Tableau Project – Crime Analysis

ABSTRACT

The Los Angeles Police Department (LAPD), officially known as the City of Los Angeles Police Department, is the law enforcement agency for the city of Los Angeles, California, United States. It has 4 major bureaus – Central, Valley, South and West Bureau and under these bureaus, there are about 21 divisions all together. The analysis is based on the LAPD crime dataset from the years 2012 to 2015. The dataset focuses on different crime categories at various locations within LA city.

DATASET

The data set shows LAPD crime data for 4 years from 2012 to 2015. It is a huge dataset with around 935K records and about 14 columns. Each row has a crime data and it is one of the cleanest dataset that requires minimal effort in data cleaning. The column consist of the fields such as date.rptd- date on which crime was reported, dr.no – the crime no, date and time occ - date and time of occurrence, area no and area name, reporting district no , crime code, crime description, status, status description, location , street and location1 – which is the latitude and longitude details.

The crime code is grouped into different crime categories ad further grouped into includes all the Part I and Part II crimes. Part I Crimes are the eight "serious offenses" for which the FBI gathers national data including Homicide, Rape, Robbery, Aggravated Assaults, Burglary, Larceny, Vehicle Theft and Arson. Part II Crimes are "less serious" offenses and include: Simple Assaults, Forgery/Counterfeiting, Embezzlement/Fraud, Receiving Stolen Property, Weapon Violations, Prostitution, Sex Crimes, Crimes Against Family/Child, Narcotic Drug Laws, Liquor Laws, Drunkenness, Disturbing the Peace, Disorderly Conduct, Gambling, DUI and Moving Traffic Violations. For this analysis we have categorized the crime code as following based on the Crime Statistics Glossary available on LAPD online website:

PART I CRIMES		PART II CRIMES	
HOM	CRIMINAL HOMICIDE	NON	NON-AGGRAVATED ASSAULTS
RAPE	FORCIBLE RAPE	VAND	VANDALISM
ROBB	ROBBERY	MISD	MISDEMEANORS
AGG	AGGRAVATED ASSAULT	MISC	MISCELLANEOUS
BURG	BURGLARY	SEX_CRIMES	SEX CRIMES

GTA	GRAND THEFT AUTO	BTFV	BURGLARY/THEFT FROM VEHICLE
ARSON	ARSON	THEFT	THEFT
GTP	GRAND THEFT PROPERTY	KID	KIDNAPPING
		OTH	OTHER

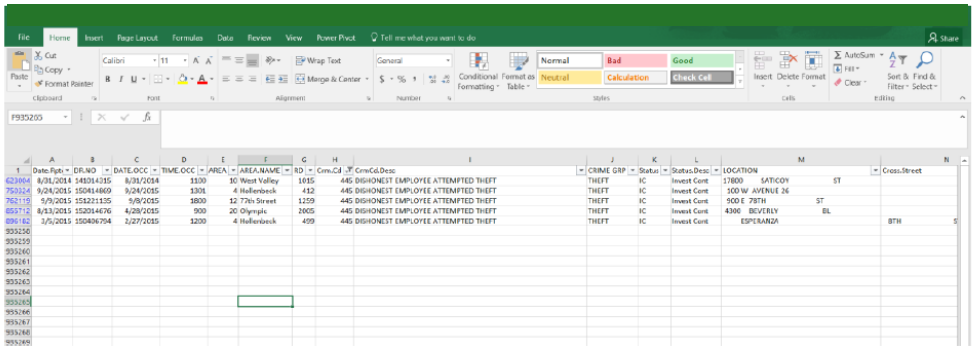
The status column gives us details on the crime status such as AA,AO,CC,IC,JO and JA. JO and JA are status updates for juvenile crime and it is this status in the dataset which helps us to distinguish between the crimes done by adults and juvenile. The location1 one has the latitude and longitude details of location of the crime. The reporting district no is a policing area defined by the Los Angeles Police Department. In the LAPD's system it belongs to a specific Basic Car Area, which is a member of one of the Division and Bureau.

DATA CLEANING

The dataset required very little cleaning to be done. The corrections made on the existing dataset are as follows:

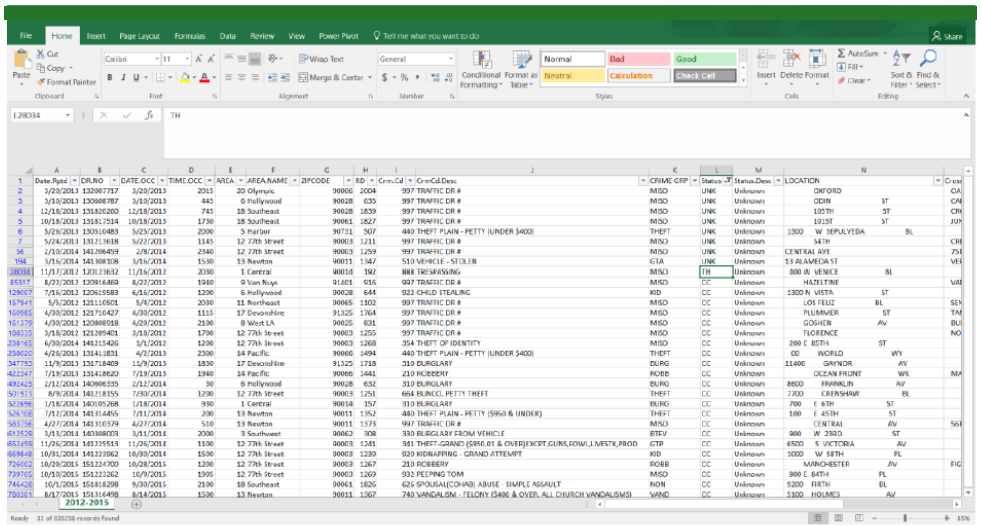
1. Missing Value – Crime description missing

   After cleaning:

