

# FIT5147

Data exploration  
and visualisation

## Crime data analysis in London

Data visualisation  
project

# Table of Contents

---

Introduction ..... 4

Design..... 5

Implementation ..... 6

User guide ..... 8

Conclusion..... 9

References ..... 10

Appendix..... 11

## List of figures

---

Figure Number	Figure Name	Page Number
Figure 1:	Screenshot of interactive map and bar chart	6
Figure 2:	Screenshot of prediction tab	7
Figure 3:	Screenshot showing working of interactive map tab	8
Figure 4:	Screenshot showing working of prediction tab	8
Figure 5:	Sheet 1	11
Figure 6:	Sheet 2	12
Figure 7:	Sheet 3	13
Figure 8:	Sheet 4	14
Figure 9:	Sheet 5	15

# Introduction

---

Crime in London, as in various cities, is very hard to measure. Police figures generally understate crime substantially and can be extremely misleading. However, with advancement in data science, which plays a major role for scrutinizing, examining and using predictive analysis and other analytical techniques for identifying potential criminal activity and just maybe, we could one day reach a full-fledged algorithm for predictive policing to stop crimes before they happen.

As of 2018, London was ranked 220<sup>th</sup> in the list of safest cities around the world with a safety index of 50.79, with substantial increases in violent and theft crime. The murder toll rose to 130 in 2018 which was at the highest in the last decade. But with a population of 8.7 million people, a lot of factors like race, religion, colour etc may come into context as to why these crimes have occurred.

In this project, I am to build an app with shiny to perform an in-depth analysis of the crime rate across various types and boroughs in London and showcase how the crime types are focused across the different boroughs and predict the future crime rate for the corresponding crime type and borough. This project can be used by politicians, police, real-estate agencies and the general public to assess how safe their borough is and to estimate which is the most likely crime to take place at a particular borough.

# Design

---

To design effective data visualisations, I used the five sheet design(FDS) methodology.

- In my first sheet, I brain-stormed the different ideas, questions, datasets, filters and performed various combinations of charts as to which may help answer my questions.
- In the second sheet, I used my ideas to design a layout and a type of interaction. I considered making a choropleth or a contour map. However, I recognized that an interactive bubble map would give a detailed view of the data. How this would work was, by using a map as the base and the bubble (size of the bubble indicating crime rate), we can click on the bubbles and retrieve information on what is the crime rate at that location. To perform more in-depth analysis, I provided a filter for crime type, so the user can visualise how the crime rate for a crime type is dispersed across London.
- In the third sheet, I thought of creating a bar or line chart to show how the crime rate is distributed across the years for a borough and crime type. I finalised my visualisation on a bar chart as it wouldn't clutter my visualisation.
- In the fourth sheet, I wanted to perform some predictive analysis as to how much we can expect the crime rate to increase or decrease for a borough and crime type. I gave the user the option to choose the smoother type and with a use of a linear model, I displayed the predicted crime rate value.
- In my final sheet, I decided to merge second and third sheet by refining my filters and charts with my interactive map. Thus, on the click of a bubble in the map and on selection of a crime type, the users can access how the crime rate is distributed on a bar chart. This is followed by an interactive prediction plot which lets users analyse whether the crime rate would increase/decrease for a given borough and crime type. I built the app using Shiny in R with an estimated time to build of a week.

# Implementation

I built the app using Shiny in R with an estimated time to build of a week. Some of the software libraries I used are,

1. shinydashboard(for creating interacting dashboard designs)
2. ggplot2(for creating basic plots)
3. plotly(for creating interactive plots)
4. leaflet(for creating interactive maps)
5. dplyr(for data manipulation)
6. flexdashboard (flexibility in designing dashboards)
7. shinythemes (provides various theme color for shiny)

It took an estimated time of a week to build the app with shiny. The reason behind using these libraries and visualisations was mainly to achieve two major tasks,

- How is the crime type dispersed around the different boroughs in London and their trend across the years?
- To predict crime rate of a crime type at a borough

I believe my choice of visualisations helps the users achieve or gain some insight to answer the above two questions. The screenshots below give us just a gist of the insights,

