**Los Angeles Police Department Big Data Analytics**

Valerie Carisma, Alan Chan, Melissa Hamamoto, Andrew Lai, Medrik Minassian

**Abstract:** The purpose of this paper is to analyze patterns within the City of Los Angeles related to law enforcement activity. Specifically, areas with least to most founded criminal activity, least to most common reasons for Los Angeles Police Department (LAPD) calls for service and find patterns from 2010 to 2018, find the sex and ethnicity of most common vehicle and pedestrian stops, and areas with most traffic collisions. The data analyzed is a compilation of twelve different datasets with a total size of 1.534 GB from the Los Angeles Police Department.

# 1. Introduction

The Los Angeles Police Department crime data has been analyzed by members of this analytics group to provide accurate information through professional visualizations. The analytics combined more than 1.5 GB of data to understand the geographic areas of the LAPD and how criminal activities are spread across the LAPD jurisdiction. The visualizations and analytics provide information on sex, location and age of the victims. The study used different charts and graphs to clearly explain trends and changes in events. The data sets spread over a period of 8 years since 2010. The data analyzed had to be filtered and cleared for creation of understandable charts for audiences. One of the challenges the analysts overcame was combining the raw data for calls for service from different years. The raw data tables had columns swapped for different categories which differed from year to year. The synching of columns was done manually after the different sets were analyzed separately and then combined for visualizations on Tableau.

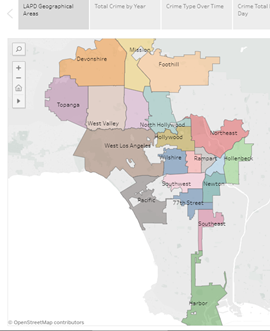
**2. Analyses**

These visualizations can easily be compared to other political, economic timelines within the same area to have a broader understanding of jurisdictional decision making. These data set visualizations will help improve LAPD-civilian relations and encourage collaboration on providing safety and better neighborhoods to the residents of LA.

Los Angeles Police Department has a total of 21 geographical areas (Area) and numerous specialized divisions. As of 2018 the area with the least crime is Foothill Area, and the area with the most crime is 77th Street Area. It is vital to understand the geographic distribution of each area and its demographics. The Foothill Area is located next to the city of Glendale which has its own Police department and is maintains a low crime rate compared to the City of Los Angeles. This causes Foothill Area’s eastern borders be influenced by the safety trends from the City of Glendale. On the other side of the spectrum is the 77th Street Area, which is in the south Los Angeles area that has historically seen higher crime rates.

Looking at the present decade since 2010, the crime rate has seen both a rise and fall. In 2010, there were more than 200,000 crimes recorded. After a two-year low in 2013 and 2014, during which period the crime rates were at 192,000 and 194,000 respectively, recorded crimes reached a high of 230,000. The year 2018 saw a decrease of 14.94% from 2017 and recorded crimes are at 196,000, third lowest in the decade. Battery-Simple Assault, Vehicles Stolen and Burglary are the dominant crime types in the City of Los Angeles. Battery-Simple Assault is 1st in seven geographical areas including 77th Street Area, which has the highest crime rate by area. Vehicle Stolen is the 1st in six areas including Foothill Area, which has the lowest crime rate by area. Burglary is the 1st in five areas mostly in Western Los Angeles, which includes West Los Angeles Area, West Valley Area and Topanga Area. This understanding of areas of criminal activity geographically can allow the LAPD to address the challenges of combating crime strategically. Geographic distribution of crime is closely related to infrastructure, law enforcement presence and economy.

