Police Shootings

Naga Tejaswi Veluri

AIT-580

Abstract

The study is all about the police shoot outs that took place between 2015 to 2019.Police shootings became a highly spoken topic after few shoot outs of specific race people caught the attention of public. This report highlights the findings from the analysis performed on data collected from Corgis. Data. And try to examine if the shootings were race biased or the biased to any factor. That is if any abnormal behaviour of any particular category of people that lead to the incident. During this analysis, R-programming language was used to data pre-processing and python was used to draw the needful graphs and SQL was used to draw the conclusions and the required analysis was made use of all the three and from the analysis, we can depict that, major incidents took place in California on men during all these years.

The result of this analysis can be used to check if there is any pattern for attacking or can be used to caution the public and take required actions or precautionary measures to mitigate the attacks. More precisely can be used to check the advancements or the changes taking place while attacking during these years

I. Introduction

With advancements in technologies, these days they are used in every field to tackle issues. Security is one such field, in which big data can play a vital role to transform the way we look into issues. The selected (Police shootings) data set is a collection of shoot outs that took place between 2015-2019.By help of this analysis, we can foresee the patterns or attacking style evolution and can be prepared to encounter the situation.

During this analysis, we try to unfold the relation between

- attack style and weapons used based on year.
- change in attacks based on factors such as race and age.
- count of shoot outs based on place and stated per years.

II. Data Set

Police Shooting data set contains data that can be mainly divided into three categories, that is incident details, victim's details and the factors that lead to the incident.

Incident details describes the place, state and date of occurrence of the event. Victim's details brief us with the name of person, age, gender and race where as factors lead to incident explain us about the types of weapons used in attack and the victims action post the attack and this data set also gives us information about the mental condition of person involved in the incident and the overall threat level of the incident and Shooting manner.

III. Literature Survey

Being a sensitive topic, there are only few works related to this topic, out of which one project gives the analysis of majority data attribute and tries to find the rates based on the categories present in data.

And also explores which category of data is mostly involved in the incidents. Here, we can get the count of total incidents in the state irrespective of years. And the author used two different graphs which represents race and age of victim.

In the other project, the developer tried to show the vary in rate of shootings by years based on race.

He used a bar graph that explain the number of persons who were affected in the attack by grouping them based on race, different colours were used to differentiate the years in the plot.

But this analysis does not give any clear picture of data attributes present in the data set. This analysis gives a very brief view of data.

There is a project, which explained about the most contributing factor that led to shooting.

As per their analysis, they did not get adequate data to analyse if there is any relation between race and occurrence of incident, but in this analysis, I would like to find the relation between race and age of individuals.

During this project, they tried to find if there is any pattern followed in states, but in the proposed analysis, I would like to see if there is any change in pattern based on years.

In this project, a box plot was used to see the age range of individual for every race, but did not explain the its change during years.

Attacking style and weapons used for attacking were not considered as impacting factors in this project.

IV. Procedure Followed

The data set was extracted in CSV format, from corgis.data. Firstly, we loaded this data into R studio and had a glimpse of the data, thereby checking for null values. The data set id well formulated for analysis and contains no null values.

Then we modified the data as per the requirement to perform analysis by changing columns names and replacing 0 with mean of column. Post the modifications, the file is saved in .CSV format.

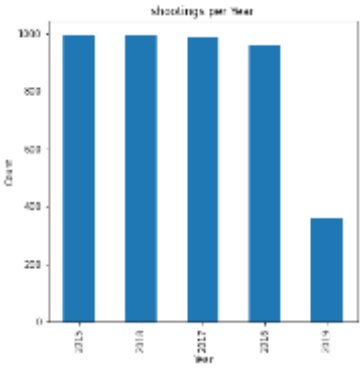
The modified file is then loaded into python for analysis, initially a generic analysis was done irrespective of years for all the contributing factors.

Then the data is made subsets based on years and in-depth analysis is done for each and every contributing factor.

The modified data is also loaded into SQL workbench, where with the help of few queries, we analysed if there are any patterns in data.

V. Out Comes

Total Number of Incidents based on Year:



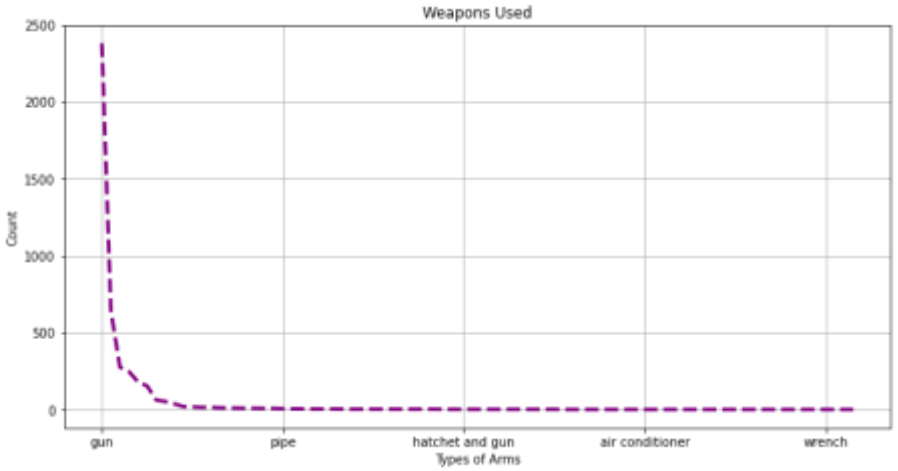
Here, we can see approximately same count of incidents took place during 2015,2016,2017 and 2018 but we can huge down fall in number of incidents that took place in 2019 throughout USA .

A) Change in attack style and weapons used based on year.

