

# MA\_304\_COURSEWORK

AHMAD\_CHEEMA\_2111168

15/04/2022

## Introduction:

Data visualization plays an important part in Data science and big data. Human mind get things more easily when they are visualized. It makes easy to analyze and interpret data. Sometimes it plots graph by using every point of dataset like scatter plot, and sometimes uses statistical summary of the data for graph plotting like histograms. It helps to explore trends in data, to group data into clusters, for data cleaning or to evaluate output models. It makes easy to see and understands the patterns in the data in less time which is very hard to get from theoretical analysis or from raw data.

Crimes exists in every country, city, district and street of the world, but effective policing can control its extent. It is very important to control the crime rate so that people can live without fear and feel save to carry out their daily routine activities. To make policing more effective and efficient, to reduce crime rate, some strategies need to be implemented. One of the way is to analyze the history of the crime reported, the location, time, age group of subjects, weapon used etc. this will help police to deal with future cases that at what location, what day and what time of the day more security is required, which weapon are used mostly at crime scenes so that police will make strategies accordingly.

In this project, we have the dataset of crimes reported in 2016 in city of Dallas. It have information regarding the location, time, race, gender of the subject, weapons used in the crime. We will visualize these variables to see the relation between them that how they are related to each other at what time and location crimes are highest, what is the most common reason behind these crimes. This information helps to reduce the crime rate in future.

## Preliminary Analysis:

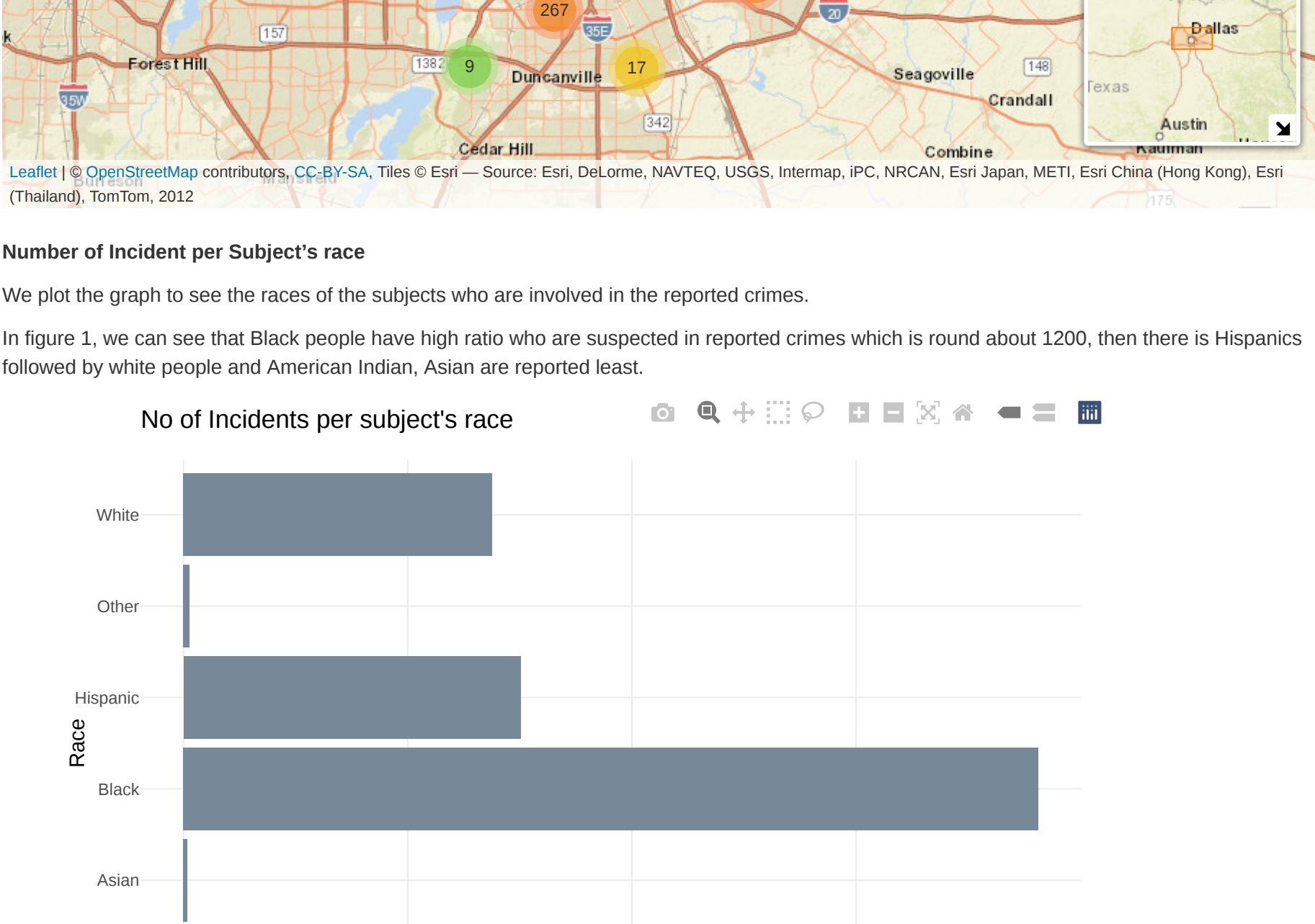
### Dataset:

