**Final Project Report**

***BOSTON CRIME DATA***

**Dataset Link:** [**https://www.kaggle.com/ankkur13/boston-crime-data**](https://www.kaggle.com/ankkur13/boston-crime-data)

**Group Members: Riddhi Kadam, Ashvit Shetty**

Our aim for the project is to prepare a data visualization representing the ***Crime Rates in Boston***.

This dataset contains data from year 2015 to 2018. With this project we aim at assisting the local police departments who are often called upon to prevent crime and keep order. It is impossible to be present everywhere at the same time. Hence this visualization which aims to represent information about where the crimes take place, what are the major types of crime that take place in Boston and much more would ease their job. Thus, providing the local authorities with more guidance on where and how they need to deploy their forces to reduce the crimes.

**Dataset and Software:**

The dataset that we are using for our project is *Crimes in Boston* containing records from the new crime

incident report system, which includes a reduced set of fields focused on capturing the type of incident

as well as when and where it occurred. The given dataset is generated and provided by the Boston Police

Department.

We have downloaded this dataset which is a csv file from Kaggle datasets.

The dataset from Kaggle is not clean and contains many null or blank values.

Our first step towards the project is to clean and prepare the dataset in order to achieve valid numbers

and visualizations. Currently we have started with the process of data cleaning.

The dataset fields provide information about the incident number, what kind of offense it was, what offense group it falls under, description of the offense, where did the offense take place (includes granular location details), when and at what time the offense took place.

The granularity of the data will help us to visualize maximum crime scenarios and help in providing a very

robust visualization.

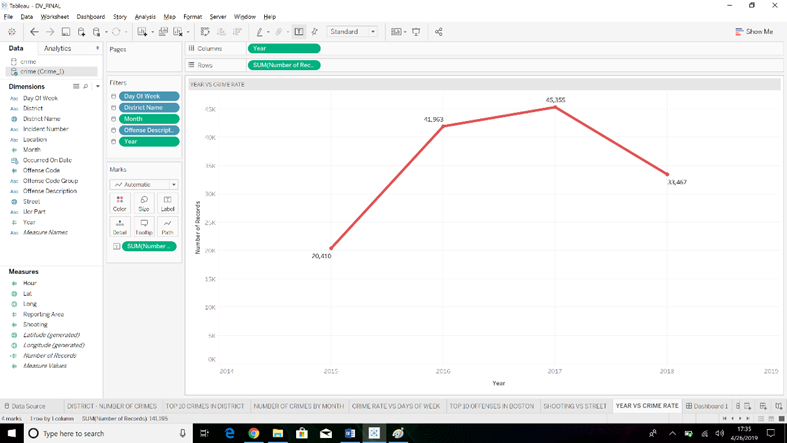
This dataset has 2,60,760 rows and 17 columns. It includes the following attributes:

* INCIDENT\_NUMBER
* OFFENSE\_CODE
* OFFENSE\_CODE\_GROUP
* OFFENSE\_DESCRIPTION
* DISTRICT
* REPORTING\_AREA
* SHOOTING
* OCCURRED\_ON\_DATE
* YEAR
* MONTH
* DAY\_OF\_WEEK
* HOUR
* UCR\_PART
* STREET
* LATITUDE
* LONGITUDE
* LOCATION

For development of this project we will be using the *Tableau Data Visualization Tool*. Our dataset is a *csv file* which works effectively with Tableau as well. The main reasons for using Tableau are its remarkable visualization capabilities, its ability to handle large amounts of data, easy to use on both developer and consumer end, high performance and impressive integration with other platforms.

