC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

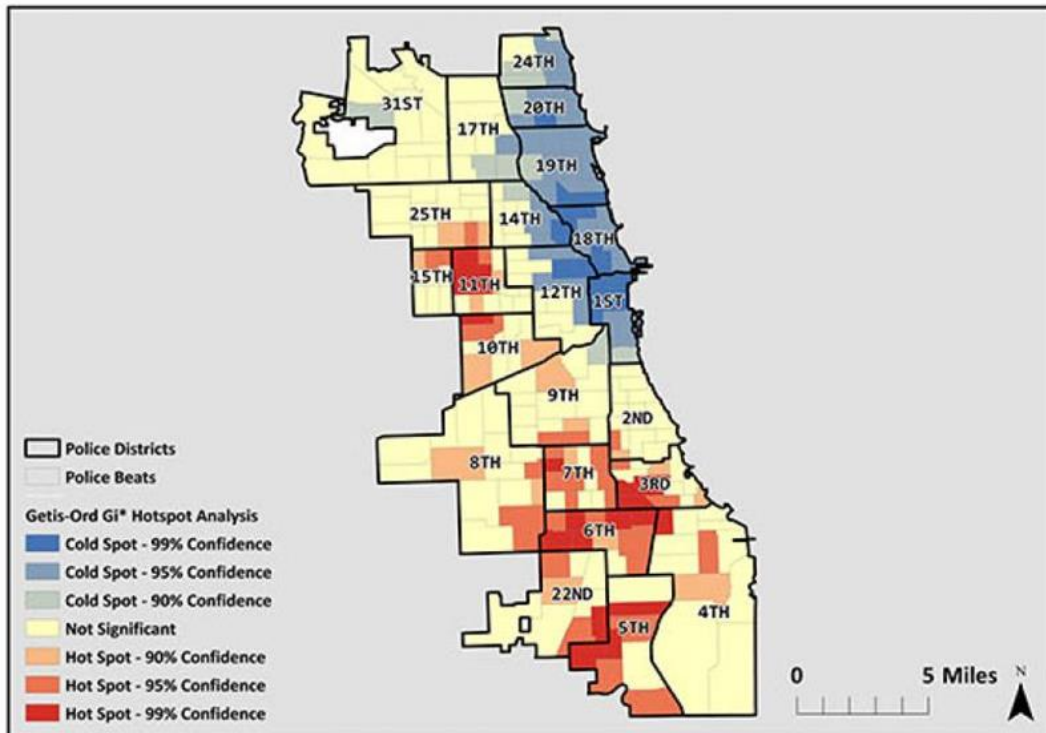


**Chicago Homicide Hot Spots, 2009-2013**  
**4:00 PM to 7:59 PM**



Author: Benjamin G. Anderson, GIS Master's Candidate, USC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

**Chicago Homicide Hot Spots, 2009-2013**  
**8:00 PM to 11:59 PM**



Author: Benjamin G. Anderson, GIS Master's Candidate, USC  
 Data Source: <https://data.cityofchicago.org>  
 Date: 1/14/2015

### Timing

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds.

Last Click: 0 seconds.

Page Submit: 0 seconds.

Click Count: 0 clicks.



Save & Continue



## USC Dornsife

Dana and David Dornsife  
College of Letters, Arts and Sciences

Please use the animated map below to answer question two. Afterward, please indicate your level of confidence in the accuracy of your response. Once finished, please click the save and continue button at the bottom of the page.

If you wish to use the time-slider to navigate between time periods, it is advisable to pause the video clip beforehand.

During which of the following time periods are the hot spots most dispersed across the city?

Between 12:00 AM and 3:59 AM



Between 4:00 AM and 7:59 AM



Between 4:00 PM and 7:59 PM



How confident are you in the accuracy of your response to question two?

Not confident



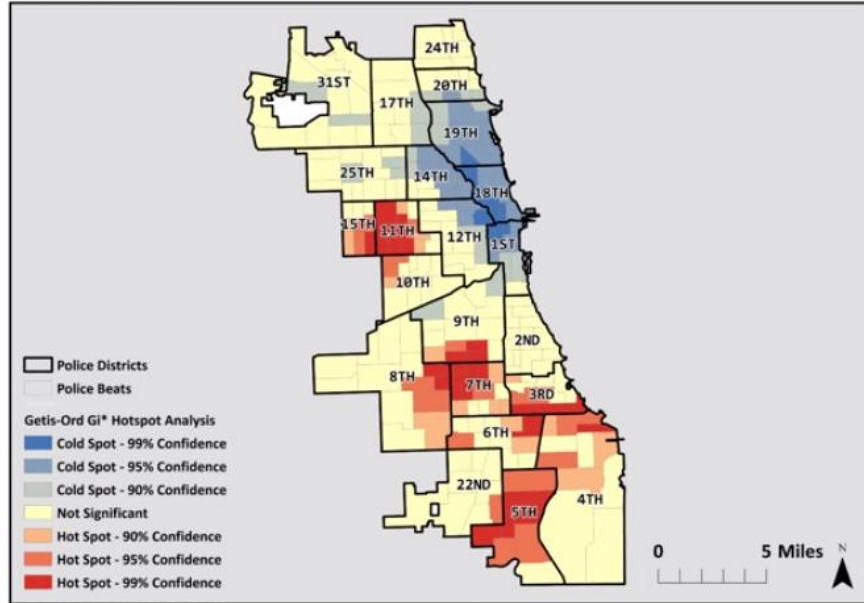
Somewhat confident



Very Confident



**Chicago Homicide Hot Spots, 2009-2013**  
12:00 PM to 3:59 PM



0:15 / 0:29

YouTube

Timing

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds.