The given dataset is of the crimes reported in Dallas in 2016. It have information of location which include city, area, division, district, sector, street name , street number and even complete address of the street where the incident is happened. It also mentioned the date and time of the occurrence of the incident.

At the same time it also provide information of the officers of dealt with the cases like his/her id, gender, race, hire date, number of years of his/her services. It also maintained details of the injuries if any officer is injured, like what type of injury he/she had, is he/she hospitalized. It includes the information of the subject of the crime incident, his sex, race, his description, detail of the injury if he/she had any and whether he/she is arrested or not.

### Analysis:

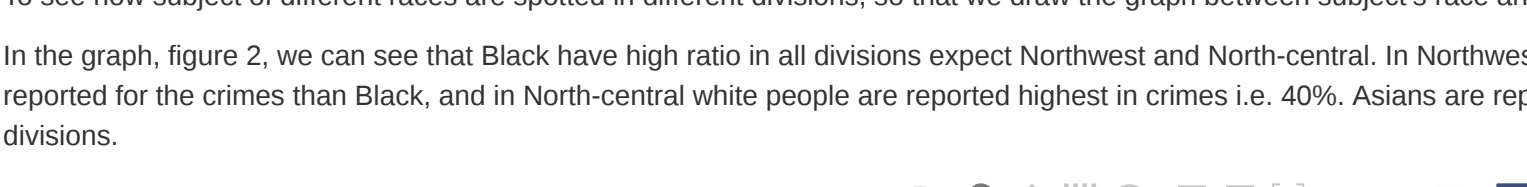
Firstly, we plot the incident on the map of Dallas city. It gives an overview that how incidents are divided into division, district and street of Dallas. More we zoom more we get the detailed overview from division to the street of the crime incident. It shows that how different regions of the city are affected by the crimes which region is highly affected which one is least. By visualizing the data on the crimes at different location, at different time helps to analyze the pattern in crime incidents



#### Number of Incident per Subject's race

We plot the graph to see the races of the subjects who are involved in the reported crimes.

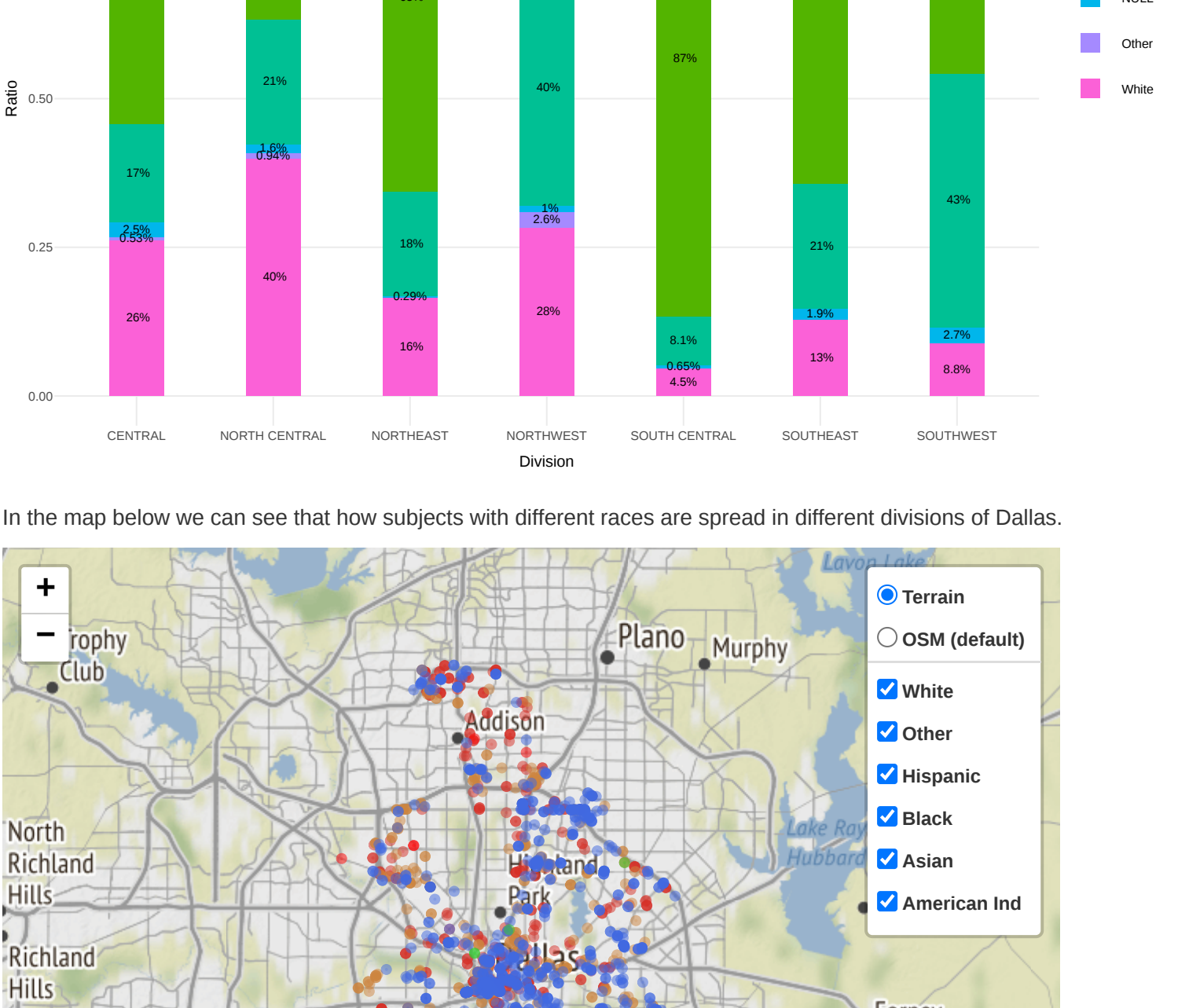
In figure 1, we can see that Black people have high ratio who are suspected in reported crimes which is round about 1200, then there is Hispanics followed by white people and American Indian, Asian are reported least.



