AN ANALYSIS OF EMERGENCY MANAGEMENT AGENCIES GIS DATA FOR PLANNING 9-1-1 SERVICES FOR THE BORDER COUNTIES OF THE LOWER RIO GRANDE VALLEY

A Thesis

by

ORLANDO NINO

BS, St. Edward's University, 1998

BSME, University of Texas Pan American, 2009

Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTERS OF SCIENCE

in

GEOSPATIAL SURVEYING ENGINEERING

Texas A&M University-Corpus Christi
Corpus Christi, Texas

August 2017

AN ANALYSIS OF EMERGENCY MANAGEMENT AGENCIES GIS DATA FOR PLANNING 9-1-1 SERVICES FOR THE BORDER COUNTIES OF THE LOWER RIO GRANDE VALLEY

A Thesis

by

ORLANDO NINO

This thesis meets the standards for scope and quality of
Texas A&M University-Corpus Christi and is hereby approved.

Richard Smith, PhD
Chair

Gary Jeffress, PhD
Co-Chair

Yuxia Huang, PhD
Committee Member

August 2017

ABSTRACT

Texas 9-1-1 entities collect data that can be used to gain knowledge on 9-1-1 call patterns and trends. It is ideal to know these issues too efficiently maximize monies appropriated to local emergency management agencies. This research will analyze 9-1-1 calls and two police departments calls for service using Geographic Information Science (GIS). This information provided by multiple emergency management agencies to better plan 9-1-1 services in the border counties of the Rio Grande Valley. The data that was used is recorded daily by these agencies. Many emergency management organizations keep their own GIS data in-house and they need to be analyzed with 9-1-1 data. Analyzing various GIS databases can yield better results, in finding correlation within each dataset, hotspots and inadequate data collection. These results can then be used by the organizations that contributed the data and shared to all stakeholders.

INDEX WORDS: 9-1-1, GIS, 9-1-1 Calls

DEDICATION

I would like to thank everyone who has inspired me to pursue this dream and making it a reality. Also, would like to dedicate this thesis to the hardest working man and woman I have ever had the privilege of knowing my parents Onesimo (Deceased) and Emma Nino for showing me with their acts how to dream big. Now I am inspired more by my beautiful wife and new baby, Illiana Alejandro – Nino and Reynaldo Elias Nino.

ACKNOWLEDGEMENTS

I would like to thank Prof. Richard Smith for his support and understanding in helping accomplish this program. I would like to thank Mr. Gomez 9-1-1 Director (Deceased), Mercedes PD, and Donna PD for being accessible and allowing me to have a sample of their data.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES PAGE

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Purpose of Study

There is a need to know how 9-1-1 calls patterns and volumes are geographically distributed in our area in order to maximize emergency services efforts and personnel. These findings will allow the more efficient use of emergency response resources. This study helps decision makers better allocate funds and resources to relieve the already burden 9-1-1 system (9-1-1 Industry Alliance, 2011). This information can also be used to petition State and Federal agencies for money to combat crime, better understand call volumes, and determine area of high calls for service. This study proposes a method of collecting and analyzing 9-1-1 calls and local emergency response agencies' data in order to determine patterns, hotspots, and if there are any correlations between 9-1-1 calls and calls for service.

For this study GIS data was collected from local law enforcement agencies to determine call locations and volume patterns. Next a series of analysis based on the data collected was conducted on the 9-1-1 call data versus the calls for service. A workflow was developed to periodically update this study. The local 9-1-1 entity provided years of raw data that has never been analyzed. This data has attached X and Y coordinates, time, and responding agency. Due that the local 9-1-1 entity, the Lower Rio Grande Development Council 9-1-1 Department, works closely with responding agencies, both sets of data can be analyzed in the same time period.

Law enforcement agencies that respond to initial 9-1-1 calls keep statistics on whether any emergency units were dispatched to where the call originated. This data is logged to electronic form from this reports can be generated. This is the data that was requested for this study. This data might have geographic coordinates or it may have attached to it an address or a caller's location. This information can easily be geocoded to determine coordinates.

9-1-1 Background

9-1-1 Emergency Services are based on the public's premise that when you dial that number help will quickly arrive. As an industry, 9-1-1 implementations are not all the same, meaning that 9-1-1 call taking is different from region to region. The difference is so great that even neighboring cities within a county cannot transfer emergency calls to one another. The industry is trying to modernize the almost forty-year-old copper wire 9-1-1 system (Harld 2012, Intrado 2014). This effort is being called NextGeneration 9-1-1. This is a multi-national effort to modernize and establish seamless communications between all emergency service providers. This effort will be based on Internet Protocol (IP) Technologies, which are web based phones and mobile devices.

Even though NextGeneration 9-1-1 is being vigorously pursued by all industry entities, it is years away from being realized (9-1-1GOV 2015, Intrado 2013, NENA 2013, Brown 2015). This study makes an effort to establish a baseline work flow and data analysis of existing technologies and practices that will remain in the local 9-1-1 entities and emergency responders for years to come. This study's results will give a keen insight to the local PSAPs (Public Safety Answering Point) on their call volume loads and hot spots. This is invaluable and it has not been done locally. Knowing where these hot spots are and the times call volumes go extraordinarily high will give Telecommunicators a better understanding on how to get help where it is needed. Having extra emergency responders at times and location where they are needed the most will save lives. Therefore, the results of this study will greatly help the local emergency respond entity that implement this study findings. Using data sets from the Lower Rio Grande

Development Council 9-1-1 Department and two local Police forces, this study will provide the results of data analysis in the hopes that it will accomplish inter-local understanding of call volumes and patterns.

This study is conducted in the understanding that the 9-1-1 industry is too vast and different depending on the country, state, or local management agency one resides in. This study is conducted under the State of Texas rules and regulations. Furthermore, it was conducted under the LRGVDC 9-1-1 Departments jurisdiction. The LRGVDC 9-1-1 jurisdiction are Hidalgo and Willacy Counties in deep South Texas, it is the Southernmost border between Mexico and The United States.

CHAPTER 2

LITERATURE REVIEW

Conceptual Framework

To understand the 9-1-1 system there are three key areas that compromise this service. The first are the entities that oversee the system, in other words policy makers and local 9-1-1 policy enforcement entities. The second are the technology providers which provide the software, hardware, data repositories, and networks. The last and most important players in the 9-1-1 services industry are the 9-1-1 Telecommunicators which wear two hats, answering and dispatching the calls as well as other law enforcement duties.

In Texas, the overseeing 9-1-1 agency is the Commission on State Emergency Communications, (CSEC). This is the state entity that is given the legal authority to collect 9-1-1 fees on every telephone communication service. This 50 cent monthly fee is collected on behalf of CSEC by the Texas Comptroller's Office and distributed to each 9-1-1 local entity based on population. This monthly fee is the source of all 9-1-1 services. CSEC is also the state agency that develops and enforces 9-1-1 rules and regulations. They work closely with the Federal Communications Commission to enforce communication devices and networks to provide 9-1-1 services.

Locally, there are three different types of overseeing agencies: Council of Governments (COGs), 9-1-1 Communications Districts, and Home-Rule Cities. At the inception of 9-1-1 in Texas, the decision was made to place local entities under local COG offices. These are regional organizations that are established to bring services otherwise local city and county governments cannot afford or provide. It also keeps duplication of services in a local setting that, if not for the

COGs, they would cost the State an enormous amount of extra money. There are currently nineteen COGs that oversee 9-1-1 entities in the state.

Communication Districts and Home-Rule Cities are large-enough cities or counties that can maintain and have enough local revenue, in the form of 9-1-1 fees, to provide their own 9-1-1 system. In all, there are thirty-two different Districts and Home-Rule Cities in Texas. Like the COG, they are the local enforcement and policy making organization that oversees, maintains, and pays for the service. They also answer to CSEC and the FCC when it comes to policies and regulations. The other three major components of these organizations are Public Education, maintenance of Certifications, and training for all 9-1-1 Telecommunicators.

The second group of 9-1-1 players are the software, hardware, and network solution providers. There are two major components of 9-1-1 services: the data flow and the voice component of an emergency call. Each step of these two components is made possible by a service, software, or hardware provided by solution providers. Multinational communications firms, such as AT&T, West Communication, TCS, Time Warner Cable and others provide most of the components needed for 9-1-1. Only the networks and technicians are locally needed to run the 9-1-1 System.

The third and most important participant in 9-1-1 are the Telecommunicators, those that answer emergency calls. This duty is assigned to local police forces, such as city police departments, county sheriff's departments, and other local agencies. By State Law, these police forces are required to answer primary emergency calls and are given the authority to develop their own rules and regulations. However, in January 2014, the Texas State Legislature passed a bill regulating telecommunicators and transferring the rule and regulation authority away from the local police forces to the State. The law established licensing with a required minimum number of training hours, courses, and examinations to be able to answer 9-1-1 calls.

Local 9-1-1 authorities work very close with police forces to maintain the 9-1-1 System so it is operational 24 hours a day, 7 days a week, 365 days a year. They only insure the system is running, and the software used is up to date. They offer the training requirements and audit the local Public Safety Answering Point (PSAP), once a year. Typically, local 9-1-1 entities go beyond what is required to insure the system is operational and the Telecommunicators are trained to offer the public high-quality of service.

That being said, the truth comes down to the performance and results of the 9-1-1 System no matter what efforts are made to insure the public gets prompt and accurate service. Due to the Telecommunicators numerous duties, many times the 9-1-1 System fall short of the goal which is getting first responders where they are needed in prompt time (Kelly 2015, Fitts 2015). A typical telecommunicator fields calls from the pubic, both emergency and non-emergency calls. They greet and instruct walk-ins, people walking into a police station with all types of problems as well as other non-9-1-1 related issues. They also are the primary radio communication operators, which involves running the State and Federal background check and license check for Police, Fire, and many times, EMS services. They also use this system to check if people have insurance or arrest warrants. It is common for Telecommunicators to support search warrants and criminal investigations.

In addition to call taking, a very important duty is to assist officers' safety by running background checks, dispatching backup, knowing officers' whereabouts, and assignments like street patrol, undercover duties or investigations. Typically, an officer will be dispatched to an incident, conducting a traffic stop, at an incident, or on call. Any one of these can be a matter of life and death if a 9-1-1 Telecommunicator (Operator/Radio Dispatcher) does not perform their duties as trained. So, due to this fact many times 9-1-1 call duties fall second in their duties and do not perform as expected.

This secondary duty, 9-1-1 call taking, is different at each PSAP. There are eighteen PSAPs, within the two counties that the LRGVDC oversees. Each one has different Standard Operational Procedures (SOPs), due to the type of management they run and the size and resources of the city. Even though they are many differences in management and practice, they all usually have the same duties and have to accomplish the same duty, which is getting help where it is needed. For this study, there were two PSAPs that allowed their data to be examined. They are Donna Police Department and Mercedes Police department.

Many times the 9-1-1 Systems fall short of the goal which is getting first responders where they are needed in prompt time (Walker,2015; Meyer, 2003; Anderson, 2013;Kelly 2011). With all the duties Telecommunicators have at hand, it is complicated even more as a region due that at any time any one PSAP must work together with any adjacent police, fire or EMS entities. There are eighteen PSAPs, within the two counties that the LRGVDC oversees. Each one has different Standard Operational Procedures (SOP's), due to the type of management they run and the size and resources of the city.

9-1-1 Call Flow

To understand the 9-1-1 system, there are three key areas that compromise this service. The first are the entities that oversee the system, in other words, policy makers and local 9-1-1 policy enforcement entities. The second are the technology providers which provide the software, hardware, data repositories and networks. The last entity is the 9-1-1 Telecommunicators which perform two duties: answering and dispatching the calls; and other law enforcement duties. Other duties that Telecommunicators are assigned is fielding public calls, taking initial police reports, helping crime victims that walk into a police station, or even keeping track of police officers schedules.

This study is based on analyzing data from three different sources. All three datasets are recorded from emergency systems. The first data set is from the LRGVDC 9-1-1 System. This is the hardware and software that records the calls being made when people dial 9-1-1. The 9-1-1 call taking solution that the LRGVDC 9-1-1 system uses is provided by Intrado's 9-1-1 company called Positron. They provide the system called Viper. This is their type of 9-1-1 System that records all the landlines and data flow that goes through the system. For this study, the data was recorded in the form of an XML file. This file then was converted to a Microsoft Excel spreadsheet to be later loaded in ArcMap for analysis.

When a 9-1-1 call is made, the voice goes through dedicated Public Telephone Lines, a reminisce of the once used long distance telephone systems. They were designed to record customers phone numbers for telephone companies to charge a toll for jumping from local telephone company to the next. The data portion of the all uses the person's telephone number to search a database with all the landline numbers in operation. They are attached to the billing and location information which is a house or business location. This database is called the Master Street Addressing Guide (MSAG). This MSAG is now incorporated to Computer phones, which are known as Voice Over Internet Protocol telephone solutions (VoIP).

For cell phones the system uses a Mobile Switching Unit, to capture the GPS coordinates of a cell phone. There are two different technologies that one has to understand for this study. The first are the Hand Based solutions that acquire the latitude and longitude from the cell phone directly. The second is a Tower Based solution that uses two or more cell phone towers to triangulate the position of the cell phone. These two technologies send the information to the Mobile Switching Center and delivers the Latitude and Longitude to the PSAP.

Both solutions take the same progression of steps to establish a coordinates. The system knows what cell phone tower a call originates from, so the coordinates of the Tower are immediately sent to the system. This is called Wireless Phase I (WRLS). When the handset or triangulation is done on the device the coordinates are retransmitted to the system and the precise position is displayed at the PSAP and the systems plots the call on a map, this is called Wireless Phase II (WPHII). This is useful when the call taker is not able to talk to the person and establish a location of the call, they can be confident that the coordinates presented are accurate and he can dispatch help.

Unlike a VoIP or Land Line 9-1-1 emergency call, a physical address is not recorded by the system, only the WRLS or WPHII coordinates. In all cases, the time of the call, the length of the call, the position and the PSAP fielding the call are recorded. Many other parameters are recorded as well but were not used for this study.

When preparing the data for it to be displayed with ArcMap the Wireless calls needed little manipulation. WPHII calls gave an exact location of a call. For many known and unknown reasons some cell phone calls never go WPHII so the calls will stay as WRLS and will plot at the tower of the originating call. The main known causes are that calls are made inside buildings and GPS functions will not work. Tower based solutions cell phones will not go into WPHII when in rural areas. There are not enough towers in close proximity to establish a location. However, due to this technological limitation, the main reasons calls of either technology don’t go WPHII is that Telecommunicators transfer the calls in such a short time that the systems do not recorded the WPHII information. Even though this is often the case, this information that often goes unrecorded is invaluable to this study.

Landline calls will always have a physical address associated to it. Every land line call is searched through the MSAG and all the physical location information is recorded by the system. In some rare occurrence there might be a mistake or omission on the address and they are corrected as part of the work flow of the local 9-1-1 entity. This is great information to have, however, when it is brought into ArcMap, all of the addresses have to be reverse Geocoded. Luckily, the LRGVDC 9-1-1 Department is also tasked as the region's authorized addressing entity much like many other 9-1-1 departments in the State. As explained in the study part of the whole analysis was to develop a geocoder and accurately plotting these calls.

Computer Aided Dispatch (CAD)

There were two CAD datasets used in this study provided by Donna PD and Mercedes PD. They both provide one year of call data recorded in their Computer Aided Dispatch (CAD) system. This is the software that they used to record all the activities they undertake during a Telecommunicators/Dispatchers shift. This includes case numbers for request for service which include police duties such as tickets, criminal cases, civil cases, fire dispatch, EMS dispatch, walk-ins and any calls relating to police functions. This also includes 9-1-1 calls that were dispatched. The CAD system records audio of all calls made to a police station including 9-1-1 calls. However, it does not record 9-1-1 data, due that the 9-1-1 Departments provide the call taking system. So the data that relates to 9-1-1 calls in directly typed in by the telecommunicator. There are many errors and omissions and some are not even recorded due to human factors.

Both sets of data from Donna PD and Mercedes PD were provided in the form of PDFs. Due to their rules, they could not provide me with electronic versions of the data. Both supervisors that provided me with the data did not know how to download reports in any other form but PDF. The information recorded into the CAD is usually the start and in many cases part of legal and criminal evidence. In a few cases it is the only evidence that they have to put criminals in jail or exonerate the innocent.

CHAPTER 3

METHODOLOGY

WORK FLOW

The following are the steps taken to complete this study. This workflow will be expanded further in the general outline of the methods section of the study.

1. Obtain Lower Rio Grande Valley 9-1-1 Department data.
2. Obtain one year of 9-1-1 call data from Donna PD, and Mercedes PD. Sample shown in Appendix A)
3. Download census tracts for Hidalgo County.
4. Determine if a LRG 9-1-1 call time and date matches the police dispatch information to determine if a call is a real emergency.
5. Determine high volume events and how they performed versus normal call volume.
6. Perform the statistics and analysis.
7. Calls per Hour day, Month
8. Clustering – Average Nearest Neighbor
9. Kernel density - Hot Spot Analysis
10. Hot Spot Analysis (Getis-Ord Gi*)
11. Conduct Spatial Autocorrelation using demographic data from the Census 2010.
12. Conduct geostatistical analysis with spatial-temporal examination.

This spatial analysis yielded what areas of the police departments jurisdictions are the ones either not using or miss-using the 9-1-1 systems. The normal 9-1-1 process is as follows. When a local resident needs emergency services to any given location within the LRGVDC 9-1-1 jurisdiction they use 9-1-1 in any of the following forms. They either use a cell phone, landline telephone or a VOIP (Voice Over Internet Protocol) phone commonly referred to as a computer phone see Supplemental A. The call is answered by their local responding agency in this case one of 16 PSAP's. In the LRGVDC jurisdiction there are 15 local police station and one fire department that answer 9-1-1 calls. Then the appropriate, law, fire or Emergency Medical Service (EMS) is dispatched to that location.

9-1-1 should only be used in two cases if a life is threatened from a crime being committed, someone is hurt, or for a medical emergency. The other is if property is threatened by fire or a crime in progress. Any call that does not result in an emergency personnel responding to the scene is considered the misuse of the system. If done deliberately, as a prank or to overload the system for criminal enterprise it is considered a crime punishable by state and federal laws (CSEC 2014).

Another type of call is one that is done unintentionally to 9-1-1 which also overburdens the system. With the majority of the calls to 9-1-1 done by cell phones, it is easy to dial 9-1-1 accidently this is known as "pocket dialing". Studies have shown that one of the main causes of non-emergency calls made to 9-1-1 and in large municipalities is the most pressing problem they face as 9-1-1 entities. It overloads the digital systems and puts stress on the Telecommunicators (Perin, 2012; Kahn, 2013). As an example the city of Ottawa received 23,000 unintentional 9-1-1 calls in one year (CBC News, 2011).

Local 9-1-1 entities are responsible for delivering these calls to the appropriate responding agency. Studies have shown that the more efficient call flow goes the better service is given the local population (Jenny, 1993). In order to maximize the system efficiency needs to be increased. In order to increase efficiency, 9-1-1 entities need to know the statistics on the calls they receive. In their normal workflow they analyze their data only using descriptive statistics.

For this study, the 9-1-1 entities' data was analyzed and compared against the calls to service calls. In this case the PSAPs are Donna PD and Mercedes PD.

Historically, there has little been done to analyze the data comparing the 9-1-1 entities' datasets versus the PSAP's datasets (Scott, 2014; Heaton, 2013; Cramer, 2012). There are some studies that have analyzed descriptive statistics on either 9-1-1 data or PSAPs data (Matrix, 2011). This study will not only take it beyond descriptive statistics but will also conduct geostatistical analysis with spatial-temporal examination. This spatial analysis will yield what areas of the police jurisdictions are the ones either not using or miss-using the 9-1-1 systems.