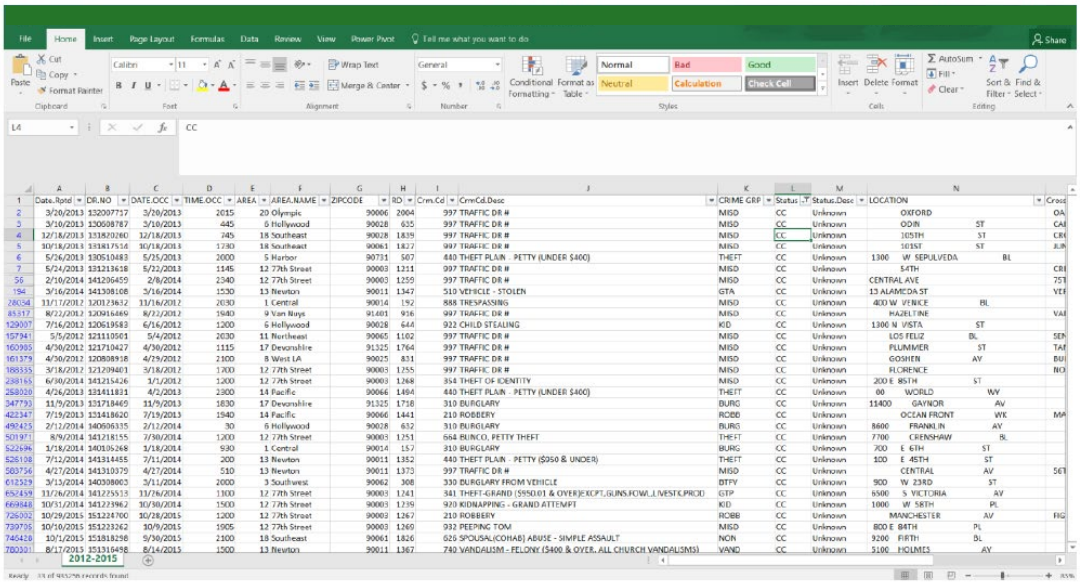


2. Mismatch in the status value. UNK is the description for status type CC. TH is unknown.

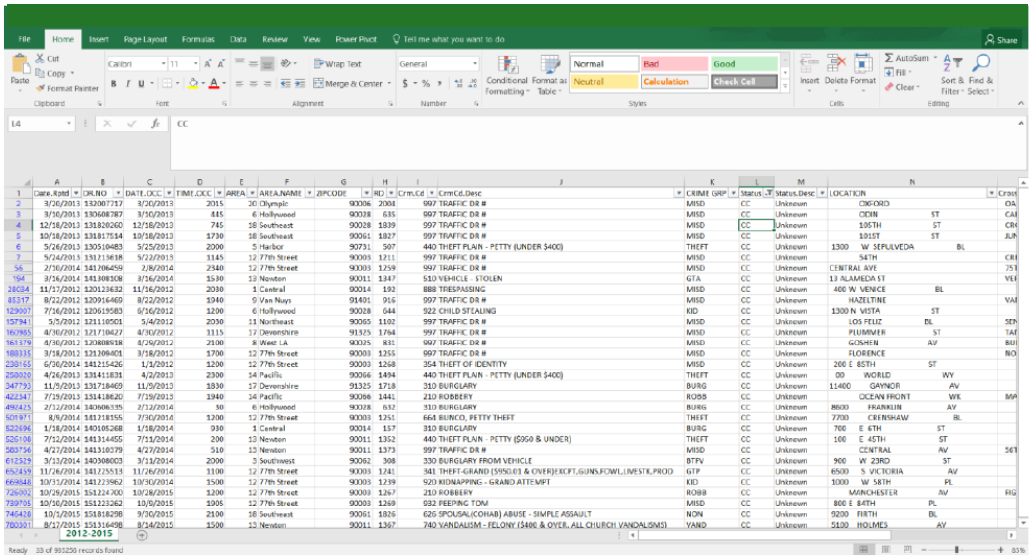
   Hence status UNK and TH has been replaced with CC because of the same description.



After cleaning:



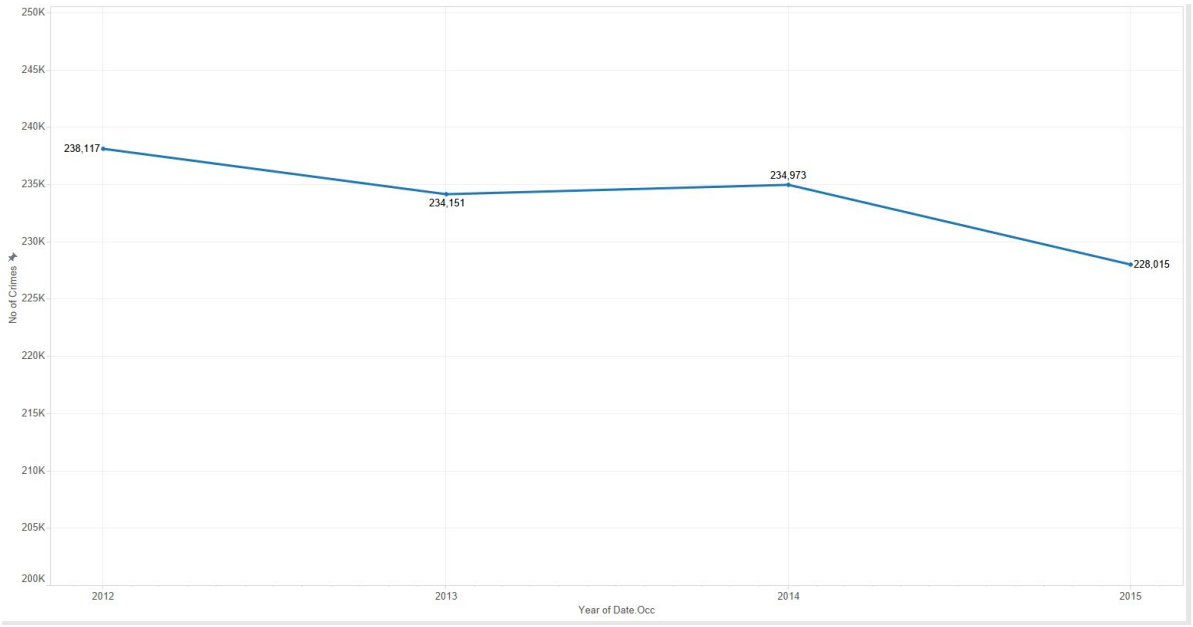
3. Addition of a new columns named Crime group and zip code. This was done to group the various crime codes which would help in portraying a better data visualization.



DATA VISUALISATION

The dataset has been analyses in such a manner that we get to know the total no of crimes committed in each year, under each bureau, the types of crimes committed and finally the status of these crimes that have been reported. These analysis helps us to understand how the LAPD have functioned over the years in order to create a safe environment for us.

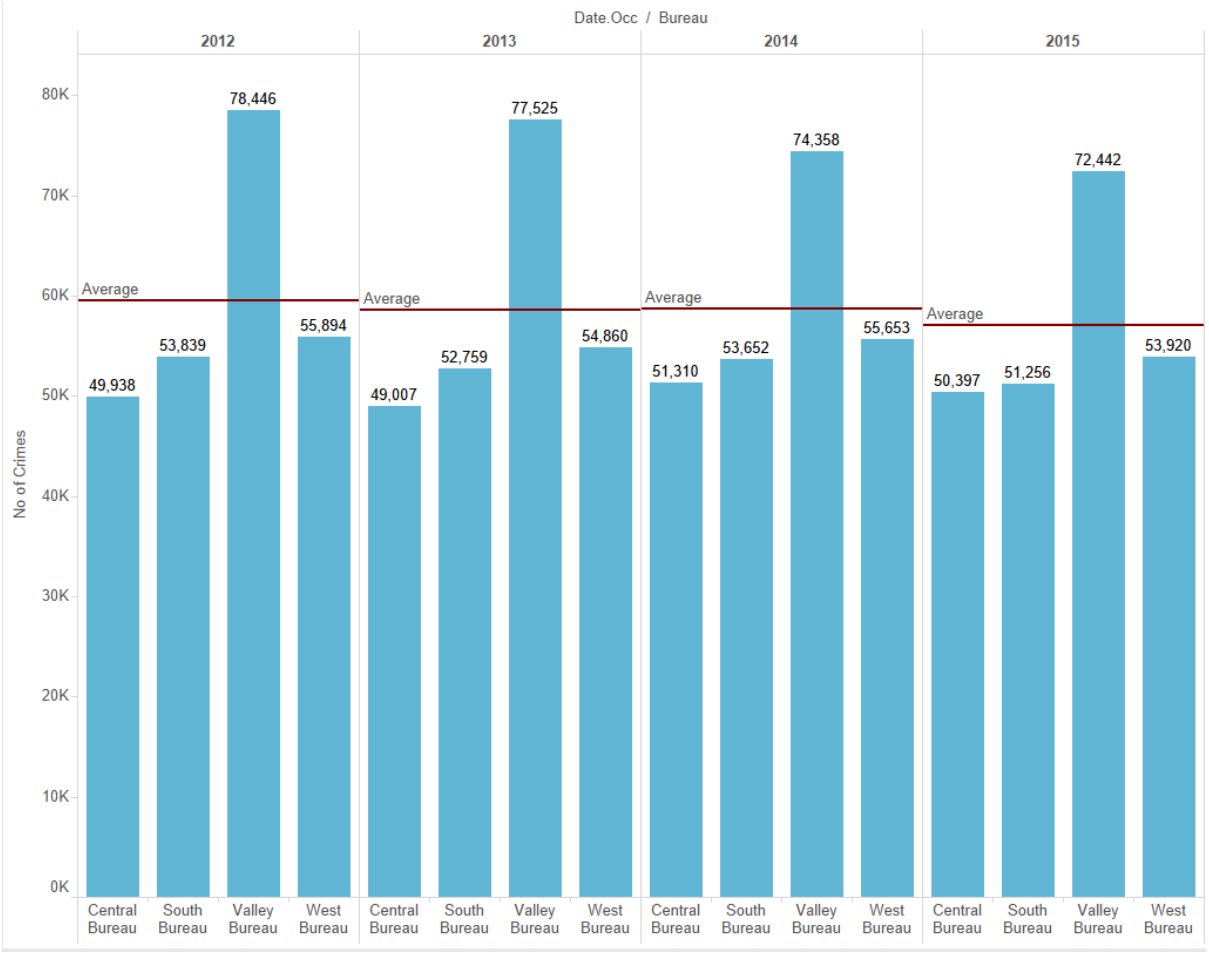
An analysis made on the no of crime of rimes over the years shows us whether the crimes have increased or decreased in number. A line graph has been used for the visual representation and it shows us the variation in crimes from 2012 to 2015. The total no of crimes was the highest in the year 2012 around 238K, though it reduced in 2013 there was an increase in 2014 and in 2015 it was the lowest around 228K.