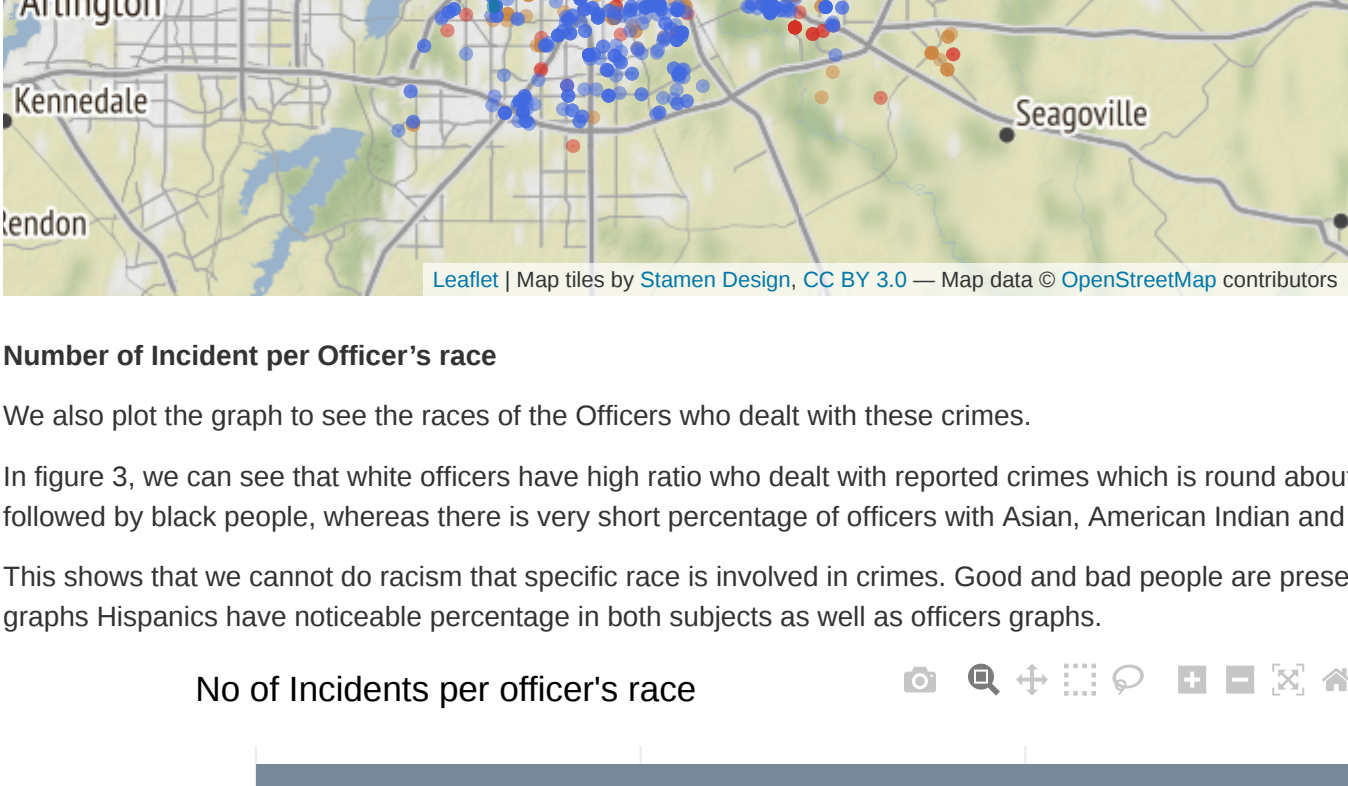
#### Subject Races' per division

To see how subject of different races are spotted in different divisions, so that we draw the graph between subject's race and divisions.

In the graph, figure 2, we can see that Black have high ratio in all divisions except Northwest and North-central. In Northwest Hispanics are more reported for the crimes than Black, and in North-central white people are reported highest in crimes i.e. 40%. Asians are reported lowest in almost divisions.



In the map below we can see that how subjects with different races are spread in different divisions of Dallas.

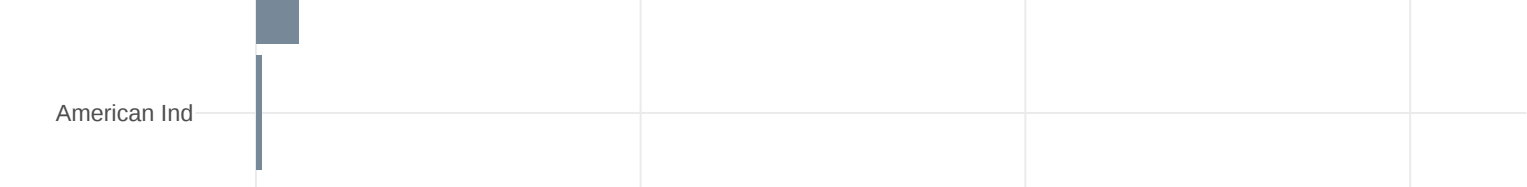