Once the data has been analyzed, it needs to be put into the proper hands. Interoperability is now, more than ever, a critical part of emergency management. It prevents escalations of disasters and protects the general public from significant danger in an emergency. Information needs to flow quickly and accurately from agencies receiving information to other agencies, from agencies and the public, and among the public. This also includes the exchange of information between information and communication systems. This data obtained in this analysis needs to be shared among the 9-1-1 entities in the area which are the PSAPs (Kuehn et al. 2011).

The general overviews of steps for this study are as follows. The first step is to collect all the data that the local 9-1-1 entity has, in this case, it is the Lower Rio Grande Valley Development Council 9-1-1 Department (Appendix B). By State rules, 9-1-1 entities are required to record all voice and data going through their circuits. This is done 365 days a year, 24 hours a day. This also goes for all police and fire agencies that receive 9-1-1 calls. For this study, the voice data will not be analyzed.

The data that was collected from the LRG 9-1-1 was for the period of six years 2008 - 2013. Due to the LRG 9-1-1 system architecture and workflow, only half the data was collected.

Half the data collected within this six years was lost due to a major system update. In other words, there was no information for eight PSAP’s in the studies time range. This lead to the limitation of asking only half, or eight, of the PSAPs for their information. From the eight PSAPs that were asked, only two agreed to give a sample of their data for this study. These two PSAPs were Donna PD and Mercedes PD.

Evaluating the data that the LRG 9-1-1 made available, there was only one year without any omissions or gaps in data. This was mid-2012 through mid-2013. This became the time range for this study. The time range from Donna PD and Mercedes PD were matched. The one-year data is a good sample size to prove the concept for this study.

Both datasets collected for this study have location information. The LRG data has three types of data due that there are three technologies used for the 9-1-1 system. The three types of technologies used for the 9-1-1 systems are cell phones, landline, and voice over IP (VOIP) phones. In terms of reported location, the technologies provide street address, State Plane Feet, and Latitude and Longitude respectively. The three types of location information were converted for to decimal degree to make it easier to enter into a geographic information system. ArcMap for Desktop 10.2 was used to do all the conversions. The Donna PD data had only address information and had to be geocoded. The Mercedes PD data only had a street location, like intersections, or block number due that they don’t keep an accurate record in their CAD system (Computer Aided Dispatch) of addresses. ArcMap 10.2, with an Advanced License, was used for the manipulation and analysis of the data acquired. To clean and prepare the data for entry into ArcMap, Microsoft Excel was used.

LRG 9-1-1 Data

The initial step for this study was to acquire data from the areas 9-1-1 entity and area PSAP's. At the start of this study, there were 16 PSAPs collecting 9-1-1 calls. These PSAPs are fifteen police stations and one fire department. A letter was drafted and sent either via email or hard copy. Due to technical issues, only eight PSAPs had data available for this study. From all the request only two PSAPs shared their information. The other six PSAPs that did not shared their data and gave no reason for not providing the data. The data from Donna PD and Mercedes PD the PSAPs that gave me a sample their data was the one used for this study.

Once the data was collected, there was an initial assessment to assess the quality of the data. The data is usually pulled out of the CAD system in PDF format. They are formatted this way so the data can't be unintentionally changed. All public information requests are handled the same way. The two PSAPs that shared their data was Donna PD and Mercedes PD.

The first data set that was obtained was the LRG 9-1-1 data set from 6/1/12 – 5/31/13. This twelve month time period was the only time frame that did not have any gaps or omissions. The three types of data collected are as follows. The first are all the 9-1-1 calls made from a landline (POTs line – plain old telephone). These calls are geocoded by the 9-1-1 system by looking up the address from the associated for that particular telephone. These address are in the required NENA standardized form seen as follows. House Number, Pre Directional, Street Name, Street Suffix, and Post Directional followed by the Community. For example, this is an address in NENA standardized form:

110 N Main St Donna, TX

2345 S Wilkinson Rd W, Mercedes, TX

All the addresses had to be reverse geocoded. The first thing was to develop an address locator using the street center line base layer that is maintained by the LRG 9-1-1. Being that they are the official addressing authority outside city limits and some small municipalities. All the address points that were obtained from the landline data set yielded 15,082 total locations. 2,070 addresses did not geocode, so they were cleaned to force a geocode. After the second cleaning, 1,803 matched yielding a total of 14,814 geocoded locations.

The second type of data from the LRG 9-1-1 were calls made from a VOIP phone which are also known as computer phones or services. Some examples of computer phones or services are OnStar telematics location data, Magic Jack, and Vonage. This data is presented to the system in the coordinate system of NAD 1983 State Plane Texas South FIPS 4205 Feet. These were converted to latitude and longitude to be merged with the other LRG 9-1-1 data. There were only 474 locations in this data.

The third type of data points are those that come from wireless cell phones. They make up 90 percent of the total call volume for the region. They are already embedded with latitude and longitude location data. Within these points, there are two types of calls those that yield a latitude and longitude: they are called Wireless Phase I (WRLS) and Wireless Phase II (WPHII) calls. These are calls that the system gets location information from. Those calls that are designated WRLS are cell phone calls that do not yield latitude and longitude data. This absent of Lat. Long information is due to the cell phones used do not record location information.

The total number of WPHII calls with Lat. and Long. are as follows:

Type	Count_Type
MOBL	40
VBUS	82
VoIP	17
VRES	71
WPH2	4459
WRLS	9697

Table1 LRG 9-1-1 Wireless Call Numbers

This might affect the analysis part due that only 45 percent of all the cell phone calls have latitude and longitude. This might not be significant due that we are going to ignore calls that are less than ten seconds and the majority of the calls from cell phones are less than ten seconds. Only those calls that are more than ten seconds will be considered for this study. Calls that are less than ten seconds in length are not long enough to receive information with the type of call. Due that there is no location information they were omitted from the study. Not having location information, the system will plot the call from the tower it was received. This would not add any significance to the study. It is a logical conclusion that very short calls, or those that are ten seconds or less, are not real calls (Having been at the PSAPs for hours of observations, calls that come into the center and, when answered, are disregarded as non-calls). These calls typically only ring once or twice. It should take at least ten seconds to do the following answer the call, ask the caller why they are calling, and then transfer the call. These calls need to be analyzed using other methods and analysis. An expansion of this study for future work. NENA and other state agencies follow the standard that 90% of all calls need to be answered within ten seconds (NENA 2006, Mission Critical Partners, 2011). Following this rule if a call does gets answered and hanged up in less than ten seconds then it was not considered as a call for this study.

For this study the calls that were less than ten seconds or less were disregarded as mentioned above. This logic is further substantiated by a 2012 Study by the New York City

Emergency Operations Department that concluded that 38% of their calls were accidental or false alarms (Zafar 2012). Of the total 10.4 million call that they received in 2010 38% of the were 19 seconds or less. They attributed these calls that they designated as short calls to be ‘butt dials’ or ‘accidental calls’. The agency does count them as calls however they designate them as not being a call that would prompt an emergency response. This study only looked at calls that yielded a geographic location and were more than ten seconds long. Standard practice at all the call center is to transfer calls to the appropriate dispatch that are within their dispatch jurisdiction. These calls are of two types. It is either a call that is not within their jurisdiction or is a type of call that they do not dispatch for example a fire call, poison control, or Border Patrol. Any call that comes into a call taking center that is not within their jurisdiction is immediately transferred. This is because it is best to dispatch when you are familiar with the responders and the area of service. There are also calls like fire, or medical that the PSAP does not handle. These types of calls are also transferred.

The following steps were taken to make the data usable. The data that Donna PD and Mercedes PD provided were in the form of a PDF. The PDF data was converted to vector format using Microsoft Excel format. From this type of data file, the data was formatted and prepared to work with ArcMap. Once in ArcMap the files were converted to shapefile format. Then all geospatial analysis and manipulations were done.

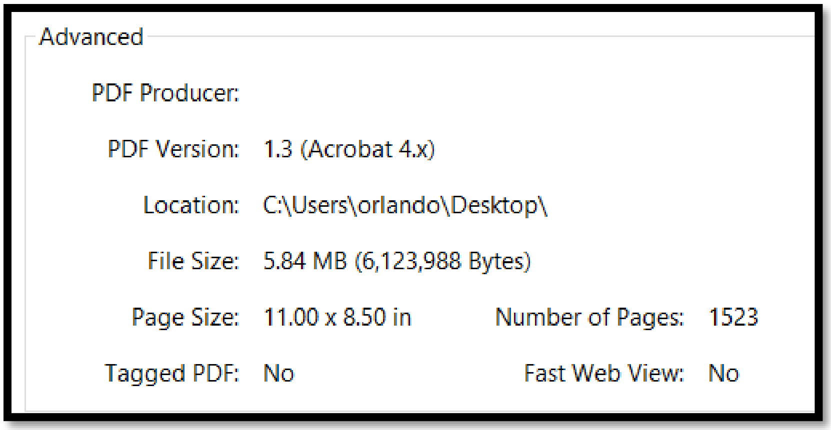


Figure 1 PDF file description for Donna PD Data Format

Donna PD Data

The total amount of records in the data that was acquired from Donna PD was 15,910 records. From these, there has to be a determination whether the calls are valid or erroneous entries. Superseding this fact is the entry has a location value. All location information is entered by the person entering data into the CAD system. If a record did not have location information, then it was discarded.

Mercedes PD Data

The Mercedes PD data also was acquired in PDF format and needed to be converted to Excel format. There were 13,842 records in the period ranging from June 1, 2012 to May 31, 2013. From these there were 11,628 records with location information. If the records did not have location information they were not used for this study.

Data Analysis

Both Mercedes PD and Donna PD datasets were converted to Excel format. Once in this format, it was imported into ArcMap. All the manipulation and geo-processing was done in the shapefile format.

The data from the 9-1-1 department was from the date range of June 1, 2012 to May 31, 2013. This one-year period was chosen due that there was complete data recorded by the system. There were 15,910 records logged from the system from Donna PD. Of those records 14,814 had location information.

Due to the large amounts of data 9-1-1 call taking generate using the spatio-temporal technique gives a better understanding on how call patters behave (Wang et al. 2005) introduced a way to store mass data sets with spatio-temporal data. There are three main models. The first is using a snapshot model that yields a map for each time interval one chooses. Each interval develops a new map for how many times one choses the interval. The second is the space time composite, takes a region of space and goes through subsequent time intervals and presents the data. It is only for that region and the object or region being analyzed does not move. So if there is a hotspot that moves over time this model does not take into account the spatial objects changing over time. The third model is a spatio-temporal object modeling it takes real world data and considers it a set of spatio-temoral atoms. These units can change in temporal and spatial dimensions. It is limited to snapshots of time and can only detect sudden changes in conditions.

Wang (2005) proposes that using census tracts might yield over–sampled results. He found that population counts of these tracts might over estimate an area due to its geographic size and it being sparsely populated. This means that small areas with high populations are not accounted for. They equate census tracts of different size equally thus yielding over estimated results. There is an assumption that we make when considering performing ordinary least-squares (OLS) regressions is that we assume that the criteria that we used to sample is equally distributed all across our study area.

Brimicombe et al. (2006) analyzed criminal data using similar techniques for this study. They ran cluster or hot-spot analysis of criminal data. They found that criminal record keeping usually does not come with accurate or location data. They found that it is typical for geocoded criminal records only have a 65 percent success rate. They proposed a method to increase this hit rate to 91 percent. The method they developed uses four steps to make available data increase it geocoded rate. They used data cleaning, text mining, specially prepared gazetteers,   and novel reconfiguration of data sets. High crime areas are considered to be a high repeat offending and high repeat victimization. These high crime rate areas are thus an accurate predictor of future victimization.

Hot spot maps can be limited due that the reader lacks the knowledge on how to interpret these types of maps. The first step in a hot spot analysis is to geocode the incidents recorded on a police force equipment. They use text address and other local identifiers and run them referencing a street center line with in an address locator geographic base file.

It is ideal that each and every call to service has useable location information. In geocoding, there has to be a location identifier to a database record. It can be in the form of a street address, or coordinates. These spatial identifiers are usually in the form of an X and Y coordinate or an alpha numeric identifier corresponding to a field in the GIS Layer.

In the past, hot spot analysis has been a useful technique that helps predict spatial patterns. However, this might have some inaccuracies due that it does not considered spatio-temporal information. In other words, it does not take time into consideration. Spato-temporal analysis take time as a factor to develop clusters and hot spot analysis. One can see the hot spot move due to time. This is more meaningful for the agencies that analyze this type of data due that they can better predict due to the time of day, or year as these patterns change and move.

Tompson et al. (2009) found that crime concentrations have spatial characteristics that are associated with attractiveness of crime opportunities in location and time where the offenders and victims interact. The author also noted that most preventive measures are made on spatial concentration (hot spots) not taking temporal changes of these patterns. These types of crime analysis found that different types of crimes occurred on the time of day also can be future work of what type of calls 9-1-1.

The following categories were downloaded from American Fact Finder (Aug. 2014). These were datasets were used at the Census Tracts level to run the Ordinary Least Squares (OLS). Ordinary Least Squares does a linear regression to model dependent variable to set of explanatory variables. The results for the four data sets are presented in Appendix C. They are also explained in the Results section of this study. The reason that the nine datasets were selected was due that there were available for the area of the study. They were also complete having no omissions or gaps.

Due to the data made available for this study the data patterns best fit the Census Tracts. There were too many small Census Blocks mostly along roadways that had zero count for the surveys. Do to this issue the Census Tracts were chosen to run OLS. The issue that was encountered later in the analysis, as explained in detail in the results section, were the few census tracts that the data for the survey span. There are only 14 tracts for the data sets for Donna and 7 for the data sets of Mercedes. Therefore this study is done as a proof of concept and an initial survey of the 9-1-1 Data and the Call for Service for both Police Departments data.

Code	Attribute
HD01_VD01	Total Population Estimates
HD011_VD01	Married Total
HD01_VD17	Total High School Diploma
HD01_VD18	Total Obtained GED
HD01_VD02	Total With Public Assistance Income
HD01VD17	Under 18 Years Old With No Health Insurance
HD01_VD33	18 to 34 Years Old With No Health Insurance
HD01_VD66	Over 65 No Health Insurance
HD01_VD05	Foreign Born

Table 2 Census Data Sets Downloads

The Census data obtained were of two types. Census data is published every ten years called the Decennial Census. This data is release for public use in the form of Summary file Data (SF 1). For the Decennial Census the Census Bureau does comprehensive house to house surveys of many criteria, the ones used are listed on Table 2. This census is done to the block level. They include all the communities in the US. Every three and five years they do the American Community Survey (ACS). This means that they go to the macro level communities only those that are over 65,000 people in population or greater (US Census 2014). For this study the 2010 Decennial Census was used to do the Hotspot Analysis and the was 2013 ACS used to run the Autocorrelation.

The Census Block data sets were used to analyze the data sets for hotspots. For all four datasets the Hot Spot Analysis (Getis-Ordi Gi*) was used. The Mercedes PD and Donna PD datasets the calls for service were variable the analyses were based on. For the Mercedes LRG 9-1-1 and Donna LRG 9-1-1 data sets the Index Variable were the number of 9-1-1 calls made. This Hot Spot Analysis yields hot or cold spots given the weighted feature stated above. For this study the calls to service and 9-1-1 calls will be used. Another method that can be used by using the points alone is the Kernel Density. This method uses points, lines or polygons with weighted attributes to determine if there is any cold or hot spots with in the data. In this case the calls for service or 9-1-1 call points are used to calculate the magnitude features using a kernel function(Eck 2008). The software fits the best fit to determine the hotspots.

CHAPTER 4

RESULTS

Data Manipulation

At the inception of this study, by observation it was a known that 9-1-1 calls have patterns and they occur at certain times of the day. This also holds true for calls to service for local police forces. This study's goal was to develop a series of steps, an algorithm, to analyze the two sets of data and truly present statistical data to determine any patters in call volumes or clustering of calls and calls to service. Assuming call patterns and peak volume times might be generally correct, assumptions are not ideal when property, crimes, medical emergencies or lives are at stake. The following are the results of the data collection and analysis performed on the datasets.

The data obtained from the two police departments was in the PDF format. Converting this to usable data consisted of much work and failure. This kept me from advancing in my study. Due to the large amount of data it is best to acquire data with the same data structures. Different organizations might be collecting the same data, however due to program field structures the data are not easily interoperable. This is a common issue that needs to be addressed as a region. A reasonable data model that can be incorporated with CAD systems and the LRGVDC 9-1-1 systems need to be developed. The time it took to be able to use the data in ArcMap could have been lessen if some type of data structure was used all across local agencies (Yaser, 1998).

As stated above for this study both data sets from Donna PD and Mercedes PD were provided in PDF format. The LRGVDC 9-1-1 data was in XML format. The PDFs that I received were close to 2400 paper pages each and were not in vector form. If they would have it would only take me a few hours to convert them to Excel format and import them into ArcGIS software. However, they were not in vector form and I had to use a CDR program to convert a raster text file over to a vector format. In this study it was converting from a raster PDF to Microsoft Excel. Both data sets exceeded 85000 records and 2250 PDF pages. I had to convert them to Excel using an online program. The online service PDFConverter.com was used to conversion from raster to vector worked, See Appendix A. The resulting product had many errors and needed to be fixed. This meant deleting duplicates, matching multiple rows to be one field and making sure each record was not compromised.

For the LRG 9-1-1 Data that was given for this study was in XML format. They were a series of large XML files multiple sets for each PSAP. The first attempt to convert the data to a usable form was to try to read the file from a SQL server. This was done due to the large number of records. Due the Microsoft SQL Server is a paid for system a full version of the software needed to be used. There was an attempt to use Microsoft SQL Server Express the free version of the software however it only limits file sizes of 1 Gig of data. The files that were extracted from the LRG911 system were 4 Gigs and up. This lead me to try open source software to really obtain the full features of the software.

Due to the cumbersome and complexity of the product it was decided not to use this solution. In both cases it took me few weeks to learn how to handle the files. Then it took me another week to learn and manipulate the data. Due that the data was obtained from the 9-1-1 system as a native Microsoft SQL file, .mdf and the accompanying .ldf it was very difficult to find and use the data. The data was acquired from the Intrado, 9-1-1 solution called Viper. (see Appendix A). Being that the Viper system is proprietary all their database schema is closely guarded and made exporting the data very difficult and incomplete.

Next I tried to open the databases using Microsoft SQL and I hit the same hurtles. In both cases I could load partial downloads of the databases but only could see the fields and not the data. The reason I could only manipulate data fields and not view that data is that Positron places an encrypted key on all its databases. Both programs MySQL and Microsoft SQL have limits on the size of the files and each database was over the 4000K limit. I had to spend another month troubleshooting the database and conceded that I had to down load small amounts of data at a time. The data acquired form the system was queried using date ranges. I had to limit myself to download one month at a time form each PSAP. Which was time consuming and I was only allowed to work on my study before and after hours of my scheduled work.

The 9-1-1 system is proprietary to Positron and one has to use its internal program MIS Reports to produce reports. These reports include a record of each call coming and going from each workstation. Each individual PSAP records the data going through its data circuits in a separate database. Once I got the raw databases I tried using MySQL a popular open source SQL database package. I had to learn and become an expert on MySQL within a few days to manipulate each individual database. However this was to no avail due that when the MIS Reports produce a database it is in the form that can only be used by Positron systems. This is done to keep information private and from being disseminated accidently to unauthorized people. This process took me weeks and further delayed my study.