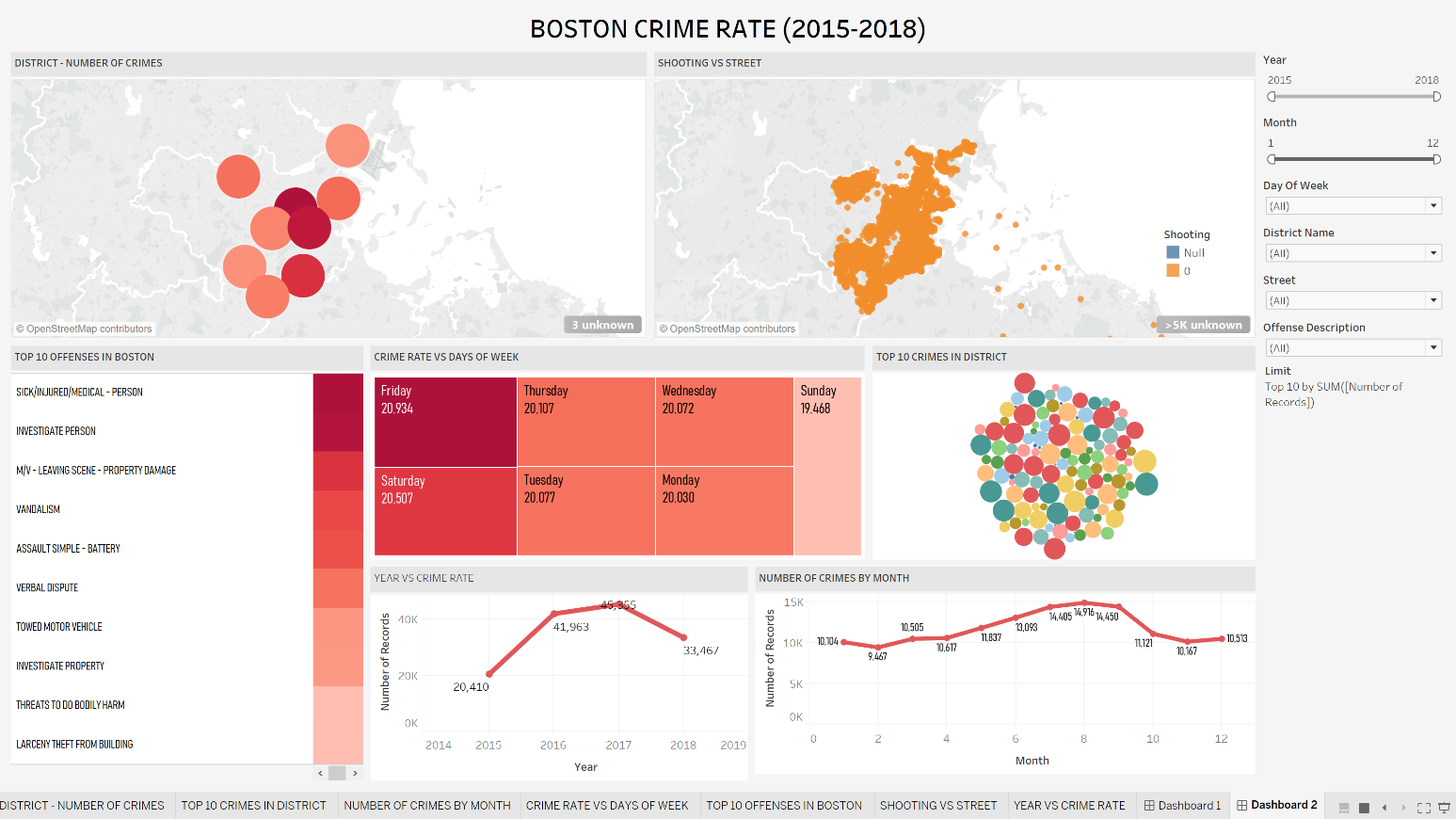
*Based on the visualizations generated, following are some questions that we tried answer.*