#### Number of Incident per Officer's race

We also plot the graph to see the races of the Officers who dealt with the crimes.

In figure 3, we can see that white officers have high ratio who deal with reported crimes which is round about 1200, then there is Hispanics followed by black people, whereas there is very short percentage of officers with Asian, American Indian and other races.

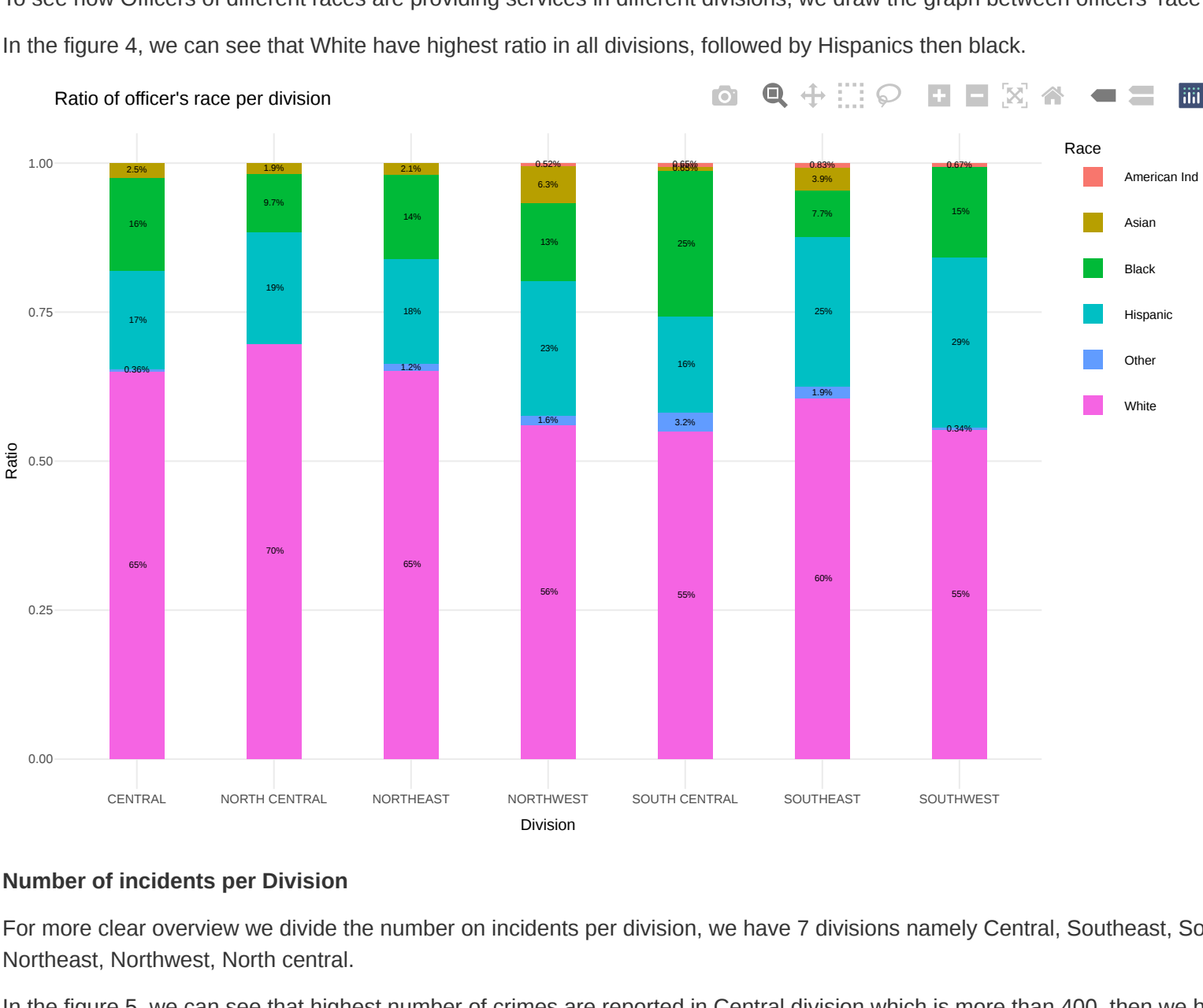
This shows that we cannot do racism that specific race is involved in crimes. Good and bad people are present everywhere, like as shown in the graphs Hispanics have noticeable percentage in both subjects as well as officers graphs.