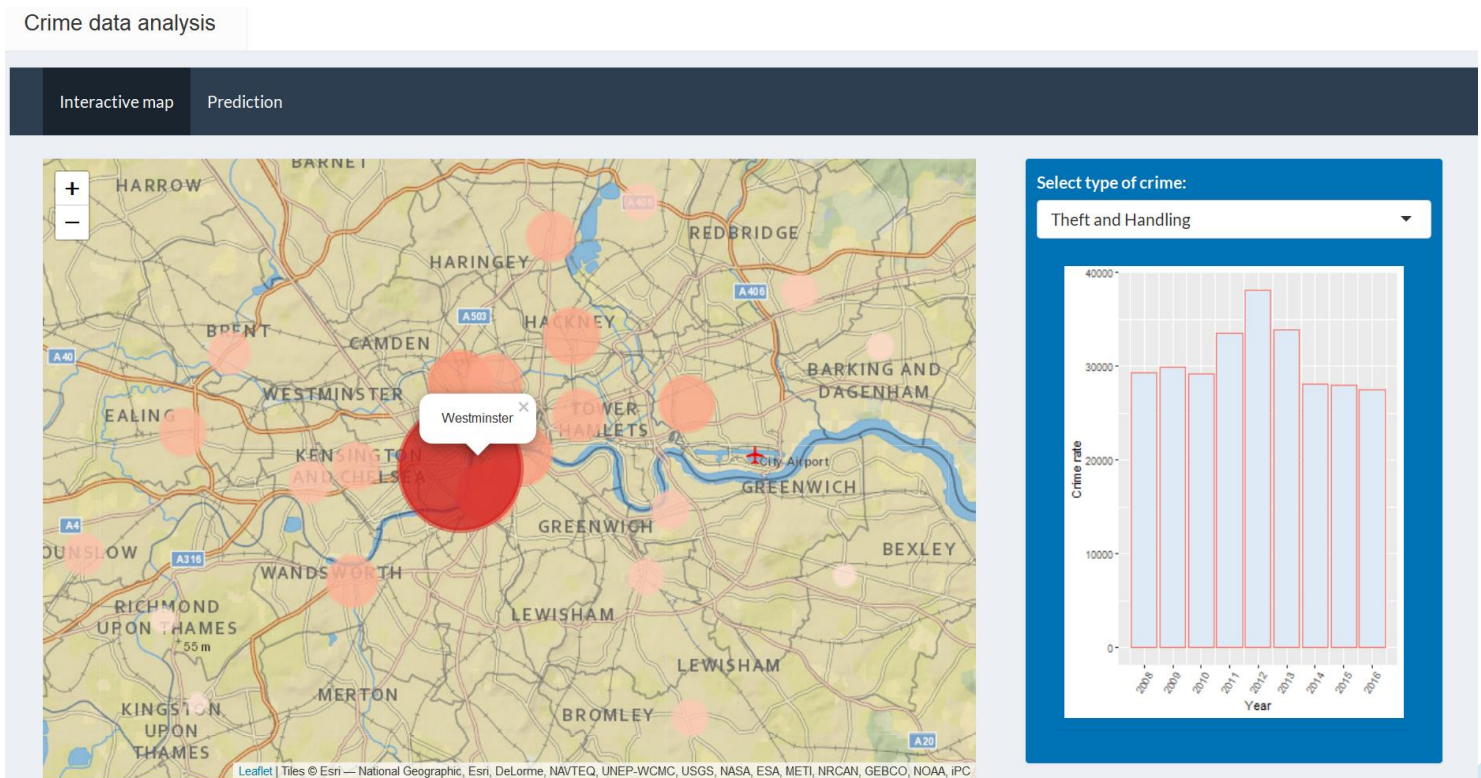


Figure 1: - Screenshot of interactive map and bar chart

As seen from the above map, theft and handling is one of the most prominent crimes in London with most occurrences at the city centre. On clicking on a bubble, as in the above screenshot, we get details of which borough we are focusing on and the crime rate for a particular crime type.

But we aren't only interested to see how the crime is dispersed but we are also interested in predicting the crime rate which is indeed of esteem importance to most of the users. The screenshot below shows how the prediction tab works,