Intrado is the Call handling suite that the LRG 9-1-1 uses. The Viper system is what records all the 9-1-1 call information in the region. The reporting is handled by the Intrado Software Power MIS. It is a proprietary system and the file and table format is not shared. Power MIS the reporting software uses a browser (HTML language) to search and select what data one wants to download. It develops reports in three forms .pdf, .xml and .xls. It would seem that the

.xls would be best however it was found that the Power MIS version 4.1 has a software bug and does not produce Excel format reports. The next best thing was to use Extensible Markup Language (XML) that can be converted to Excel. All the data from the LRG911 department was downloaded as a XML file and then converted to Excel format (Intrado 2014).

Once the data was in Excel format it was imported to ArcMap for initial statistical Downloading the Census Data for this study was straightforward. The data was obtained by using America Fact Finder the website the US Census Bureau provides publicly to disseminate census data (US Census Bureau). When all the data was downloaded and analyzed with ArcMap it was found that 114 Census tracts that have either 9-1-1 calls plotted or calls for service were within those tracts. These 114 Census tracts were used for the study area.

The initial descriptive statistics yield 51,847 total data points. Mercedes PD had almost twice the number of calls for service than 9-1-1 calls, however being at the call taking center for hours of observation and in asking the Telecommunicators it is their policy to log all administrative activities into their CAD system. This data includes general calls to the PSAP from the public that are non-emergencies or that do not require some emergency personnel to be dispatched. For example, a person calling to know a city departments number, or calling for advise on an insurance claim. That person calling gets an incident number. In contrast larger police forces only allow their police officers to issue incidents number. Each incident number corresponds to an entry into their CAD system and the result is perceived calls for service however if one would omit these types of calls the calls to service and 9-1-1 calls would be closer in magnitude. Due to many missing data cells these types of calls were not taken off the analysis. This issue became evidently clear in the Ordinary Least Square analysis due that it failed.

Number of calls per Source				
Department	Lat. & Long.	GeoCode Address	GeoCodeFeet	Total Data Points
DonnaPD		12144		12144
Donna911	12106	485		12591
MercedesPD	3822	22058		25880
Mercedes911	17730		502	18232

Table 3 Total Number of Data Points Location

Once the Mercedes PD data was converted to Shapefile format it was determined that the time of the call was striped. Initially when the PDF were acquired from the PSAP the time of the call was transferred the Excel file. When the Excel was imported into ArcMap the Hours:Minutes:Seconds format was not imported into the shapefile. So for this data set only the location and the date of the incident was useful. Multiple attempts were made to try to make the date import in to the shapefile. Once attempt was to split each component within the Excel however it did not import into the Shapefile. The Excel was also converted into text and ASCII format to see if it import however, this was also to no avail. Finally, a decision was made to move forward in the study without the time stamp.

For Mercedes PD the average calls per day were 58. Like all other entities they are measured in a 24-hour cycle starting at 12 midnight. The Max was 160 and the Min was 22. The most obvious statistic was that the Maximum number of calls was 160. There are also some in the 150 all the way to 100. Upon verifying these high numbers in a single day it was found that these days with usually high number of calls, more than 100, had duplicates. This was the cause of high numbers in certain days. The Average and Daily average are in reasonable range, they can be seen in the Figure below.

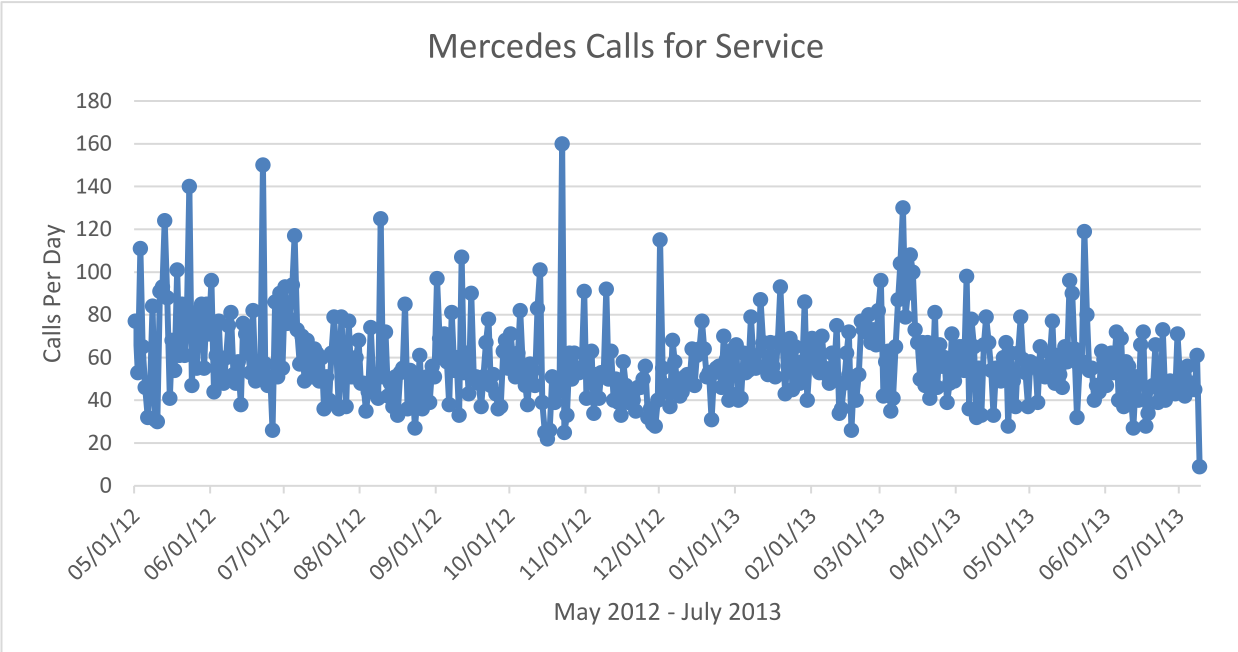


Figure 2 Total number of Data Points Location Per Day Mercedes PD

The Mercedes LRG 9-1-1 data set consisted of an average date call volume of 50. The maximum was 174 and the minimum was 18. There are also a few high incidents of calls that are due to the same issue we saw in the Mercedes PD calls for service, which are duplicates. This data set had the Date and Hour:Minute:Second format that was not very useful to compare the PSAP’s data to the 9-1-1 calls. This issue needs further work to be able to compare the datasets to each other. Then one can run a simple script to determine if a call for service came within seconds of a 9-1-1 call.

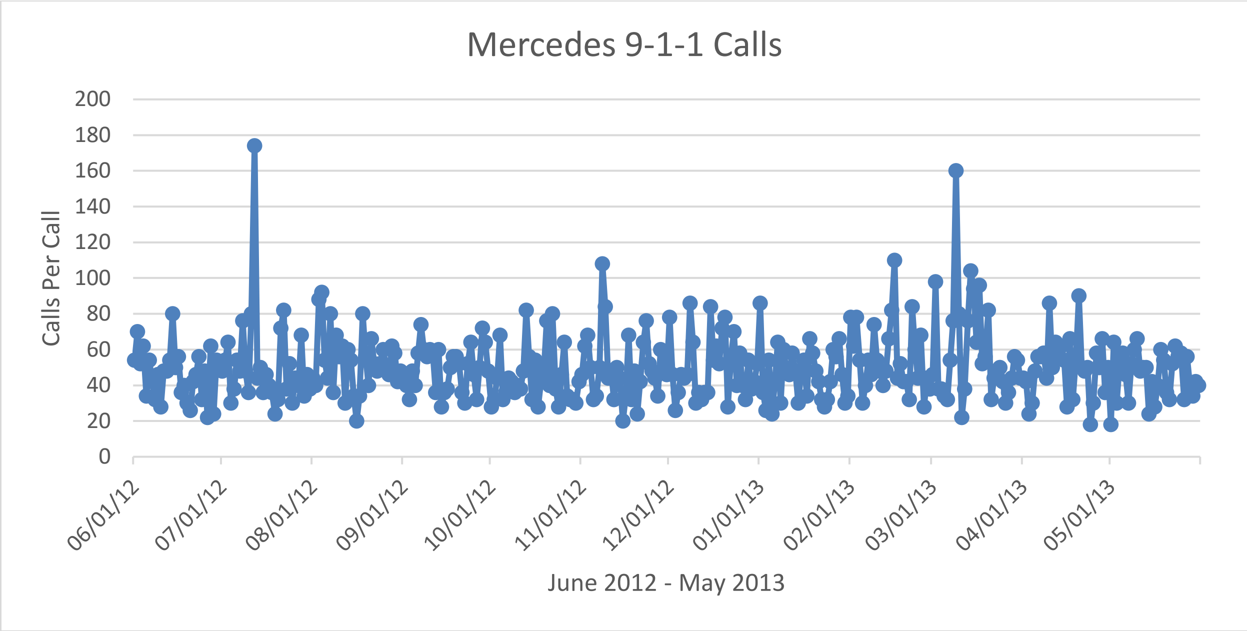


Figure 3 Total number of Data Points Location Mercedes 9-1-1

The Donna PD data set consisted of an average date call volume of 33. The maximum was 66 and the minimum was 13. There are also a few high incidents of calls that are due to the same issue we saw in the Mercedes PD calls for service, which is duplicates. This data set had the Date and Hour:Minute:Second format that has also been solved. This issue was tried to be solved using two different computer systems and was not able to be solved. It needs further work to be able to compare the datasets to each other. Below is a graph plotting each day and the number of calls.

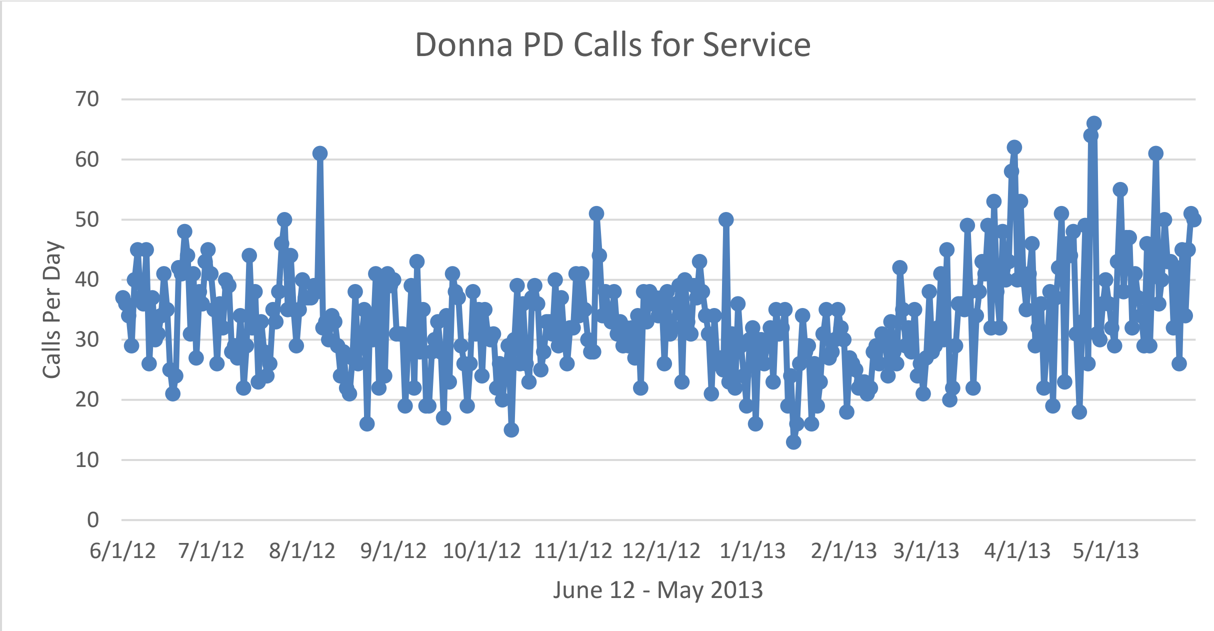


Figure 4 Total number of Data Points Location Donna PD

The Donna PD data set consisted of an average date call volume of 29. The maximum was 57 and the minimum was 11. This is the most consistent data set of the four. They all don't stray much above or below the average. There are none that go beyond max which is 57 calls in one day.

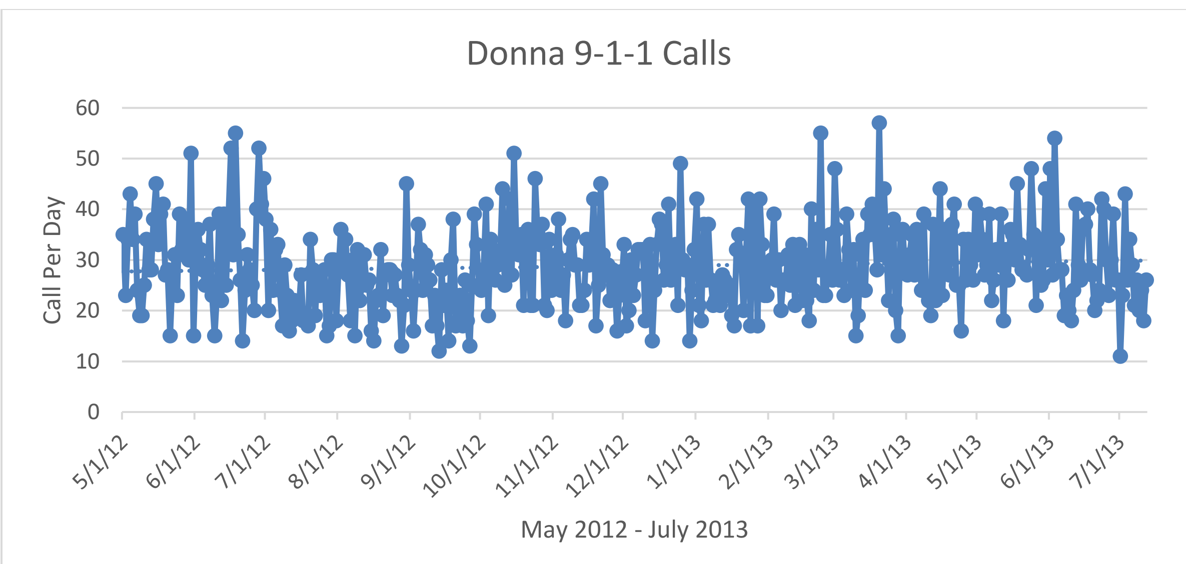


Figure 5 Total number of Data Points Location Donna 9-1-1

Hotspot Analysis

Before the Hotspot Analysis was conducted it is ideal to know if there is any clustering for any of the data sets. An Average Nearest Neighbor analysis was conducted and the results are shown in Appendix B. This analysis takes the average distance of adjacent features and determines if there is any clustering. For all four datasets all showed significant clustering. Due that the clustering was significant the Hotspot Analysis was done.

The Hotspot Analysis was done on all four datasets, Donna PD, Donna LRG 9-1-1, Mercedes PD and Mercedes LRG 9-1-1. The Kernel Density Tool in ArcMap yield good results summarized in the following screen captures. They all yield some hotspots that gave a great insight on how 9-1-1 calls and calls for service fall within certain areas. The first was Donna LRG 9-1-1 datasets which are the 9-1-1 calls within the dates of the study. All hotspots are towers due to WRLS not having Lat. and Long. Coordinates so, it plots at the tower location.

As seen in the overview of the Donna (Figure 6) PSAP jurisdiction there are only four significant hot spots. As mentioned before these are attributed to calls that are WRLS that do not have location information so they are plotted at the tower that picked up a wireless 9-1-1 Call. Figure 7, shows the two significant hotspots, within them there are the location of two cell phone towers. The LRG 9-1-1 Department provided the location of all cellular towers in the region. The cell phone tower location exactly matches the center of the hotspots.

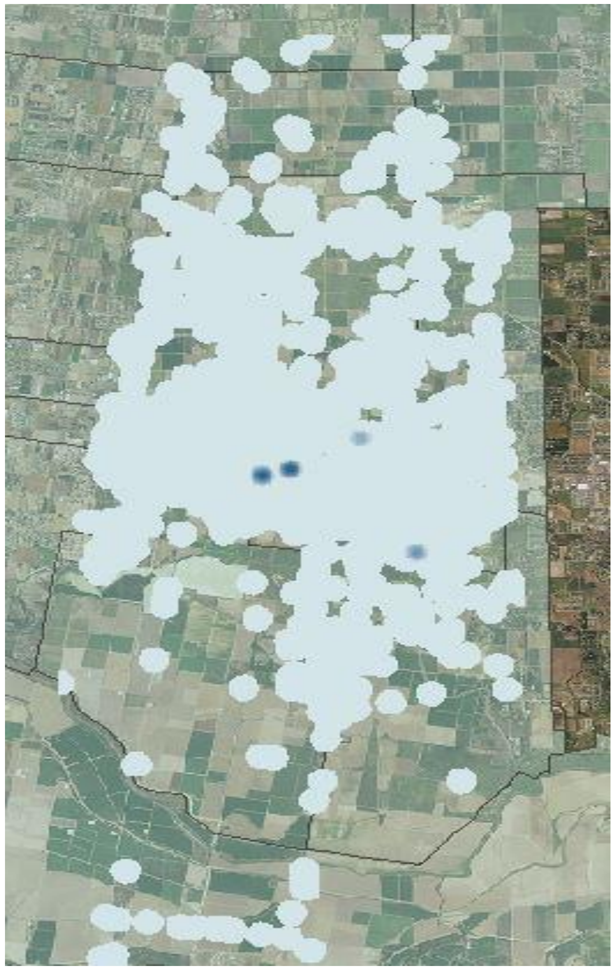


Figure 6 Overview of Donna LRG 9-1-1 Calls

Figure 7 Zoomed into two hotspots Donna LRG 9-1-1

The same Kernel hotspot analysis was done to the Donna PD calls for service. This resulted in more wide-spread hotspots as seen in Figure 8. Most are seen to be at intersections due that most CAD entries are done by entering the block number and not the exact address where an incident occurred. A few were further investigated using Google Earth Pro to get a spatial view of the aerials provided to the public. A zoomed in capture (Figure 9), one can see several points of interest. This is the farthest West of all the clusters. In this intersection one can see Main Ave. and Interstate 2 (I2). This is the downtown area that leads to the heart of the city.

Another is seen in Figure 10 which is the Waterburger on Salinas Blvd. and I2. This is significant due that to the North just a couple of blocks is the cities' Walmart. To the South just a few blocks is the cities' high school. All around the intersection one can see that calls for service are the most significant. Not a surprise due that it is the busiest intersection in the city.