1. **How has crime changed over the years?**



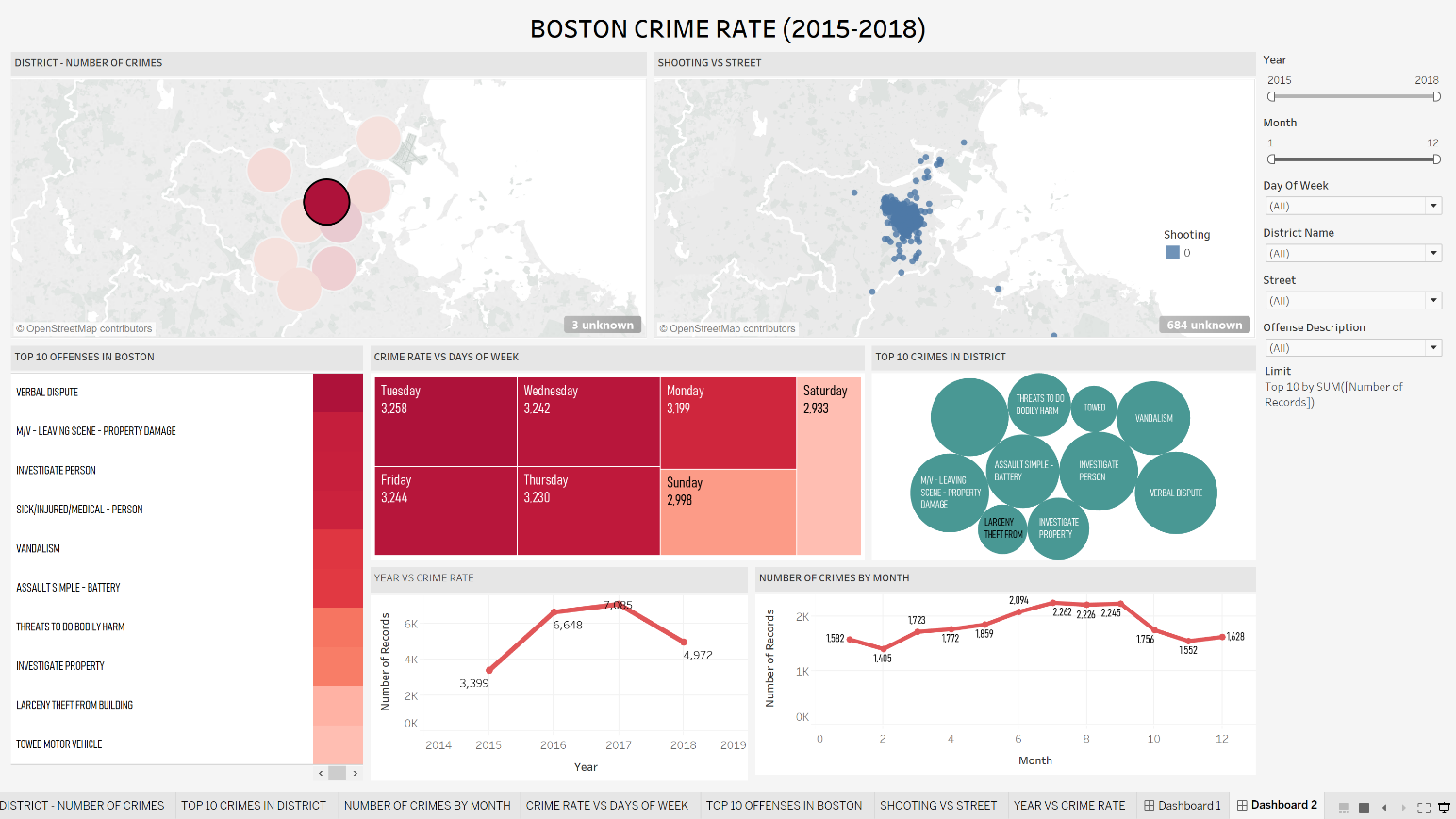
**Answer:** The given visualization provides a graphical view of the crime rate that are committed in the past couple of years. This helps us answer the above question by showing us that 2017 was the year with highest number of crimes recorded and 2015 was the year with lowest number of crimes recorded.

