IDENTIFYING VARIOUS CRIMES USING HOTSPOT TECHNIQUES IN THE CITY OF HOUSTON TX

SARAH SIDDIQUI

SAM HOUSTON STATE UNIVERSITY

TABLE OF CONTENTS

ABSTRACT

The purpose of this report to identify crime hotspot such as Burglary, Robbery and Motor vehicle theft in the city of Houston Texas. Hotspot analysis techniques are extremely helpful for police to identify such areas of crime being committed and the possible way to respond. There are several methods and theories have been used to identify crime density areas and places where these crimes occurring most frequently in the passage of time. Each method has its strength and different ways to display results. The reported crime data are from the city of Houston from January 2019 to December 2019. This research has applied three different techniques to analyze different patterns of crime in the Houston area. The Kernal density estimation and Getis Ord Gi* could be identified as the methods that are helpful to examine rate of crime in study area, while the thematic mapping is a great technique to identify spatial distribution of crime within the geographical regions.

CHAPTER 1 INTRODUCTION

Crime is not consistently spread across areas. Some areas have high concentration of crimes and some have low and even absent in some places. In some places people might lock their cars and hide personal belongings and in other places they don't. Key component of dealing with crime is to recognize where the specific crime taking place. This crime event distribution is identified as inherent geographical quality explained by by Chaine and Ratcliffe, 2005 and explained by different theories such as Ecology of Crime (Brantingham and Brantingham, 1984) or routine activities (Cohen and Felson, 1979), and others. The places where crime events are densely distributed are called hotspots. Hotspot analysis is at the epicenter of analysis of crime and hotspot mapping is growing wide attention among crime mapping. These crime maps can be utilized by law enforcement agencies to better allocate their resources to control or minimize the crime rates in the areas. There are many different mapping techniques to identify different patterns of crime such as point mapping to observe the geographic distribution of crime, utilizing geographic boundaries such as Census tract, police beats and shaded them with different colors to show the intense and low rates of crimes or represent distribution of crime as a continuous surface that show volumetric densities of crime in the geographic regions.

CHAPTER 2 DATA AND GEOCODING

2.1 The Study Area and the Spatial Data

The study area of this research is the territory of Houston Police District (HPD) which is the key law enforcement agency of city of Houston. On the geographical extent Houston police department extends from -95.784602°W to -95.000783°E and from 30.126094°N to 29.519338°S.

Crime report data were available on the website of Houston police Districts. In this research crime data for Burglary, Robbery, and Motor Vehicle Theft has been used from the year January 2019- December 2019. The Crime Data consist of number of offenses, Incidents, Street address and zip code. In order to do the Analysis Geocoding has been done using ArcGIS online world Geocoding Service available in ArcGIS Pro. The census tract data has been downloaded from Houston Galveston Area council dataset. The census tract data was for entire Texas so using select by location tool to only select polygons in the Houston to narrow down the data for further analysis.

CHAPTER 3 METHODOLOGY

3.1 Thematic Mapping

Thematic mapping is a popular technique for representing any spatial distribution. These boundaries are usually administrative or political boundaries such as beats, census blocks, polling districts etc. Burglary Data has been joined spatially with the Houston Police districts polygon to get the total count of incidents within each district. This spatially joined data has symbolized using Graduated Colors to show the districts with light to dark shade that represents high and low counts of Burglary. In ArcGIS Pro several classification methods has provided. They include natural breaks, equal interval, quantile, standard deviation, manual classification, etc. Choosing an appropriate classification method and the corresponding class boundaries is important in crime analysis research. Different classification schemes will place crimes into different categories once the classification specified the thematic map can be produced based on crime count associated with each district. The map displays data with light to dark shaded colors. Light colored districts are cold spots wile the dark shaded colors represent hotspots of the Burglary within the Houston Police district.

Its common that crime rates are high in densely populated areas because large number of people working and living which are potential victims to criminals. For example, downtown area has high number of crimes compared to a suburb due to its large amount of people living and visiting these areas. Similarly shopping districts are also have a high chance of having crime in the area as well. To assess this factor crime rate has been calculated by dividing the number of reported burglary incidents by the total population and the result is multiplied by 100,000.