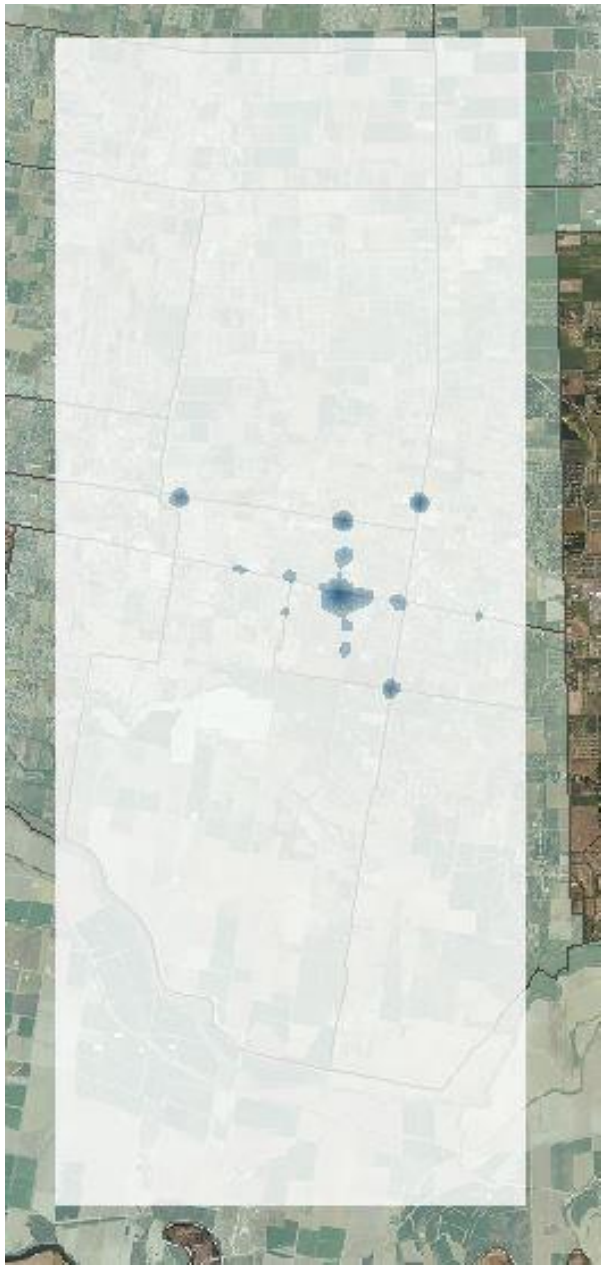


Figure 8 Donna PD calls for service overview

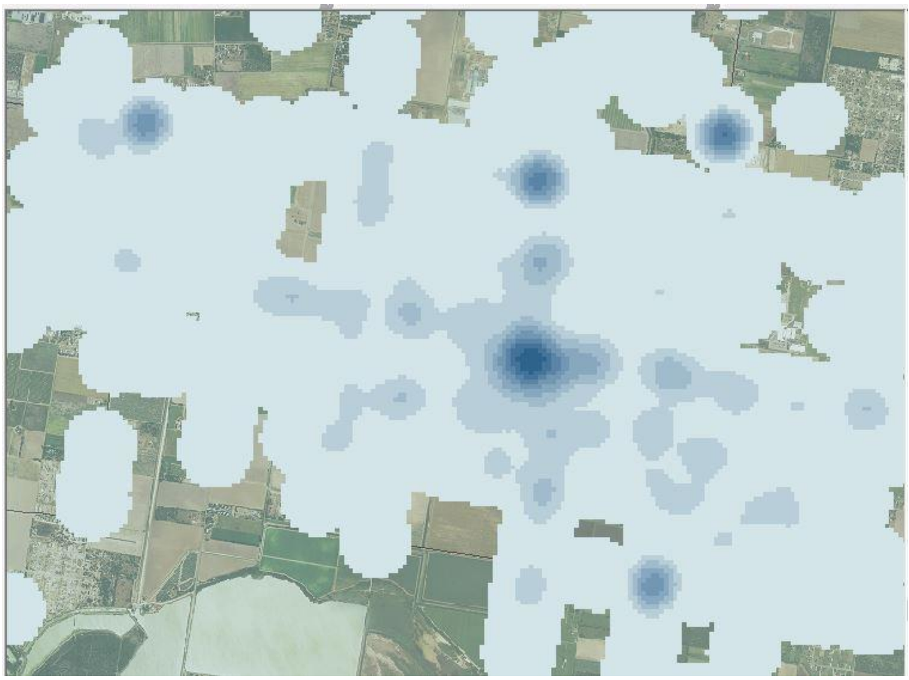


Figure 9 Zoomed in Donna PD Calls for Service

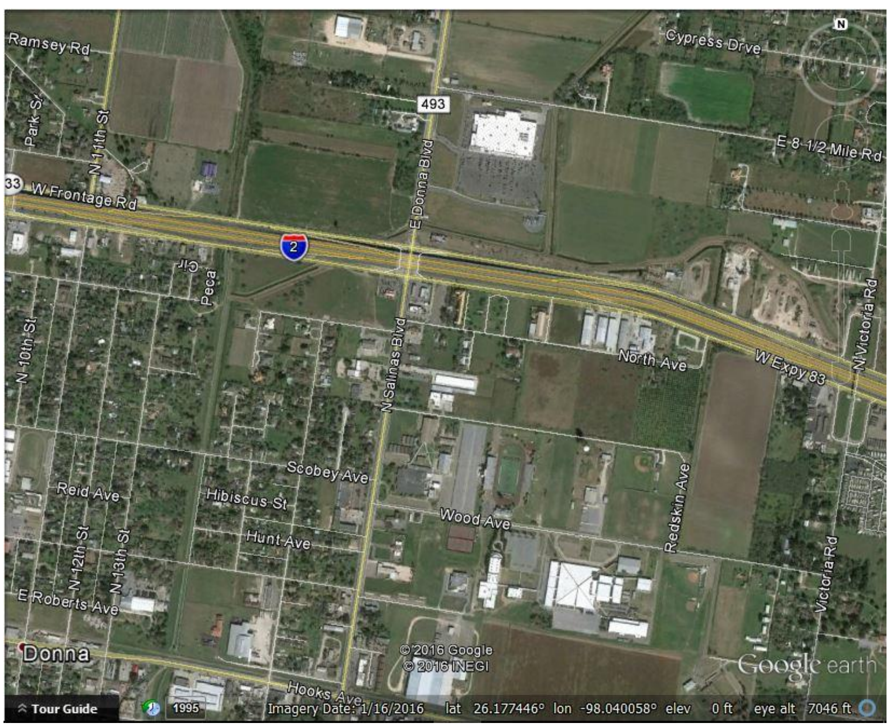


Figure 10 Whataburger on Salinas Blvd. and I2Donna, TX

When analyzing the results from Mercedes LRG 9-1-1 Data the same patterns were observed as in the 9-1-1 calls in Donna PD. However, one hotspot the most Eastern one as seen in Figure 12 it did not have a tower at the center. At closer inspection in Google Earth Pro one can see that its along I2 and the Simon Outlets. It seems the large number of local and Mexican Nationals that frequent these Outlets produce a hotspot due to the number of calls. What type of calls has to be further investigated.

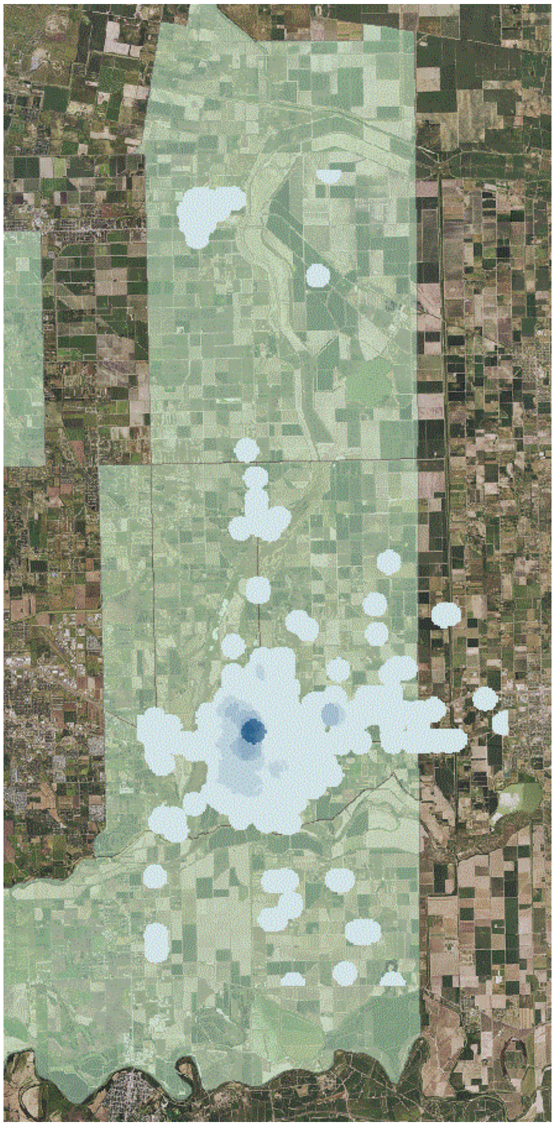


Figure 11 Overview of Hotspots of Mercedes LRG 9-1-1 Emergency Calls

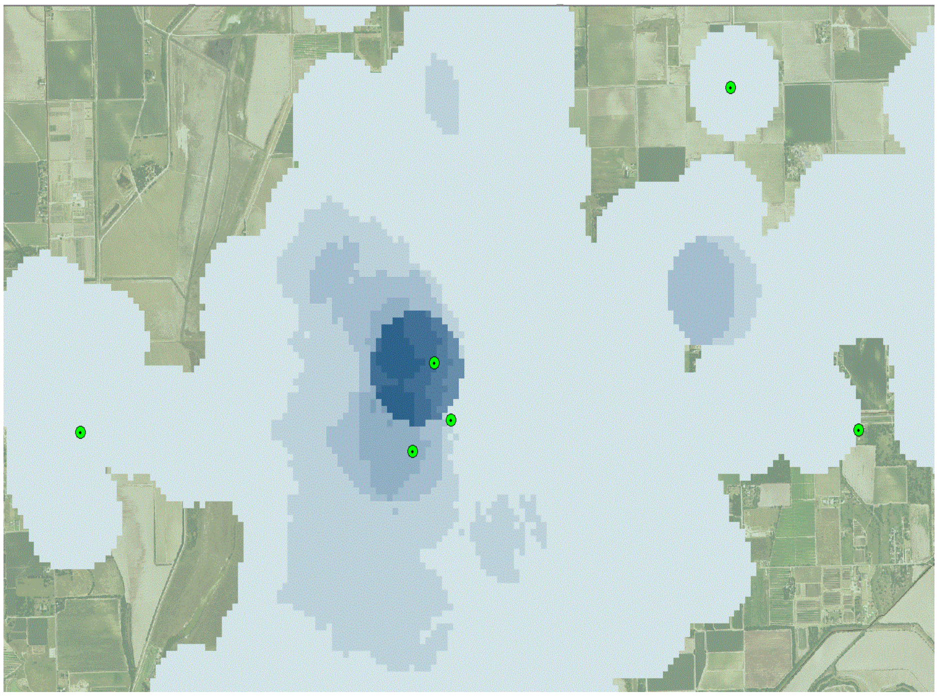


Figure 12 Zoomed in Area of Downtown Mercedes

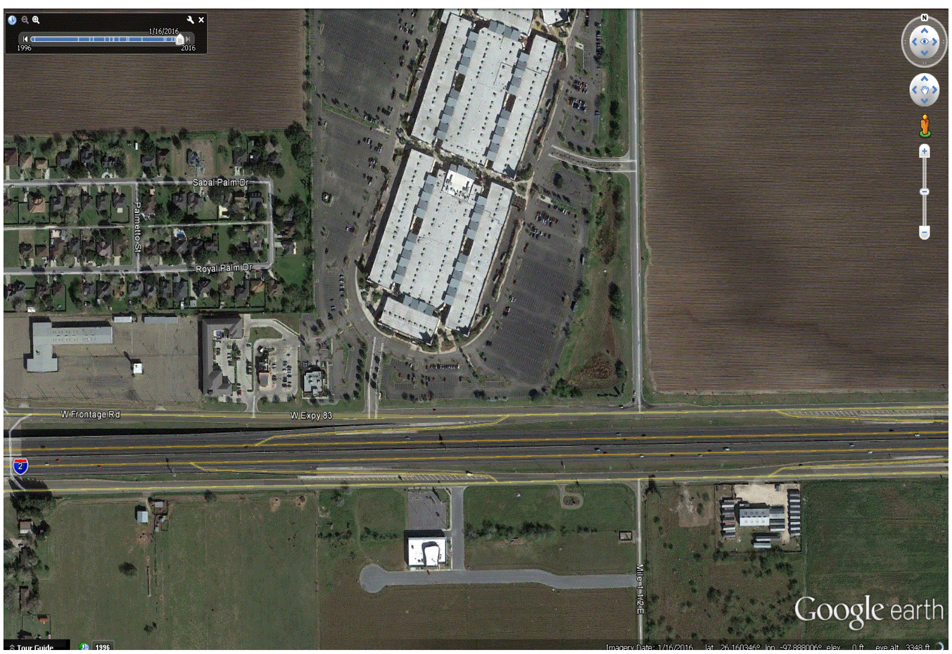


Figure 13 Mercedes, TX Simon Outlets on I2

As before the same analysis was done to the Mercedes PD calls for service dataset. The results were almost identical to the Mercedes LRG 9-1-1 calls. The Simon Outlets hotspot was more significant and was at the same spot. This is why this study is so important. The local Police Forces, EMS and Fire Rescue entities can use this information to station personnel closer to these areas or have local businesses or residents be more vigilant in these areas.

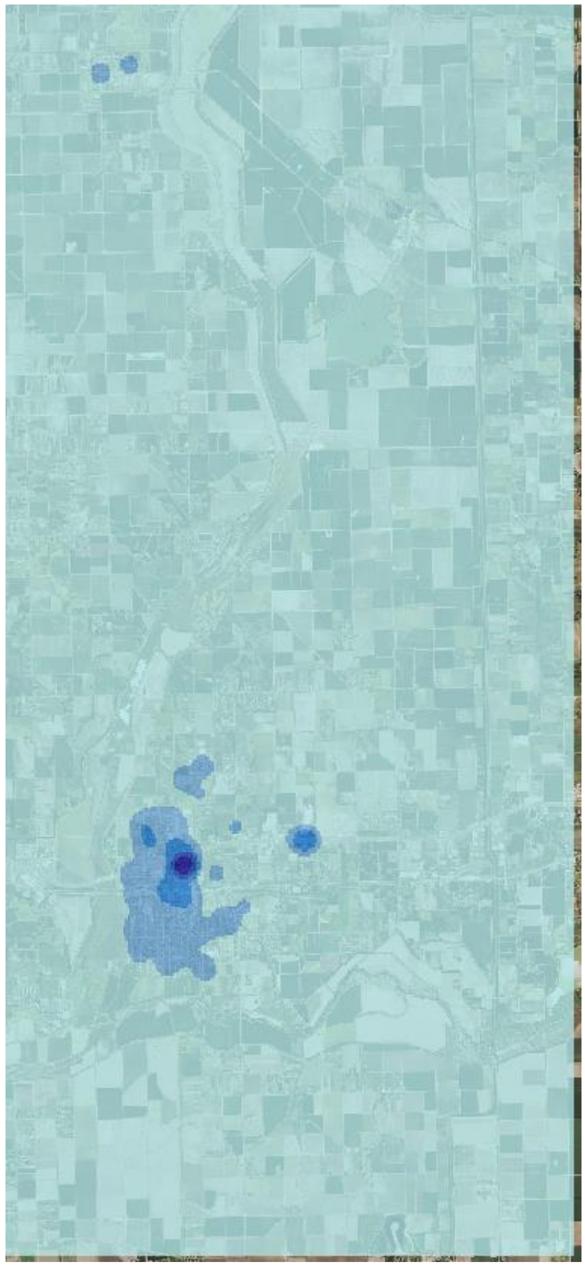


Figure 14 Overview Hotspot Map of Texas Ave Mercedes, TX

Temporal Spatial analysis could not be done for this study due that there was no Time stamp in the shapefile. The day of the call was maintained in the data however the time of day for each call was lost due to the conversion. When the data was import to ArcMap the time stamp format was lost. However, the theory and tools are available with ArcMap and with good data format, no duplicates and no omissions Temporal Spatial Analysis can be easily be done. Then a better insight would emerge as one can see a hotspot during certain times of the year. month or even time of day.

A problem with the study is that many points that were attempted to be imbedded to the Census blocks failed. This was due that most calls for service and 9-1-1 calls fall along the roads and streets. These are the demarcations and boundaries used by the US Census Bureau to determine the boundaries of Tracts and Blocks. A large number of points failed trying to Spatially Merge the points to the census blocks using ArcMap. Another problem is that some Census Blocks have ID numbers but no attributes to them because no one lives along those corridors. As seen in Figure 15 the Calls for Service run along Interstate 2 within a Census Block that no one lives in. So when statistical analysis are attempted for example Hot Spot Analysis (Getis-Ord Gi*) the blocks are incomplete that the script fails.

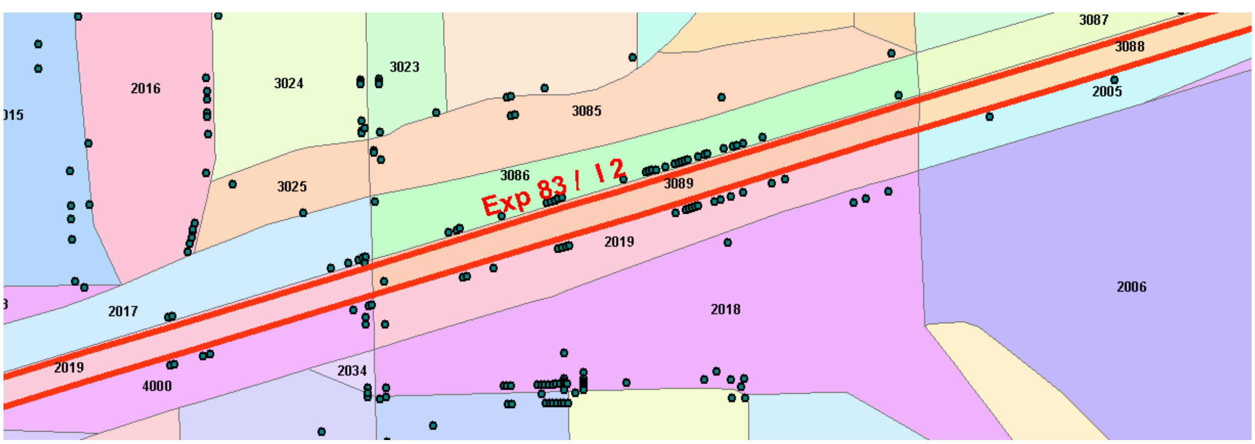


Figure 15 Mercedes PD Calls For Service Along I2

Ordinary Least Squares

In the final statistical analysis, the four sets of data were modeled by Regression. Using Ordinary Least Squares Regression (OLS), the data was investigated to see if there were any spatial patterns and any predictions based on spatial relations. To start the project all files were converted and assigned the same Projected Coordinate System. The original data came with a projection of NAD_1983_HARN StatePlane_Texas_South_FIPS_4205_Feet, so all data was standardized to this projection. The data that was used as explained earlier in this study were those available from the US Census. In the next steps the calls from each of the four data sets were examined for an optimal statistical analysis. Then were also examined to determine how well the results explained the model.

The analysis was done to model a dependent variable to known explanatory variables by the use of ArcMap embedded Ordinary Least Squares Tool. For this study the Independent Variable were the Calls to Service and 9-1-1 calls. The Dependent Variable were all four datasets were treated to attributes in Table 2. These are as follows. Total Population Est., Married Totals, Total Obtained GED, Total with Public Assistance Income, Under 18 years old with no Health Insurance, 18-34 Years Old with No Health Insurance, Over 65 No Health Insurance and Foreign Born. These attributes were chosen due that they are relevant to the study at hand. It is vital due to the demographics that the elderly, foreign born new immigrants and those needing medical services are using 9-1-1 as it should be. All detail findings are shown in Appendix B. Trying to model these nine variables is involved, however if done correctly we can get keen insight on how calls relate to people's real world conditions.

In order for the best model to be developed a series of checks were done. The first was to graph out the four data sets and the Explanatory Variables to see if there was a good trend to apply to the model. The Scatter Plot Graphs did not yield any significant conclusions so the Explanitory Regression tool from ArcMap 10.1 was done on all four data sets.