1. **Is it possible to predict where or when a crime will be committed?**



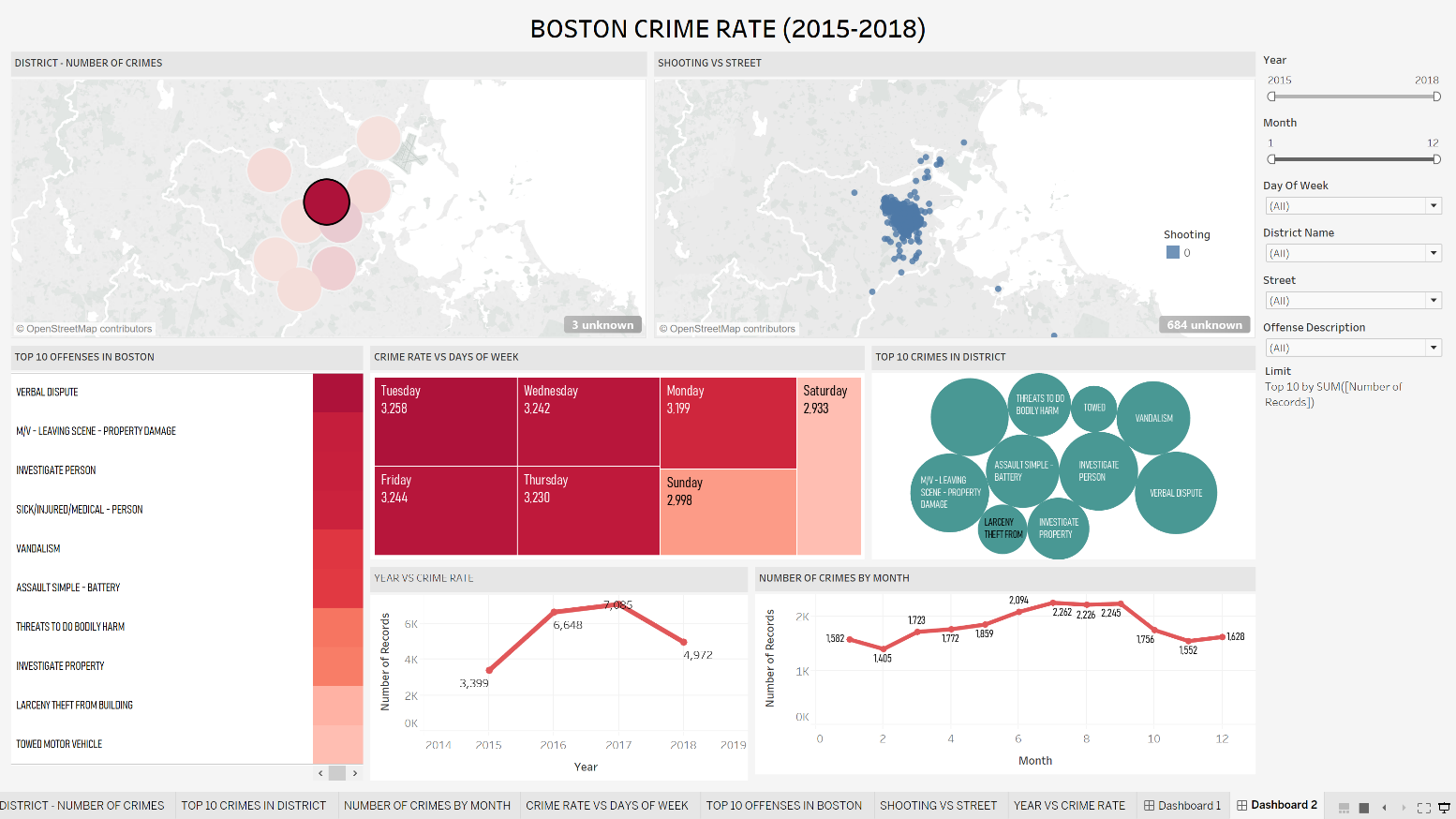
**Answer:** The answer to the above question depends on a lot of factors. It depends on the location as every location has a different crime rate and secondly, we would have to look into the type of offence and which offence is more recorded in a given location. We can make use of the above given dashboard to answer the question considering the above factors and also considering the months when the crime recorded are the highest and the day of the week when the crimes are committed the most. These factors will help any user predict when and where a crime will be committed in the future.