#### Officer's Races per division

To see how Officers of different races are providing services in different divisions, we draw the graph between officers' race and divisions.

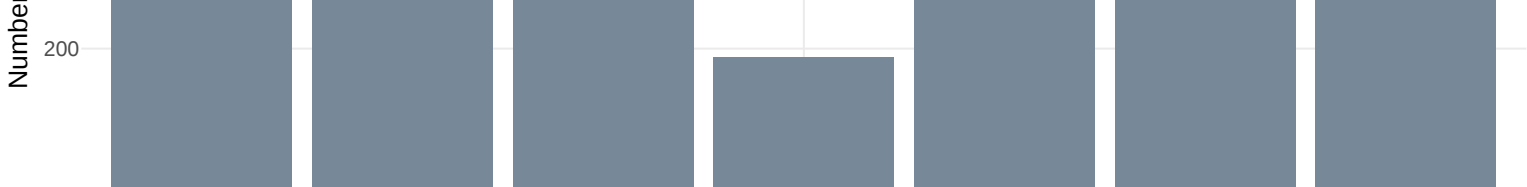
In the figure 4, we can see that White have highest ratio in all divisions, followed by Hispanics then black.



#### Number of crimes per Division

For more clear overview we divide the number on incidents per division, we have 7 divisions namely Central, Southeast, Southwest, South central, Northeast, Northwest, North-central.

