

## LA Crime Analysis - Linyan Dai, Jing Li

**URL:** [https://public.tableau.com/profile/jing.li5125#!/vizhome/LACrime\\_15762061642310/Story1?publish=yes](https://public.tableau.com/profile/jing.li5125#!/vizhome/LACrime_15762061642310/Story1?publish=yes)

### Dataset

This dataset reflects incidents of crime in the City of Los Angeles dating back to 2010. This data is transcribed from original crime reports that are typed on paper and therefore there may be some inaccuracies within the data. We used Python to clean data to make more accurate analysis.

### Analysis Goal

Our team aims to explore the following three topics:

- What is the trend or time pattern of crime incidents in LA? Does seasonality exist?
- What is the difference between each crime types in terms of victim age, sex, descent and premise?
- When it comes to incident location, does overall crime related to traffic collision?

The target audiences are the manager and policemen in LA Police Department.

### Analysis Story

From 'Trend and Seasonality', we can see the number of crime incidents are fluctuated in past years but the overall trend is increasing. Seasonality does exist because there is a drop in every February and an increase after that until August. And at the end of each year, there will be an increase as well. It is clear that Feb has the least crime incidents according to the heatmap. Focusing on the year dimension, the color of 2016-2018 are apparently deeper which means the incidents overall number are increasing.

From 'Occurrence Time', we can see most of the crime incidents happened at 16pm to 21pm, but the peak is at 12pm on every day. Comparing the weekdays, we found the least crime incidents time range on weekend are around 2 hours later than other weekdays. It is reasonable to indicate people sleep later on weekend and have more activities so that the incidents happening time is postponed as well.

From 'Occurrence Map', we provided analysis from three different dimensions. We can see the geographical distribution of different kinds of crimes and sex, whether there is maldistribution in certain types of crime and whether the number of certain types of crime changes with time. Most kinds of crime have a uniform distribution of men and women, but male victims are significantly more than female victims in 'Robbery' and 'Burglary', and female victims are significantly more than male victims in 'Raping'. The second part is explainable, but the first part is quite interesting since female is supposed to be an easier target for a robber. This is an interactive map that audiences can explore what interests them most by clicking interaction button.

From 'Crime Type vs. Age & Descent', we can see that there are no significant differences among victims' age and crime type. But we find something interesting between crime type and victim descent.

'Hispanic/Latin/Mexican' is clearly the biggest victim of many crime types, and 'Black' comes second. Also, 'Simple Assaults' is the most frequent crime type.

From 'Crime Type vs. Premise', we can see that most kinds of crime happened most frequently on the street. Burglar, robbery, and some kinds of theft happened most on single-family dwelling. Also, all crime types happened more to single-family dwelling than multi-unit dwelling. This could mean that single family is an easier target if the population of single family is not a lot larger than multi-unit family.

From 'Crime vs. Traffic', two maps on the top can show some correlation between the number of traffic collisions and the number of crimes, but the bottom graph shows that there is not much correlation between traffic and crime. We arrange the line graph of crime amounts from highest to lowest, if traffic is really related to crime, it should also be from highest to lowest.

### Key Takeaways

- Seasonality exists for the number of crimes happened.
- 12 p.m. is the peak time that crime happened the most in LA.
- Male victims are significantly more than female victims in 'Robbery' and 'Burglary', and female victims are significantly more than male victims in 'Raping'.
- 'Hispanic/Latin/Mexican' is the biggest victim group in many crime types.
- Crimes happened most frequently on the street and then single-family dwelling.
- There is no significant correlation between crime incident and traffic collision.

### Design Choice

- Dashboard 1: Line chart is very suitable for showing time series to see the overall trend. In order to explore the seasonality, we use color coding to quarter so that it's easy to compare the trend of each quarter in each year. To vividly show the difference of incidents number in every month, a heatmap is utilized.
- Dashboard 2: It's quite straightforward by using color and size encoding in heatmap to see the frequency clearly.
- Dashboard 3: A map with filters and interaction is used to represent the change of incidents in different area in the past years in terms of different crime type. In order to show the type clearly, we used python to cluster the small type to eleven big types using official type categories. Audiences can choose the crime type they want to explore and see the trend change by clicking the play button on the filter in the right bottom.
- Dashboard 4: Boxplot can represent distribution very clearly. For analyzing descent, we used packed bubbles. Crime type is represented by color and the number of different victim descent is represented by the bubble size.
- Dashboard 5: Bar chart are utilized to show the top location which is most likely for crime incidents happening. Interaction can help audiences choose the top n premise in each type easily.
- Dashboard 6: Since we want to show whether there is a correlation between crime and traffic in different areas, we chose to put two data set on two maps so we can see straightforward whether they are correlated. But we can be fooled by the map, so we also made a line graph to compare two data in one chart more precisely.