From this it was found that both data sets for Mercedes failed. It is strongly suspected to be that there are only 7 Census Blocks to work with and the minimal recommended features is 30. The Donna PD and Donna 911 data sets did obtain a result however it is viewed with caution. With the Spatial Autocorrelation all yielded high clustering which makes the model not as reliable as one likes. Looking at the results for both sets of data it was found that the highest Adjusted R2 values and the three models that have the largest Jarque-Bera p-values is the ideal to run.

The results with the highest Adjusted R2 Value and Jarque-Bera P-values were choosing five sets of variables out of the nine. This held true for both sets of Donna data sets. For Donna PD Calls for Service the following 5 variables were Married Total HD011_VD01, Total High School Diploma HD01_VD17, Total Obtained GED HD01_VD18, Total With Public Assistance Income HD01_VD02, and Over 65 No Health Insurance HD01_VD66. For Donna 9-1-1 the results yielded the best 5 variables were Married Totals HD011_VD01, 18 to 24 Years Old With No Health Insurance HD01_VD17, Total With Public Assistance Income HD01_VD02, Total Obtained GED HD01_18 and Over 65 No Health Insurance HD01_VD66.

To begin summarizing the results both the Mercedes PD data and Mercedes LRG 9-1-1 datasets fail to complete the analysis. This might be due that the data has too many omissions as found previously. It might also have been that the Census Tracts are too large in geographic area to get an accurate analysis. There are only 7 Tracts for this data set way too few to run a proper analysis. This means that the variable downloaded from the Census need to be changed. However, I believe that it's a combination off all three issue. The results were omitted from Appendix D.

The analysis for both the Donna PD and Donna LRG 9-1-1 data was a success and offers important results that can be used petition for funding, and better use personnel to streamline emergency services. The Donna LRG 9-1-1 data resulted in four variable having positively statistically significant correlation. These were Married Totals HD011_VD01, 18 to 24 Years Old With No Health Insurance HD01_VD17, Total With Public Assistance Income HD01_VD02, and Total Obtained GED HD01_18. The Over 65 No Health Insurance HD01_VD66 resulted in nonsignificant rob probability coefficient. This model indicates that people are Married Totals have a negative correlation to the volume of 9-1-1 Calls in the area of study. All other had a positive correlation.

Now checking the model in order for the results to be trusted was done and found that the data can be trusted. There were six things that were checked, the first was if the coefficient signs were as expected. They were all positive with the exception of Married Total this was lead to believe that statistic is not that strong. The second check was to make sure that the Variance Inflation Factor (VIF) was less than 7.5. Only two factors were above the Married Total and Total High School Diploma which contributes to bias and yield the same trend. The third check was to for statistically significant Coefficients and it was determined that the Koenker (B)Statistic was not significant one can determine that the significant Probabilities help the model. In this case the statistics stand due that the Koenker (BP) Statistic is not significant. The fourth check was to determine if the Jarque-Bera Statistic was statistically significant and it was found that it was not this also indicates that the model is trusted. The fifth check was the models performance. This is done by checking the Adjusted R-Square value to be from 0 to 1. This explains how well the dependent variable explains the model. In this case, it is does well by accounting for 0.72 of the variables. The last check is to determine by using Spatially Autocorrelated Regression Residuals are not clustered. As shown above all four sets of data had significant clustering. Knowing this the model still performs well and can be trusted.

The Donna PD OLS results can also be seen on Appendix D. This model does perform as well as the Donna 9-1-1 Calls however it is a start and a good benchmark for future models. The results yielded four explanatory variables to be statistically significant. These are Married Total HD011_VD01, Total High School Diploma HD01_VD17, Total Obtained GED HD01_VD18, and Total With Public Assistance Income HD01_VD02. Those that are Over 65 No Health Insurance HD01_VD66, the model determined that they were not significant. Again the six checks were done and they are as follows.

The first check was to determine if the signs were as expected and they were. The only variable that had a negative correlation was Married Total HD011_VD01. The second check was to determine if the Variance Inflation Factor (VIF) were less than 7.5 and three of the variables passed this check. The third check was to statistically significant coefficients by checking the Koenker (BP) Statistic was significant. The model found that it was not so the model variables are statistically significant thus making reliable. The fourth check was to determine the Jarque-Bera test to be not statistically significant and the model determined that it was not. This check determines if the residuals are normally distributed and they are making the model non-biased. The fifth check is to see if the Adjusted R – Square value is between 0 and 1, ideally being greater than .5. For this model it was determined that the Adjusted R-Square value as .579 which is good performance. This tell us that the model explains more than half of the explanatory variables. The last check is to determine if there is clustering of the residuals. After running the spatial autocorrelation as the same for Donna 9-1-1 calls it was determined that there is clustering. However, due to the results of the project it can be said that the model can be trusted. After running both data sets through the OLS tool it is determined that there are patterns in the data. The full results can be seen on Appendix D.

CHAPTER 5

CONCLUSIONS

As a region the Lower Rio Grande Valley needs to have all data forms interoperable. This means be able to exchange data from a single record or whole days of data for easy analysis or investigations. There are already useful data formats already developed like 9-1-1 NextGen format (NENA 2014). As this study showed two different entities the LRG 9-1-1 and two PSAPs cannot interchange their among each other. The conclusions of this study have shown this. It also has proven that analysis can be done however having the same data structure and data file format would tremendous ease the process and cut on the time it would take to run the analysis.

Large data sets provided for this study made it difficult to conduct the study. There has to be a way to provide the data for future work or update this study in a manageable file size and type. What is proposed is a limit of 1 gig files in the form of XML or Excel. This would make the data statistics easily summarized in Excel and then use ArcMap to conduct the geospatial analysis.

Hotspots attributed to towers having a large number of calls on their location are due to WRLS calls that plot to the tower due that they have no geographic coordinate information only addressing information. Those 9-1-1 calls that fail to obtain a geographic coordinate are plotted using the address that is automatically assigned the address for that tower. True hotspots as seen in the results can be attributed to major intersections and places with a large congregation of people. Overall the calls follow the roads due that they are plotted to the nearest known address which is the block.

Even with these limitations the Hotspot Analysis proved very valuable to identify areas of high calls to service and/or 9-1-1 calls for emergency help. The steps in this study have been proven that this method can yield the valuable information easily. Good knowledge has been gathered from the algorithm proposed and with a few recommended implementations can make this method a golden standard to follow. These steps are easily replicated and followed to get a good sense of where the calls for service are plotted.

What was found from the Hot Spot Analysis is that the LRG 9-1-1 data yields hotspots at locations of wireless towers. There are some that are actual hotspots as seen when it comes to large commercial areas, or other public areas where masses of people congregate. The PSAP calls for service also yield hotspots for the same reason, which is large number of people getting together at a small geography location.

When attempting to conduct some modeling Ordinary Least Squares, both data sets Mercedes PD and Mercedes 9-1-1 failed. The Mercedes PD data can be easily understood why it failed due that the data had many missing or incomplete information, and that the data had too many duplicates. However, the Mercedes LRG 9-1-1 data set failed as well. This failure is unknown and might be several issues. One hypothesis is that many of the calls are WRLS, meaning that they have the coordinates for the tower in which the call originated from. These clusters might have not allowed the ArcMap OLS script to yield no results. These results can be seen in Appendix C.

CHAPTER 6

FUTURE WORK

Cooperation with local police departments that field 9-1-1 calls need to have common codes to input into their CAD systems. There is a need to standardize codes for all PSAPs in order to make their data more interoperable. The data structure for each type of data also must be defined. For example, the use of Vehicle can describe a truck, minivan, car or wagon, specific terms need to be used across all agencies. This would facilitate the algorithm proposed in this study much easier and faster. So, each CAD system would record incidents with the same data structure and codes. Then each system would have to be able to export the data in the same format and share across the region.

A geometric network of roads needs to be developed with the purpose of conducting this study. There is already effort to address this issue across the area however its purpose is to address and not support studies like this one. This is due to the assumption that the majority of calls for service and 9-1-1 calls fall along streets. The reason that the calls fall along the streets is because the addressing is done to the block address and the calls plot along these street ranges. The Roads Layer which can serve as our geometric network needs to be manipulated to facilitate the replication of this study.

Analyses of the LRG 9-1-1calls with no location information need to be investigated. This problem needs to be solved due that if there are real calls for emergency services lives will be at stake. Further OLS modeling can be done on other attributes found in the census data such as Minority Populations, Housing Units, and Race. This can yield other insights that were not evident in this study.

Having proven this study through it sets the framework to expand the analysis to larger cities and even do the whole region with its three counties. Donna and Mercedes can be considered somewhere in the middle, Midsized semi-urban. This study must be conducted in larger urbanized areas and small rural communities. With its proximity to the border it allows a unique set of demographics that can model as the future communities in the United States.

References

1. 9-1-1 Industry Alliance. Industry Council for Emergency Response Technologies “The Overloaded 9-1-1 System. Status Report. October 5, 2011.
2. 9-1-1.GOV, “2014 National 911 Progress Report” National Highway Traffic Safety Administration National 911 Program, Washington, DC, March 2015.
3. Abbas, R., Michael, K., and Michael, MG.(2009). ‘Demonstrating the Potential for Covert Policing in the Community: Five Stakeholder Scenarios.’ University of Wollongong.
4. U.S. Census Bureau; Census 2010, Excel Downloads; generated by Orlando Nino; using American FactFinder; <http://factfinder2.census.gov>; (September 2014).
5. Anderson, Rob. "Dial 911 from a wireless phone? Emergency responders may not be able to locate you." GEO World Oct.-Nov. 2013: 26.
6. Barr, JL. Et. Al., Dept. of Homeland Security. (2010). ‘Gap Assessment in the Emergency Response Community.’ Richland Wa. Pacific Northwest National Laboratory.
7. Beconyte, Giedre, A. Eismontaite, D. Romanovas. (2012). ‘Analytical Mapping of Registered Criminal Activities in Vilnius City.’ *Geodesy and Cartography.* Volume 38(4): 134-140.
8. Brimicombe, J. Allan, Lily C. Brimicombe and Yang Li. “Improving geocoding rates in preparation for crime data analysis” International Journal of Police Science & Management. 2006, 9 (1), 80 – 93.
9. Brown, Justine, “Kansas Latest to Implement Statewide Next-Gen 911,” Government Technology. Solutions for state and local government. July, 2015, http://www.govtech.com/public-safety/Kansas-Latest-to-Implement-Statewide-Next-Gen-911.html.
10. CBC News. "Pocket dialers keep 911 dispatchers busy." The Canadian Broadcasting Corporation 11 Jan. 2011.

11. Champion J. Dean and Richard D. Hartley. (2010). *Statistics for Criminal Justice and Criminology, 3ed Edition*. Columbus, Ohio, Prentice Hall.
12. Cheng, T., Williams, D. (2012). "Space-time analysis of crime patterns in central London." International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, B2.
13. "Computer analysis of 911 calls from Calif. wildfires offers potential early warning system." Science Letter 11 Mar. 2008: 4400.
14. Cramer, Dan, Brown A. Albert and Gongzhu Hu. Predicting 911 Calls Using Spatial Analysis. Springer-Verlag Berlin Heidelberg. 2012.
15. CSEC - Commission on State Emergency Communications. http://www.csec.texas.gov/9-1-1# 2014.
16. Eck, John etal. Mapping Crime: Understanding Hot Spots. National Institute of Justice. 2008 Washington DC.
17. Fitts, S. Alexix. "When 9-1-1 Operators Can't Find Their Callers." The Atlantic. Nov. 19,2015.
18. Harld. "Turning Off The Phone System? What Do You Mean We're Turning Off The Phone System?" Wetmachine. June 4, 2012. Web Blog.
19. Heaton, J. Matthew etal.(2013)." An Analysis of an Incomplete Marked Point Pattern of Heat-Related 911 Calls". Journal of the American Statistical Association: Vol. 110, Is. 509: 123-135.
20. Intrado Inc. Public Safety - Call Handling http://www.intrado.com. Industry Website. 2014.
21. Intrado Inc. (2013) "NEXT-GENERATION 9-1-1: THE ESSENTIAL GUIDE TO GETTING STARTED An Introduction to Next-Generation 9-1-1 Terminology, Concepts and Considerations." 2013. Online Publication.
22. Jenny, Donna L. An Evaluative Policy Analysis of Municipal 911. Order No. EP21573 Kutztown University of Pennsylvania, 1993 Ann Arbor.

23. Kahn, David. "Phony 911 callers shutting down critical emergency communications." Fire Chief 1 Aug. 2013.

24. Kelly, Caitiin. "Making the right call: dialing 911 might be one of the most important things you do, so be prepared to do it right." Boys' Life Feb. 2011: 32.

25. Kelly, John, Brendan Keefe. "9-1-1 deadly flaw: Lack of location Data." USA Today, 24 March 5, 2015.

26. Kuehn, A., Kaschewdky, M., Kappeler, A., Spichiger, A., and Riedi, R. Interoperability and Information Brokers in Public Safety: An Approach toward Seamless Emergency Communications. Journal of Theoretical and Applied Electronic Commerce Research. Vol. 6: Iss. 1, pp. 43-60, 2011.

27. Matrix Consulting Group. "Report on the Consolidated 911 / Dispatch Feasibility Study." July 2011.

28. Medina, M. Richard, L.K. Siebeneck, G.F. Hepner. (2011). 'A Geographic Information Systems (GIS) Analysis of Spatiotemporal Patterns of Terrorist Incidents in Iraq 2004–2009.' *Studies in Conflict & Terrorism,* 34:862–882.

29. Meyer, Dan. "911 calls can fail to connect. (News)." RCR Wireless News 13 Jan. 2003: 4.

30. Microsoft Online source. Features Supported by the Editions of SQL Server 2014. http://msdn.microsoft.com/en-us/library/cc645993.asp.

31. Mission Critical Partners. "Recommendations for Establishing and Maintaining a Quality Assurance Program Related to PSAP Quality Assurance." March 2011.

32. Murray, Jeff. "9-1-1 cell callers can be tough to find." Star Gazette. Dec. 24,2014.

33. Murray, T. Alan, T.H. Grubesic, R. Wei, E.A. Mack. (2001), 'A Hybrid Geocoding Methodology for Spatio-Temporal Data. *Transactions in GIS,* 15(6): 795-809.

34. http://www.pdfconverter.com/pdf-to-excel-online-free/

35. National Emergency Number Association "9-1-1 Statistics." *National Emergency Number Association*. 2016. https://www.nena.org/?page=911Statistics.

36. NENA: Call Answering Standard/Model Recommendation. Doc. 56-005. June 10, 2006.

37. NENA: The 9-1-1 Association. "NG9-1-1 Project," *National Emergency Number Association*, 2013, http://www.nena.org/?NG911_Project.

38. Perin, M. (2012). Butt calls, eye-rollers, disbelievers and scowlers. Officer.Com, http://pathfinder.utb.edu:2048/login?auth=utblib&url=?url=http://search. proquest.com/docview/1423367459?accountid=7119.

39. Piett, T. (2015). Security personnel and the importance of understanding 9-1-1 dispatch operations. Security, 52(1), 28-28,30.

40. Riedel, Marc. (2000). *Research Strategies for Secondary Data*: *A Perspective for Criminology and Criminal Justice*. Thousand Oaks: Sage Publications, Inc.

41. Scott, Jason, Annette Patricia Strickland, Karen Warner, and Pamela Dawson. 2014. "Describing and predicting frequent callers to an ambulance service: analysis of 1 year call data." Emergency Medicine Journal: EMJ 31, no. 5: 408-414.

42. Tompson, Lisa and Michael Townsley. "(Looking) Back to the Future: using space–time patterns to better predict the location of street crime". International Journal of Police Science & Management 2009: 12 (1), 23-41.

43. Underwood, Sarah. (2010). 'Improving Disaster Management.' *Communications of the ACM*, 53, 19-20.

44. U.S. Census Bureau. A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know U.S. Government Printing Office, Washington, DC, 2008.

45. Wang, H, Tang, X, Lei, B., and L.Zhai. Spatio-temporal data model: An approach based on version-difference.

46. Wang, Fahui. “Job Access and Homicide Patterns in Chicago: An Analysis at Multiple Geographic Levels Based on Scale-Space Theory” Journal of Quantitative Criminology, 2005, 21 (2), 195 – 217.

47. WALKER, LAUREN. 2015. "Not-Even-Close Calls." Newsweek Global 164, no. 15: 19.

48. Yaser Bishr. 1998. Overcoming the semantic and other barriers to GIS interoperability, International Journal of Geographical Information Science, 12:4, 299-314, DOI: 10.1080/136588198241806.