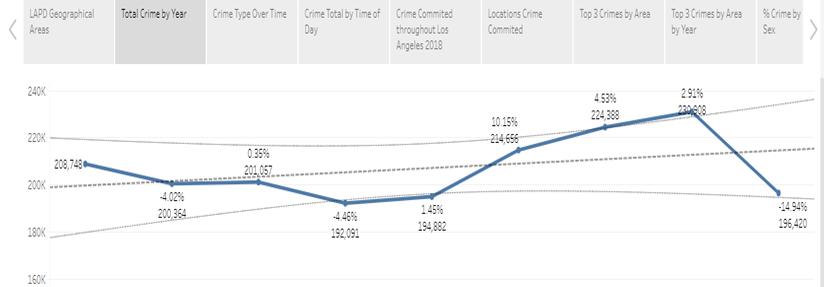
The following are a few select charts from the complete analyses. Each chart is followed by an explanation on the purpose it serves and relevant data it provides.



**Figure 1. Geographical Distribution of LAPD areas of service  
Source: Tableau Visualizations**

Figure 1 is the first visualization, which distributes LAPD’s areas of service geographically providing a clear understanding to viewers which areas fall under the jurisdiction of LAPD and where calls for service are recorded. There are unmarked areas within the LAPD areas of service, these are areas that have their own separate data and are served either by the Los Angeles County Sheriff's Department (LASD) or smaller police departments, for example Beverly Hills Police Department and San Fernando Police Department. Understanding the geographical distribution allows us to better comprehend other data and visualizations.

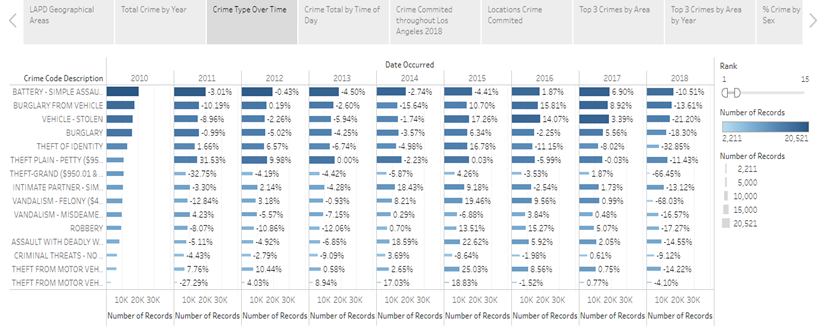
Figure 2 shows the crime rate in the City of Los Angeles from 2010 to 2018. As seen on the chart, the crime rate has moved up and down over the years, 2010 recorded 208,748 while it reached a low of 192,091 it moved up to a high of 230,908 in 2017 and as of November 2018, the crime rate stand at 196,420 a 14% decrease from 2017.



**Figure 2. Crime Rate 2010-2018**

**Source: Tableau Visualizations**

As we understand the crime rates, we can then go on to study the same per number of crimes over years and how they have changed over time. Figure 3 shows crimes ranked over years from 1st to 15th. Each year after 2010 has the change in percentage up to 2018. For example, we can see that Battery-Simple Assault has an average change of 4.2%, which is the lowest change among all types of crimes. The highest change rate was for grand thefts above $950, which was at 15.42%.