3.2 Kernel Density Estimation

Kernel density estimation (KDE) is regarded as the most suitable spatial analysis technique for visualizing crime data (McGuire and Williamson, 1999; Williamson et al. , 1999, 2001 ; Chainey et al. , 2002 ; Chainey and Ratcliffe, 2005 ; Eck et al. , 2005). Point data (offences) are aggregated within a userspecifi ed search radius and a continuous surface that represents the density or volume of crime events across the desired area is calculated. A smooth surface map is produced, showing the variation of the point / crime density across the study area, with no need to conform to geometric shapes such as ellipses. To analyze incidents of Robbery Kernel density tool has been used to generate smooth output raster. The Input was Robbery crime and the processing extent has the input of census tract of Houston, Robbery data has been selected for Raster analysis output. After the tool ran the density of Robbery crime has displayed on the map. There are several classification methods like Equal Interval and Quantile etc. To better cartographic look the data classified into 5 classes and Qunatile method has selected.

The output raster map has been converted into census tract so we can get crime rate per tract using raster data. Zonal statistics tool as table has been used to do the conversion. This tool calculates statistics on values of a raster within the zones of another datasets. This tool Summarizes the values of a raster within the zones of another dataset and reports the results to a table. In the tool parameters the input raster was the Census tract layer, the zone field was set to ID,input value raster was set to Kernel density raster, for statistics type value set to MEAN that’s the average raster pixel value for each area of consideration. The output table was joined with the

Census tract table based on ID field and this joined layer exported into a shapefile. This layer was symbolized with the MEAN data. So, the Kernal density map converted into tract map. The map displayed clustering of Robbery crime on the tract layer.

3.3 Hotspot Analysis Getis Ord Gi*

The Hot Spot Analysis tool calculates the Getis-Ord Gi* statistic for each feature in a dataset. The resultant Z score tells you where features with either high or low values cluster spatially. This tool works by looking at each feature within the context of neighboring features. A feature with a high value is interesting but may not be a statistically significant hot spot. To be a statistically significant hot spot, a feature will have a high value and be surrounded by other features with high values as well. The local sum for a feature and its neighbors is compared proportionally to the sum of all features; when the local sum is much different than the expected local sum, and that difference is too large to be the result of random chance, a statistically significant Z score results. The Gi* statistic returned for each feature in the dataset is a Z score. For statistically significant positive Z scores, the larger the Z score is, the more intense the clustering of high values (hot spot). For statistically significant negative Z scores, the smaller the Z score is, the more intense the clustering of low values (cold spot).

For analyzing Motor Vehicle Theft cases, this tool has been used in which we have set the conceptualization of spatial Relationship to Fixed distance band. We have set the Analysis field Incident from Vehicle theft data to construct hot and cold spots.

CHAPTER 4 RESULTS

The hotspot techniques based on different theories and different concepts and set of parameters so the resulting outputs mainly the hotspot maps differ from each other. The Hotspot map generated from the Thematic mapping shows high and low areas of Burglary crime rate, Kernel density map shows results in the form of raster and zonal statistics based on the census tract, Getis Ord Gi* map shows clusters of hot and cold spots of Motor Vehicle theft crime in the Houston area. Burglary rate is extremely high on west division of Houston Texas which is the high population district of Houston, Central division has moderately low and Northeast division has significantly low rates of Burglary. Thematic maps outputs are important map outputs because it represents geographic regions for political and administrative purpose. For example, if some police officer has a responsibility to combat with crime rate in a group of geographic area so the inspector will most likely want the table of crime counts by categories and the boundary of geographic area.

The Kernel density smoothing map has provided Robbery crime incident in raster format. And then convert this raster format into census tracts data so that we can get the crime rate per tract. The first method of continuous smoothing surface is visually appealing and allow to understand hotspots easily on the map. West and central part of the Houston has high density of Robbery incidents. The next output has generated from converting raster into the tracts

using Zonal statistics table tool that provide Robbery crime within each tract. The mean values are easy to read and helps law enforcement agencies to understand the map. The increasing mean values represent hotspots of Robbery in the tract data.

The third map shows Motor Vehicle theft cases in the Houston area. The hotspot Getis Ord Gi* provided map in a clustered point that represent Motor Vehicle Theft. The Gi* returned for each Feature in the dataset is a Z score. For statistically significant positive Z scores, the larger the Z score is, the more intense the clustering of high values (hotspots). For statistically significant negative Z score, the smaller the Z score is the more intense of the clustering of low values cold spots. P values are the probability of hotspots or cold spots on its spatial distribution. Northwestern Part of Houston has indicated high Z values clusters and are hotspots of Motor Vehicle Theft incidents.