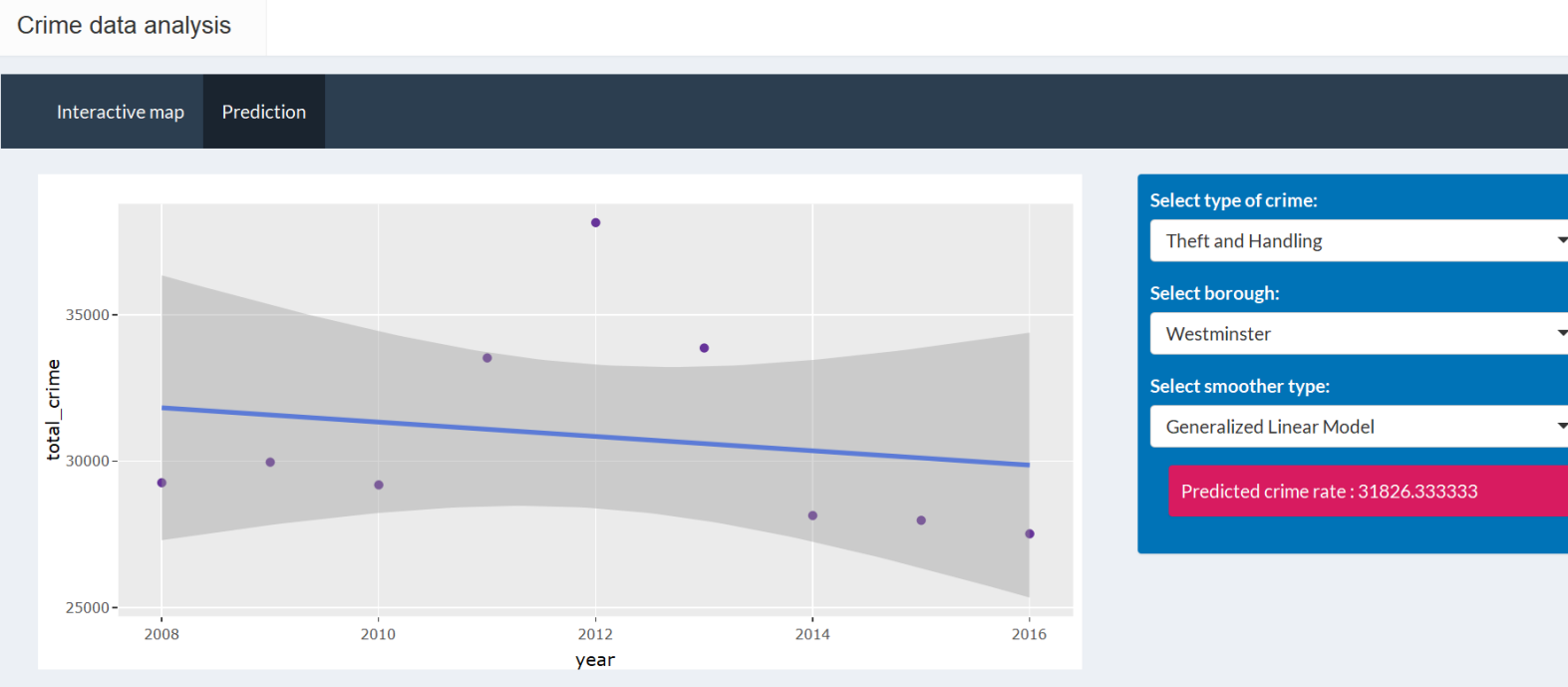


Figure 2: - Screenshot of prediction tab

As seen above, we can get the predicted crime rate at Westminster for theft and handling. This can be done for different crime types and boroughs to get an estimate as to how the crime rate differs.

# User guide

- The application consists of two tabs, named interactive map and prediction
- The interactive map tab displays an interactive bubble map, a filter to choose the crime type and a bar graph showing the crime trend for that borough and crime type. The following screenshot illustrates the working,