**Figure 3. Crime type over time  
Source: Tableau Visualizations**

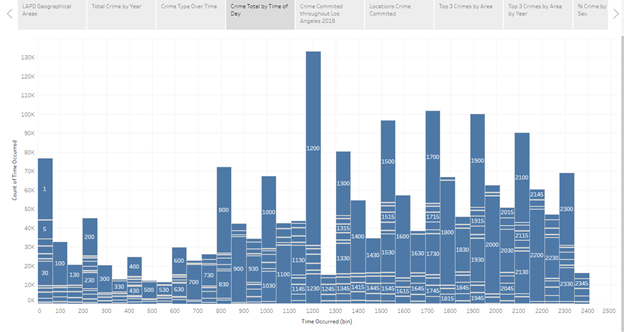
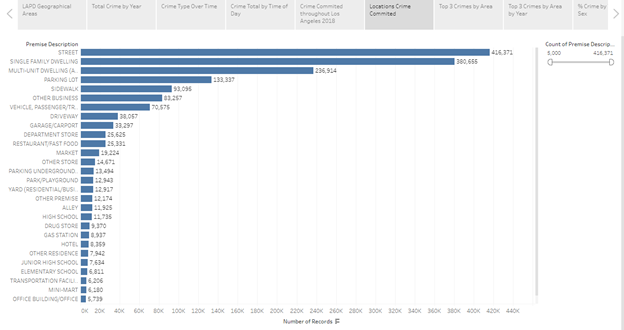
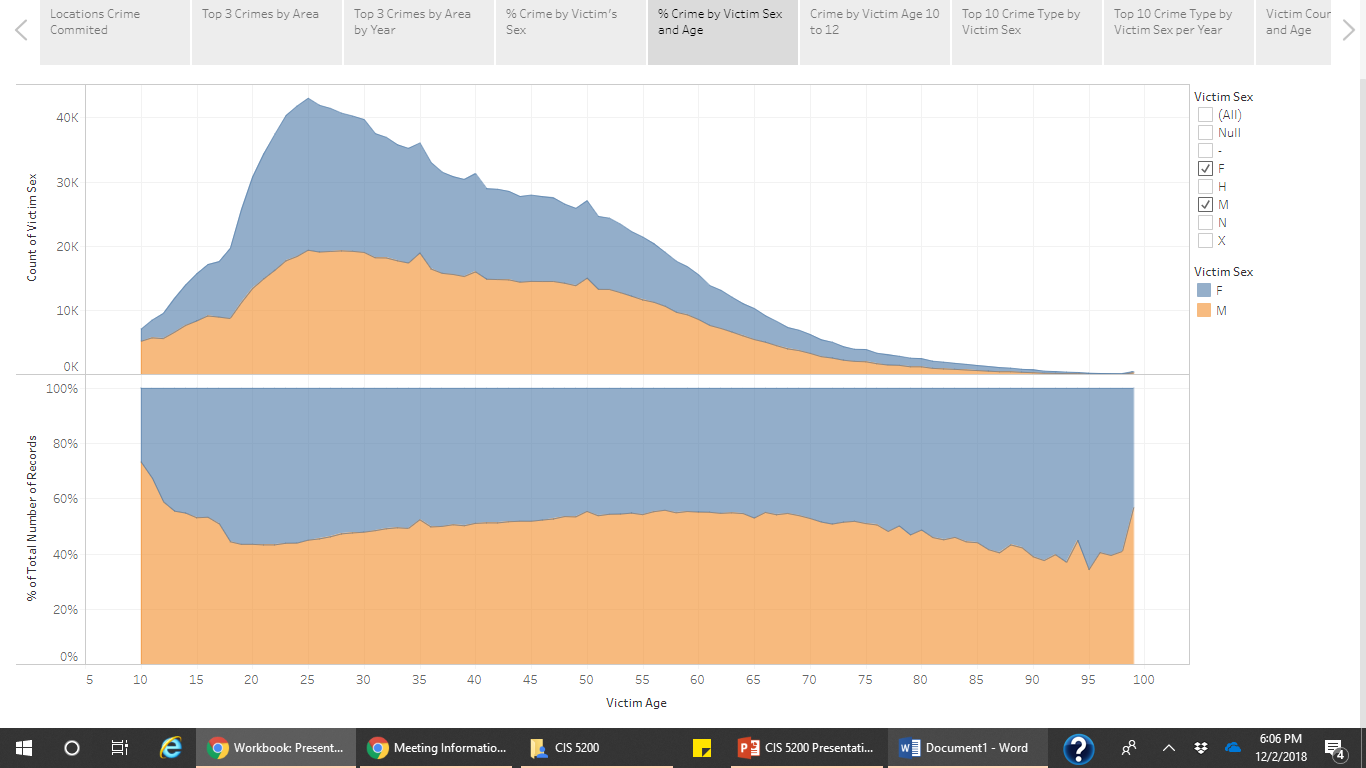
**  
Figure 4. Crime Total by Time of Day  
Source: Tableau Visualizations**