CHAPTER 5 List of Figures

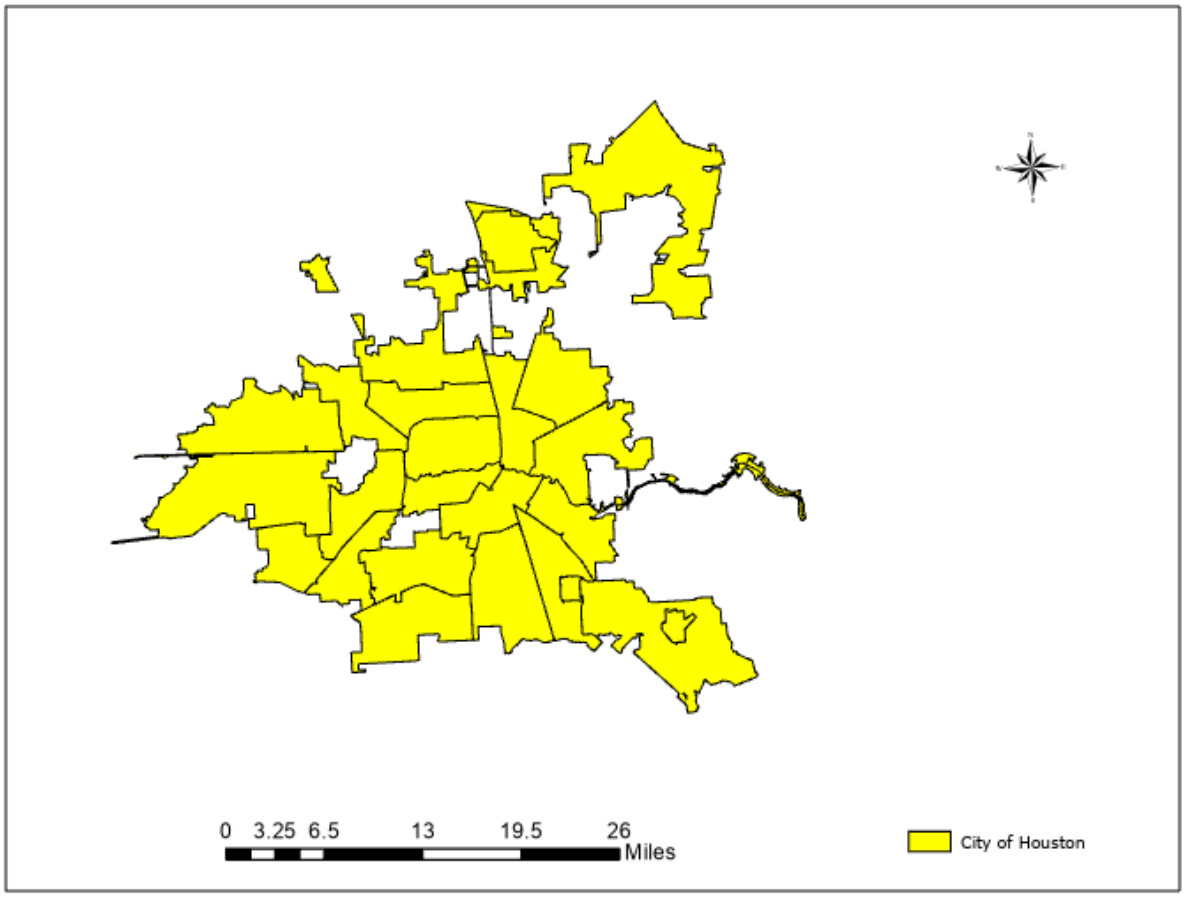

Figure 1: Study Area City of Houston

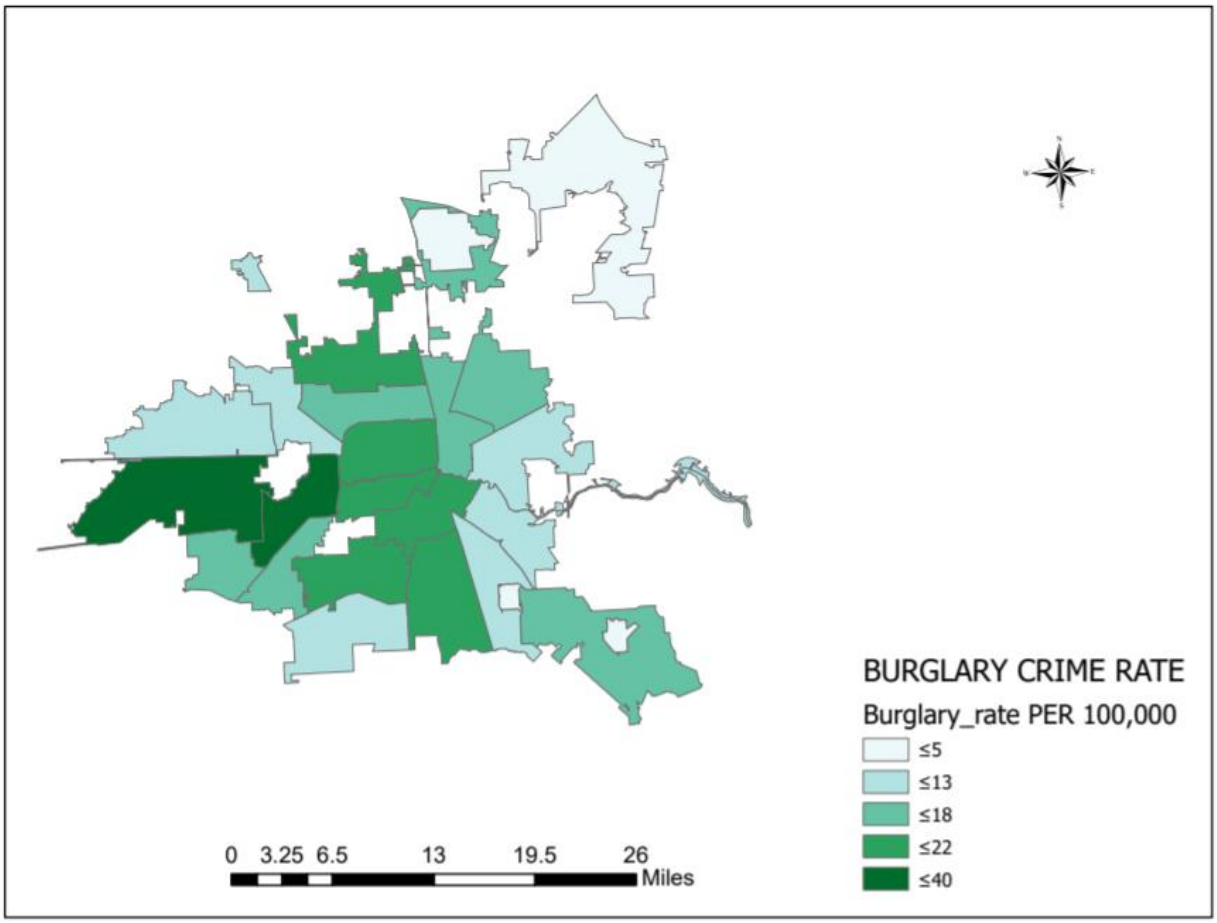