Factors such as Weapons used during attack, victims action post attack(fleeing or not), threat level of incident and the manner attack are used to depict the attacking style of a person.

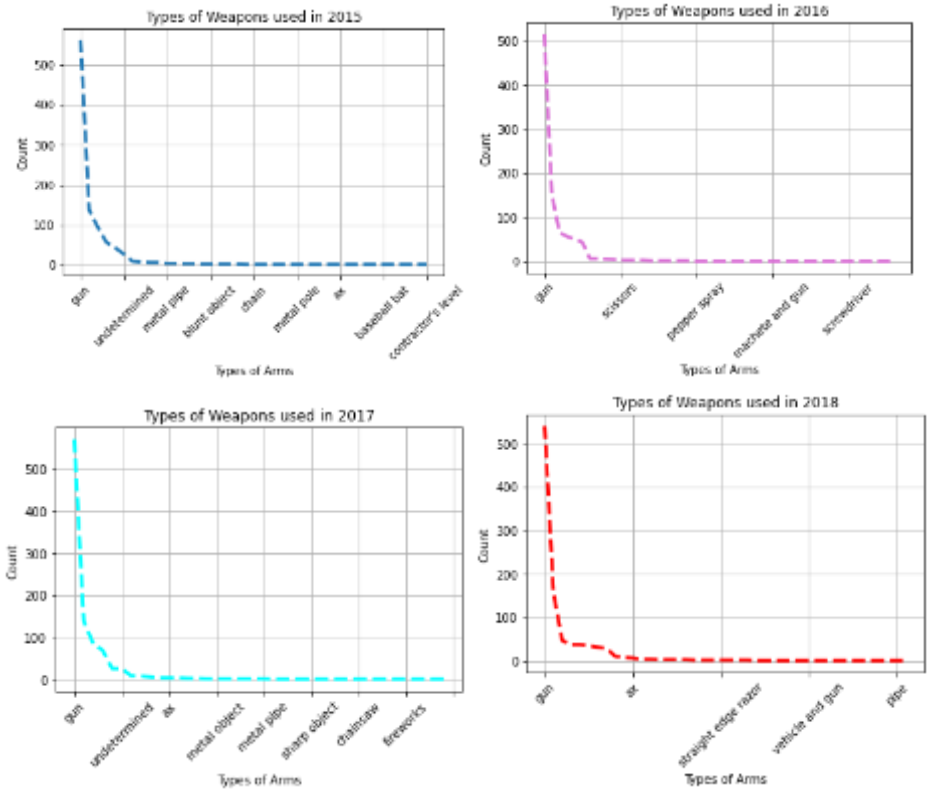
Weapons Used:

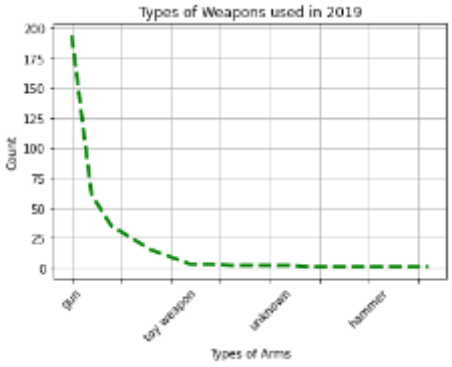
Generic Analysis:



Here, we observe guns are widely used during the attack and followed by pipe and hatchet and other tools, irrespective of time of incident and place where it took place. If We get into in-depth analysis, that trying to figure out if there is any change in usage of weapons during year.

Year- Wise Analysis:



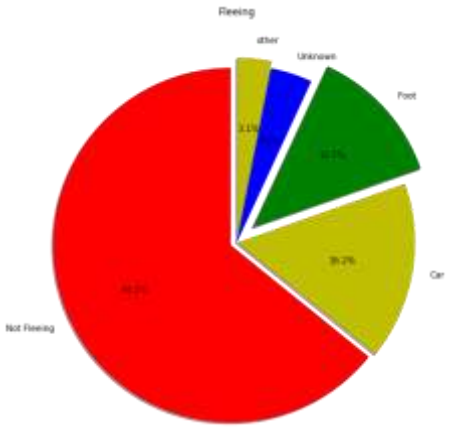


From the above analysis, we see guns are mostly used in attacks during all the years. Then comes metal pipes, scissors and ax and toy weapons, etc in different years.

We can conclude that there is pattern of usage of guns that is being followed through years.