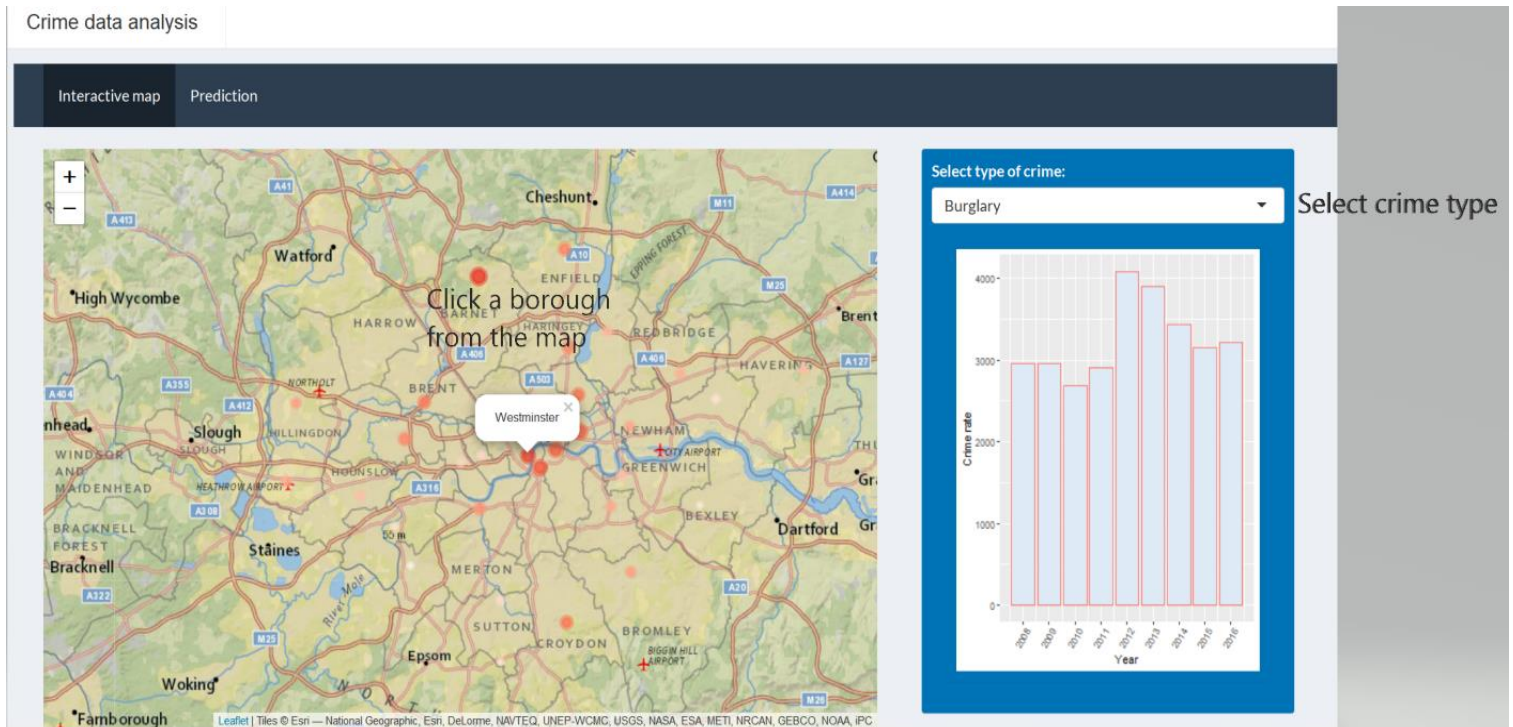


Figure 3: - Screenshot showing working of interactive map tab

- The prediction tab displays a prediction plot for crime rate predictions with three filters, crime type, borough and smoother type. This has a predicted crime rate in a box which shows the predicted crime rate. The screenshot below illustrates this,