Figure 2: Thematic Map showing Burglary Crime rate per 100,000 Population

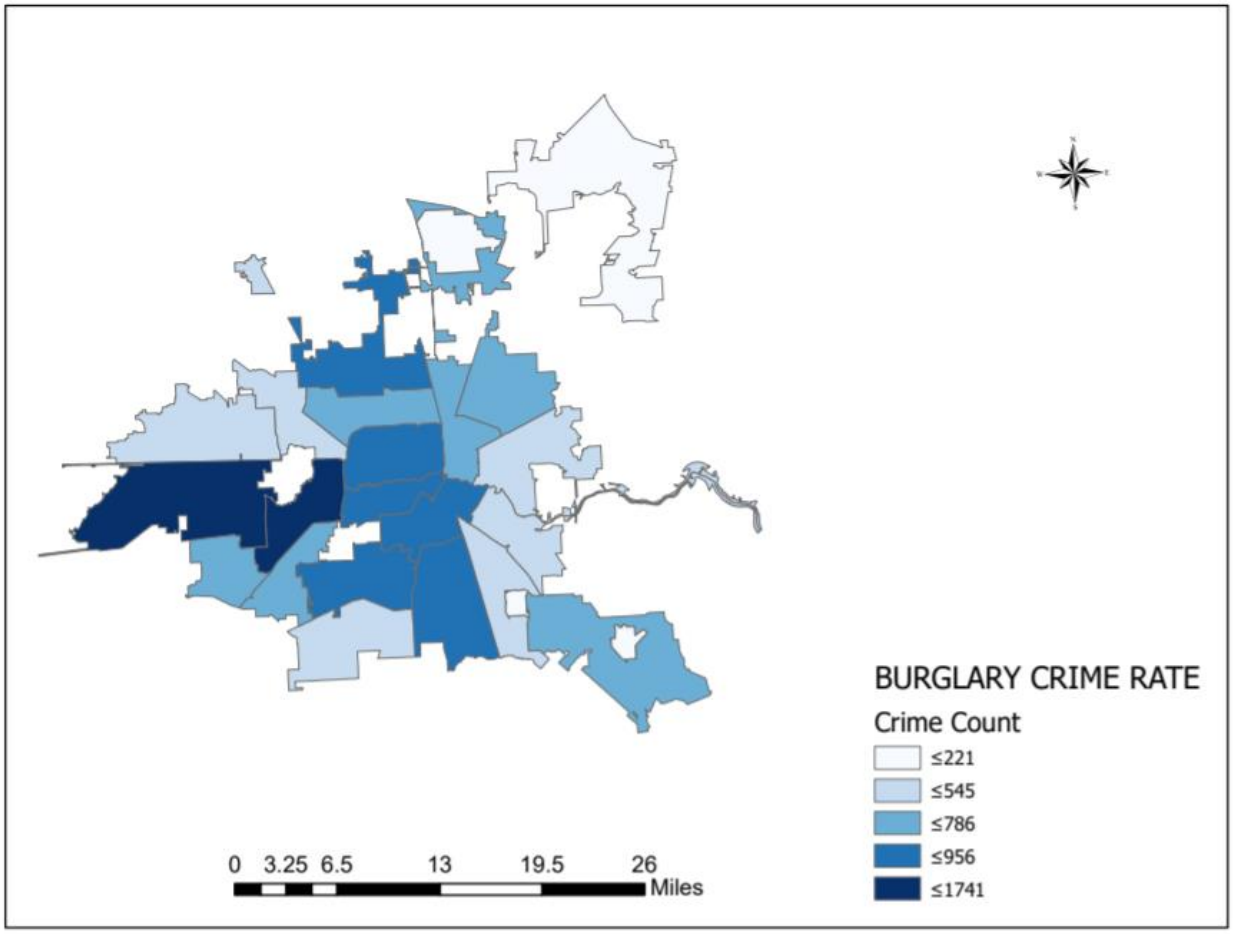