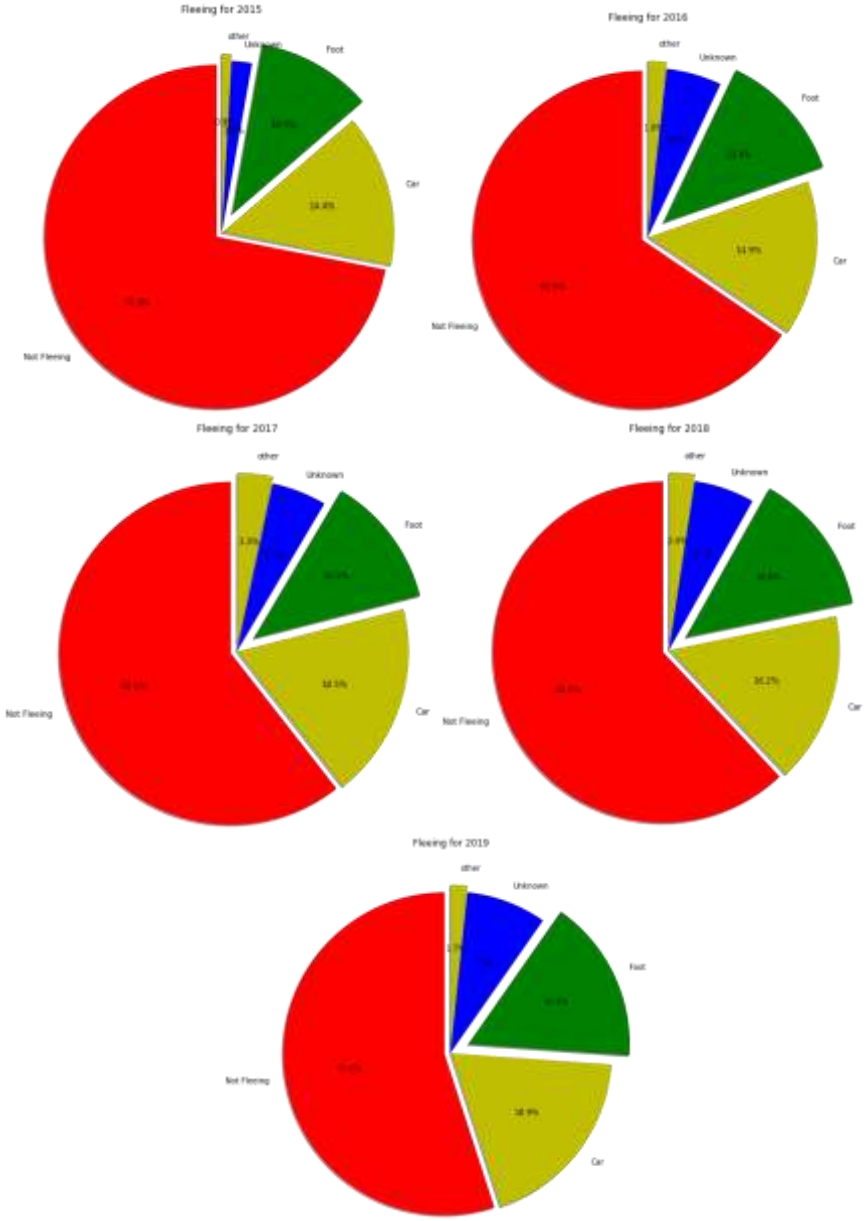
Fleeing:

Generic Analysis:



In general, when attacked by police, majority of the victims did not try to flee but few tried to escape by car or through feet irrespective of years.

Year-Wise Analysis:



Here we can see decline in the rate of non-fleeing victims from 2015 to 2018, and a sudden raise in the same category in 2019.

And also, we can see sharp raise in proportions in fleeing by car and foot.

To mitigate fleeing, the attacking manner is chosen. That is if the rate of fleeing increases, then victims are tasered after shooting.

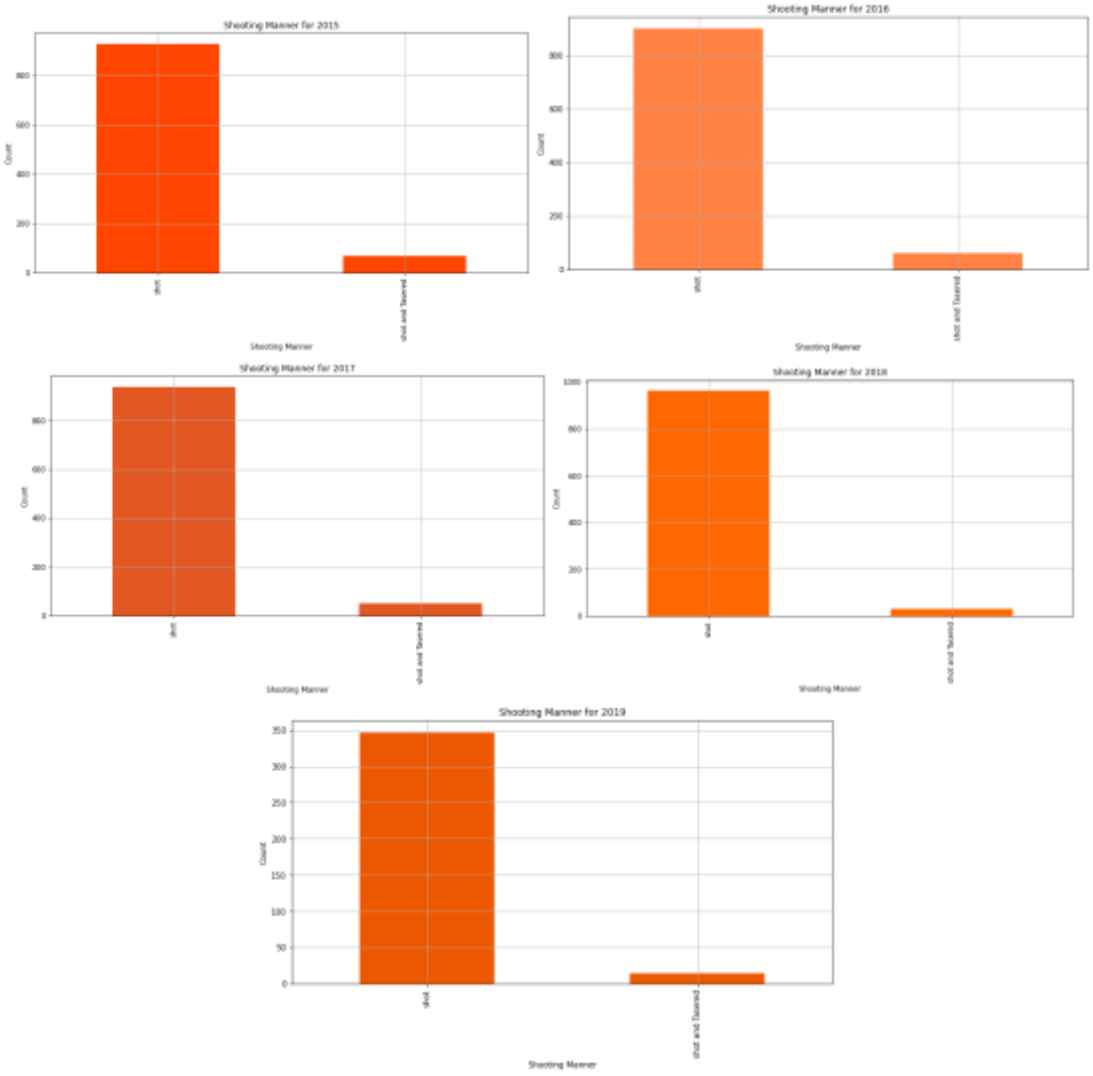
Now let’s analyse change in rate of shooting manner.