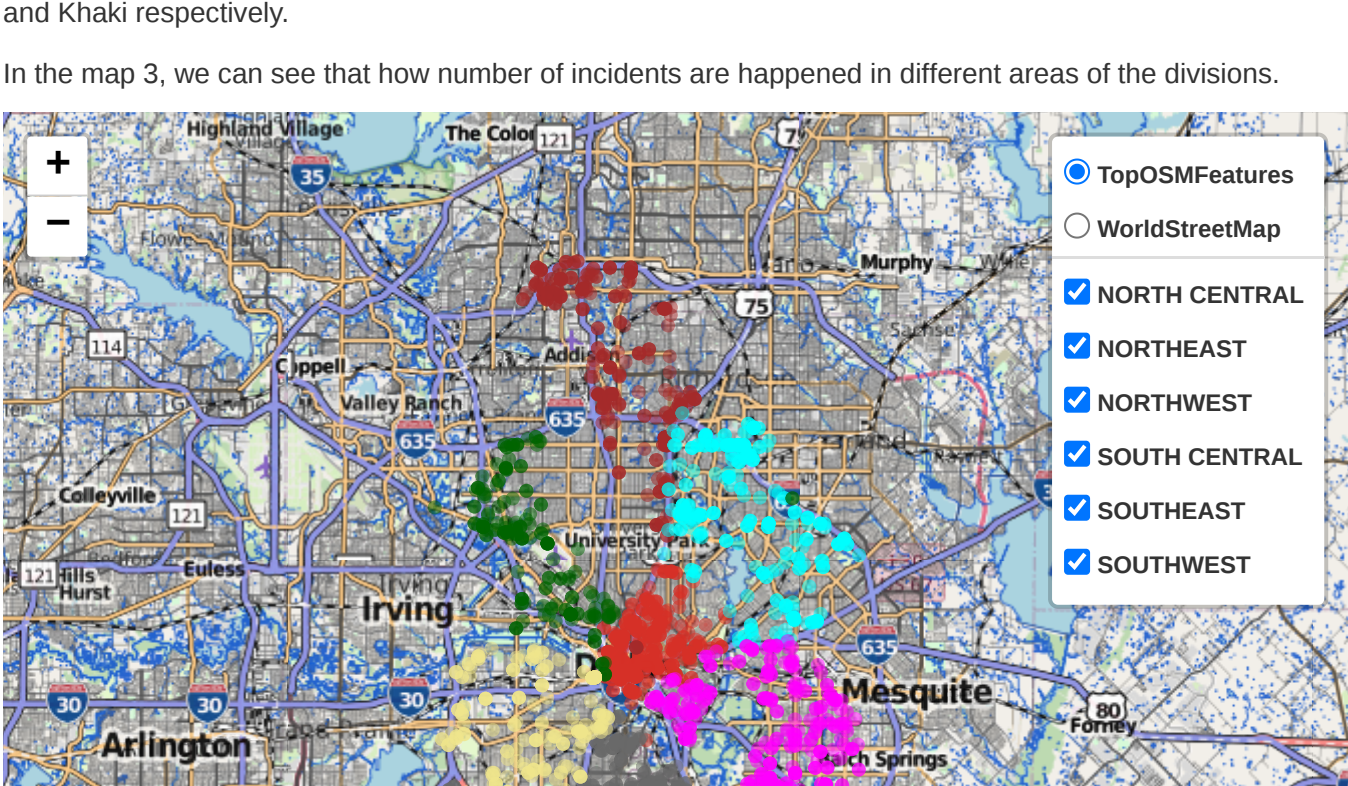
In the figure 5, we can see that highest number of crimes are reported in Central division which is more than 400, then we have southeast followed by Northeast, Southwest, South-central, North-Central almost have same number of crimes. Lowest number of crimes are reported in Northwest.



We also show number of incidents in different division on the map of Dallas city in the graph below.

Here Brown color represents Central division, North central is represented by Maroon color, Cyan color is used for Northwest, and Dark-green shows incidents in Northwest. South-Central is represented by Grey color, whereas Southeast and Southwest are represented by colors Fuchsia and Khaki respectively.

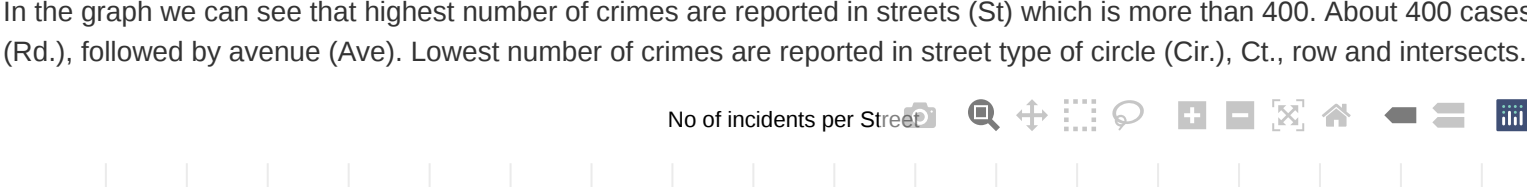
In the map 3, we can see that how number of incidents are happened in different areas of the divisions.



#### Number of incidents per Street type

For more clear and detailed overview we divide the number on incidents per street type to see the crime rate at smaller level, we have a different type of streets like avenue (Ave), boulevard (Blvd), street (St), Parkway (pkwy), freeway (Fwy), highway (Hwy), Expressway (Expy) etc.

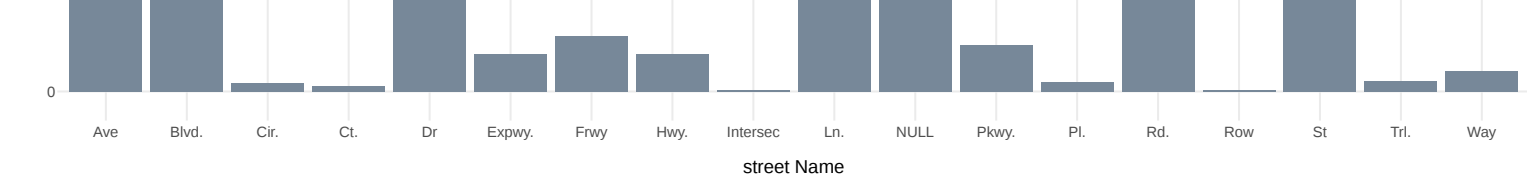
In the graph we can see that highest number of crimes are reported in streets (St) which is more than 400. About 400 cases are reported on roads (Rd.), followed by avenue (Ave). Lowest number of crimes are reported in street type of circle (Cir.), Ct., row and intersects.