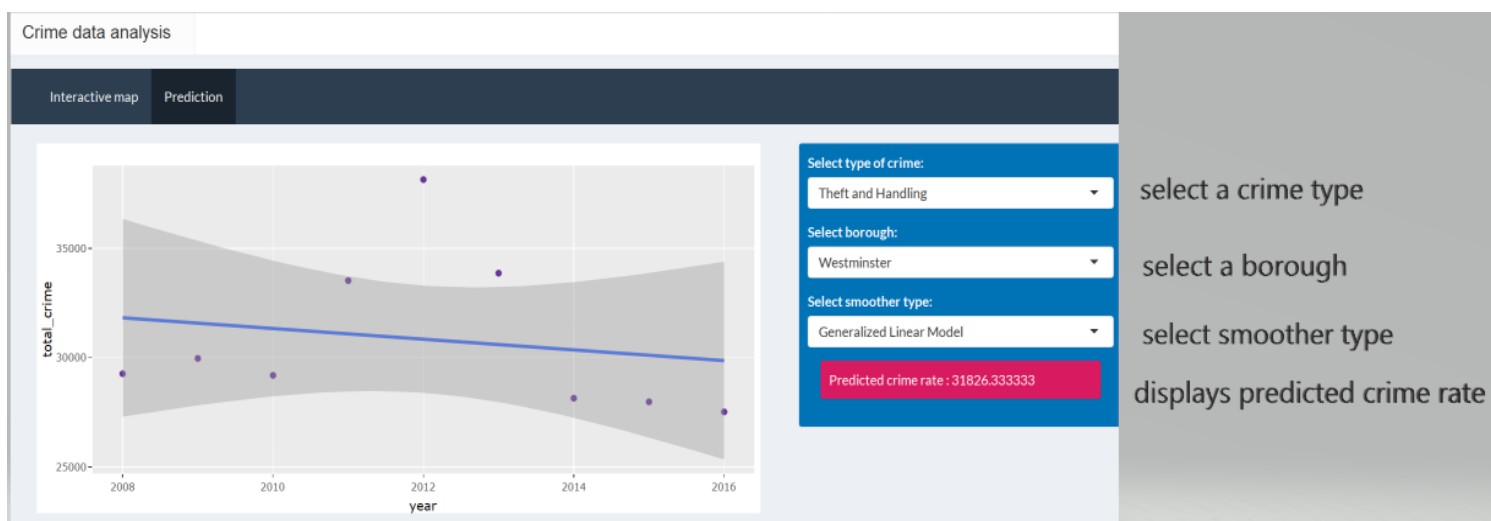


Figure 4: - Screenshot showing working of prediction tab



# Conclusion

---

From the above app, we get the following insights,

- Burglary is concentrated more around the city centre and the northern suburbs especially at Barnet, but the crime rate is gradually decreasing as compared to Hackney which is experiencing a rise in these crimes. Although more occurrences have been recorded around the city centre, the rate is gradually decreasing except for Tower Hamlets.
- Criminal damage is more concentrated on the southern suburbs especially around Croydon. But it has experienced a substantial decline over the years, as is the case for most of the suburbs in London.
- Drugs related crimes are increasing Lambeth and although there was a significant rate at Brent, over the years it has decreased.
- Robbery is more prominent at the city-centre and eastern suburbs. However, robbery related crime has been increasing at Haringey and Hackney which are the northern suburbs of London.
- Theft and handling is more prominent in the city centre especially at Westminster which is a popular touristic destination which may be one of the factors as to why the rate is so high. Theft and handling rate has been consistently high across the boroughs which could contribute to its ranking of 220<sup>th</sup> in the safest cities in the world.
- Violence against the person has been mainly centred at the city centre, western and southern suburbs of London and at Haringey.
- Other crime types have increased significantly in Hillingdon.

# References

---

1. Tools Used-Shiny, bigQuery, jupyter
2. [www.r-graph-gallery.com](http://www.r-graph-gallery.com)
3. [www.rstudio.github.io/shinydashboard/](http://www.rstudio.github.io/shinydashboard/)
4. <https://stat.ethz.ch>

# Appendix

SHEET-1  
BRAIN STORMING

## London Crime

Audience :- London police, General public

### IDEAS

- \* PREDICT WHICH CATEGORY OF CRIME IS LIKELY TO OCCUR IN EACH BOROUGH
- \* CHECK IF CRIME TYPES ARE RELATED
- \* MOST COMMON TYPE OF CRIME IN EACH BOROUGH
- \* TREND OF CRIME RATE IN EACH BOROUGH ACROSS THE YEARS
- \* RECOGNIZE SAFEST AND DANGEROUS BOROUGH IN LONDON

### DATASET

- \* BOROUGH (32 BOROUGHs)
- \* YEAR (2008-2016)
- \* CRIME TYPE
  - ↳ BURGLARY
  - ↳ ROBBERY
  - ↳ DRUGS
  - ↳ THEFT & HANDLING
  - ↳ VIOLENCE
  - ↳ CRIMINAL DAMAGE
  - ↳ OTHERS
- \* Number of Occurrence
- \* LATITUDE
- \* LONGITUDE