Last Click: 0 seconds.

Page Submit: 0 seconds.

Click Count: 0 clicks.



Save & Continue

Please use the map series below to answer question three. Afterward, please indicate your level of confidence in the accuracy of your response. You will need to scroll down to see all of the maps. Once finished, please click the save and continue button at the bottom of the page.

During which of the following time period does the 1st district contain the most cold spots?

Between 4:00 AM and 7:59 AM



Between 8:00 AM and 11:59 AM



Between 8:00 PM and 11:59 PM



How confident are you in the accuracy of your response to question three?

Not Confident



Somewhat Confident



Very Confident



\*The static map series was displayed again here in the survey interface but has been omitted here.

#### Timing

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds.

Last Click: 0 seconds.

Page Submit: 0 seconds.

Click Count: 0 clicks.



Save & Continue

Please use the animated map below to answer question four. Afterward, please indicate your level of confidence in the accuracy of your response. Once finished, please click the save and continue button at the bottom of the page.

If you wish to use the time-slider to navigate between time periods, it is advisable to pause the video clip beforehand.

Which of the following statements best characterizes the overall distribution of homicide cold spots in Chicago across all time periods?

The cold spots tend to be clustered in the Northeastern districts.      The cold spots tend to be clustered in the Southeastern districts.      There are no significant clusters of cold spots.




How confident are you in the accuracy of your response to question four?

Not confident

Somewhat confident

Very Confident




\*The animated map series was displayed again here in the survey interface but has been omitted here.

Timing

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds.

Last Click: 0 seconds.

Page Submit: 0 seconds.

Click Count: 0 clicks.



Save & Continue

# USC Dornsife

Dana and David Dornsife  
College of Letters, Arts and Sciences

Please use the map series below to answer question five. Afterward, please indicate your level of confidence in the accuracy of your response. You will need to scroll down to see all of the maps. Once finished, please click the save and continue button at the bottom of the page.

---

During which of the following time periods are the hot spots most clustered in the southern half of the city?

---

Between 12:00 AM and 3:59 AM



Between 4:00 AM and 7:59 AM



Between 4:00 PM and 7:59 PM



How confident are you in the accuracy of your response to question five?

---

Not confident



Somewhat confident



Very Confident



\*The static map series was displayed again here in the survey interface but has been omitted here.

## Timing

---

**These page timer metrics will not be displayed to the recipient.**

First Click: 0 seconds.

Last Click: 0 seconds.

Page Submit: 0 seconds.

Click Count: 0 clicks.

0%  100%

Save & Continue

Please use the animated map below to answer question six. Afterward, please indicate your level of confidence in the accuracy of your response. Once finished, please click the save and continue button at the bottom of the page.

If you wish to use the time-slider to navigate between time periods, it is advisable to pause the video clip beforehand.

During which of the following time periods does the 24th district contain the most cold spots?

Between 4:00 AM and 7:59 AM



Between 8:00 AM and 11:59 AM



Between 8:00 PM and 11:59 PM



How confident are you in the accuracy of your response to question six?

Not confident



Somewhat confident



Very Confident



\*The animated map series was displayed again here in the survey interface but has been omitted here.

#### Timing

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds.

Last Click: 0 seconds.

Page Submit: 0 seconds.

Click Count: 0 clicks.



Save & Continue

Static map series are better than animated maps for communicating spatiotemporal crime patterns to an audience.

Very Strongly Agree   Strongly Agree   Agree   Disagree   Strongly Disagree   Very Strongly Disagree

Animated maps are better than static time-series maps for communicating spatiotemporal crime patterns to an audience.

Very Strongly Agree   Strongly Agree   Agree   Disagree   Strongly Disagree   Very Strongly Disagree

Static maps are easier to understand because animated maps are too fleeting and have too many moving parts that are hard to perceive accurately.

Very Strongly Agree   Strongly Agree   Agree   Disagree   Strongly Disagree   Very Strongly Disagree

Animated maps are easier to understand because they utilize time itself to demonstrate temporal change.

Very Strongly Agree   Strongly Agree   Agree   Disagree   Strongly Disagree   Very Strongly Disagree

For quality assurance purposes, please select very strongly disagree.

Very Strongly Agree   Strongly Agree   Agree   Disagree   Strongly Disagree   Very Strongly Disagree

0%  100%

Save & Continue



Please use the space provided below to provide any additional feedback you may have.

Please use the space provided below to discuss your map reading strategies and any additional information you may wish to provide.

0%  100%

Save & Continue



**USC Dornsife**

Dana and David Dornsife  
College of Letters, Arts and Sciences

Please wait while you are redirected.

[Click here if you are not automatically redirected.](#)

0%  100%