Figure 3: Thematic Map showing Burglary Count in each Houston Police Districts

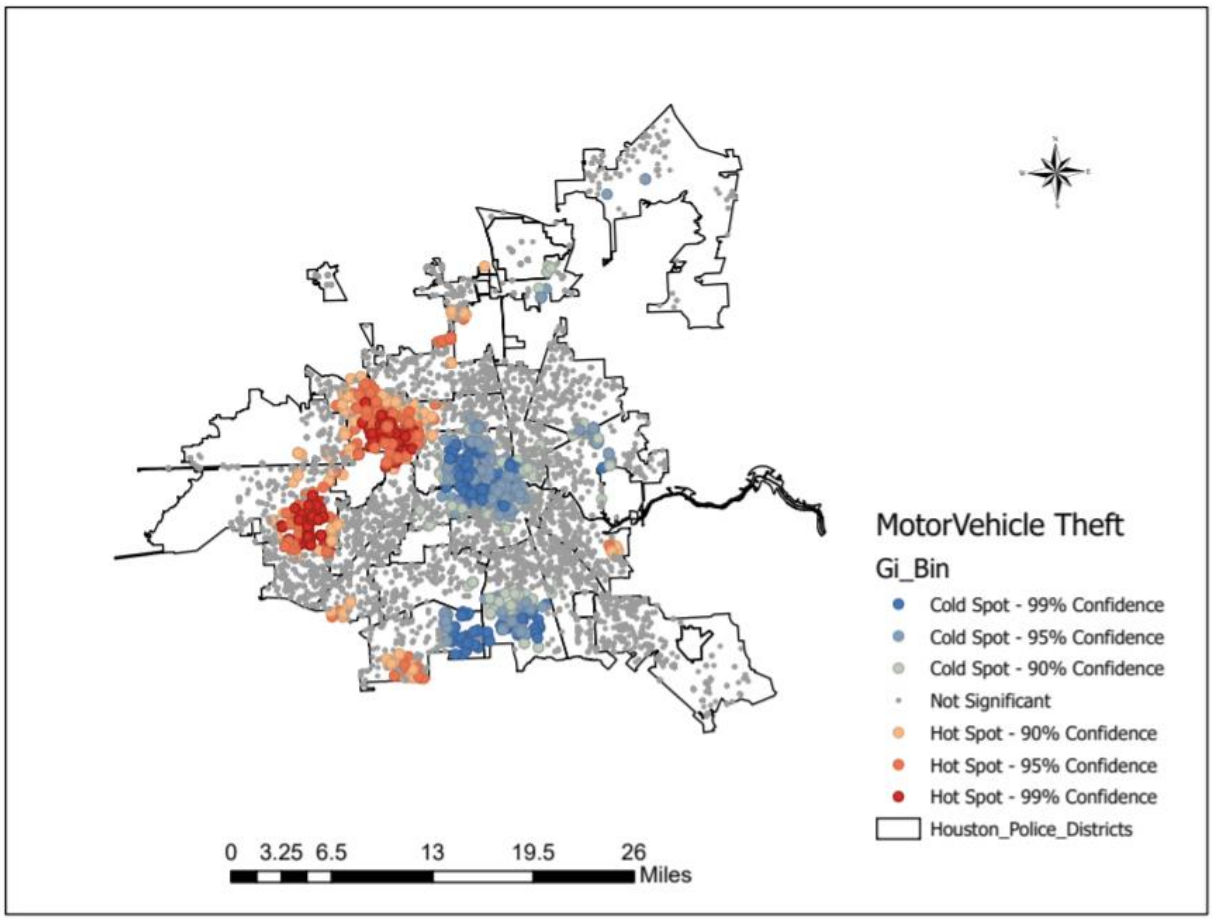

Figure 4: Hotspots of Motor Vehicle Theft in Houston