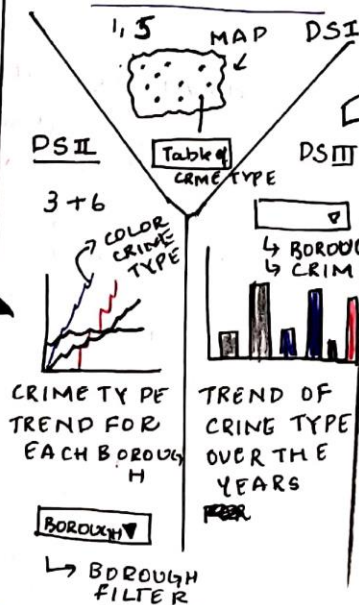
### FILTER

- year
- BOROUGH
- CRIME TYPE

### IDEAS

- \* 1
- \* 5
- \* 3
- \* 6
- \* 2
- \* 4

### CATEGORIZE

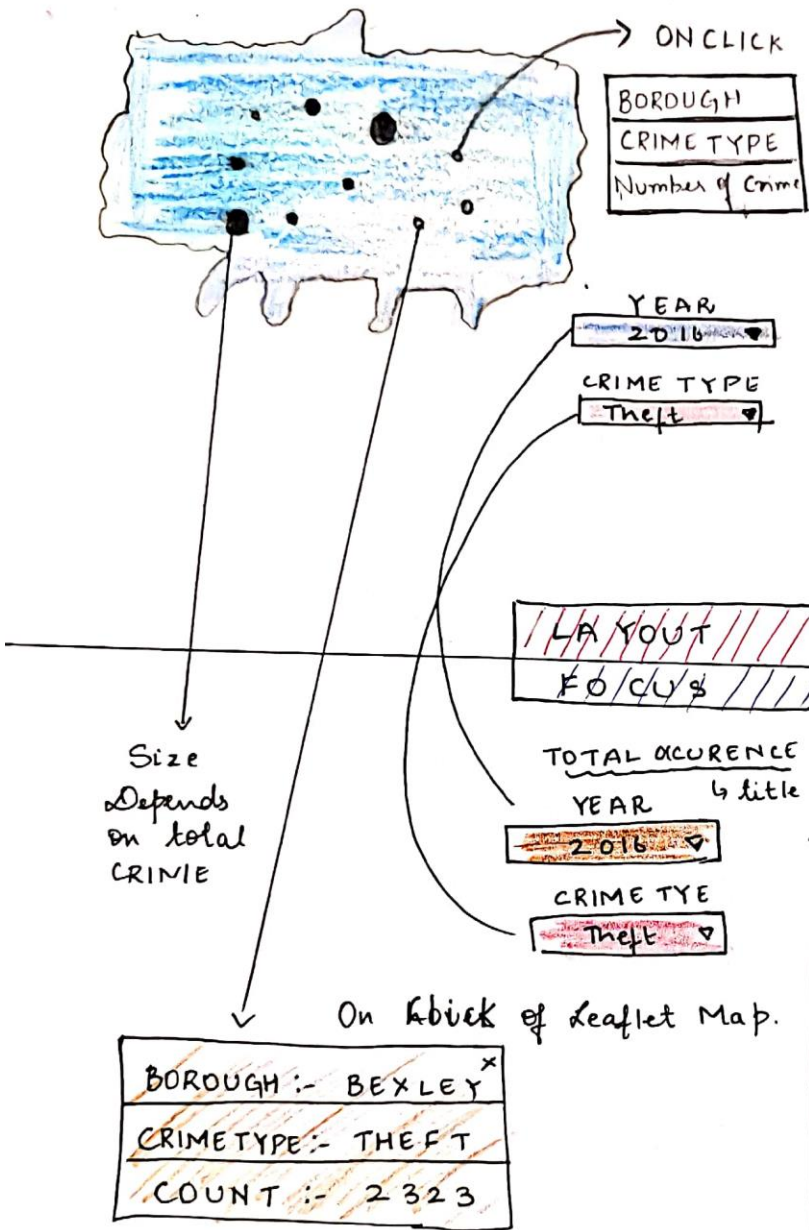


### COMBINE AND REFINES DS I



Figure 1 :- Sheet-1

# Sheet-2



TITLE :- Interactive Map.

Author :- Mohaab.

Date :- 31/05/2019

SHEET :- FDS-2 Initial Design

TASK :- Visualise Crime type

- The interactive map is appended with bubbles.
- The color indicates CRIME TYPE and Size of bubble indicates Occurrence on. Hover displays information about Borough, Crime type & Number of Crime.
- Dropdowns for Year and CRIME TYPE selection.

OPERATIONS

DISCUSSIONS

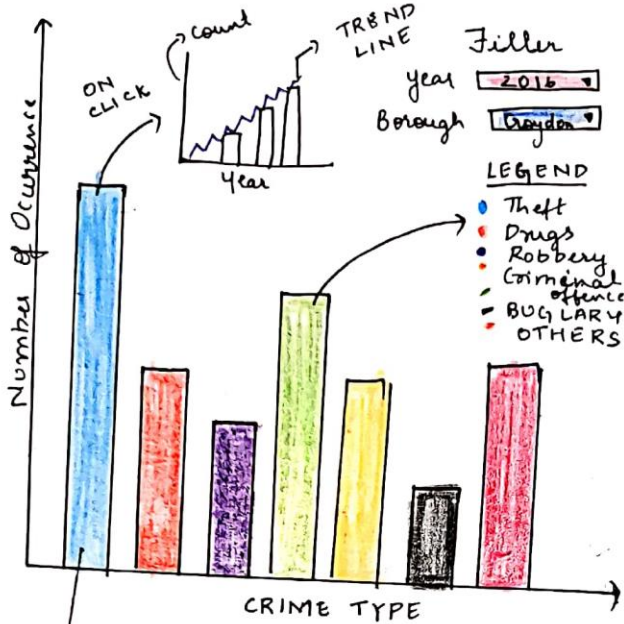
POSITIVE

Regions shown on map.  
By changing filters, can get more detailed information on year & CRIME TYPE

Negative - Information map is cluttered.

Figure 2 :- Sheet-2

## SHEET-3



TITLE :- Interactive Bar and line chart

Author :- Mohaab

SHEET :- FDS-3

TASK :- Visualize crime type.