Attacking Manner:

Generic Analysis:

From the general analysis, we can determine that there is very negligible rate of shot and tasered shootings to understand this better, we perform a year wise analysis.

Year-wise Analysis:

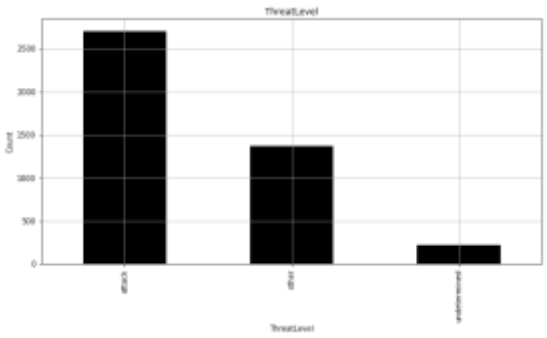


The year wise analysis also gives us the same picture, we see no raise or fall patterns in the attacking manner during the years.

To estimate the attacking patterns, we have to even check the change in levels of threat levels during the years.

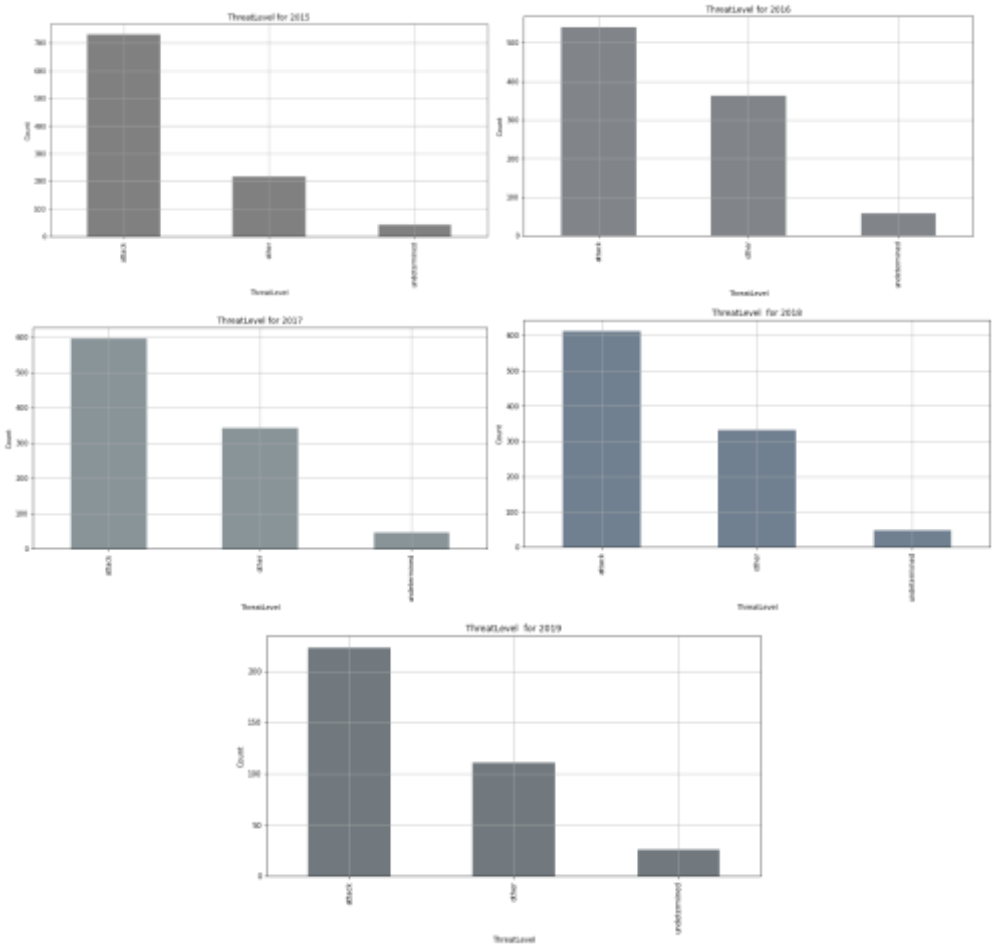
Threat Level:

Generic Analysis:



From this, we can say that most of the incidents took place as they were attacking. To understand it bit clear, let's have yearly view of data.

Year-Wise Analysis:



We can observe that most of the incidents took place throughout the years were because they tend to be in attacking in nature.

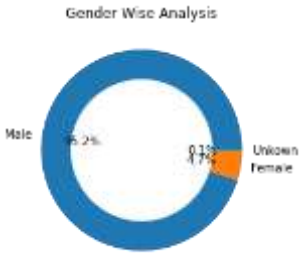
We can notice few falls and raise in undetermined and other factors for attack with disturbed proportion of attacking natured incidents.