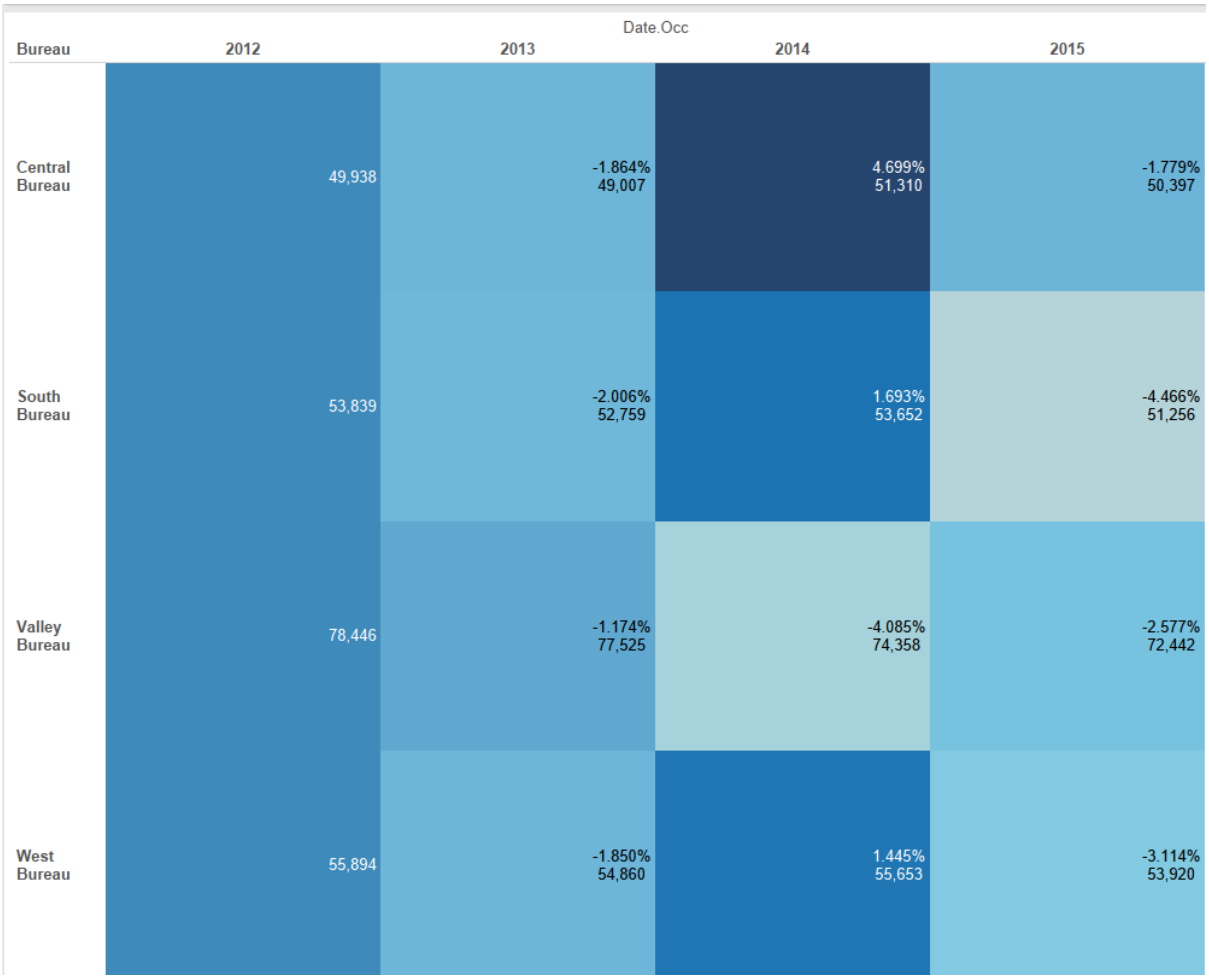
The above graph provides us with the basic idea about the crimes in LA. Hence another graph was plotted using bar graphs which shows us the crimes over the years under each bureau. This graph helps us to understand which bureau has the maximum crime rates and which has the minimum crime rates. A reference line has been plotted for each year. The reference line is the average of total crimes over the year. It shows us which bureaus have crimes more than the average value. The 21 areas names are grouped under the 4 bureaus using the create group option in tableau.

Bureau	Area/Division Name
Central	Central
	Hollenbeck
	Newton
	Northeast
	Rampart
Valley	Devonshire
	Foothill
	Mission
	North
	Topanga
	Van Nuys
	West Valley
South	77th Street
	Harbor
	Southwest
	Southeast
West	Hollywood
	Pacific