Figure 4 features an hourly breakdown of the time of day criminal activities occur in the Los Angeles area. Each column is broken up into sections representing the time of day a crime occurred with the most frequently recorded times highlighted for simplicity. Based on this chart, we can conclude that between 2010 and 2018 most crimes occurred during the noon hour. Specifically, over 100,000 crimes were reported to have occurred around noon – the highest number of crimes reported at any time. Overall, criminal activities were most likely to occur in the afternoon and evening (between 12pm and 11pm) and least likely to occur before dawn (between 3am and 6am).



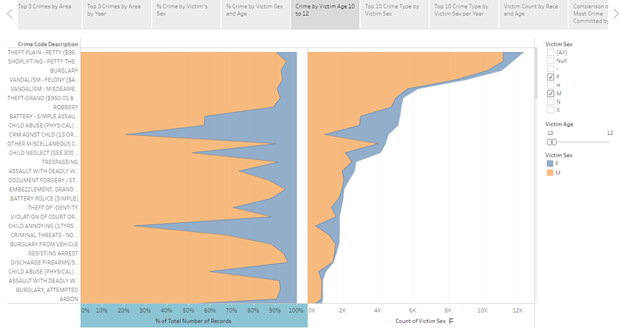
**Figure 5. Locations of Criminal Activity  
Source: Tableau Visualizations**

Every Los Angeles crime report includes a premise code and premise description to more accurately describe the scene of a crime. Examples of premise codes and descriptions include 404.0 – Department Store and 701.0 – Hospital. For simplicity, only the premise description was used to develop the Locations of Criminal Activity visualization (Figure 5), a horizontal bar graph that presents all crime locations ordered by most to least frequent. Based on this visualization, the crimes in Los Angeles between 2010 and 2018 mostly occurred on the street or in a single-family dwelling; about 800,000 crimes were reported at these top two locations. When this graph is compared with other graphs, we can start to see potential correlations. For example, battery and burglary are commonly reported crimes as shown in later graphs. When we see that single-family homes are the second most common crime location, we could conclude that more research could be done on domestic violence and break-ins in the LA area.



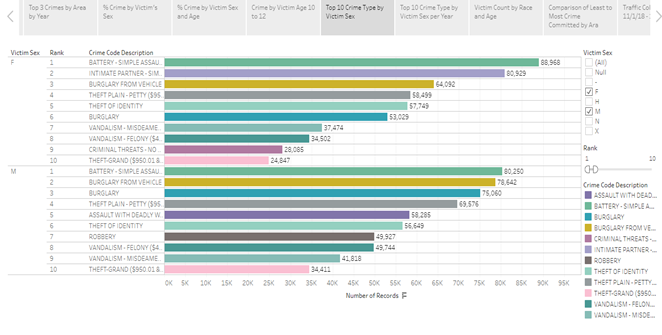
**Figure 6. Crime Percentage by Victim Sex & Age** **Source: Tableau Visualizations**

Figure 6 is made up of two stacked area charts visualizing the spread of crime victims based on age and sex. Along the X axis of both charts are victims’ ages making data trends related to age easily visible. Both charts also feature an orange shaded area for men and a blue shaded area for women. The top chart shows the number of crime victims by gender and age while the bottom chart shows the percentage of crime victims by sex and age. A quick insight this visualization provides is – overall – Los Angeles has more female crime victims - especially between the ages of 25 and 35. In addition, we can see from the second chart that men roughly make up about 40% of crime victims while women make up the remaining 60%.