B) Rate of change in attacks based on factors such as race and age.

When coming Victim's analysis, we mainly focus on three main factors, they are age, gender and race of victim to determine if the shootings are biased.

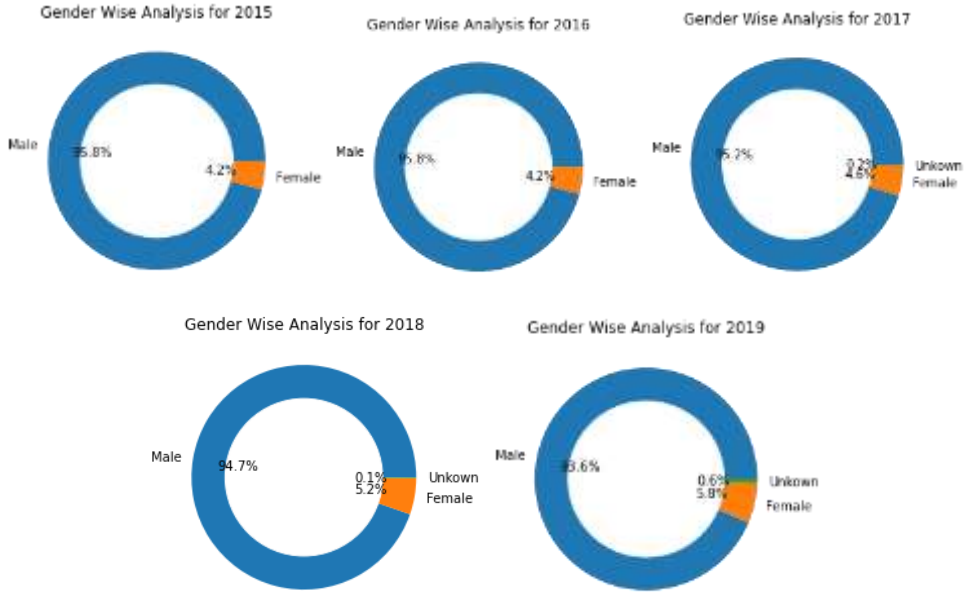
Gender-Based Analysis:

General Analysis:



We notice males are majorly involved in the incidents. To understand the change in rates, year wise analysis is done.

Year-Wise Analysis:

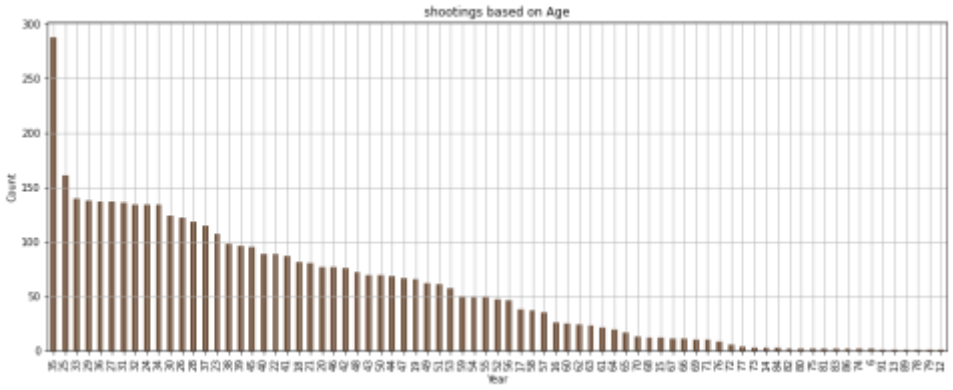


From this analysis, we notice a very slight change in proportions during years. That is we can observe increase in rate of female victims rate as years passes.

To understand what age of people are mostly involved in shootings, let's analyse the age factor.

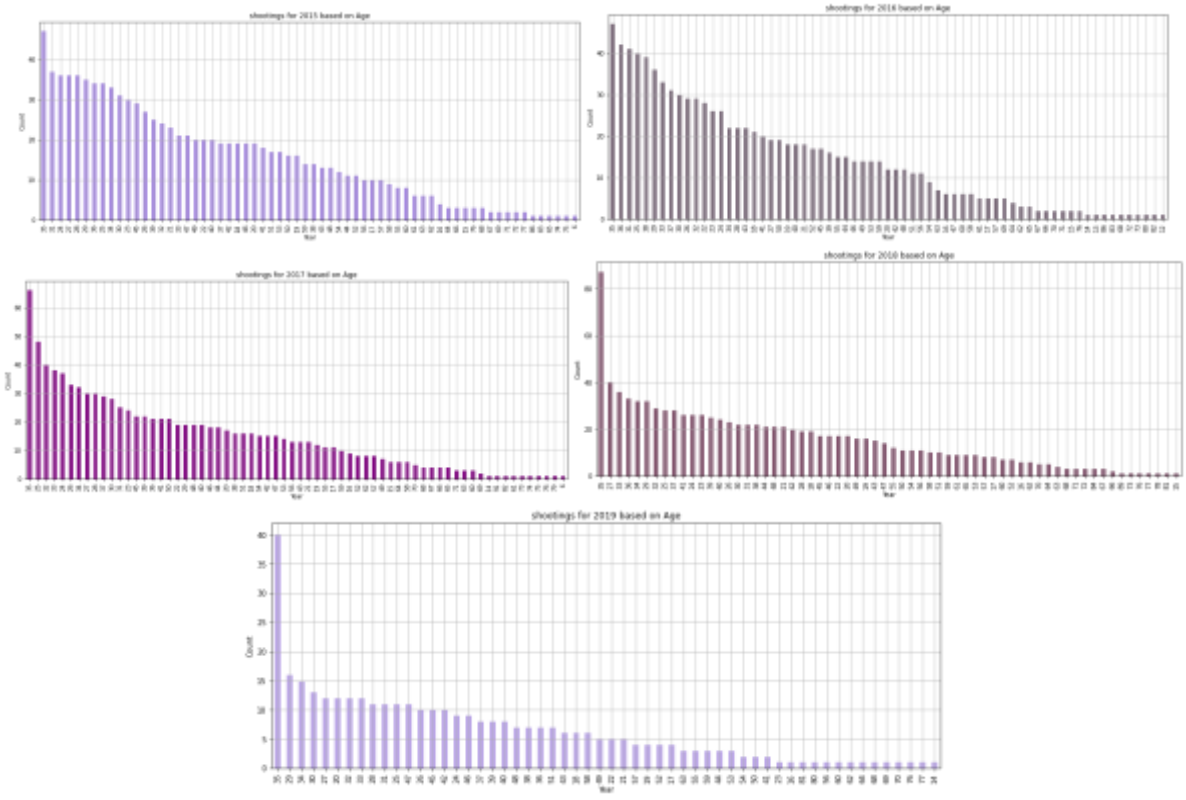
Age Analysis:

Generic- Analysis:



In generic analysis, we can notice young adults aging between 25-35 are mostly involved in shootings to understand the change in pattern, data is analysed based on year.

Year-Wise Analysis:



During 2015, we can observe victims ranging from 6 to 86 are involved in incidents with majority of young adults' involvement (25-35 age group).

While in 2016, with constant rate of involvement of 25-35 age group, we could see a narrow cut in the range of ages. That is with youngest victim being 15 years old and the oldest aging 86.

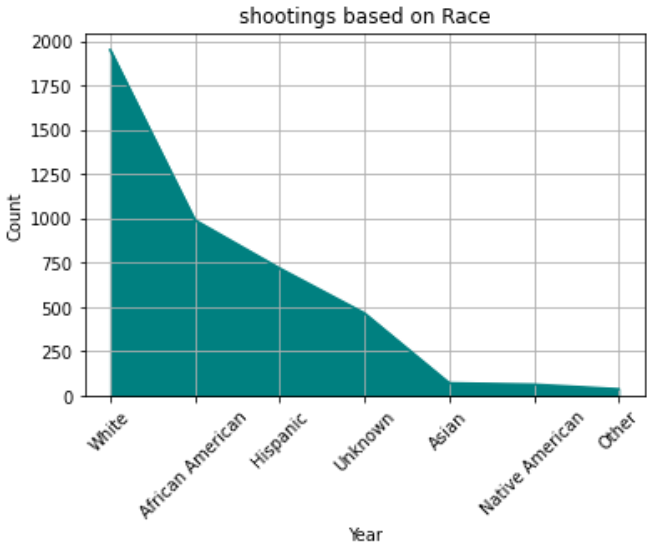
In 2017, we could notice a 91-year-old involved in the shooting activity and on the other hand a 6-year minor too with constant rate of 25-35 age group's involvement.

Where as a similar scenario of 2016 repats in 2018 with victims age ranging from 15 to 89. When coming to 2019, we see youngest being 14 and oldest being 81 involved in shooting with a same pitch of 25-35 age group.

Race-Wise Analysis:

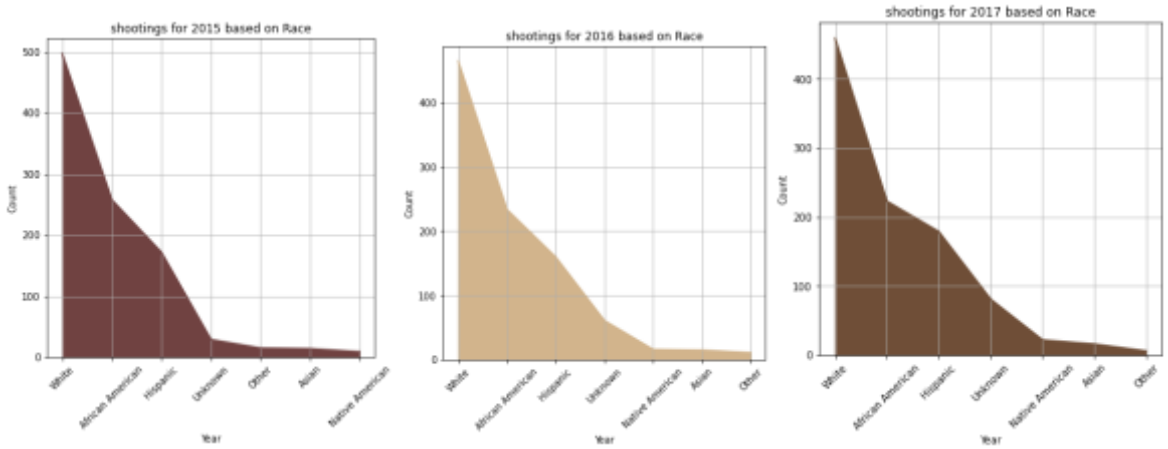
In general, police are always accused of showing racism during the shootings. In this analysis, we try to uncover if there is any such pattern involved.