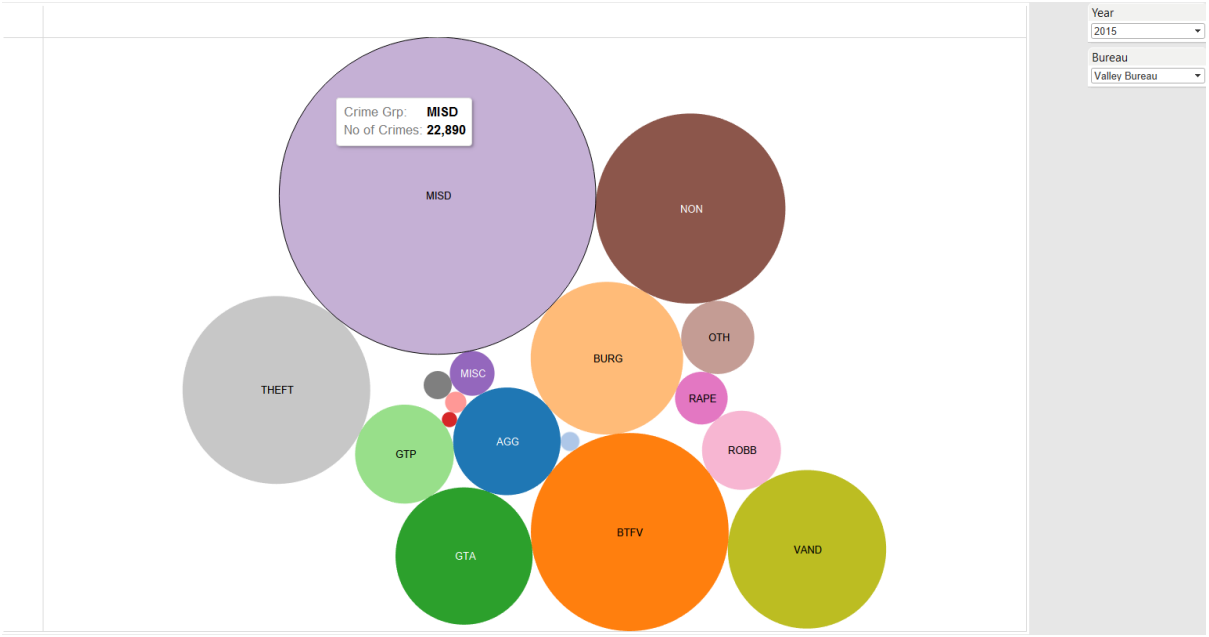
	West LA
	Wilshire
	Olympic



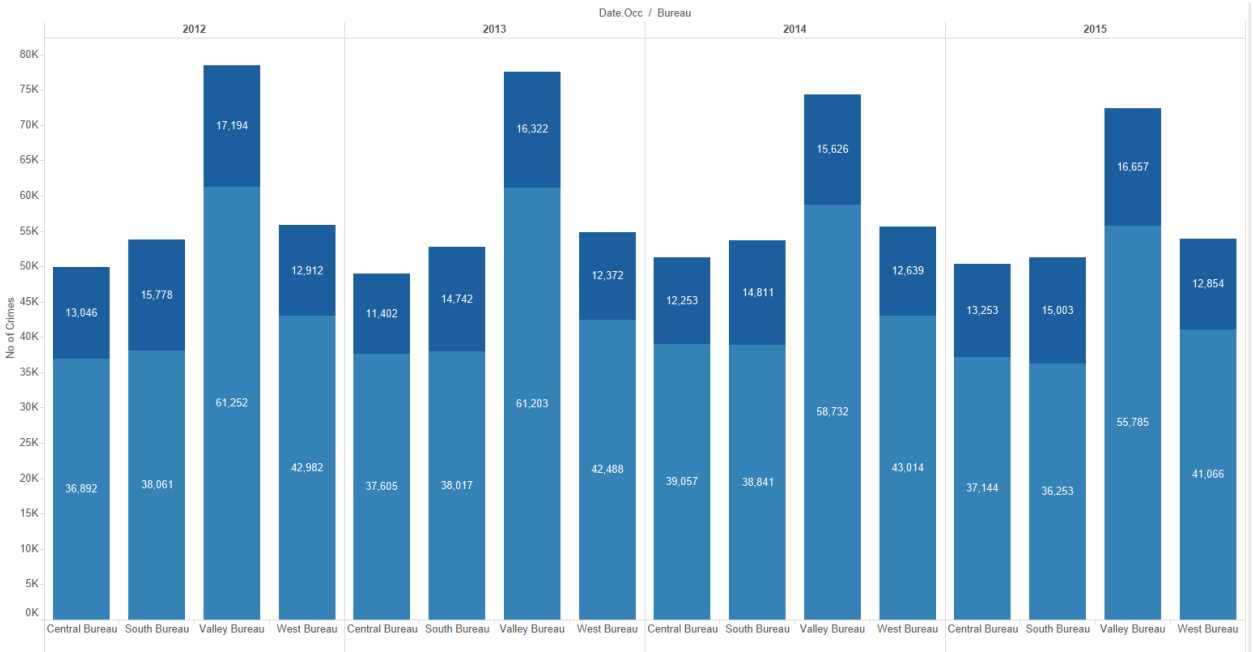
In order to get the exact increase and decrease over the years under each bureau we can perform difference in percentage calculation. Where the value of one year is percentage difference relative to the previous year. Different color shades have been provided and hence this helps us to determine at a single glance which bureau in which year has the reduced the most compared to the previous year. It is evident that south bureau has reduced in 2015 and Central bureau has increased in crimes in 2014. Such visualizations make data analysis easier.



Analysis made between the crime group, bureau and year. A colorful bubble chart has been used for visualization. The chart shows us which type of crime has been committed the most under each bureau in each year. Here the year and the bureau is set as a filter and is made available as a drop down using the show filter option. The below graph depicts the total no of crimes committed under each category for the year 2015 and Valley Bureau. We can choose multiple bureaus and multiple years as it depends on how we would like to represent the data. Here the total no of crimes is provided in the tooltip as it is difficult to represent the count on the bubble chart due to the various size of the bubble.

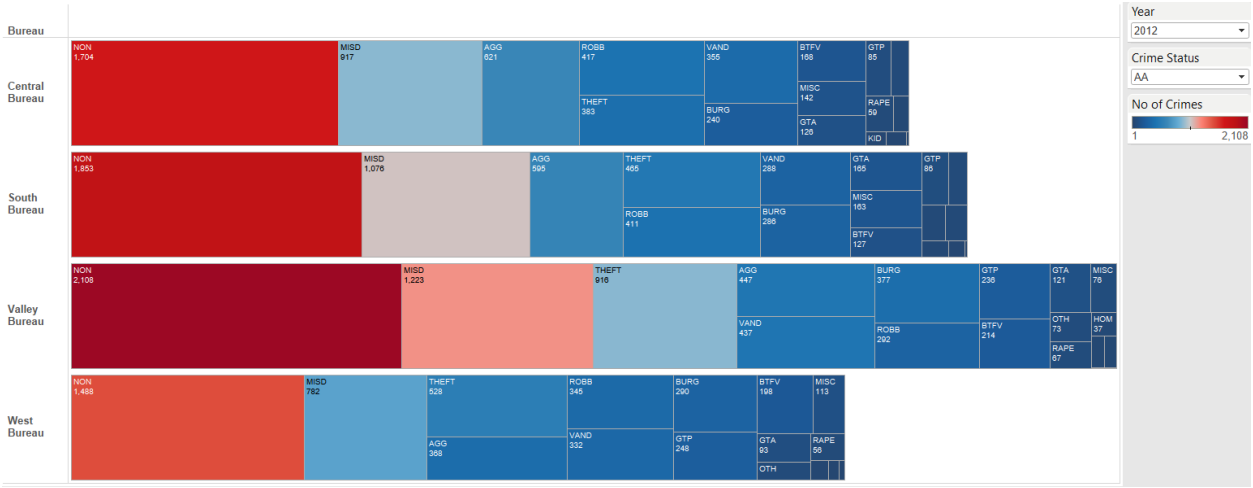


The LAPD segregates the crimes as Part I and Part II crimes based on the degree of offense. An analysis has been made to examine the count of Part I or Part II crimes under each bureau. A stacked bar graph is used to visually represent the above analysis.

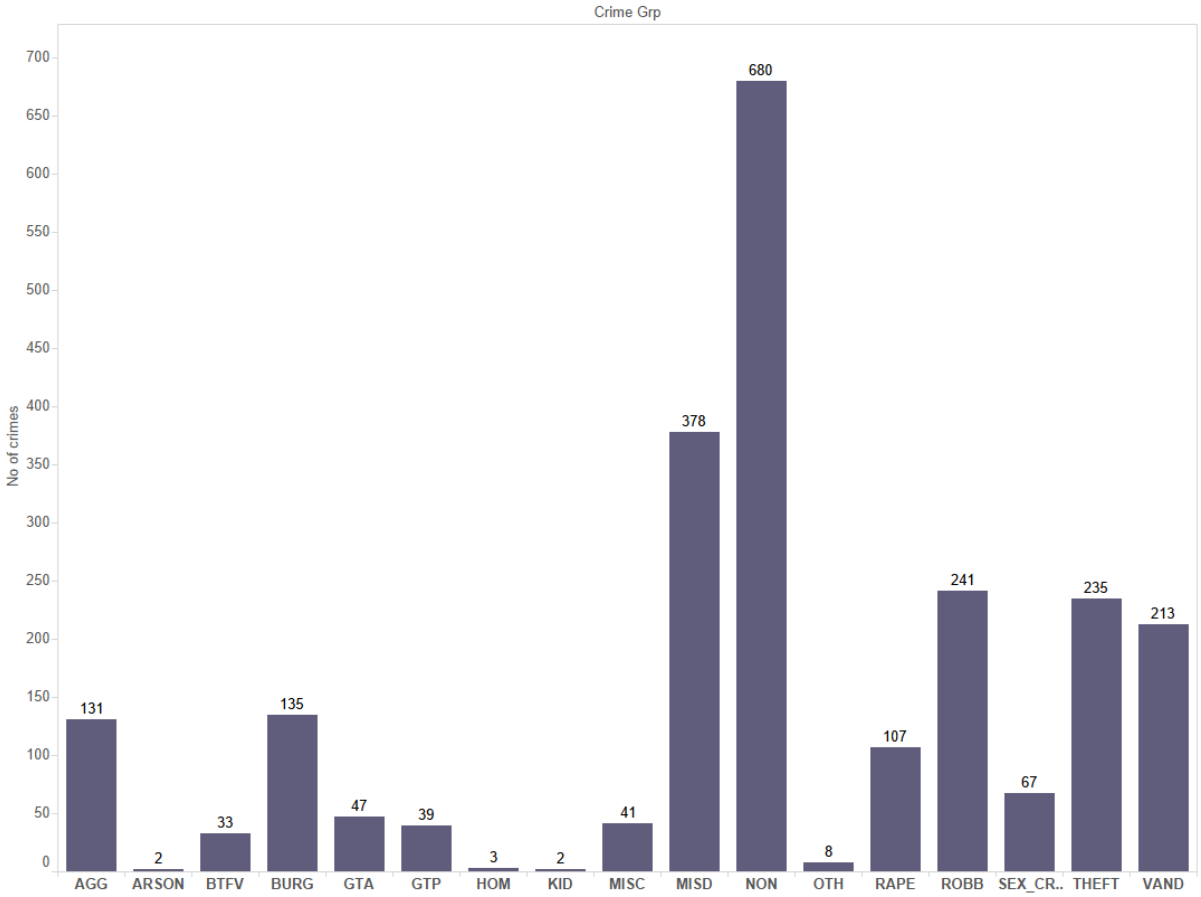


Crime status helps us to determine the status of the crimes recorded. Hence an analysis is made between the crime types, bureau, total no of crimes, the year and the crime status. We have used a tree map in order to display the desired visualization of the analysis. In the tree map we are able to see the distributions of adult

arrest made in the year 2012 under each category. The maximum no of adult arrest was made for NON crime category in valley bureau.