****

**Figure 7. Crime towards Victims Ages 10-12  
Source: Tableau Visualizations**

While reviewing the Crime Percentage by Victim Sex & Age visualization (Figure 6), we noticed one major deviation from the trend – for the youngest victims, boys made up a significant percentage of the total victims. To analyze this data in more detail, we created the Crime towards Victims Ages 10-12 visualization (Figure 7). Like the previous visualization, the two stacked area charts show the percentage and number of child victims. However, these charts display this information based on the crime code description, so we can easily see the type of of crimes these young people are victims of. The most common crimes include petty theft, petty shoplifting, burglary, and vandalism. While these crimes could be associated with low-value items and therefore assumed to be standard delinquent behavior, others could see this same information and associate this trend with gang activity. Through additional research into where and when these crimes are occurring, we could gather more insight into which scenario is more likely to be true.

****

**Figure 8. Top 10 Crime Type by Victim’s Sex  
Source: Tableau Visualizations**

Figure 8’s horizontal bar graph displays the top 10 crime types based on number of records and victim’s sex. Based on this information, female victims suffer far more from criminal threats and battery cases where the assailant is an intimate partner, supporting the hypothesis that women are the primary victims of domestic violence. Men and women in Los Angeles suffer almost equally from identity theft and simple battery cases. Men suffer much more from assault with a deadly weapon, burglary from a vehicle, and robbery – all crimes that could be linked to gang activity.

3. Conclusion

The above analyses provide visualizations on several raw data sets, which give insight on different dimensions of crimes and crime-related activities. These are locations, times of day, types of crimes and categories of victims. These details can assist LAPD to better understand their areas and demographics. Through a better understanding, LAPD can identify areas and allocate resources accordingly. It can identify victims and provide more efficient services. For example, realizing that most victims of simple assault are youngsters can enable LAPD to bring in counselors or resources to help such young teenagers and children deal with crimes and assaults in a better way.

**3.1. Limitations**

One of the limitations of the data set was that the raw data were not recorded in the same format over the years and therefore the analysts must refine the data sets and create a consecutive order for the 2010 to 2018 data sets.   
  
**3.2 Future Works**  
The recommended future work that can be done on the data sets or use the results for future analytics are comparative analytics between different results, such as comparing sex, age and ethnicities of victims. The same can be done in relation to geographic distributions to compare demographics and crime related activities. In the next two years, analysts can use a larger data set for the 2010-2020 decade, break down analytics in two five-year periods and then compare the differences. Also, analysts can compare crime rates and activities with other data sets with economic and political data, such as comparing crime rates with economic trends during the same time period. Security increase and crime rates can be analyzed during politically eventful times such as local, statewide and national elections.

**References**

Mrandrewlai. “Mrandrewlai/Lapd-Crime-Data.” GitHub, github.com/mrandrewlai/Lapd-Crime-Data/blob/master/5200 Project.

“Crime Data from 2010 to Present | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/Crime-Data-from-2010-to-Present/y8tr-7khq.

“LAPD Calls for Service 2010 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2010/iy4q-t9vr.

“LAPD Calls for Service 2011 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2011/4tmc-7r6g.

“LAPD Calls for Service 2012 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2012/i7pm-cnmm.

“LAPD Calls for Service 2013 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2013/urhh-yf63.

“LAPD Calls for Service 2014 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2014/mgue-vbsx.

“LAPD Calls for Service 2015 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2015/tss8-455b.

“LAPD Calls for Service 2016 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2016/xwgr-xw5q.

“LAPD Calls for Service 2017 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2017/ryvm-a59m.

“LAPD Calls for Service 2018 | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/LAPD-Calls-for-Service-2018/nayp-w2tw.

“Traffic Collision Data from 2010 to Present | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/Traffic-Collision-Data-from-2010-to-Present/d5tf-ez2w.

“Vehicle and Pedestrian Stop Data 2010 to Present | Los Angeles - Open Data Portal.” Data.lacity.org, data.lacity.org/A-Safe-City/Vehicle-and-Pedestrian-Stop-Data-2010-to-Present/ci25-wgt7.