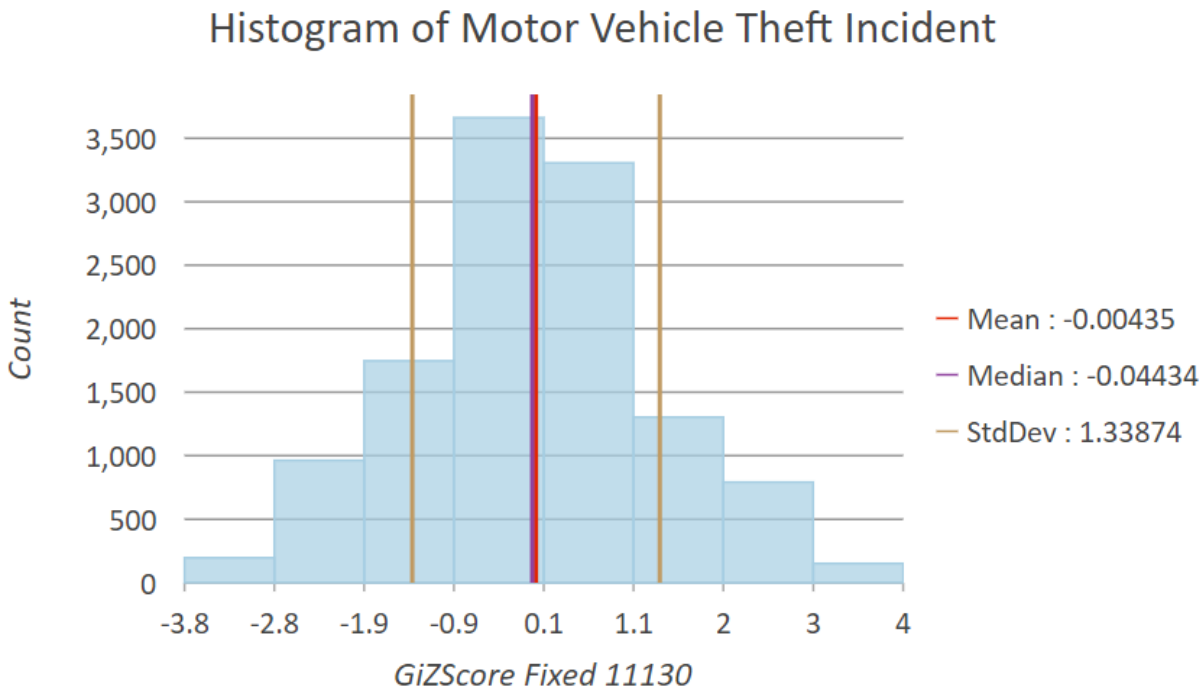

Figure 5:Histogram of Motor Vehicle Theft Incident shows GiZscore values with the count