49. Zafar, Aylin. “Almost 40% of New York’s 9-1-1 Calls Are ‘Butt Dials’”. *Time* (May 12, 2012): http://newsfeed.time.com/2012/05/10/almost-40-of-new-yorks-911-calls-are-butt-dials/

Appendix A
Sample Data

Call Detail

Created :	05/05/2010 17:37:40:543
PSAP :	LRGVDC EPD Host

Type :	Incoming
Trunk :	San Juan PD 9111
Line :	San Juan PD 9111
Phone :	(866) 866-5006
DNIS :	
Incident Type :	No Value
Call Taker :	Position 81 - San Juan PD
Position :	POSITION81
ALI Status :	1 : GoodAliWithoutXY
TTY Info :	No

Duration (hh:mm:ss.sss)

Setup Time : (Incl. ANI Decoding)		00:00:02.000
Process Time :		00:01:30.907
Answer :	00:00:14.422	
Talk :	00:01:16.485	
Hold :	00:00:00.000	
Trunk Seizure Time :		00:01:32.907
Call Taker Answer Time :		00:00:14.422

Caller Information

Caller Name :	ONSTAR TELEMATICS CALL CENTER
Location :	, VOIP CALL, SAN JUAN, TX
Phone :	(866) 866-5006

Call Information

Customer Name :	ONSTAR TELEMATICS CALL CENTER
Location :	, VOIP CALL, SAN JUAN, TX
Community :	SAN JUAN
State :	TX
County :	
Zip Code :	
Common Place :	
Cross Street :	VERIFY CALLERS NUMBER VERIFY LOCATION
Class of Service :	VOIP
ESN :	713
ESN Detail:	VOIP CALLER
Phone :	(866) 866-5006 Ext :
Main Number :	
Atom :	
Cell :	956-211-04
Sector :	
X Coordinate :	-098.155932
Y Coordinate :	+026.201642

Associated Calls

Date/Time	Call Type	Line	Phone Number	Call Taker / Position
05/05/2010 17:39:12:283	Incoming	San Juan PD 9111	(866) 866-5006	Position 81 - San Juan PD / OBJECTSERVER8

Page 4 of 51 Generated on : 07/07/2010 15:02:38

Figure A.1: Sample Data from LRGVDC 9-1-1 Dept. Intrado Data report

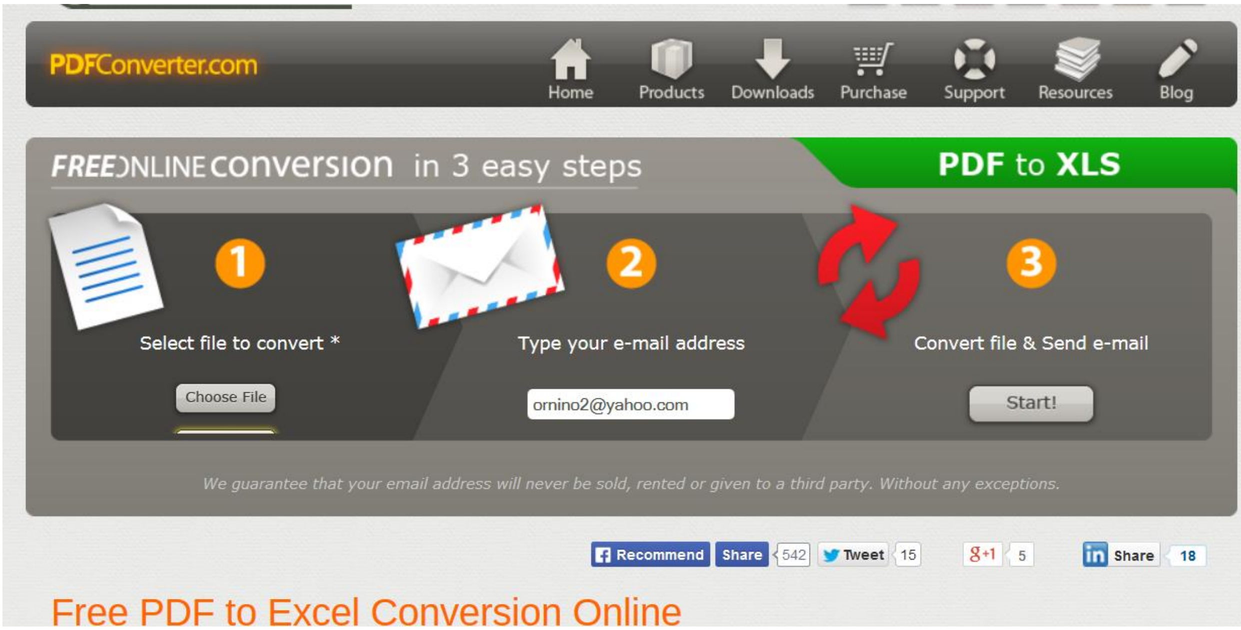


Figure A.2: PDF Converter.com main dialog used to convert PDF's to Excel .xls

Calls For Service Report - Donna Police Department

Sorted by Call_Number, Call_Date, Call_Time

Call Date : 06/01/2012 07:00 - 07/31/2013 07:00

Call Number Case Number	Date	Time	Agency	Call Type	Caller	Address	Zone
12-006476	06/01/2012 Sent To Dispatch- 06/01/2012 07:29:01	07:28:41	1 - DONNA PD	TRAFFIC STOP	PEREZ, ERIC	518 N MAIN ST DONNA, TX 78537	ZONE 2
12-006477 12-006477	06/01/2012 Sent To Dispatch- 06/01/2012 08:10:42	08:06:49	1 - DONNA PD	BURGLARY OF VEHICLE	HERMOSILLO, CANDELARIO	418 ROSA PRIEGO DONNA, TX 78537	
12-006478 12-006478	06/01/2012 Sent To Dispatch- 06/01/2012 08:13:03	08:11:52	1 - DONNA PD	RUNAWAY	OLVERA, HORTENCIA	212 ELM ST DONNA, TX 78537	
12-006479 12-006479	06/01/2012 Sent To Dispatch- 06/01/2012 08:35:25	08:35:05	1 - DONNA PD	BURGLARY OF VEHICLE		422 ROSA PRIEGO DONNA, TX 78537	ZONE 1
12-006480	06/01/2012 Sent To Dispatch- 06/01/2012 08:40:34	08:40:13	1 - DONNA PD	TRAFFIC STOP		400 N DANIEL SALINAS BLVD DONNA, TX 78537	ZONE 2

Figure A.3 Sample Data for Donna PD Data

MERCEDES POLICE DEPARTMENT

CAD Summary - All Units All Shifts

From 05/01/2011 00:00 To 06/30/2012 23:59

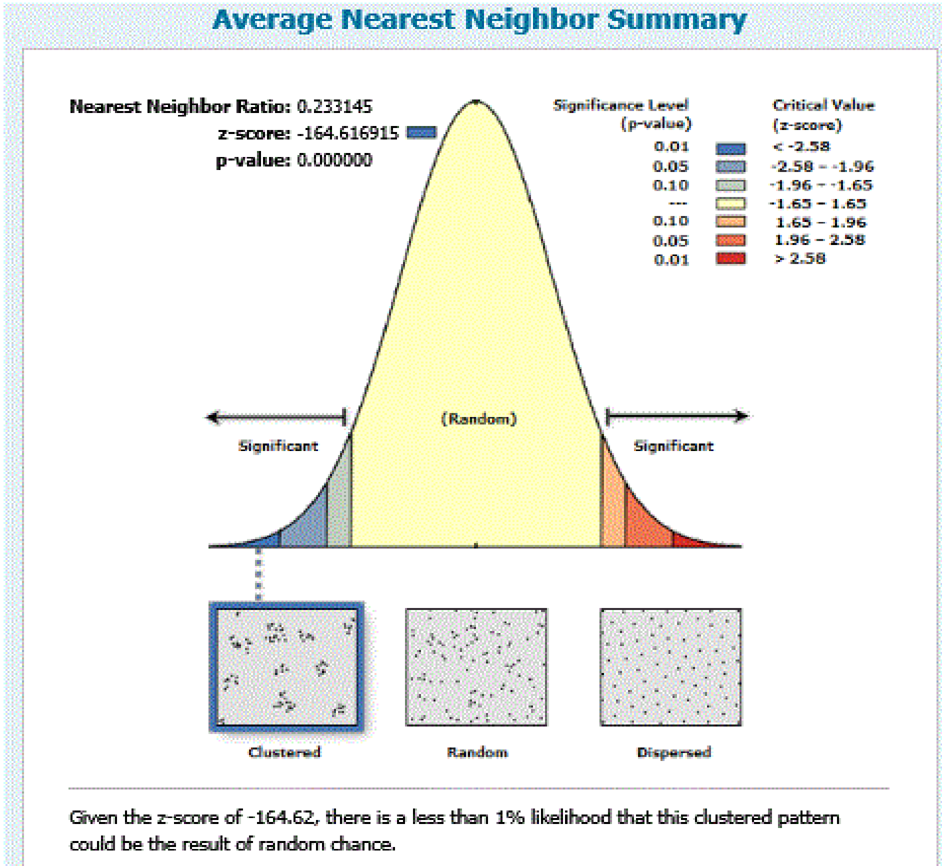
Printed: 10/28/2013 11:58

Date	Number	Agency No	Unit ID	Type	Description	Disp	Beg	End	Total
04/04/12	1203471	9045788	4969	CFS	INFORMATION CALL	001	21:56	22:14	22
04/04/12	1203472	9045789	207	CFS	PHONE HARASSMENT	001	22:01	22:04	3
04/04/12	1203473	9045790	7274	CFS	AREA CHECK	001	22:20	22:44	24
04/04/12	1203474	0	MALL1	CFS	BURGLAR ALARM	CB	22:42	22:43	2

Number	Location	Address	Cross	Intersection	Landmark
1203471		227 S VIRGINIA			
1203472	MPD	316 S OHIO	ST 4TH		
1203473	HI-RISE	501 S TEXAS			
1203474	OUTLET MALL	5001 E EXPRESSWAY 83			

Figure A.4 Sample Data for Mercedes PD Data

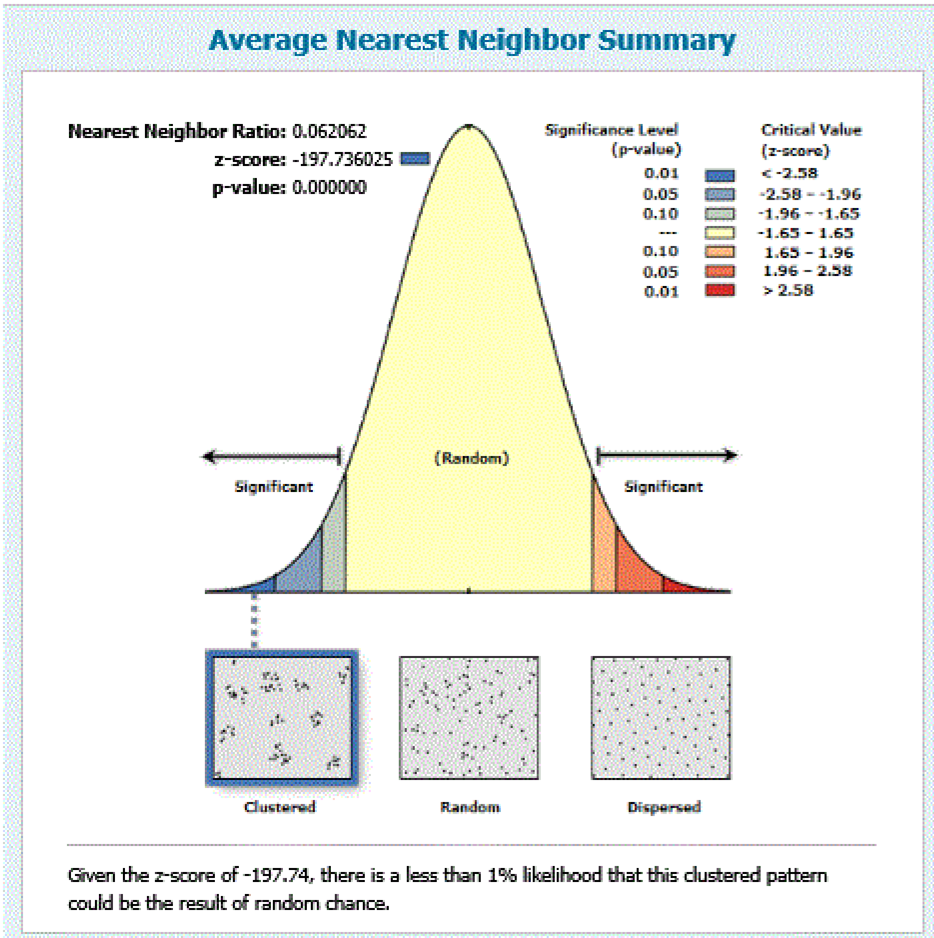
Appendix B
Average Nearest Neighbor Result Graphs



Average Nearest Neighbor Summary

Observed Mean Distance:	0.000131 Degrees
Expected Mean Distance:	0.000563 Degrees
Nearest Neighbor Ratio:	0.233145
z-score:	-164.616915
p-value:	0.000000

Figure B.1 Donna LRG 9-1-1 Nearest Neighbor Results

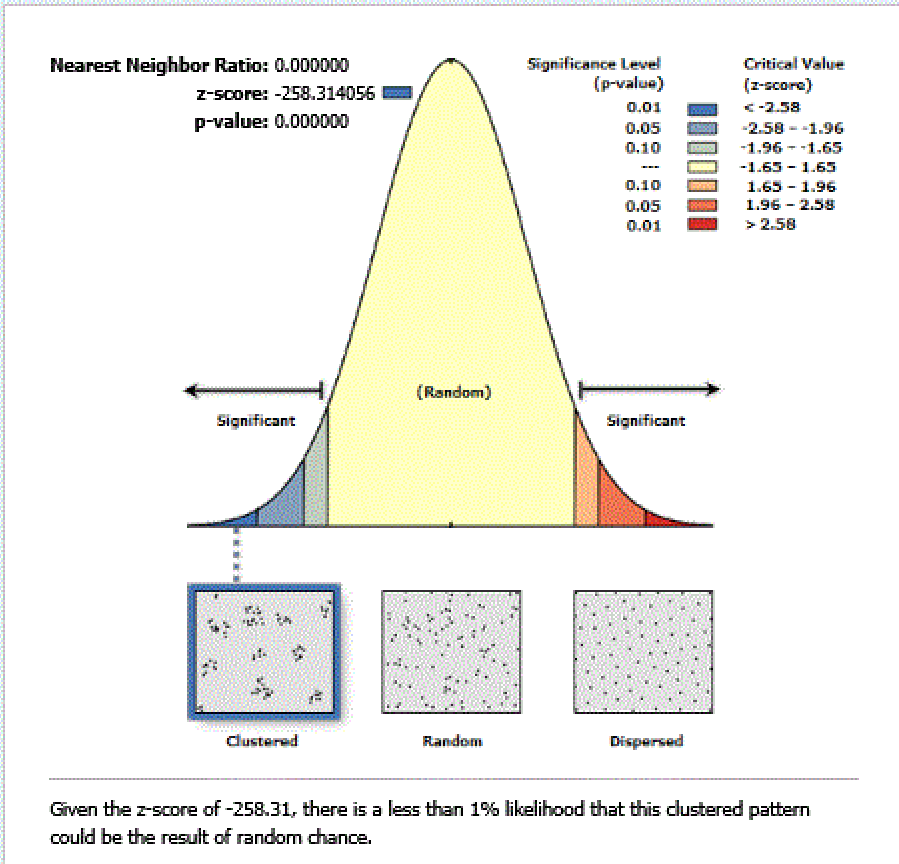


Average Nearest Neighbor Summary

Observed Mean Distance:	11.289425 US_Feet
Expected Mean Distance:	181.904988 US_Feet
Nearest Neighbor Ratio:	0.062062
z-score:	-197.736025
p-value:	0.000000

Figure B.2 Donna PD Nearest Neighbor Results

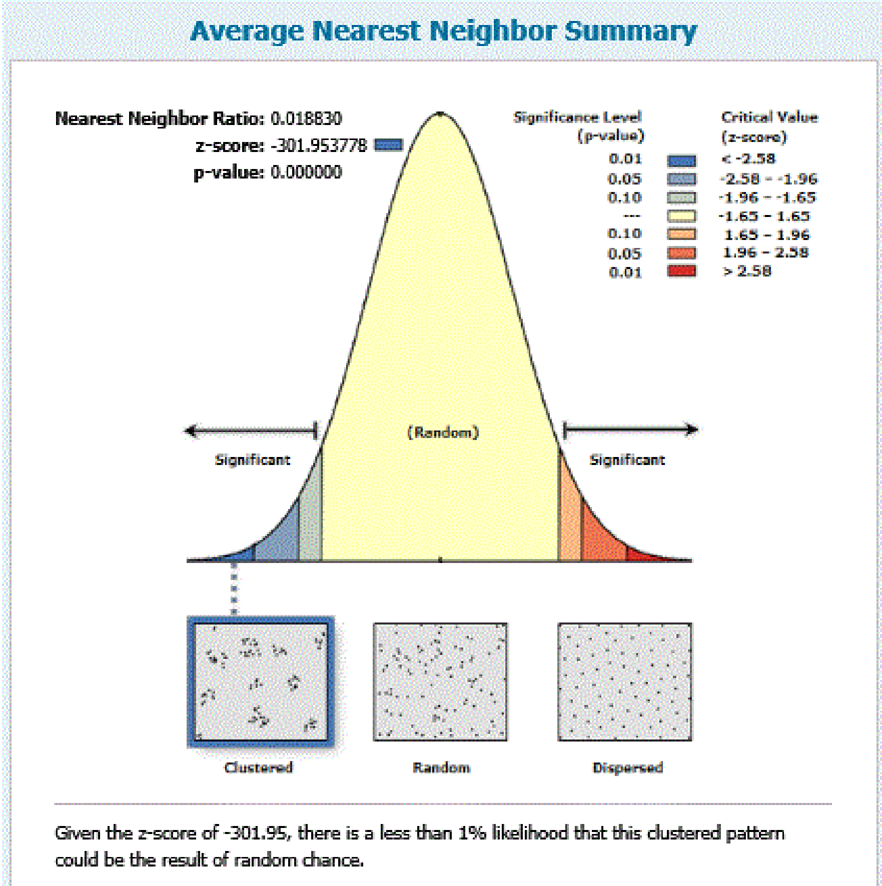
Average Nearest Neighbor Summary



Average Nearest Neighbor Summary

Observed Mean Distance:	0.000000 US_Feet
Expected Mean Distance:	160.549865 US_Feet
Nearest Neighbor Ratio:	0.000000
z-score:	-258.314056
p-value:	0.000000

Figure B.4 Mercedes LRG 9-1-1 Nearest Neighbor Results



Average Nearest Neighbor Summary

Observed Mean Distance:	3.209897 US_Feet
Expected Mean Distance:	170.466580 US_Feet
Nearest Neighbor Ratio:	0.018830
z-score:	-301.953778
p-value:	0.000000

Figure B.5 Mercedes PD Nearest Neighbor Results

Appendix C
Explanatory Regression Results

Donna LRG 9-1-1 Explanatory Regression Results

Messages
Executing: ExploratoryRegression fDonna911Tract CallNumb HD01_VD01
;HD011_VD01;HD01_VD17
;HD01_VD18
;HD01_VD02
;HD01VD17
;HD01_VD33
;HD01_VD66
;HD01_VD05
F:\170708_OLSNew\ResultsFinal\OLSExploratoryRegression\Donna911ExpReg.txt
F:\170708_OLSNew\ResultsFinal\OLSExploratoryRegression\Donna911ExpRegTable 5 1 0.5
0.05 7.5 0.1 0.1
Start Time: Thu Jul 13 01:33:31 2017
Running script ExploratoryRegression...
WARNING 000845: A minimum of 30 input features for analysis is ideal.
****
Choose 1 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2 AICc JB K(BP) VIF SA Model
0.15 245.32 0.10 0.20 1.00 0.39 -HD01_VD05
*
0.12 245.82 0.02 0.29 1.00 0.45 -HD011_VD01
0.10 246.17 0.02 0.29 1.00 0.46 -HD01_VD01

Passing Models
AdjR2 AICc JB K(BP) VIF SA Model

****
Choose 2 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2 AICc JB K(BP) VIF SA Model
0.35 244.37 0.20 0.17 2.47 0.21 +HD01_VD02
* -HD01_VD05
**
0.18 247.69 0.44 0.07 5.67 0.62 -HD011_VD01 +HD01_VD17

0.17 247.79 0.07 0.52 1.02 0.72 -HD011_VD01* +HD01_VD18

Passing Models

AdjR2 AICc JB K(BP) VIF SA   Model

****
Choose 3 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
0.49 244.68 0.25 0.63 2.88 0.68 +HD01_VD18
* +HD01_VD02
** -HD01_VD05
***
0.38 247.34 0.49 0.25 4.49 0.43 +HD01_VD17
+HD01_VD02
** -HD01_VD05
**
0.37 247.66 0.32 0.40 10.58 0.37 +HD01_VD01
+HD01_VD02
** -HD01_VD05
**
Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model

****
Choose 4 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP) VIF   SA   Model
0.64 244.77 0.92 0.29 8.43 0.07 -HD011_VD01*** +HD01_VD17
** +HD01_VD18
*** +HD01_VD02
***
0.54 248.34 0.87 0.31 8.64 0.38 -HD01_VD01
*** +HD01_VD17
** +HD01_VD18
** +HD01_VD02
***
0.48 249.86 0.40 0.93 4.50 0.97 +HD01_VD17
+HD01_VD18
+HD01_VD02
** -HD01_VD05
***
Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model

****
Choose 5 of 9 Summary

Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
0.72 248.18 0.81 0.28 10.34 0.65 -HD011_VD01*** +HD01_VD17
*** +HD01_VD18
*** +HD01_VD02
*** +HD01_VD66
*
0.64 251.73 0.64 0.47 11.44 0.30 -HD01_VD01
*** +HD01_VD17
*** +HD01_VD18
*** +HD01_VD02
*** +HD01_VD66
*
0.63 252.38 0.88 0.28 13.85 0.31 -HD011_VD01*** +HD01_VD17
** +HD01_VD18
*** +HD01_VD02
** +HD01VD17

Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model

Writing Results to Output Table....
F:\170708_OLSNew\ResultsFinal\OLSExploratoryRegression\Donna911ExpRegTable.dbf
****
************** Exploratory Regression Global Summary (CALLNUMB) **************