1. **Which areas of the city have evolved over this time span?**



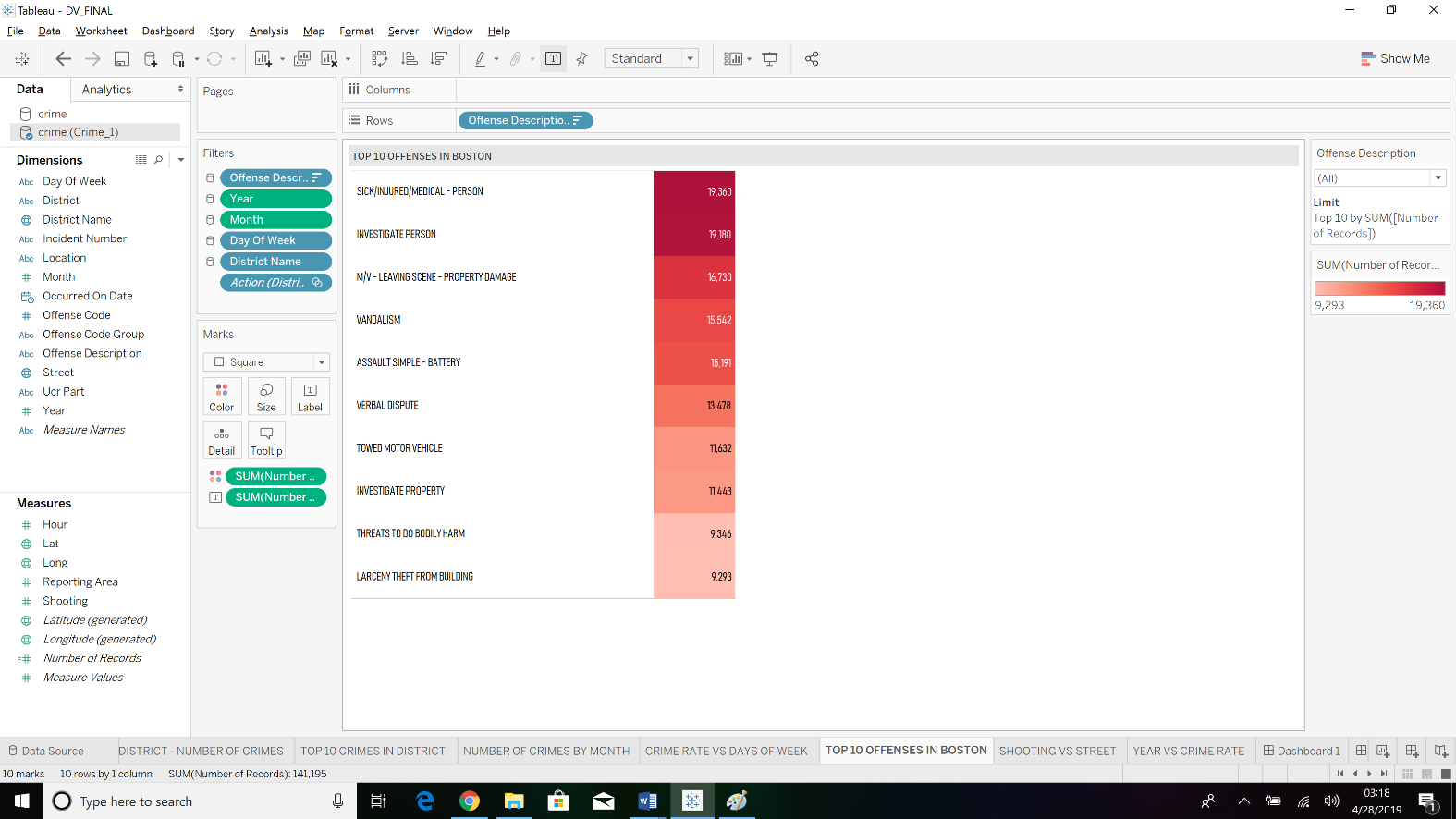
**Answer:** In terms of crime rate the district of Roxbury has evolved to a greater extent with a reduction of crime records from the year 2017 to 2018. The total number of records recorded were 22104 with number of crimes committed in 2017 were 7085 which reduced to 4972 in 2018.