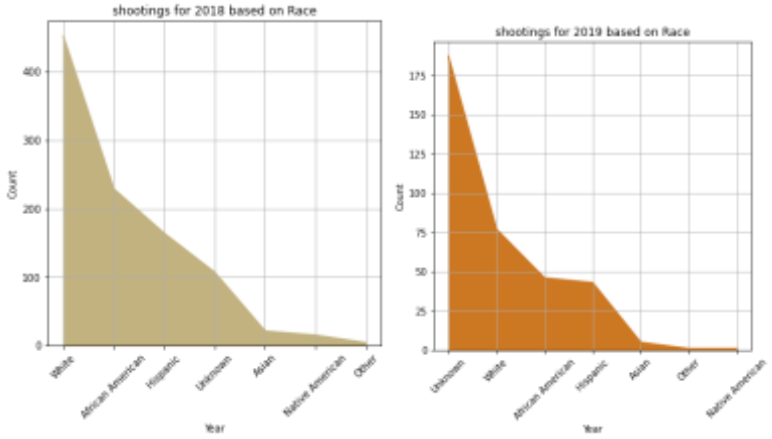
Generic Analysis:



Here, we see most of the victims are white skinned, to have a clear picture to check and analyse if there are any patterns involved, year wise analysis is performed.

Year-Wise Analysis:



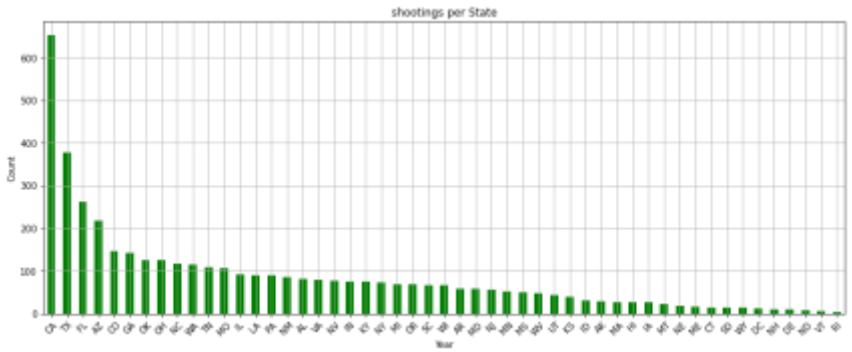


From the given data, we can notice that whites are majorly involved in shooting and least being Native American's and Asian's.

C) Alter in count of shoot outs based on place/ States per years.

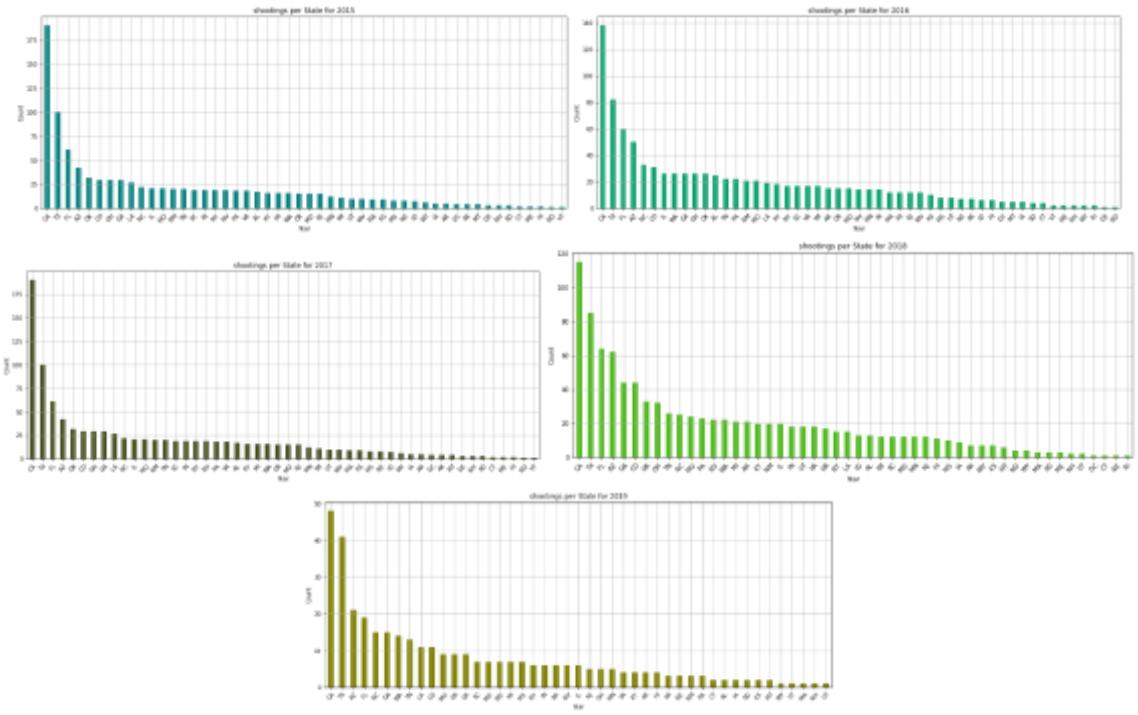
Place of where the incident took place also plays a vital role to estimate the safety- level of the state.

Generic-Analysis:



We can see that most number of incidents takes place in CA,TX and FL to see if there is any change in count or any patterns, year wise analysis is made.

Year-Wise Analysis:



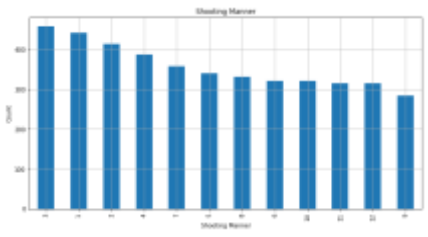
Year wise analysis also gives us the very same picture , with CA, TX and FL being the states with more shooting rate.

To understand if there are any specific months, where there is growth in crime rate analysing month wise data play a key role.