Percentage of Search Criteria Passed

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	381	10	2.62
Max Coefficient p-value	< 0.05	381	2	0.52
Max VIF Value	< 7.50	381	131	34.38
Min Jarque-Bera p-value	> 0.10	381	242	63.52
Min Spatial Autocorrelation p-value	> 0.10	18	17	94.44

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
HD01_VD02	35.58	1.23	98.77
HD01_VD05	17.79	93.25	6.75
HD011_VD01	7.36	94.48	5.52

HD01_VD18
6.75 0.00 100.00
HD01_VD17
5.52 7.98 92.02
HD01_VD01
3.68 54.60 45.40
HD01_VD33
0.61 53.37 46.63
HD01VD17
0.00 22.09 77.91
HD01_VD66
0.00 27.61 72.39

Summary of Multicollinearity

Variable VIF Violations Covariates
HD01_VD01
121.98 142 HD011_VD01 (98.46), HD01VD17
(98.46), HD01_VD33
(87.69), HD01_VD05
(86.15), HD01_VD17
(41.54)
HD011_VD01 73.15 132 HD01_VD01
(98.46), HD01VD17
(90.77), HD01_VD33
(75.38), HD01_VD05
(61.54), HD01_VD17
(43.08)
HD01_VD17
12.03 48 HD01_VD05
(52.31), HD011_VD01 (43.08), HD01_VD01
(41.54), HD01_VD33
(21.54), HD01VD17
(20.00)
HD01_VD18
2.93 0 --------
HD01_VD02
4.01 0 --------
HD01VD17
9.76 106 HD01_VD01
(98.46), HD011_VD01 (90.77), HD01_VD33
(49.23), HD01_VD05

(41.54), HD01_VD17
(20.00)
HD01_VD33
14.39   110   HD01_VD01
(87.69), HD011_VD01 (75.38), HD01_VD05
(60.00), HD01VD17
(49.23), HD01_VD17
(21.54)
HD01_VD66
1.67   0   --------
HD01_VD05
21.80   95   HD01_VD01
(86.15), HD011_VD01 (61.54), HD01_VD33
(60.00), HD01_VD17
(52.31), HD01VD17
(41.54)

Summary of Residual Normality (JB)

JB   AdjR2   AICc   K(BP)   VIF   SA   Model
0.981975 0.519559 255.904875 0.295626 24.629695 0.245732 -HD01_VD01
** +HD01_VD17
** +HD01_VD18
** +HD01_VD02
** +HD01_VD33

0.962335 0.499446 256.479011 0.171758 41.871519 0.589308 -HD01_VD01
+HD01_VD17
+HD01_VD18
* +HD01_VD02
** -HD01_VD05

0.936405 0.494331 256.621365 0.394301 14.161619 0.765682 -HD01_VD01
** +HD01_VD17
** +HD01_VD18
** +HD01_VD02
** +HD01VD17

Summary of Residual Spatial Autocorrelation (SA)

SA   AdjR2   AICc   JB   K(BP)   VIF   Model

0.969886 0.484892 249.862562 0.402952 0.931133 4.496018 +HD01_VD17
 +HD01_VD18
 +HD01_VD02
** -HD01_VD05
***
0.765682 0.494331 256.621365 0.936405 0.394301 14.161619 -HD01_VD01
** +HD01_VD17
** +HD01_VD18
** +HD01_VD02
** +HD01VD17

0.720198 0.170369 247.788873 0.065737 0.519936 1.019545 -HD011_VD01* +HD01_VD18

Table Abbreviations
AdjR2 Adjusted R-Squared
AICc Akaike's Information Criterion
JB   Jarque-Bera p-value
K(BP) Koenker (BP) Statistic p-value
VIF   Max Variance Inflation Factor
SA   Global Moran's I p-value
Model Variable sign (+/-)
Model Variable significance (* = 0.10, ** = 0.05, *** = 0.01)

Completed script ExploratoryRegression...
Succeeded at Thu Jul 13 01:33:32 2017 (Elapsed Time: 1.00 seconds)

Donna PD Explanatory Regression Results

Messages
Executing: Exploratory Regression fDonnaPDTract CallNumb
HD01_VD01;HD011_VD01;HD01_VD17
;HD01_VD18
;HD01_VD02
;HD01VD17
;HD01_VD33
;HD01_VD66
;HD01_VD05
F:\170708_OLSNew\ResultsFinal\OLSExploratoryRegression\DonnaPDExpReg.txt
F:\170708_OLSNew\ResultsFinal\OLSExploratoryRegression\DonnaPDExpRegTable 5 1 0.5 0.05
7.5 0.1 0.1
Start Time: Thu Jul 13 01:26:38 2017
Running script ExploratoryRegression...
WARNING 000845: A minimum of 30 input features for analysis is ideal.
****
Choose 1 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2 AICc JB K(BP) VIF SA Model
0.06 245.02 0.28 0.09 1.00 0.72 -HD01_VD05

0.04 245.27 0.23 0.09 1.00 0.68 -HD011_VD01
0.03 245.43 0.22 0.11 1.00 0.62 -HD01_VD33

Passing Models
AdjR2 AICc JB K(BP) VIF SA Model

****
Choose 2 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2 AICc JB K(BP) VIF SA Model
0.24 244.90 0.80 0.01 8.05 0.06 -HD011_VD01* +HD01_VD17

0.08 247.43 0.46 0.08 2.69 0.91 +HD01_VD02
-HD01_VD05
*
0.05 247.93 0.60 0.05 6.04 0.34 -HD01_VD01 +HD01_VD17

Passing Models
AdjR2 AICc JB K(BP) VIF SA Model

****

Choose 3 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2 AICc JB K(BP) VIF SA Model
0.31 247.17 0.65 0.09 65.27 0.00 +HD01_VD01* -HD011_VD01** +HD01_VD17
**
0.28 247.72 0.73 0.04 11.19 0.00 -HD011_VD01** +HD01_VD17
** +HD01_VD02
**
0.26 248.10 0.68 0.06 9.42 0.37 -HD011_VD01** +HD01_VD17
* +HD01_VD66

Passing Models
AdjR2 AICc JB K(BP) VIF SA Model

****
Choose 4 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2 AICc JB K(BP) VIF SA Model
0.47 248.57 0.57 0.14 82.32 0.00 +HD01_VD01** -HD011_VD01** +HD01_VD17
** +HD01_VD18
*
0.44 249.20 0.81 0.09 12.52 0.01 -HD011_VD01*** +HD01_VD17
*** +HD01_VD18
** +HD01_VD02
***
0.38 250.74 0.63 0.10 34.93 0.12 -HD011_VD01*** +HD01_VD17
** +HD01_VD18
** +HD01_VD33
***
Passing Models
AdjR2 AICc JB K(BP) VIF SA Model

****
Choose 5 of 9 Summary
Highest Adjusted R-Squared Results
AdjR2 AICc JB K(BP) VIF SA Model
0.58 252.27 0.57 0.14 313.69 0.63 +HD01_VD01*** -HD011_VD01*** -HD01_VD33
** +HD01_VD66
** -HD01_VD05
**
0.58 252.32 0.68 0.67 14.52 0.75 -HD011_VD01*** +HD01_VD17
*** +HD01_VD18
** +HD01_VD02

** +HD01_VD66
*
0.52 254.17 0.60 0.08 82.63 0.22 +HD01_VD01** -HD011_VD01*** +HD01_VD17
*** +HD01_VD18
** +HD01_VD66
**
Passing Models
AdjR2 AICc JB K(BP) VIF SA Model

Writing Results to Output Table....
F:\170708_OLSNew\ResultsFinal\OLSExploratoryRegression\DonnaPDExpRegTable.dbf
****
************* Exploratory Regression Global Summary (CALLNUMB
) **************

Percentage of Search Criteria Passed

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	381	3	0.79
Max Coefficient p-value	< 0.05	381	5	1.31
Max VIF Value	< 7.50	381	104	27.30
Min Jarque-Bera p-value	> 0.10	381	332	87.14
Min Spatial Autocorrelation p-value	> 0.10	18	13	72.22

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
HD011_VD01	30.67	93.87	6.13
HD01_VD17	22.70	4.91	95.09
HD01_VD01	9.82	41.72	58.28
HD01_VD02	4.91	9.82	90.18
HD01_VD18	3.68	0.00	100.00
HD01_VD33	2.45	58.90	41.10
HD01_VD05	2.45	85.28	14.72
HD01_VD66	1.84	0.00	100.00
HD01VD17	0.00	60.12	39.88

Summary of Multicollinearity

Variable VIF Violations Covariates
HD01_VD01 366.17 148 HD011_VD01 (98.46), HD01_VD33
(98.46), HD01VD17
(95.38), HD01_VD05
(92.31), HD01_VD17
(87.69)
HD011_VD01 158.80 143 HD01_VD17
(98.46), HD01_VD01 (98.46), HD01_VD33
(95.38), HD01VD17
(86.15), HD01_VD05
(75.38)
HD01_VD17
28.21 99 HD011_VD01 (98.46), HD01_VD01 (87.69), HD01_VD33
(58.46), HD01_VD05
(52.31), HD01VD17
(50.77)
HD01_VD18
4.01 0 --------
HD01_VD02
4.32 0 --------
HD01VD17
8.84 115 HD01_VD01 (95.38), HD011_VD01 (86.15), HD01_VD33
(69.23), HD01_VD05
(53.85), HD01_VD17
(50.77)
HD01_VD33
63.97 140 HD01_VD05
(98.46), HD01_VD01 (98.46), HD011_VD01 (95.38), HD01VD17
(69.23), HD01_VD17
(58.46)
HD01_VD66
1.51 0 --------
HD01_VD05
25.32 118 HD01_VD33
(98.46), HD01_VD01 (92.31), HD011_VD01 (75.38), HD01VD17
(53.85), HD01_VD17
(52.31)

Summary of Residual Normality (JB)

JB AdjR2 AICc K(BP) VIF SA Model
0.994885 0.093112 263.061414 0.108273 143.642806 0.635069 +HD01_VD01 -HD011_VD01
+HD01_VD02
+HD01VD17
-HD01_VD05
*
0.992661 0.193786 254.396337 0.217877 143.010749 0.636159 +HD01_VD01 -HD011_VD01
+HD01_VD02
-HD01_VD05
*
0.974168 0.159373 254.981515 0.081699 127.562308 0.438758 +HD01_VD01** -
HD011_VD01** +HD01VD17
-HD01_VD05
**

Summary of Residual Spatial Autocorrelation (SA)

SA AdjR2 AICc JB K(BP) VIF Model
0.909563 0.084804 247.425222 0.456648 0.077557 2.685795 +HD01_VD02
-HD01_VD05
*
0.749838 0.579074 252.315544 0.679846 0.669966 14.519303 -HD011_VD01***
+HD01_VD17
*** +HD01_VD18
** +HD01_VD02
** +HD01_VD66
*
0.719684 0.056748 245.021668 0.275489 0.090142 1.000000 -HD01_VD05

Table Abbreviations
AdjR2 Adjusted R-Squared
AICc Akaike's Information Criterion
JB Jarque-Bera p-value
K(BP) Koenker (BP) Statistic p-value
VIF Max Variance Inflation Factor
SA Global Moran's I p-value
Model Variable sign (+/-)
Model Variable significance (* = 0.10, ** = 0.05, *** = 0.01)

Completed script ExploratoryRegression...
Succeeded at Thu Jul 13 01:26:39 2017 (Elapsed Time: 1.00 seconds)

Appendix D
Ordinary Least Squares Results

Donna LRG PD OLS Results

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	741.548680	628.241415	1.180356	0.271827	575.223906	1.289148	0.233441	--------
HD011_VD01	-1.872533	0.406661	-4.604653	0.001796*	0.369240	-5.071316	0.000986*	14.519303
HD01_VD17	8.087175	2.064528	3.917203	0.004505*	1.761679	4.590607	0.001829*	11.046565
HD01_VD18	8.893903	3.531367	2.518544	0.035825*	3.284080	2.708187	0.026689*	1.532500
HD01_VD02	9.446337	3.355424	2.815244	0.022630*	2.347644	4.023752	0.003891*	2.222451
HD01_VD66	16.034735	8.125517	1.973380	0.083855	5.157874	3.108788	0.014483*	1.195796

OLS Diagnostics

Input Features:	fDonnaPDTract	Dependent Variable:	CALLNUMB
Number of Observations:	14	Akaike's Information Criterion (AICc) [d]:	252.315544
Multiple R-Squared [d]:	0.740969	Adjusted R-Squared [d]:	0.579074
Joint F-Statistic [e]:	4.576861	Prob(>F), (5,8) degrees of freedom:	0.028705*
Joint Wald Statistic [e]:	31.301904	Prob(>chi-squared), (5) degrees of freedom:	0.000008*
Koenker (BP) Statistic [f]:	3.194907	Prob(>chi-squared), (5) degrees of freedom:	0.669966
Jarque-Bera Statistic [g]:	0.771777	Prob(>chi-squared), (2) degrees of freedom:	0.679846

Notes on Interpretation

* An asterisk next to a number indicates a statistically significant p-value (p < 0.05).

[a] Coefficient: Represents the strength and type of relationship between each explanatory variable and the dependent variable.

[b] Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistically significant (p < 0.05); if the Koenker (BP) Statistic [f] is statistically significant, use the Robust Probability column (Robust_Pr) to determine coefficient significance.

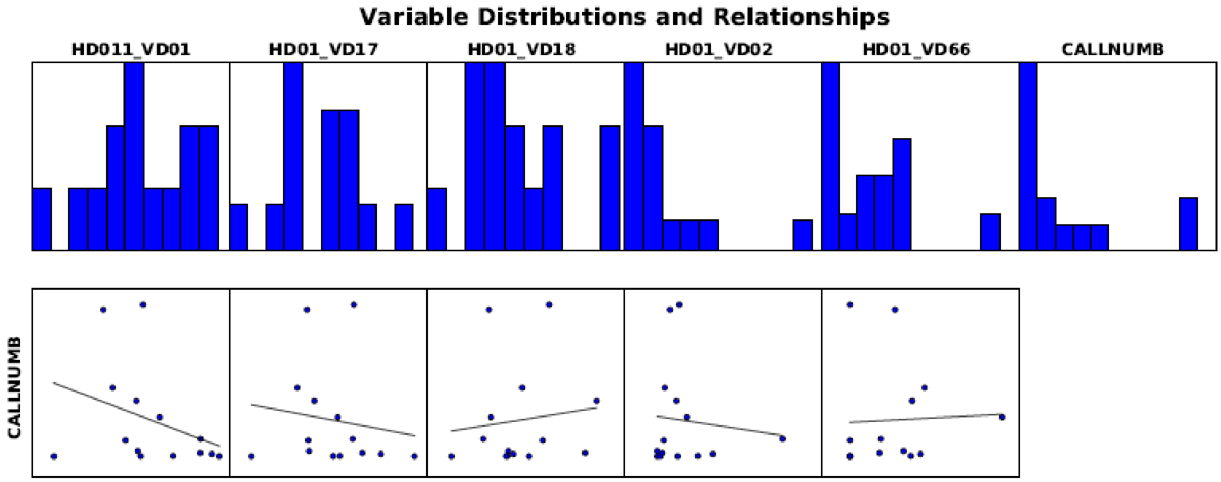
[c] Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redundancy among explanatory variables.

[d] R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.

[e] Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.05); if the Koenker (BP) Statistic [f] is statistically significant, use the Wald Statistic to determine overall model significance.

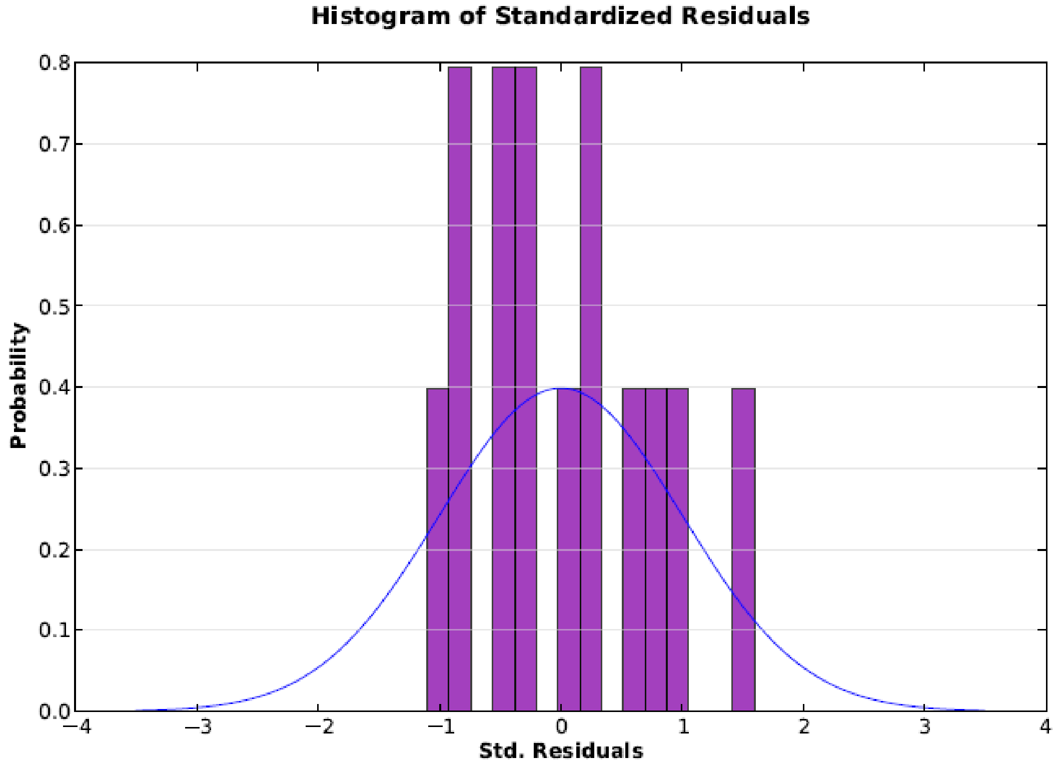
[f] Koenker (BP) Statistic: When this test is statistically significant (p < 0.05), the relationships modeled are not consistent (either due to non-stationarity or heteroskedasticity). You should rely on the Robust Probabilities (Robust_Pr) to determine coefficient significance and on the Wald Statistic to determine overall model significance.

[g] Jarque-Bera Statistic: When this test is statistically significant (p < 0.05) model predictions are biased (the residuals are not normally distributed).