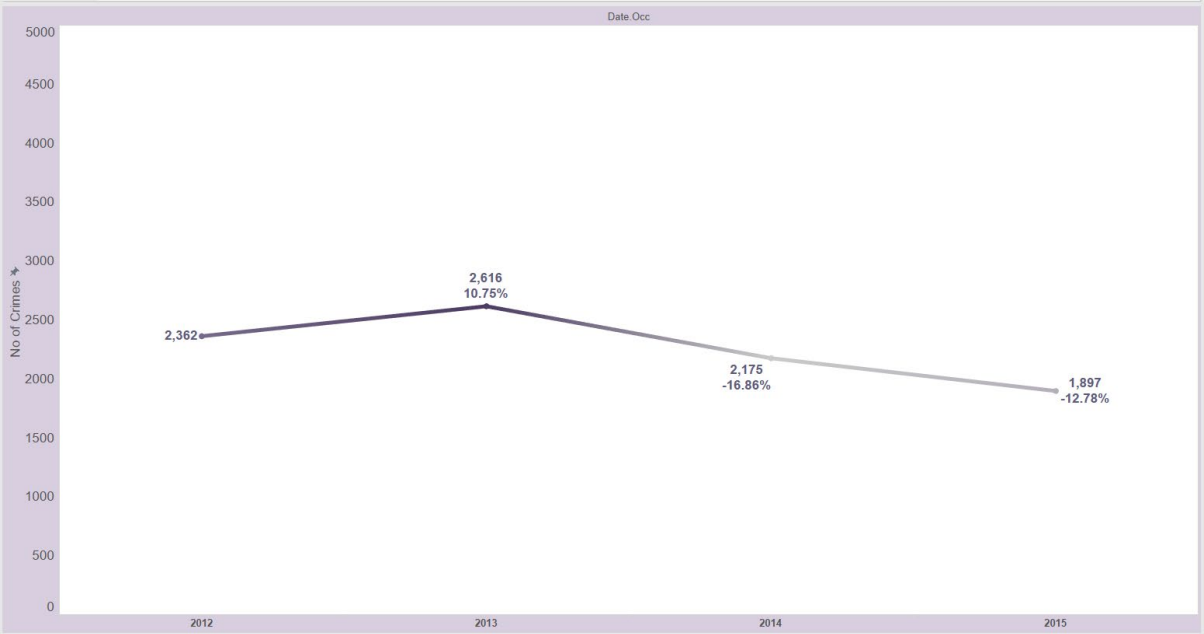


An analysis was made to determine the total amount of juvenile crimes in LA city and also in which category is the juvenile crime prevalent. The below graph helps us to analyze it visually.



Compare the rate of juvenile crimes over the years and to find out if there was a increase or decrease in juvenile crimes over the years. We have used a line graph to display the analysis visually. The graph shows that

after 2013 there has been a gradual decrease in the juvenile crimes in LA city. We have also determined the percentage difference relative to the previous year. This helps us to see that in 2014 the percentage decrease in crime was more.

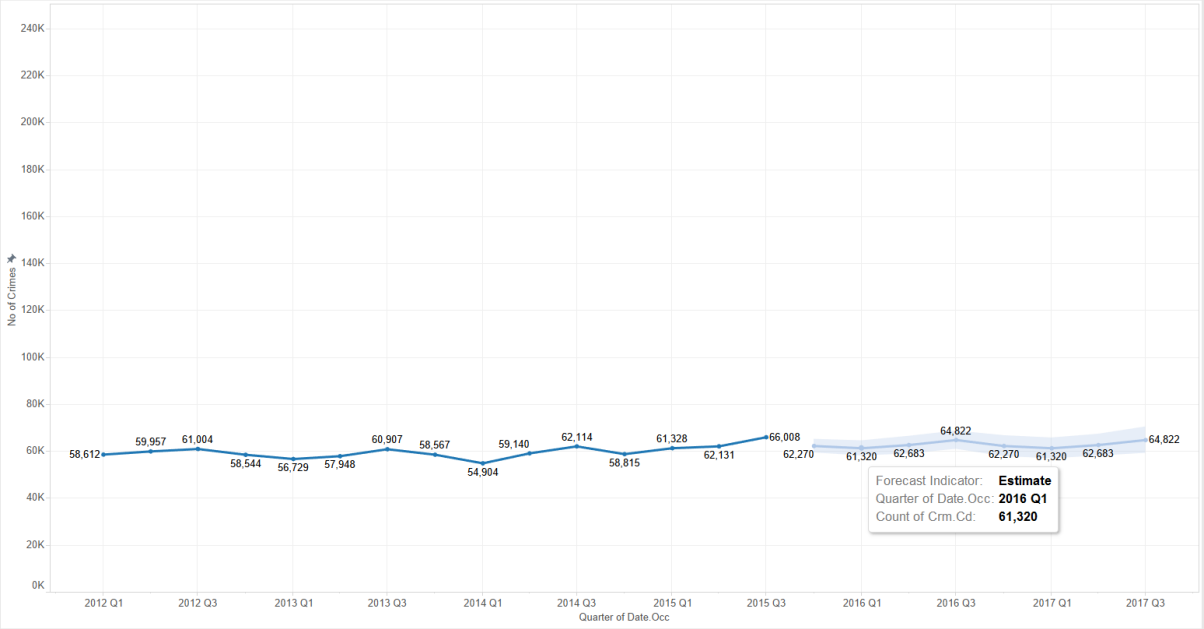


In order to understand which bureau ranks in the maximum juvenile crime a simple text graph is utilized to display the rank based on no of juvenile crimes over the years with respect to each bureau.

Date.Occ

Bureau	No of Crimes				Rank based on Maximum Juvenile Crimes			
	2012	2013	2014	2015	2012	2013	2014	2015
Central Bureau	352.0	458.0	364.0	300.0	4.0	3.0	3.0	3.0
South Bureau	867.0	911.0	689.0	693.0	1.0	1.0	2.0	1.0
Valley Bureau	774.0	904.0	799.0	619.0	2.0	2.0	1.0	2.0
West Bureau	369.0	343.0	323.0	285.0	3.0	4.0	4.0	4.0

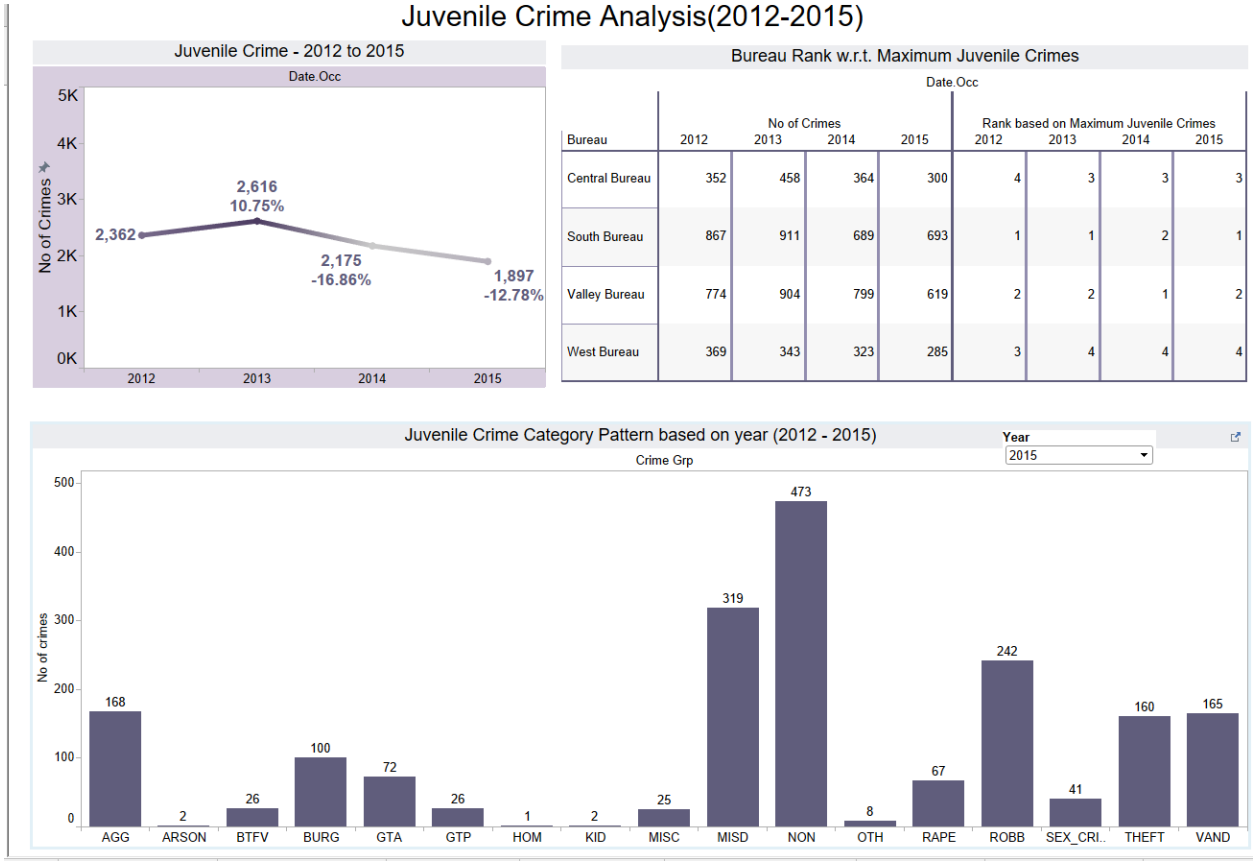
Finally, no analysis is complete without a prentice analysis of the dataset. Hence a forecast of the crimes is made based on the quarters (QI to Q3) of the years 2012 to 2015. Since year wise there was only 4 values of total crimes available, it was insufficient to make a predictive analysis. Hence predictive analysis was done based on quarters than years.