#### Number of Incidents per Day

After visualizing and analyzing the number of incidents in different divisions, we also analyze that how the ratio of the number of incidents varies in different days of the week.

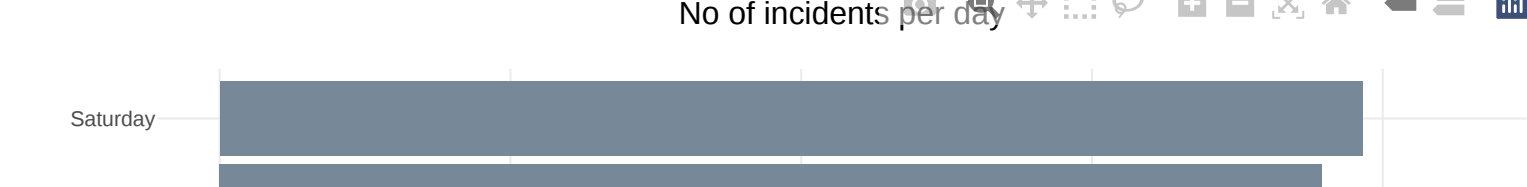
The graph below shows that Sunday have the highest rate of crime while Monday have the least. Tuesday and Thursday have almost same rate. Whereas crime rate gradually starts increasing from Wednesday, mid of the week, followed by Thursday, Friday, Saturday and have highest rate on Sunday. It shows that crimes incidents are usually high during weekends as compare to week days.



#### Incidents on different days per Division

We plot the graph to show the proportion of incidents on different days of the week.

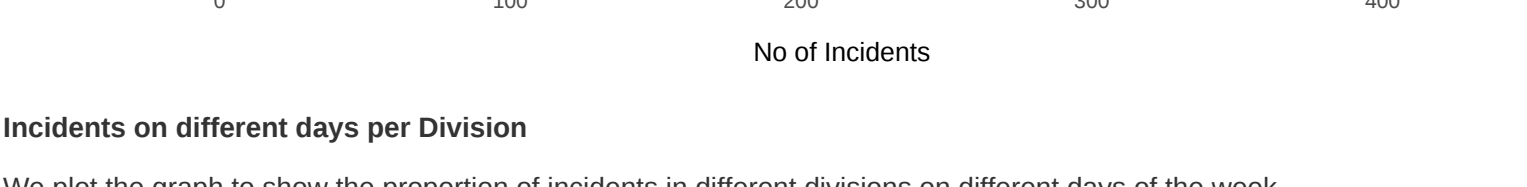
As shown in the graph below, figure 8, Southeast and Northwest have lowest crime rate on Wednesday and highest on Sunday. South-central highest crime rate is recorded 17% on Thursday and lowest 12% that is on Monday. In Northeast highest crime rates are recorded on Friday (24%) and lowest on Tuesday (9.1%). North-central have lowest and highest rate on Saturday and Sunday respectively. Whereas Central division have least crime rate on Thursday i.e. 11% and highest on Saturday i.e. 20%.



#### Number of Incident at different time of the day

Incident happens on different time of the day. So we divide the 24 hours of the day into groups namely morning, evening, afternoon, night, where 5am to 12pm is grouped as morning, 12pm to 5pm is denoted as afternoon, 5pm to 9pm is in evening group and 9pm to 4am represent night.

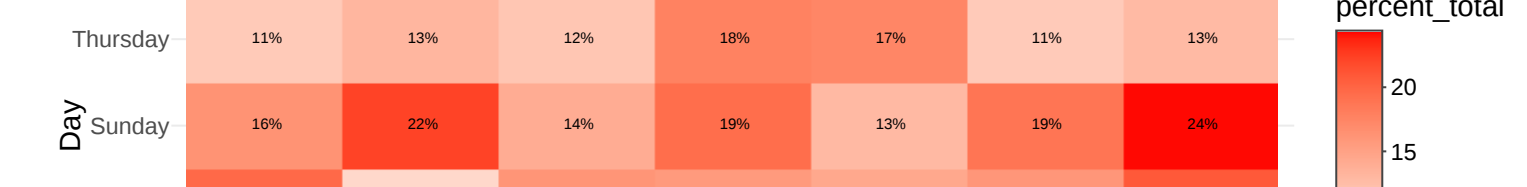
In the figure 9, we can see that most of the crimes happened at night time and lowest at morning. As the time passes, it is noticed that crime rate increases gradually from morning to night where crime rate is highest.