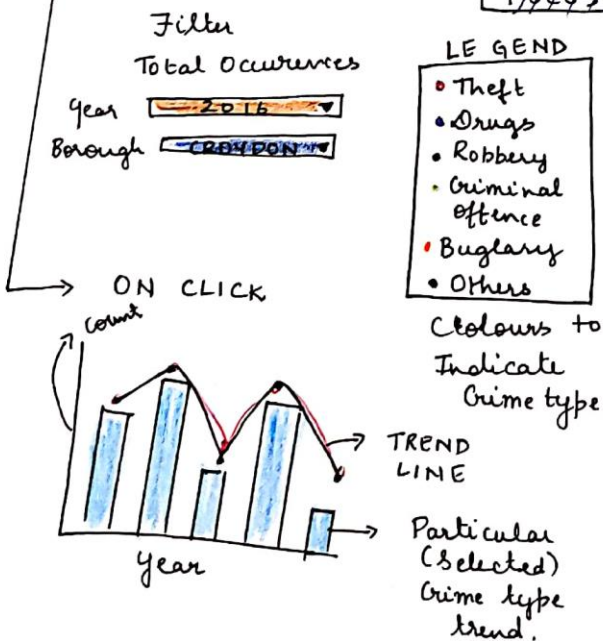
- By Using filters we can easily select year and borough to Visualize crime data in a particular point of time.
- On click of bar trend of particular crime type across the years is shown.

LAYOUT

OPERATIONS

FIELDS

DISCUSSIONS



### Advantages

- \* Easy to filter data
- \* Easy to Access information
- Graph shows detail information on each crime type of borough & also trend can be seen

### Disadvantages

Can't process Generalized View, instead many Subviews

Figure 3:- Sheet-3



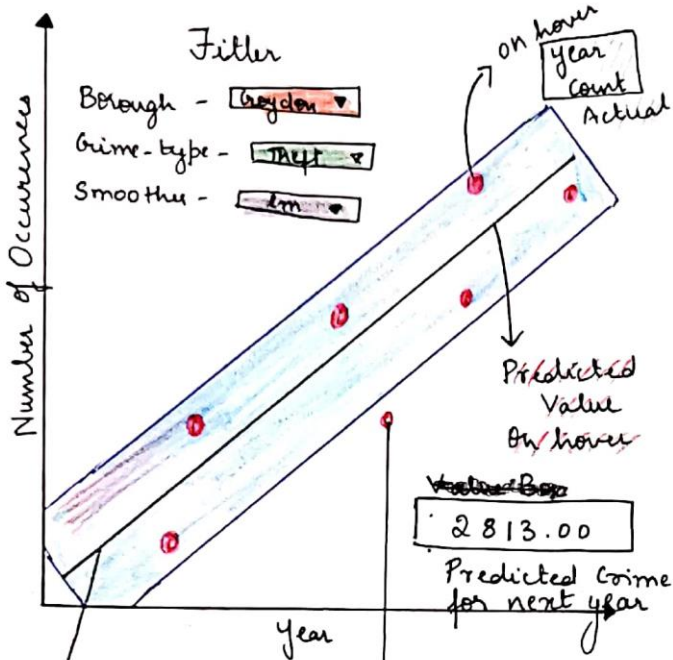
Sheet - 4

Title:- Prediction using scatter plot & trend line

Author :- Mohaab

Sheet :- FSD-4

Visualize crime type



\* Above graph, we can get a sense of Prediction in terms of a particular borough and for same type, how much, can we predict in the rate of occurrence for that crime type borough

\* On hover gives the value in a particular year

\* Various filters for Borough, Crime & Smoother

Operations

Disadvantages

### Positives

gives us rich information on how much we can expect the crime rate to rise or fall for particular type in each borough

### Disadvantage

Difficult to predict values based on smoothing methods.