DASHBOARD

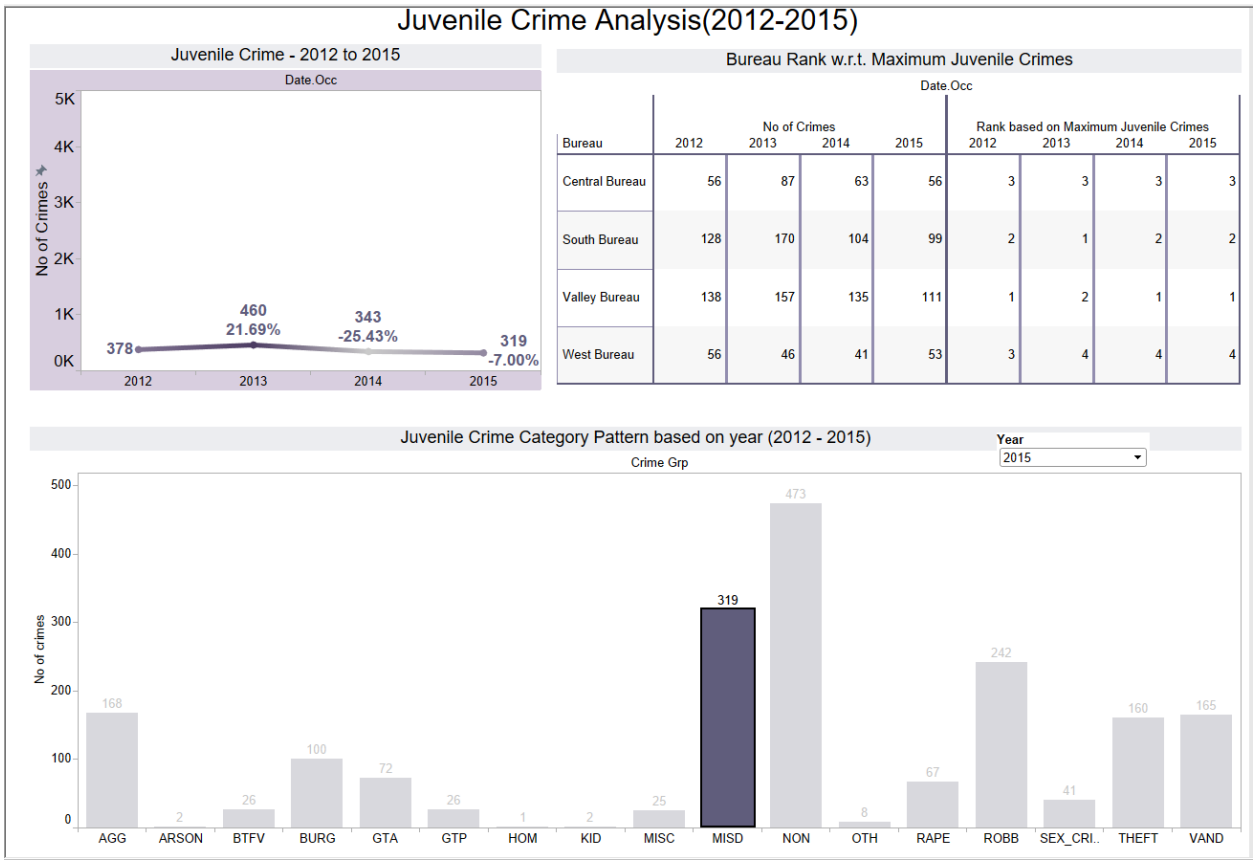
The LAPD crime dataset has been analyzed based on the total no of crimes and the total juvenile crimes under various crime categories, bureau, and year. A proper analysis with respect to the juvenile crimes helps us to understand the pattern of crimes prevalent in juveniles and by studying this pattern we can prevent children from committing crimes in the future. Thus helping such unfortunate children in building a promising future. From the below mentioned dashboard we can which bureau has the maximum crimes in the year and the different categories of crimes committed in the year 2015. The line graph tells us that year 2015 has the least no of juvenile crime. From these 3 graphs on the dashboard we can determine that which bureau had the maximum juvenile crime and what was the overall juvenile crime, it also helps us to compare the juvenile crimes per bureau with the maximum juvenile crimes in that year. The other graph helps us to understand what type of crime category was the most predominant in the year 2015.

Without Filter:



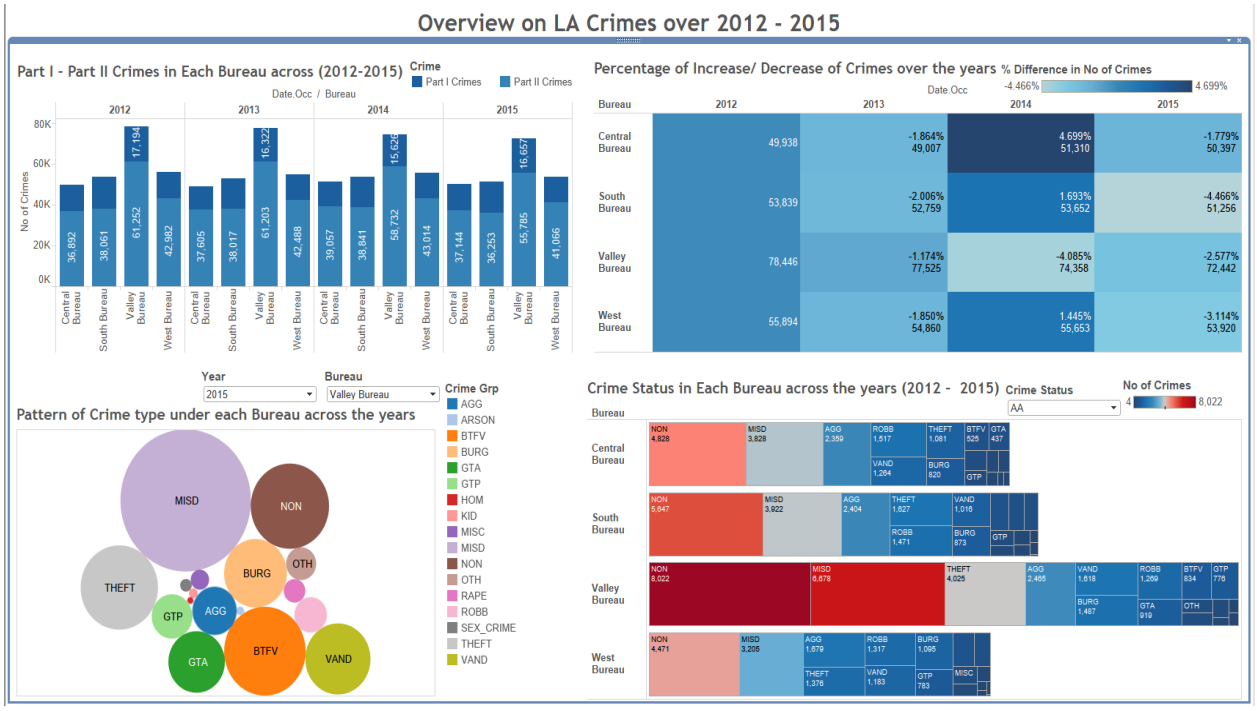
Now when we apply a filter on the Juvenile crime category and choose MISD we will obtain a different perception on the other two graphs. The line graph will show the no of MISD crimes committed over the years and the percentage difference relative to the previous year and the Bureau rank wise graph will show the no of MISD crime committed is each bureau and also the rankings over the years. This visual analysis provides us with a transparent understanding on the juvenile crimes than the data.