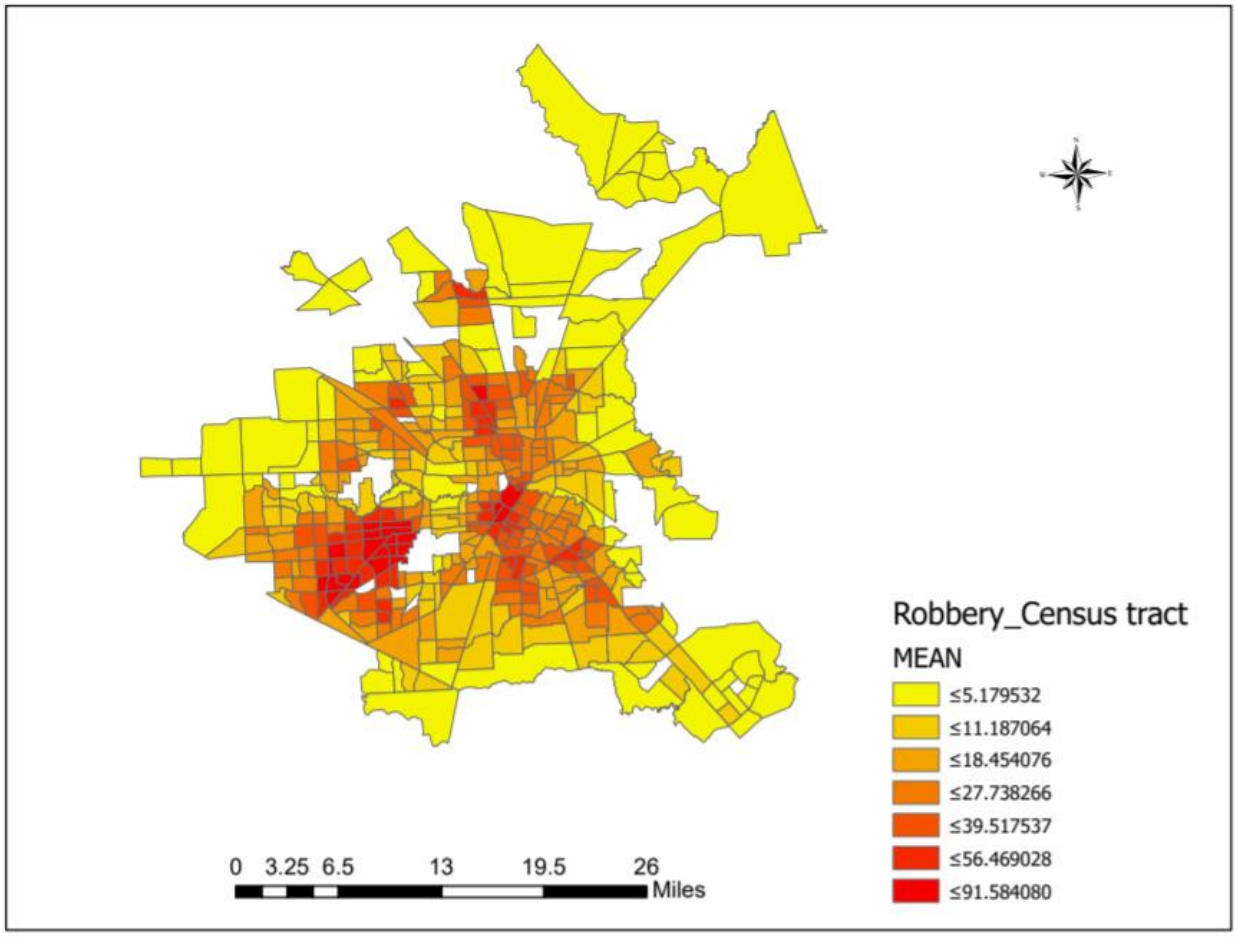

Figure 6: Map showing Robbery Hotspots in Census Tract by using Zonal statistics tool.