1. **In which area most crimes are committed?**



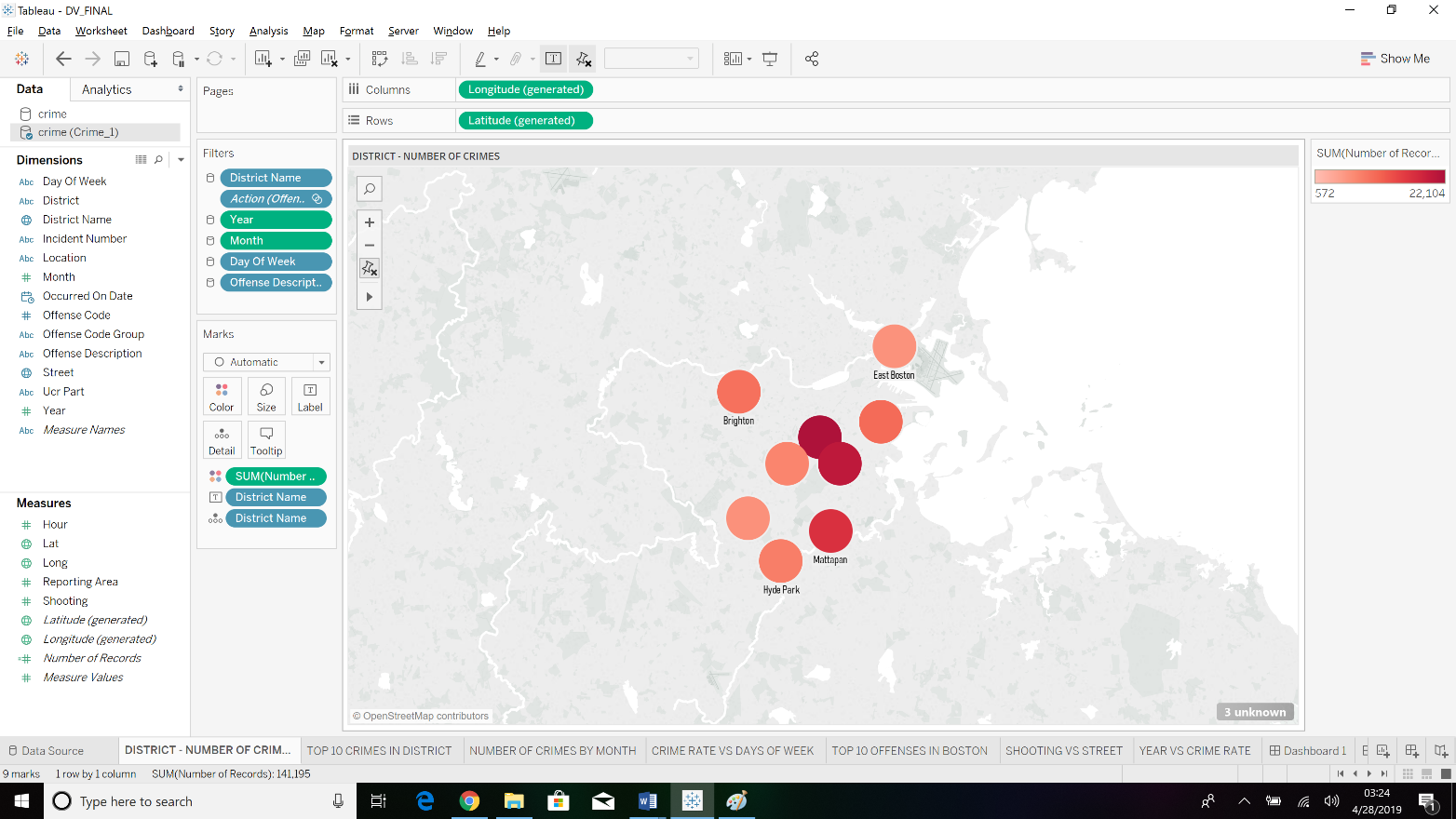
**Answer:** The district of Roxbury has the highest number of crime records recorded from the year 2015 to 2018. It has a total of 22104 crimes recorded. Major crimes are considered to be committed on Tuesday’s. And the above visualization provides with the information related to the offences committed and the trend over months.