Month-Wise Analysis:

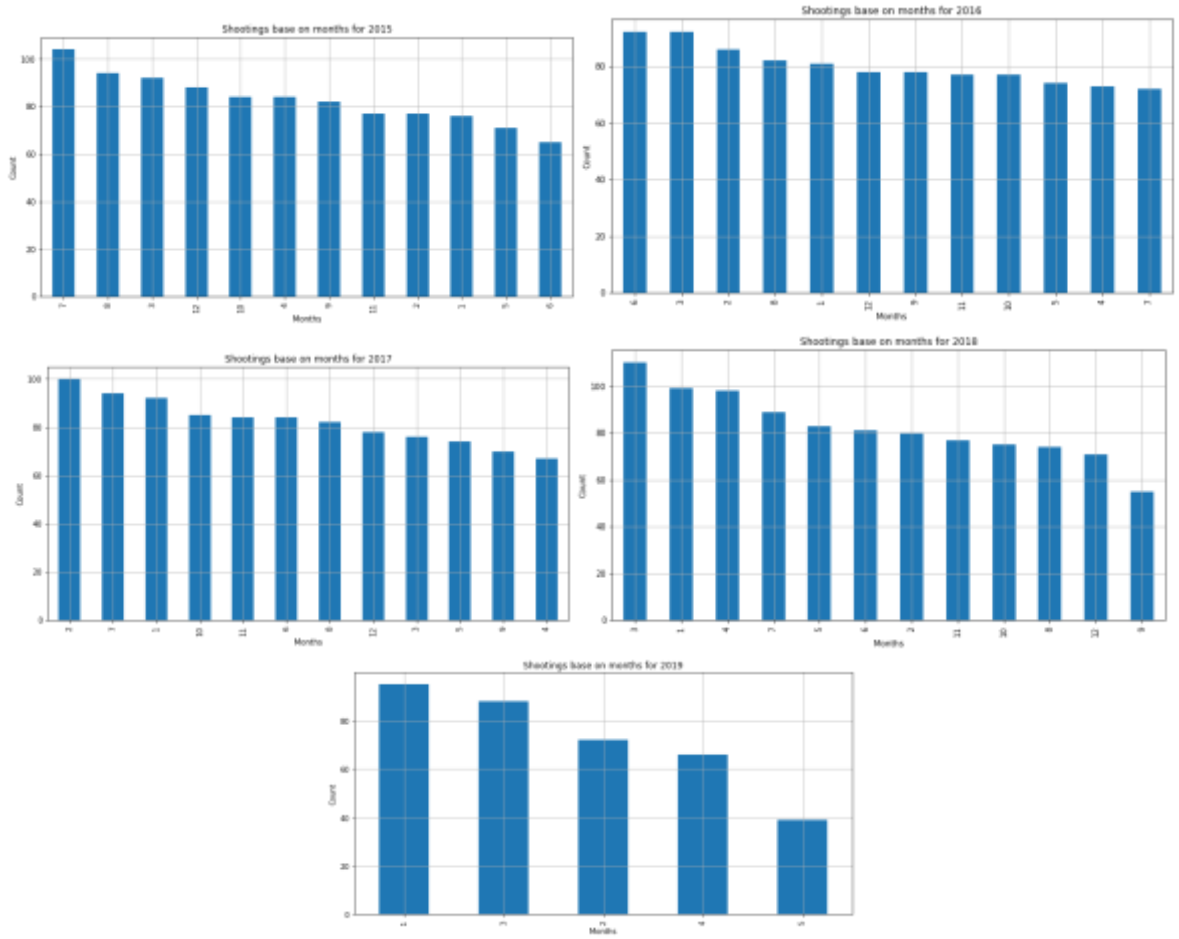
By analysis this factor, we try to uncover if there are any specific months with highest crime rate.

Generic-Analysis:



As the data is collected only till June of 2019, generic analysis does not answer our question.

Year-Wise Analysis:



During 2015 and 2016, we can observe most of the incidents took place in middle of year, and we can see a change in trend from 2017, where majority of the shootings took place in beginning of the year.

VI. Conclusion

A) Change in attack style and weapons used based on year.

From the above analysis, we can conclude that we can see a pattern of usage of guns with attacking natured victims lead to most of the shootings. One good note is that most of the victims did not try flee which lead to very few shot and tasered cases.

B) Rate of change in attacks based on factors such as race and age

We observe white, males and age group of 25-35 are mostly involved in the shooting activity.

We also notice slight raise in proportion of female involvement as years passes.

C) Alter in count of shoot outs based on place/ States per years

We notice CA, TX and FL are the states with highest shoot outs and during 2015 and 2016, we notice higher activities during mid of year but there is quite opposite scenario when it comes to 2017,2018 and 2019. We see more incidents in the beginning of the year.

VII. Limitations and Future Study

The data set contains data ranging only from 2015 to Mid 2019, which restricted us to find the patterns. And most of the data being categorical data disabled us to perform any statistical models on data.

A better ranged data set helps the analyser to perform and unfold the patterns and the relation among the data.

VIII. References

Works Cited

corgis-edu. (n.d.). *corgis-edu*. Retrieved from corgis-edu: https://corgis-edu.github.io/corgis/csv/police_shootings/

statista.com. (n.d.). *statista.com*. Retrieved from statista.com: https://www.statista.com/statistics/585152/people-shot-to-death-by-us-police-by-race/

towardsdatascience.com. (n.d.). *towardsdatascience.com*. Retrieved from towardsdatascience.com: https://towardsdatascience.com/an-examination-of-fatal-force-by-police-in-the-us-db897d97085c

washingtonpost.com. (n.d.). *washingtonpost.com*. Retrieved from washingtonpost.com: https://www.washingtonpost.com/graphics/investigations/police-shootings-database/