With Filter :

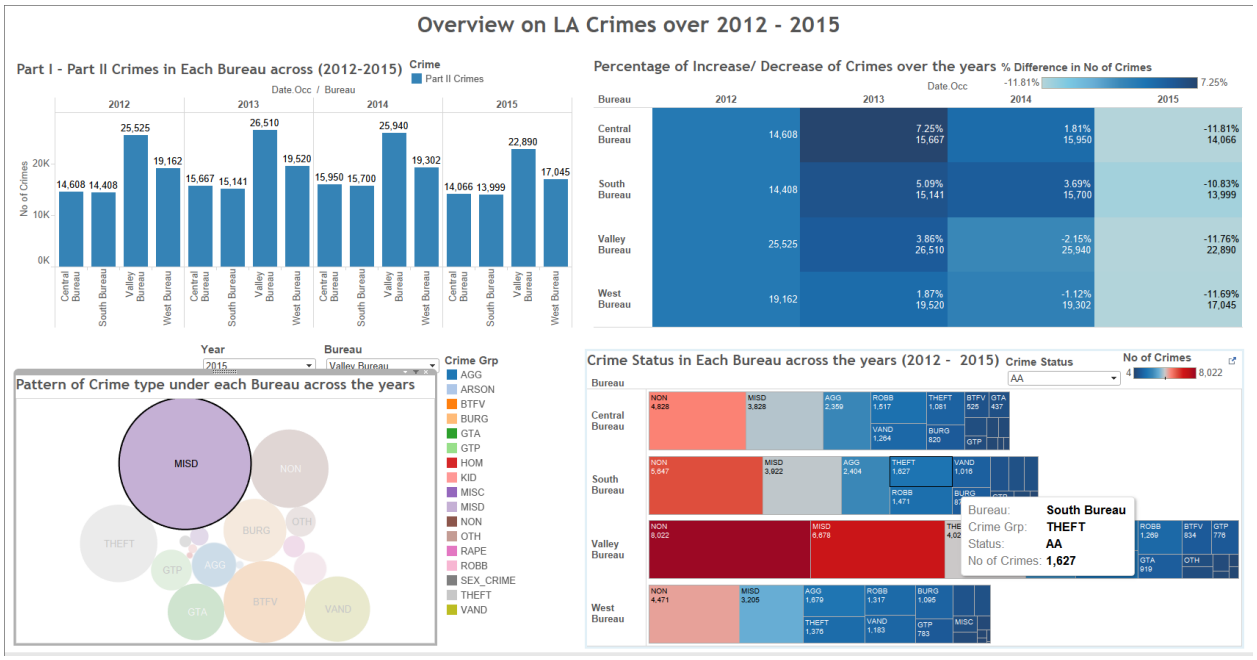


The next dashboard consists of Crime patterns based on crime category, status, bureau and total no of crimes over the years. In the dashboard action a filter is set on the bubble chart and except for the tree map remaining two charts are checked as the target sheets on which the filter will work. Now when we choose each crime category in the bubble chart, the corresponding value of that crime category will appear in the other two chars. Suppose we choose MISD category in the bubble chart then we get only the part II data over the years across each bureau in the first graph because MISD is part II crime. Similarly the second graph the difference in percentage of MISD relative to the previous year and also the total MISD under each year across each bureau.

Without Filter :



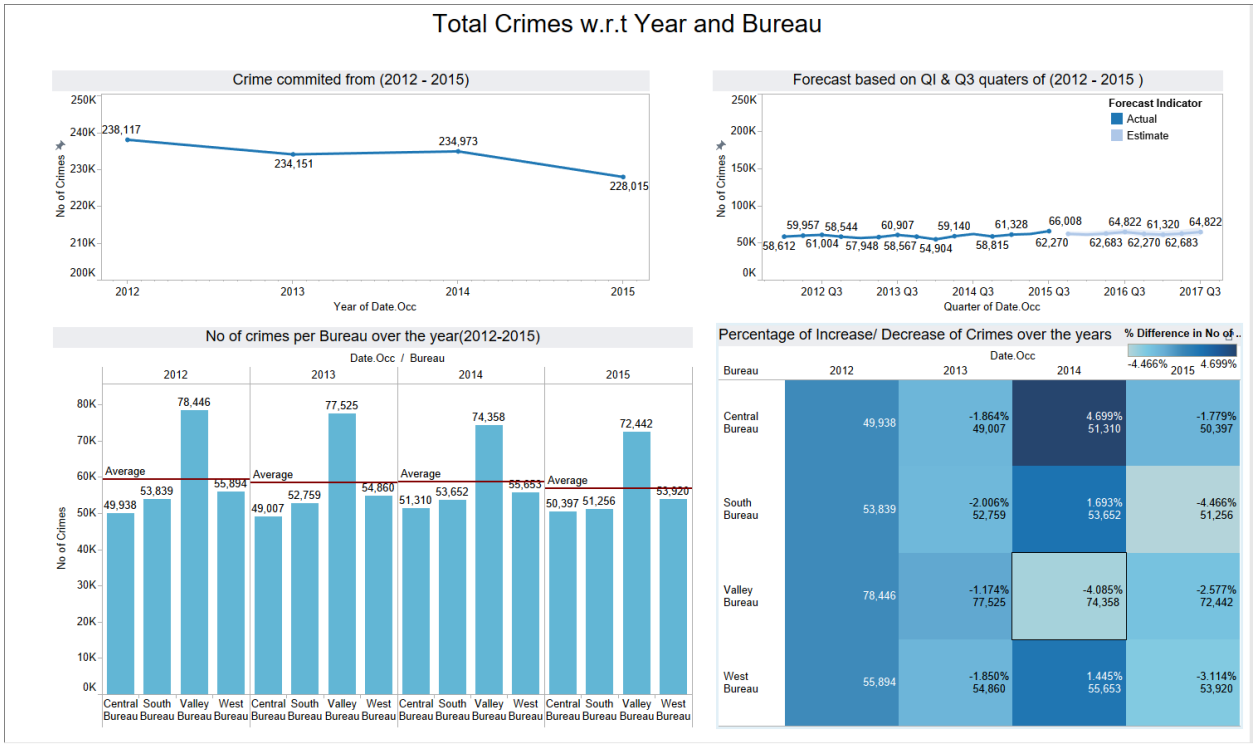
With Filter :



STORY TELLING

LAPD crime dataset consist of all crimes registered from 2012 to 2015 including fields such as date.rptd- date on which crime was reported, dr.no – the crime no, date and time occ - date and time of occurrence, area no and area name, reporting district no , crime code, crime description, status, status description, location , street and location1 – which is the latitude and longitude details. Some additional research was conducted in order to fetch data that was required for the proper visual analysis, such as grouping the area of crime w.r.t the bureau, then grouping crime codes to crime category and then further grouping into Part I and Part II crime. These groupings have made the visualization of such a huge dataset appeasing and interesting. Some of the interesting analysis made on this data are discussed below with supporting images from Tableau story.