1. **What the top crimes/offense that take place in Boston?**



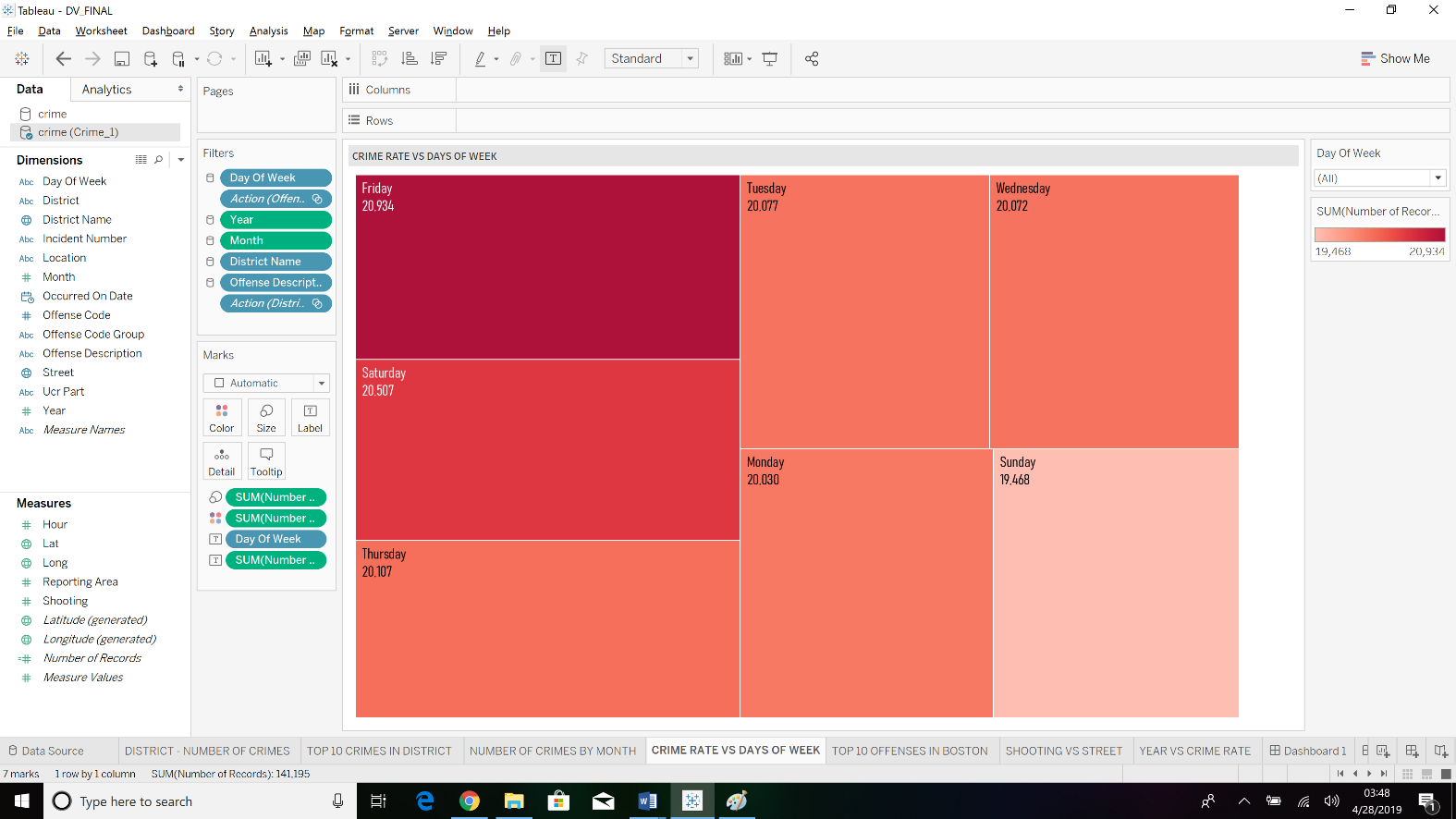
**Answer:** The above visualization provides the user information about the top 10 crimes that are committed in the city of Boston. Sick/Injured/Medical person is seen on the top of the list with a total crime record count of 19360 followed by Investigate person with crime record count of 19180 and so on.