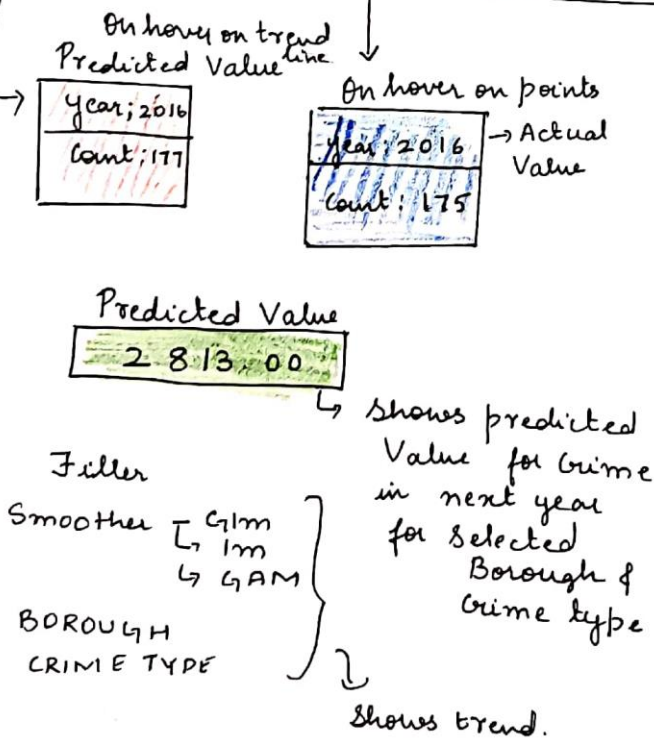
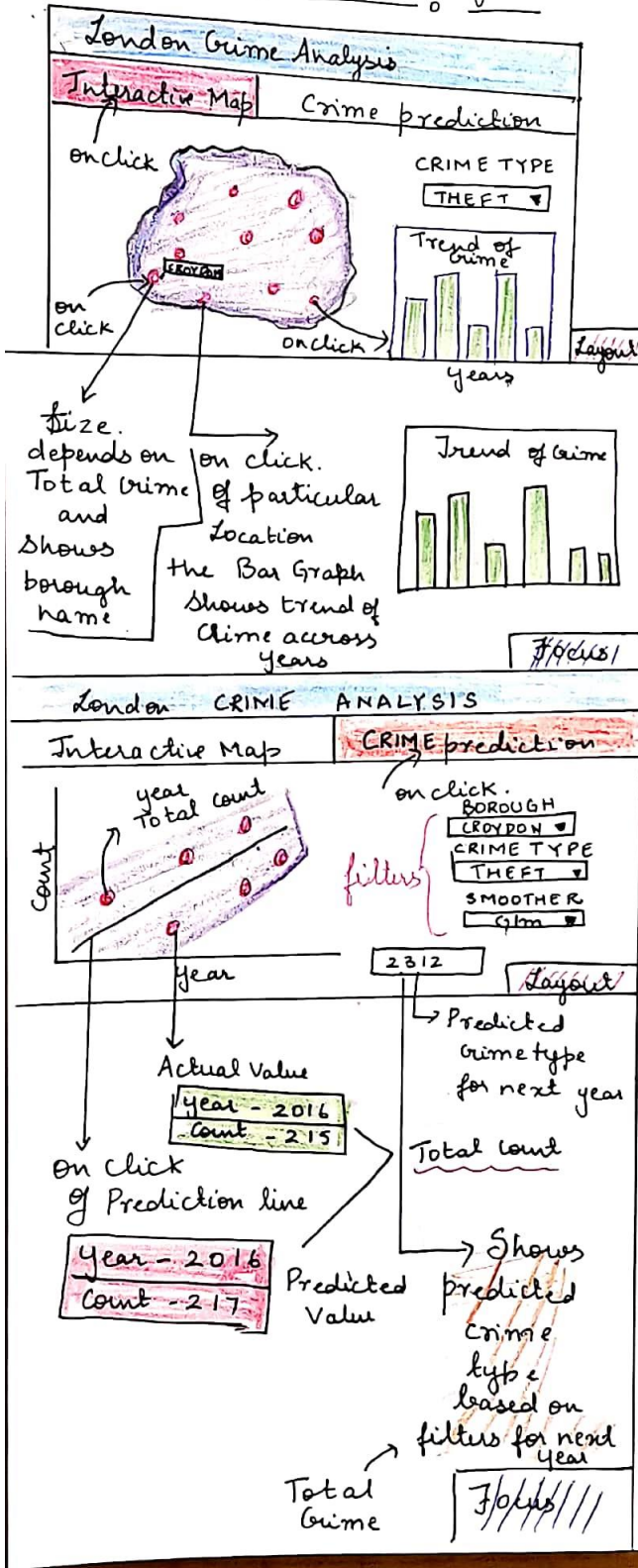


Figure 4 :- Sheet 4

# Sheet - 5 Final Design



Title :- Interactive Dashboard

Author :- Mohaab.

Sheet :- FSD - 5

Visualize crime type of predict crime

- \* On hover of Map (leaflet) trend of crime in particular Region is shown
- \* On click of map points location is displayed.

Second tab Crime prediction

- \* On hover of points actual value is shown
- \* On hover on trend line predicted value
- \* Select filters for Borough, Crime, type of Smoother.

operation

Discussion

\* ALGORITHMS :-

- \* Linear Regression for crime for prediction

\* Design

- \* Size of bubble indicate total crime.
- \* Different colours for each crime type
- \* Different points (scatter plot) has colours for each crime type
- \* Actual Value box & prediction Box of different colour

\* Libraries - Shinydashboard, leaflet, plotly, ggplot, Shiny.

\* Time to build - 1 week

\* Tools - Shiny - R

Figure 5 :- Sheet