#### Incident At different time of the day per division

To analyze the pattern how crime rate increase or decrease during different time of the day in different division we plot the graph between the different time of the day and divisions.

It is shown in graph, figure 10, that in all divisions except South-central crime rate is highest at night, in South-central more crimes are reported in the evening than night. While Southwest, South-central, Northeast and Central have lowest crime rate in the morning, whereas Southeast, Northwest and North-central have least crime rate in the afternoon.

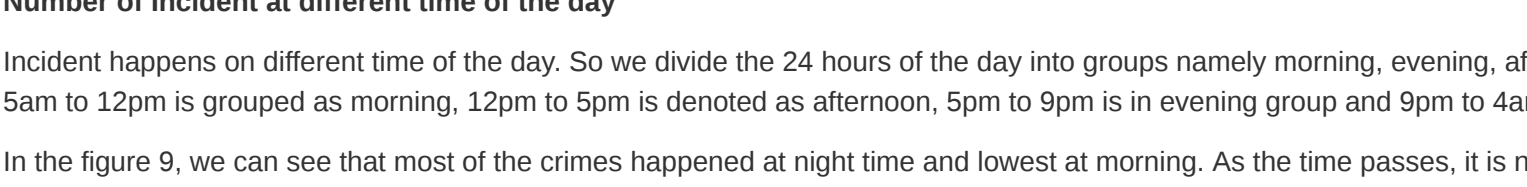


#### Number of Incident per Subject Description

Subject description can also play an important part to analyze the trends in crime incidents. It helps to reach at the root cause of the crime incident, it also give information about the reason to charge the subject, in future it helps the police and government to properly deal with the cases by taking necessary actions to avoid these type of incidents in future.

There are number of description that are recorded in the incidents like different drugs, mentally unstable with any kind of weapons.

In figure 11, we can see that huge part of the subjects who are charged are mentally unstable. After mentally unstable most subjects are charged for alcohol or other type of drugs. Some of the subjects are spotted with gun or other kind of weapons. It is also noticed that about 300 subjects are not detected with any of these description.



#### Incident with different Subject description per division

We plot the graph to see how the subjects are charged with different subject description in different division.

In the figure 11, we can see that subject with subject description mentally unstable, suspect with alcohol and other unknown drugs are highly reported in Central division. Marijuana is reported highest incidentally in Northeast. Subject with subject description suspected with gun are mostly belong to Southeast, and suspected with other weapons than gun have high ratio in North-central.



## Results and Discussion

On the provided dataset we do some data analysis to get some specific patterns in the crime incidents. From the analysis we get that mostly subjects who are charged for the crime incident belongs to black race in all divisions except Northwest and North-central where subjects belong to Hispanics and white race respectively. We also do analysis of race by using officers' race data, which shows that mostly officers belong to white race in all divisions.

By analyzing and visualizing we get very interesting patterns in the location and time of the crime incidents. Central division have the highest crime rate whereas lowest crime rate is recorded in northwest. Streets (St.), roads (Rd.), avenues (Ave.) and drive (Dr.) street types have noticeable crime rate. Streets have the highest crime rate among all of them. We also see crime rate in high on weekends as compare to week days, highest crime rate is recorded on Sunday while lowest on Monday. In almost all divisions crime rate is highest on weekends (Saturday and Sunday) or near to weekend (Friday) except South-central which have highest rate on Thursday. Crime rate varies during different times of the day, highest number of incidents are recorded in night and lowest in morning. All divisions have high crime rate during night time expect South-central, it has highest rate during evening.

From the analysis we get results that mostly subjects are mentally unstable and have alcohol or other type of drugs and they mostly belong to Central, while marijuana rate is high in Northwest and subjects with guns and other type of weapons are mostly spotted in southeast and North-central.

## Conclusion

In world crime is a very big issue, and it is very important to deal with it to keep the peace of the world. To deal with crime, police needs to be more efficient. For efficient and effective policing there are certain necessary steps and actions are needed. By getting the trends and patterns from the past incidents, police will understand what necessary and important steps should take to avoid or at least overcome these type of incidents in future. Like Central division and streets have highest number of crime incidents, and on weekends crime rate is high as compare to week days and crime rate gradually increase from morning to night so police should increase the security in these circumstances or stay more alert during such situations. From the analysis we get that most subjects who are charged mostly black and mostly officers in all divisions are belong to white race. In Central mostly subjects are mentally unstable or have drugs, whereas in Southeast and North-central mostly subjects are suspected with gun and other weapons so police should use supplies and do arrangements accordingly.