1. **Which are the districts that are highly prone to crime?**



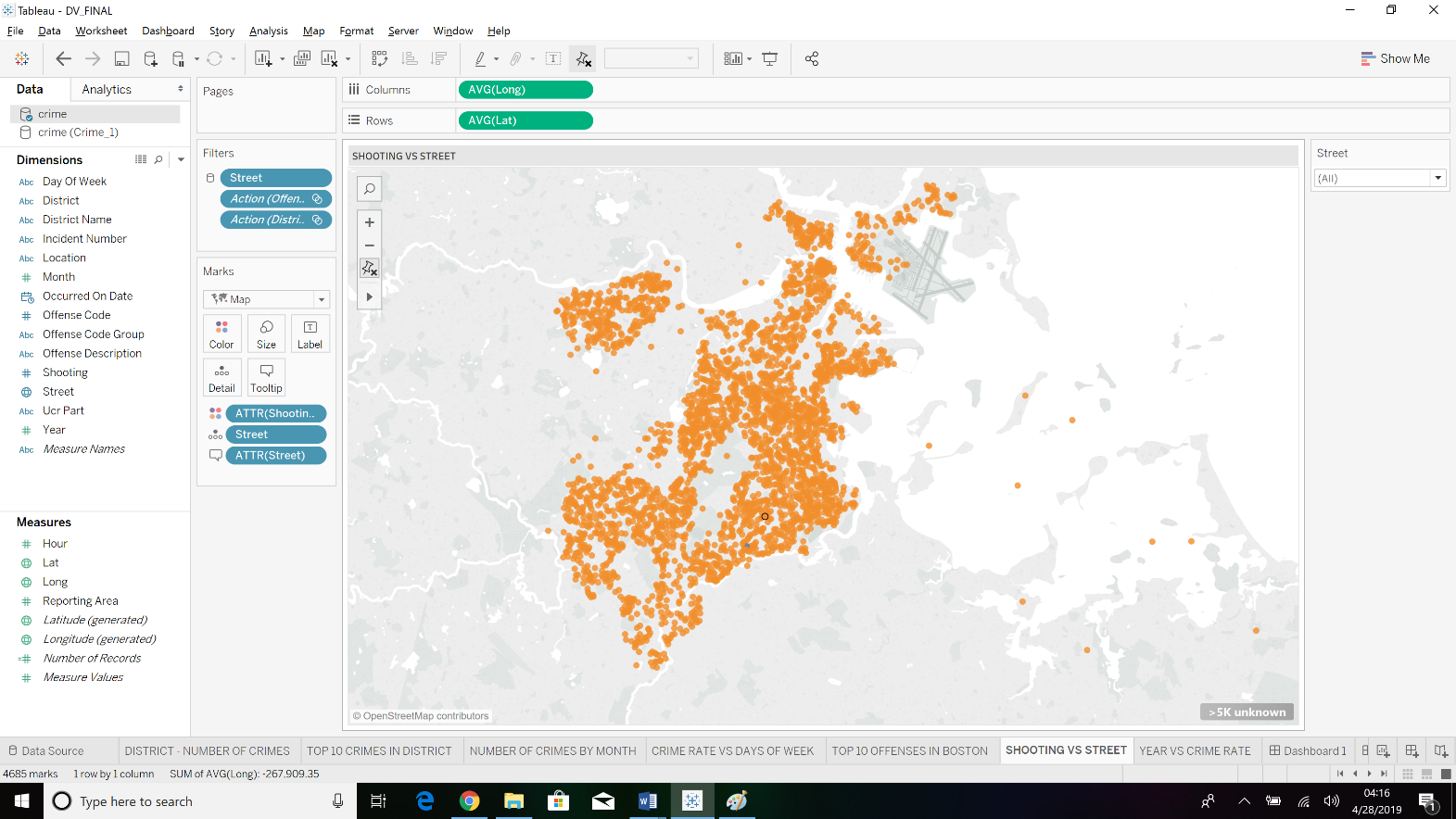
**Answer:** The given visualization answers the following question giving the different districts and the number or crime record count. The visualization uses color scheme to indicate the count of records. The darker the shade the more number of records registered in that district. The highest number of records recorded are in the district of Roxbury that is 22104 and the lowest number of records recorded are in the district of East Boston that is 5972.

1. **Can we predict on which days crimes have taken place the most?**



**Answer:** The given visualization helps us to answer our above question. Based on districts or for the entire city of Boston we can check based on previous records as to which days the crime has been committed the most. Based on our given dataset the crime has been committed the highest on Friday and thus this could help the users predict the future crimes as well.

1. **How intense is the situation of open-shooting/gunfire in Boston?**



**Answer:** The given visualization maps the shooting to their appropriate streets. This helps the users to understand and predicts which are the major streets around the city of Boston that has reported shooting/gunfire. This could help the officials to provide better safety.

These were few of the questions that we focused our visualization on. We can further get the detailed visualization by using the different interactive filters (we can filter either based on year or month. We could also filter based on the streets or districts or even the offence description). This could help the officials to filter based on their requirements for a particular day or district or year.

**Conclusion:** Based on the visualization we have generated this could help the police officials to help predict the future crimes as well as help them to deploy their officials in the respective locations that have high crime rate. This could indeed help them reduce the overall crime rate and also work efficiently.