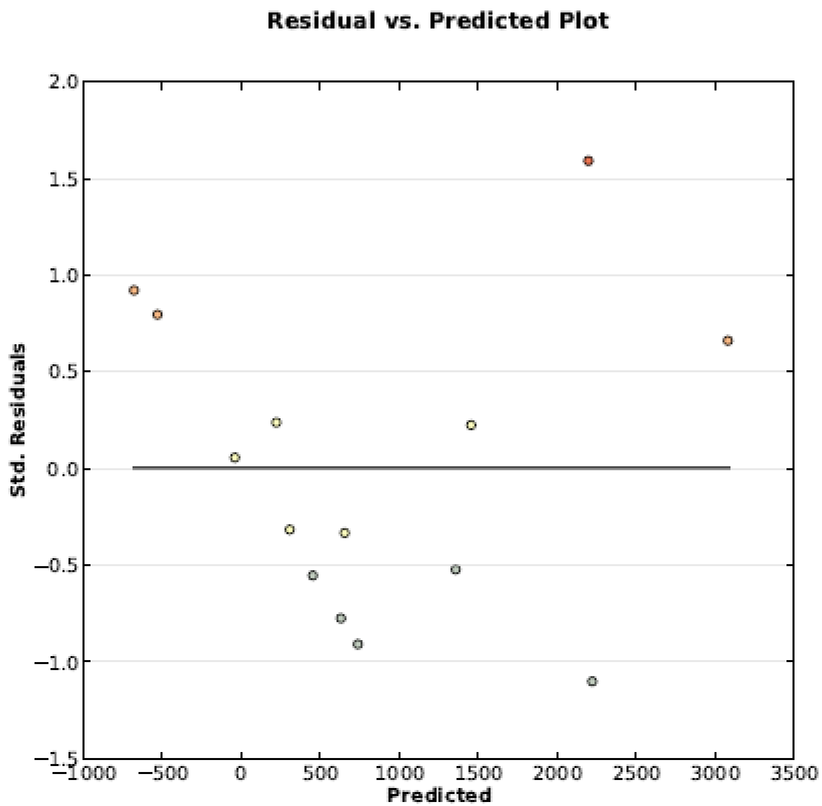


The above graphs are Histograms and Scatterplots for each explanatory variable and the dependent variable. The histograms show how each variable is distributed. OLS does not require variables to be normally distributed. However, if you are having trouble finding a properly-specified model, you can try transforming strongly skewed variables to see if you get a better result.

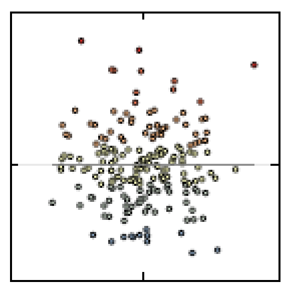
Each scatterplot depicts the relationship between an explanatory variable and the dependent variable. Strong relationships appear as diagonals and the direction of the slant indicates if the relationship is positive or negative. Try transforming your variables if you detect any non-linear relationships. For more information see the Regression Analysis Basics documentation.



Ideally the histogram of your residuals would match the normal curve, indicated above in blue. If the histogram looks very different from the normal curve, you may have a biased model. If this bias is significant it will also be represented by a statistically significant Jarque-Bera p-value (*).



This is a graph of residuals (model over and under predictions) in relation to predicted dependent variable values. For a properly specified model, this scatterplot will have little structure, and look random (see graph on the right). If there is a structure to this plot, the type of structure may be a valuable clue to help you figure out what's going on.



Random Residuals

Ordinary Least Squares Parameters

Parameter Name	Input Value
Input Features	fDonnaPDTract
Unique ID Field	OBJECTID
Output Feature Class	None
Dependent Variable	CALLNUMB
Explanatory Variables	HD011_VD01
	HD01_VD17
	HD01_VD18
	HD01_VD02
	HD01_VD66
Selection Set	False

Donna LRG 9-1-1 OLS Results

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	280.536939	770.624930	0.364038	0.725249	673.627600	0.416457	0.687989	--------
HD011_VD01	-1.890935	0.346877	-5.451308	0.000607*	0.400040	-4.726862	0.001533*	10.338563
HD01_VD17	7.551050	1.834754	4.115566	0.003433*	2.051860	3.680099	0.006281*	8.302089
HD01_VD18	14.652060	3.317100	4.417129	0.002295*	2.615437	5.602146	0.000500*	1.545686
HD01_VD02	12.464905	2.908713	4.285367	0.002733*	2.376440	5.245200	0.000790*	2.140892
HD01_VD66	14.068582	7.372089	1.908358	0.092728	3.958069	3.554405	0.007517*	1.291724

OLS Diagnostics

Input Features:	fDonna911Tract	Dependent Variable:	CALLNUMB
Number of Observations:	14	Akaike's Information Criterion (AICc) [d]:	248.180388
Multiple R-Squared [d]:	0.829719	Adjusted R-Squared [d]:	0.723294
Joint F-Statistic [e]:	7.796249	Prob(>F), (5,8) degrees of freedom:	0.006093*
Joint Wald Statistic [e]:	84.903823	Prob(>chi-squared), (5) degrees of freedom:	0.000000*
Koenker (BP) Statistic [f]:	6.302137	Prob(>chi-squared), (5) degrees of freedom:	0.277920
Jarque-Bera Statistic [g]:	0.423479	Prob(>chi-squared), (2) degrees of freedom:	0.809175

Notes on Interpretation

* An asterisk next to a number indicates a statistically significant p-value (p < 0.05).

[a] Coefficient: Represents the strength and type of relationship between each explanatory variable and the dependent variable.

[b] Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistically significant (p < 0.05); if the Koenker (BP) Statistic [f] is statistically significant, use the Robust Probability column (Robust_Pr) to determine coefficient significance.

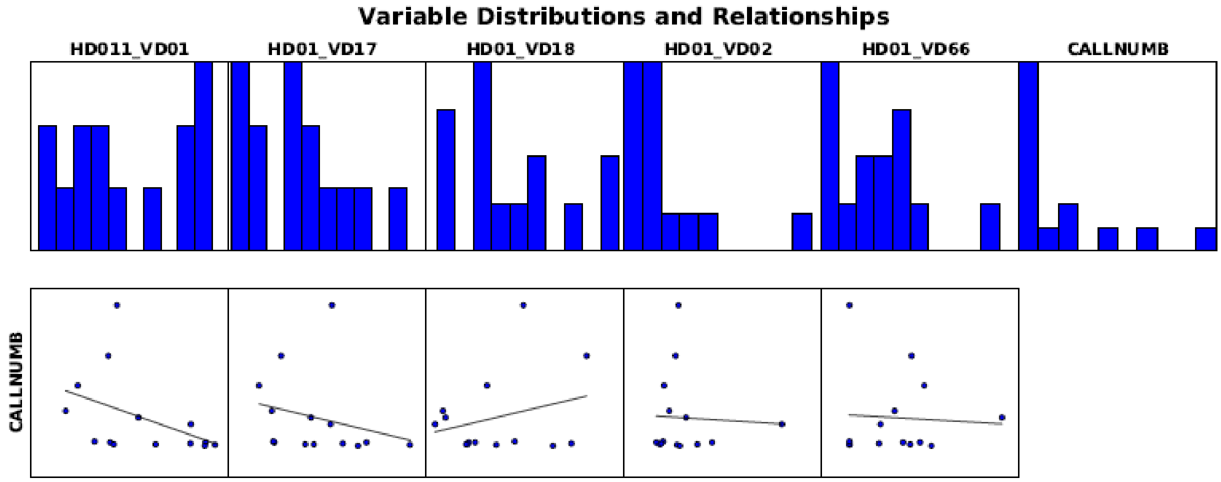
[c] Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redundancy among explanatory variables.

[d] R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.

[e] Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.05); if the Koenker (BP) Statistic [f] is statistically significant, use the Wald Statistic to determine overall model significance.

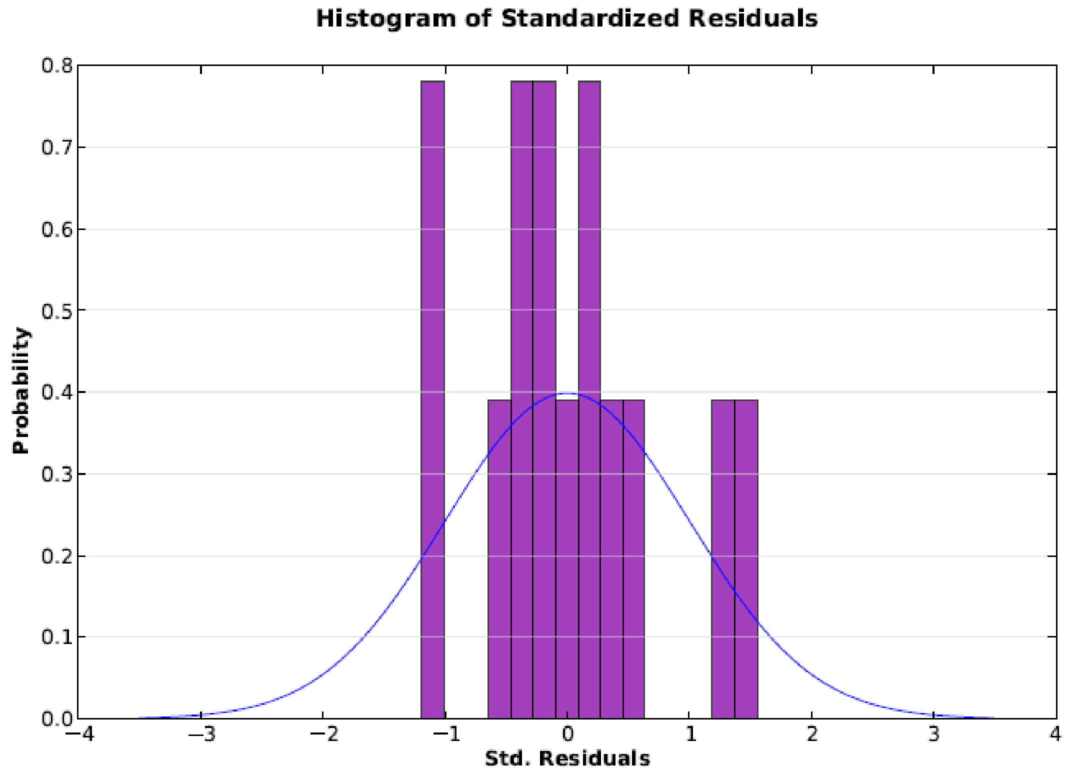
[f] Koenker (BP) Statistic: When this test is statistically significant (p < 0.05), the relationships modeled are not consistent (either due to non-stationarity or heteroskedasticity). You should rely on the Robust Probabilities (Robust_Pr) to determine coefficient significance and on the Wald Statistic to determine overall model significance.

[g] Jarque-Bera Statistic: When this test is statistically significant (p < 0.05) model predictions are biased (the residuals are not normally distributed).

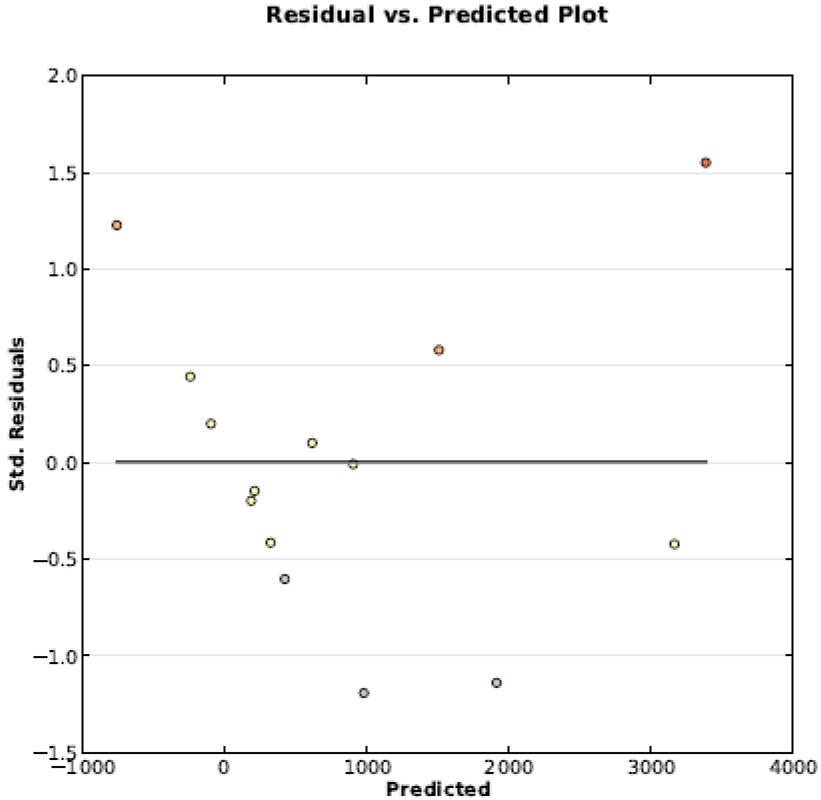


The above graphs are Histograms and Scatterplots for each explanatory variable and the dependent variable. The histograms show how each variable is distributed. OLS does not require variables to be normally distributed. However, if you are having trouble finding a properly-specified model, you can try transforming strongly skewed variables to see if you get a better result.

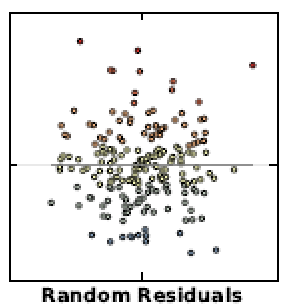
Each scatterplot depicts the relationship between an explanatory variable and the dependent variable. Strong relationships appear as diagonals and the direction of the slant indicates if the relationship is positive or negative. Try transforming your variables if you detect any non-linear relationships. For more information see the Regression Analysis Basics documentation.



Ideally the histogram of your residuals would match the normal curve, indicated above in blue. If the histogram looks very different from the normal curve, you may have a biased model. If this bias is significant it will also be represented by a statistically significant Jarque-Bera p-value (*).



This is a graph of residuals (model over and under predictions) in relation to predicted dependent variable values. For a properly specified model, this scatterplot will have little structure, and look random (see graph on the right). If there is a structure to this plot, the type of structure may be a valuable clue to help you figure out what's going on.



Ordinary Least Squares Parameters

Parameter Name	Input Value
Input Features	fDonna911Tract
Unique ID Field	OBJECTID
Output Feature Class	None
Dependent Variable	CALLNUMB
Explanatory Variables	HD011_VD01
	HD01_VD17
	HD01_VD18
	HD01_VD02
	HD01_VD66
Selection Set	False

Supplemental A

9-1-1 Data Background

This study is based on analyzing data from three different sources. All three datasets are recorded from emergency systems. The first data set is from the LRGVDC 9-1-1 System. This is the hardware and software that helps record the calls being made when people dial 9-1-1. The 9-1-1 call taking solution that the LRGVDC 9-1-1 system uses is provided by Intrado's 9-1-1 company called Positron. They provide the system called Viper. This is their type of 9-1-1 System that records all the landlines and data flow that goes through the system. For this study only the data was recorded in the form of an XML file. This file then was converted to an Microsoft Excel form to be later uploaded to Arcmap for analysis.

When a 9-1-1 call is made the voice goes through dedicated Public Telephone Lines, a reminisce of the once used long distance telephone systems. They were designed to record customers phone numbers for telephone companies to charge a toll for jumping from local telephone company to the next. Depending on how many one jumped to talk the other person a the other end they companies charged more. The data portion of the all uses the persons telephone number to search a database with all the landline numbers in operation. They are attached to the billing and location information which is a house or business location. This database is called a the Master Street Addressing Guide (MSAG). This MSAG is now incorporated to Computer phones, which are known as Voice Over Internet Protocol telephone solutions (VoIP).

For Cell phone the system uses a Mobile Switching Unit, to capture the GPS coordinates off a cell phone. There are two different technologies that one has to understand for this study. The first are the Hand Based solutions, that acquire the Lat. and Long. from the cell phone directly. The second is a Tower Based solution that uses two or more cell phone towers to triangulate the position of the cell phone. These two technologies send the information to the Mobile Switching Center and delivers the X and Y to the PSAP

Both solutions take the same progression of steps to establish a Lat. and Long. position. The system knows what cell phone tower a call originates from, so the Lat and Long. coordinates of the Tower are immediately sent to the system. This is called Wireless Phase I (WRLS). When the handset or triangulation is done on the device the coordinates are retransmitted to the system and the precise position is displayed at the PSAP and the systems plots the call on a map, this is called Wireless Phase II (WPHII). This is when the telecommunicator is not able to talk to the person and establish a location of the call, they can be confident that the coordinates presented are accurate and he can dispatch help.

Unlike a VoIP or Land Line 9-1-1 emergency call a physical address is not recorded by the system only the WRLS or WPHII coordinates. In all cases the time of the call, the length of the call, the position and the PSAP fielding the call are recorded. Many other parameters are recorded as well but were not used for this study.

When preparing the data for it to be displayed with ArcMap the Wireless calls needed little manipulation. WPHII calls gave an exact location of a call. For many known and unknown reasons some cell phone calls never go WPHII so the calls will stay as WRLS and will plot at the tower of the originating call. The main know causes are that calls are made inside buildings and GPS functions will not work. Tower based solutions cell phones will not go into WPHII when in rural areas. There are not enough towers in close proximity to establish a location. However, due to this technological limitations the main reasons calls of either technology don’t go WPHII is that Telecommunicators transfer the calls in such a short time that the systems does not recorded the WPHII information. Even though, this is the case the information recorded is invaluable to this study.

Landline calls will always have a physical address associated to it. Every landline call is searched through the MSAG and all the physical location information is recorded by the system. In some rare occurrence there might be a mistake or omission on the address and they are corrected as part of the work flow of the local 9-1-1 entity. This is great information to have however when it is brought into ArcMap all the addresses have to be Reversed GeoCoded. Luckily, the LRGVDC 9-1-1 Department is also tasked as the regions authorized addressing entity much like many other 9-1-1 departments in the State. As explained in the study part of the whole analysis was to develop a geocoder and accurately plotting these calls.

Computer Aided Dispatch

The other two sets of data that came from two different Police Departments. They both provided one year of call data recorded in their Computer Aided Dispatch (CAD) system. This is the software that they used to recorded all the activities they undertake during a Telecommunicators/Dispatchers shift. This includes case numbers for request for service which include police duties such as tickets, criminal cases, civil cases, fire dispatch, EMS dispatch, walk-ins and any calls relating to police functions. This also includes 9-1-1 calls that were dispatched. The CAD system records audio of all calls made to a police station including 9-1-1 calls. However, it does not record 9-1-1 data, due that the 9-1-1 Departments provide the call taking system. So the data that relates to 9-1-1 calls in directly typed in by the dispatchers. There are many errors and omissions and some are not even recorded due to human factors.

Both sets of data were provided to me in the form of PDF’s. Due to their rules they could not provide me with electronic versions of the data. Both supervisors that provided me with the data did not know how to download reports in any other form but PDF's. This is so because they never do. The information recorded into the CAD is usually the start and in many cases part of legal and criminal evidence. In a few cases it is the only evidence that they have to put criminals in jail or exonerate the innocent. As explained in the study this caused many hours of work and failure that kept me from advancing in my study.

In both data sets the PDF's that I received were close to 2400 pages each and were not in vector form. If they would have it would only take me a couple of hours to work them to Excel format and import them into ArcGIS software. However, they were not in vector form and I had to use a CDR program to read text and convert them to vector format. Both data sets exceeded 85000 records and 2250 PDF pages. I had to convert them to Excel using an online program. Lucky it worked and the product was off and needed to be fixed. This meant deleting duplicates matching four rows to be one record and making sure each record was not compromised.

The 9-1-1 system is proprietary to Positron and one has to use its internal program MIS Reports to produce reports. Each PSAP records the data going through its data circuits in a separate database. Once I got the raw databases I tried using MySQL a popular open source SQL database package. I Had to learn and become an expert on MySQL to manipulate each individual database. However this was to no avail due that when the MIS Reports produce a database it is in the form to be only used by other Positron systems. This is done to keep information private and from being disseminated accidently to unauthorized people. This process took me weeks and further delayed my study.

Next I tried to open the databases using Microsoft SQL, and I hit the same hurtles. In both cases I could load partial downloads of the databases but only could see the fields and not eh data. Both programs also are limited in the size of the files and each database was over the

4000K limit. I had to spend like another month troubleshooting the database and conceded that I had to down load small amounts of data at a time. The data acquired form the system was queried using date ranges. I had to limit myself to download one month at a time form each PSAP. Which was time consuming and I was only allowed to work on my study before and after hours of my scheduled work.