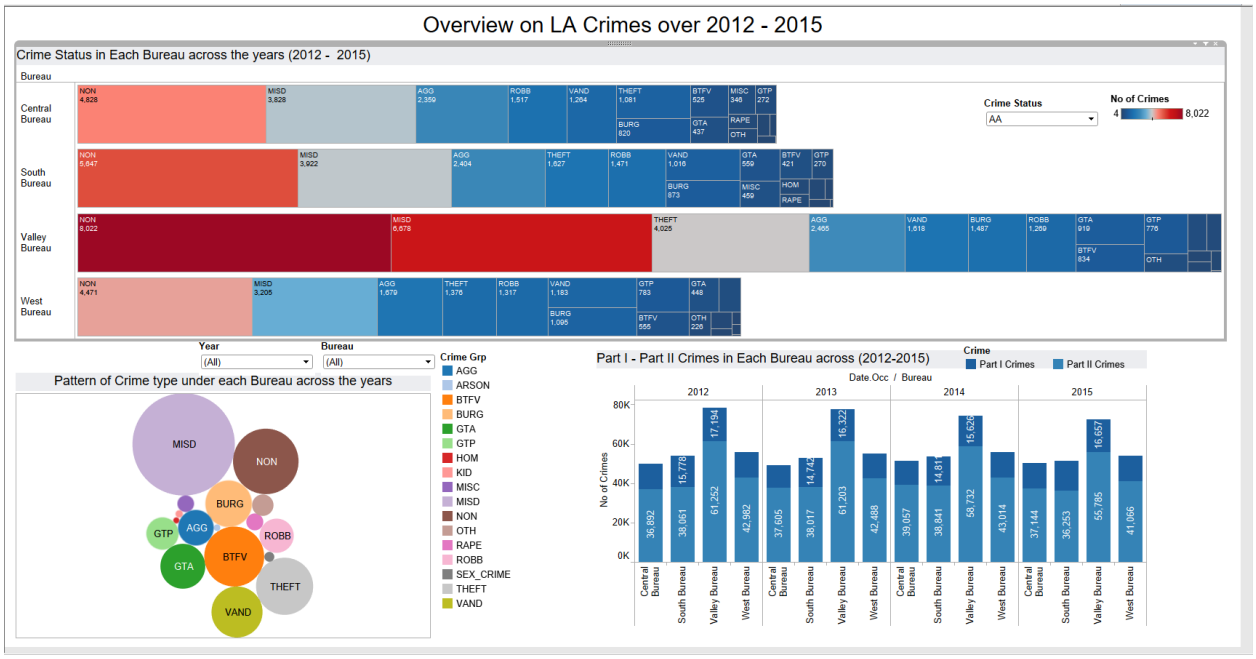
Crime awareness is crucial in today's world and one must always be aware of the criminal activities within the area they live. This is one of the main reason that stirred an interest for data analysis on LAPD crime. The crimes committed can be measured against factors like years, crime types, area, bureau and so on. Based on the data available a simple line chart is plotted displaying the total no of crimes over the years and also a line chart showing predictive analysis based on quarters Q1 and Q3 value over the years is also plotted for the next 2 years. This chart forecast the crime rates over the next 2 years which would be a great help for the law and enforcement team as it makes them well prepared for the future requirements. Below there are graphs plotted bureau and years. This provides us with information on how many crimes are reported at each bureau in a year and we can also analyze fluctuation of the crime rates in each bureau. When is plot a reference line which is the average value of the total no of crimes in that year, we can determine which bureau has crimes above and below the reference line. We can further perform the same function over the years too. Once we have the knowledge about crime rates in our area with respect to year and bureau, it is interesting to find the percentage difference in crime rates relative to the previous years. It tells us by what percentage there is an increase or decrease in the crime rates over the years under each bureau.



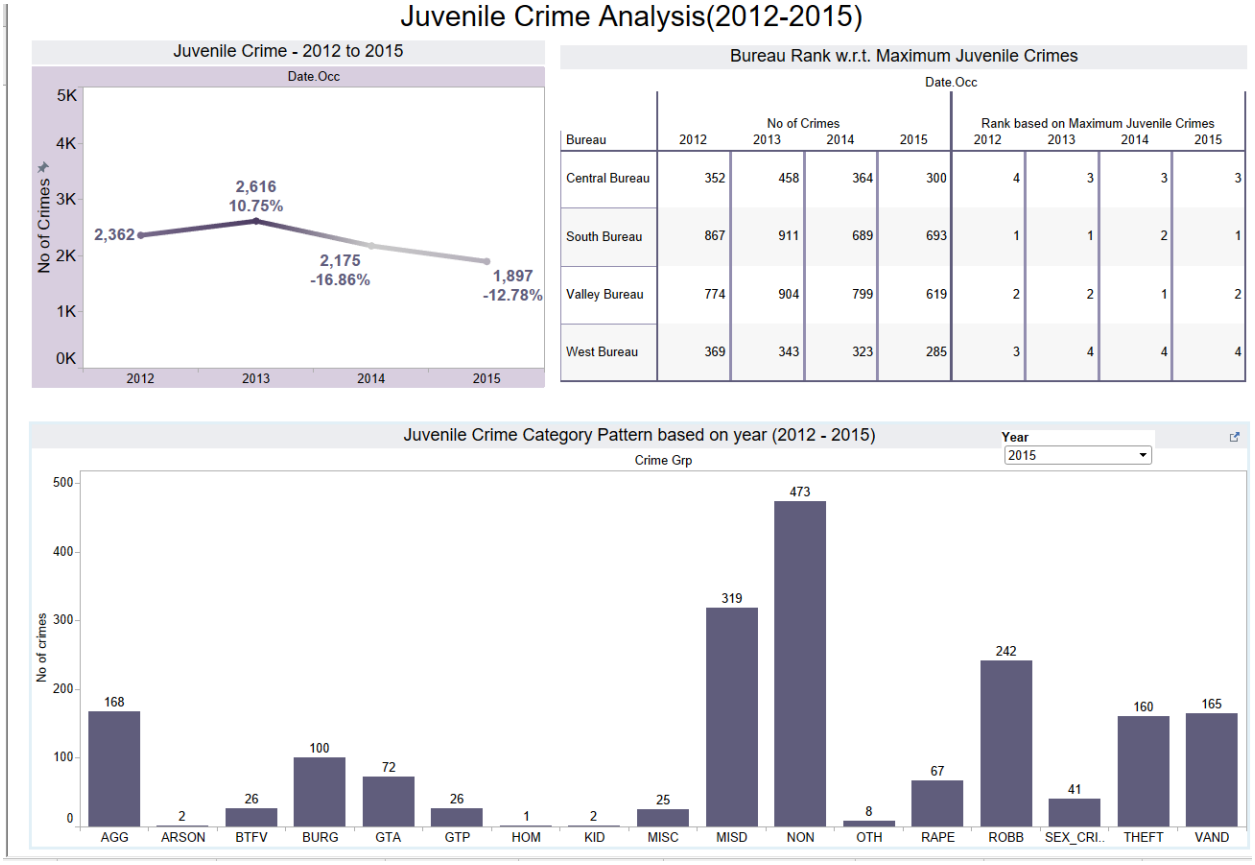
Crime Pattern Theory is a way of explaining why crimes are committed in certain areas. This theory has inspired me to analyze the dataset based on the crime patterns. Visually displaying all the crime codes in the dataset would snatch away the essence of data visualization and hence crime codes are grouped into crime categories and further grouped into Part I and Part II crimes. The grouping pattern has been discussed earlier in this document. The below graphs depicts the crime pattern in LA city over the years 2012 to 2015. The first graph displays the crime status for each crime category under each bureau for each year. This graph illustrates the actions and the efforts taken by each bureau in order to solve the crimes reported at their division. The status AA resembles an adult arrested for the crime. From the graph we can see the various crime categories under each bureau and the no of adults arrested in each crime categories. The color variation in the tree graph makes it a cake walk to determine which bureau has the maximum arrest under various crimes categories. The status is a dropdown filter with the following values – AA,AO,JA,JO,CC,IC.

The bubbe chart shows the no of crimes committed per crime categories under each bureau for each year. This gives us an idea about the most prevalent crime type in the LA city and whether these have decreased or increased over years. The stacked vertical bar chart is a representation of Part I and Part II in each bureau. The grouping of crime categories into Part I and Part II is done based on the mapping provided in LAPD online website. This provides us with the information on the no of violent crimes reported in these 4 bureaus over

the years. It is a releif to see that though the no of crimes on an average is more than 40k the violent crimes or Part I crimes constitute a smaller region. These three analysis potrays the crime pattern under different criterias.



Another major issue our society is facing today is Juvenile crimes. It is unfortunate to see young minds committing crimes voluntarily or unvoluntarily and thus putting their future at stake. An analysis on the Juvenile crimes will provide us the insight about the future of our society. The below graphs represent juvenile crimes committed over the years 2012 to 2015. The total no of juvenile crimes are measured against various factors like crime category, bureau, and years. A ranking is performed on the bureau with maximum no of crimes over the years, this will give us a better insight on understanding which bureau has taken efforts to minimize juvenile crimes over the years. In general without applying an filters on crime category from those 3 graphs on the dashboard we can determine that which bureau had the maximum juvenile crime and what was the overall juvenile crime, it also helps us to compare the juvenile crimes per bureau with the maximum juvenile crimes in that year. The juvenile crime category graph helps us to understand what type of crime category was the most predominant in the year 2015. When we apply a filter on the crime category field we will be able to determine determine the bureau ranking for the selected category field and we will be able to compare the percentage difference of crime relative to the previous years.



Crime and criminal activities can never be eradicated from our society but it can be curbed to a greater extend by the efforts from the law and enforcement department and the citizens. We as citizens do have responsibility in creating a safe living enviroment and one of the steps that we should diligently follow is to report crime if we come across one. It is not necessary that we should be the victims but we can prevents other people from being victims but reporting crimes on the spot.