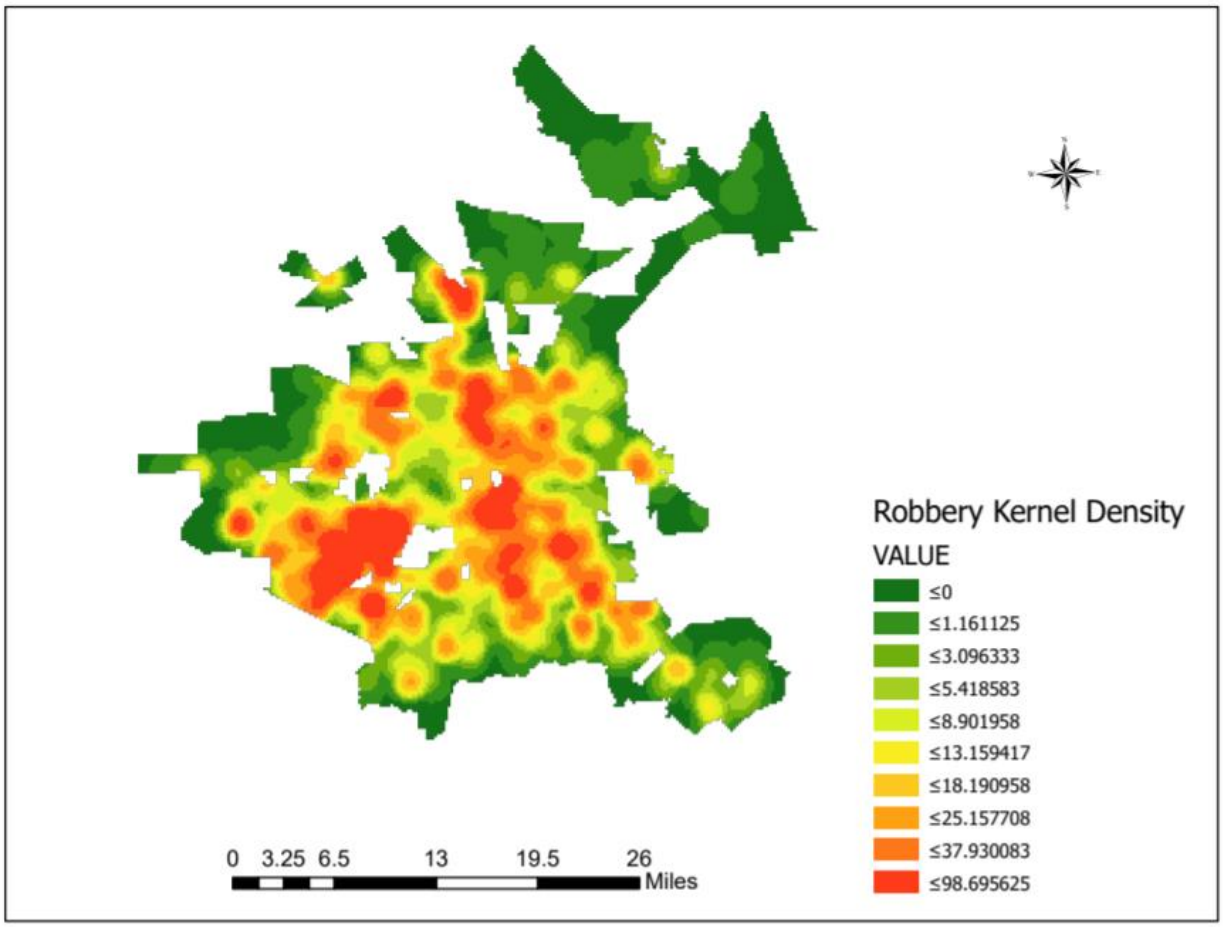

Figure 7: Kernel Density smoothing map showing Robbery hotspots

CHAPTER 6 Conclusion

The purpose of this study is to Identify Various Crime Types in Houston by using hostspot techniques. The results in this study shows that hotspot analysis techniques can be effective in targeting different types of crimes for example thematic mapping is effective to identify crimes rates within geographical boundary while Kernal density mapping is visually appealing and easy to read and understand the map. Getis Ord Gi* is also an effective technique to marked hotspots and cold spots based on Z values and P Values is the probability of hot spot occur or cold spot on its spatial distribution.

Overall results indicate that Crime rate of Robbery was extremely high comparatively Burglary and Motor Vehicle Theft in the year of 2019. These maps can play important roles for law enforcement agencies to target the areas where crime rates are high, and they can create better strategies to minimize these crimes.

References:

https://www.e-education.psu.edu/geog884/sites/www.e-education.psu.edu.geog884/files/image/lesson2/Chainey%20et%20al.%20%282008%29.pdf

https://www.esri.com/arcgis-blog/products/arcgis-pro/public-safety/introducing-the-new-crime-analysis-tools-in-arcgis-pro/

https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html

https://www.researchgate.net/publication/337631809_How_Is_the_Confidentiality_of_Crime_Locations_Affected_by_Parameters_in_Kernel_Density_Estimation