Title: MM 802 - Visualization Mini-project

Date: March 18, 2019

Name: Kushal Mahalingaiah and Harsh Sharma

Visualization Title: VanWatch - Helping locate crime corridors in Vancouver

1 Abstract

Migrating to a new city is a daunting task, especially when it is a big one and you are unaware of the security risks and threats involved in your new neighborhood. Even though with the Open Data initiative the data for all the breaches is readily available, understanding it is time consuming, difficult, and boring. In this assignment we visualize and analyze the crime rate in various neighborhoods of Vancouver, one of the most important cities of Canada. We provide a map based visualization for the different types of crimes committed all across the city between 2003 and 2019 and a graphical analysis to help identify the trends.

2 Introduction

When moving to a new place, the first ting that most of us do is, inquire about the crimes in the locality. One of the ways of acquiring these details is via the Open Data Catalogue of that particular city but in this time and age no one has the patience to scroll through an excel sheet containing hundreds of thousands of rows. Our main objective for this mini-project was to assist the user in making an informed decision about the safety of their neighborhood.

During our brainstorming sessions, we came across a map based analysis of the crimes in Chicago [1]. Being inspired by the clean and intuitive concept, we have created a graphical user interface (GUI) that provides the end user to find out useful information like the type of crime occurred in a neighborhood in a specific year or month. The user can filter the data to find answers to questions like, *which locality had the most break-ins in the year 2015* or *which street had the most theft of cars in the month of June*. We have provided filters for month, year and type of crime to answer such queries.

We have obtained the data of the crimes that took place in the city of Vancouver from the Vancouver Open Data Catalogue [2] to create these visualizations. The data was extracted on 2019-02-28 and last updated on 2019-03-18. It contains 590,738 records from 2003-01-01 to 2019-01-04. The crimes have been categorized as follows

- Break and Enter Residential/Other
- Mischief
- Break and Enter Commercial
- Theft from Vehicle

- Vehicle Collision or Pedestrian Struck (with Injury)
- Vehicle Collision or Pedestrian Struck (with Fatality)
- Theft of Vehicle
- Theft of Bicycle
- Offense against a person
- Other Theft

In order to protect the privacy, the data for ”Offense against a person” has been anonymized and there are no locality details available. Since, for the purpose of this assignment, those details are essential we have filtered out all the crimes in this category.

We are storing the data as comma separated values and processing it using R. Our dataset contains the following information

1. Type or Category of Crime - detailed above
2. Year, Month, Day, Hour and Minute when the crime occurred
3. Hundred block radius information of the crime occurrence, with block or avenue or street
4. Neighborhood name in Vancouver
5. Coordinates in Urchin Tracking Module (UTM) Zone 10

The latitude and longitude details are calculated by us from the UTM zone and appended to the data frame.

3 Mini-Project Status

While developing this mini-project, our aim was to create community awareness about the policing activity in Vancouver by providing the information in an aesthetic manner. We intended to create graphs and maps that would be easy to decipher for a layman. We have attained this by creating an intuitive map based layout with easy to use filters. To analyze the trends in crime, we have constructed graphs for various types of crimes in the city and are thus able to provide all the necessary information in an elegant manner.

In order to achieve our goals in the best possible manner, we divided the work equally and coordinated together on the dataset gathering and review process. The mini-project can be broadly divided into following six parts. The names of the member responsible is mentioned alongside

1. Dataset gathering and review - *Harsh and Kushal*
2. Dataframe generation for dataset preprocessing - *Harsh and Kushal*
3. Shinyapp server logic for plotting - *Harsh*

4. Shinyapp UI for the application - *Kushal*
5. leaflet map plotting - *Kushal*
6. ggplot2 graphs plotting - *Harsh*

4 Development Environment

With more and more users accessing the internet via mobile devices, it is crucial for the developers to make the website responsive and adaptive to any screen. The framework for our website consists of a client machine that interacts with the GUI and a server that processes the user queries by connecting with the datafile.

We have developed the application using the R programming language and used Shiny [3] as our server to interact and process the data. Our development environment consists of RStudio - a free and open-source integrated development environment for R [4]. The details of the libraries used are provided in the next section.

5 Development Work

R is a programming language that is widely used among statisticians and data miners for data analysis and visualization. There are several libraries that have been developed to assist this process. For the purpose of this assignment we have used the following libraries

- **leaflet** - This is one of the *most popular open-source JavaScript library for creating interactive maps. It is used by websites ranging from The New York Times and The Washington Post to GitHub and Flickr, as well as GIS specialists like OpenStreetMap, Mapbox, and CartoDB.*[5] It creates widgets that has features like zooming, panning, etc.
- **ggplot2** - *ggplot2 is a system for declaratively creating graphics, based on The Grammar of Graphics* [6]. It is one of the most elegant and aesthetically pleasing graphics framework in R. It helps to create complex plots by using data in data frames.
- **readr** - We use this library to read and parse data from the csv file [7].
- **dplyr** - After reading the data into data frames, we use dplyr to manipulate and summarize the data [8].
- **rgdal** - rgdal provides bindings to the 'Geospatial' Data Abstraction Library, a library for reading and writing geospatial data [9].
- **shiny** - The web server (explained in the previous section)
- **rsconnect** - This is a deployment interface for Shiny applications [10]. We utilized this to deploy our app to https://hsharma.shinyapps.io/mm802a3/ and https://kushalhm.shinyapps.io/vanwatch/

5.1 How to use VanWatch?

The GUI for VanWatch has been developed to make the usage intuitive and simple. On first load the screen loads with the map displaying total crime count in various areas of the city between 2018 and 2019. The user has the ability to zoom in or out further to see details upto the street level

The second tab on the screen, ”Crime Statistics”, shows the trends for variuos types of crimes across the years, months and hours. We provide the user with the ability to see into detail a particular type of crime by selecting the checkboxes provided in the left panel.

5.2 Output Images

We have added a few screenshots from our application in the appendix A below that depict the details provided in the previous subsection visually. The first three figures show the map based visualization at various zoom levels. When hovered over the marker, we also display the crime details. The remaining five figures show the trends graphically and exhibit the use of various types of filters available to the user.

5.3 Limitations

We have developed the application keeping in mind the KISS principle - *Keep It Simple Silly!*. This has enabled us to keep the UI minimalistic and natural but also has a few limitations. To keep the layout uncluttered, we have not shown different crimes using different markers which makes it a little difficult for the user when visualizing data across several years. Also, the trends that we have visualized graphically are the ones that we found to be most significant. However, that might not be true for all of our users.

The data available to us was only till January 2019 and hence, our trends are limited to that particular month and year. We need to manually update the database to provide the latest information to our users.

6 Concluding Remarks

As mentioned in the earlier sections, our objective with this mini-project was to make the end user aware about their locality and the types of crimes committed so that they are better prepared. We have been able to accomplish this by creating this dashboard and hosting it online. The citizens of Vancouver have now access to a portal where they can see and understand the breaches that took place in their city over the years. With this mini-project, we have managed to provide information that was available as boring excel sheets on the government websites in an aesthetically pleasing manner.

During the development, we faced several challenges that we have highlighted below. Our main hurdle was the selection of the tech stack. Since both of us had experience (horrible) with servers like Apache Tomcat and we wanted to learn something new, we decided to dig into Python and R based frameworks. We did a thorough review of all the available server and finally decided to use Shiny with R because of it is elegance and simplicity. Our code is completely platform independent and can be run out of the box. The developed application is fully responsive and runs seamlessly on both mobile and desktop devices. During the development, we came across the existing website of the Vancouver police department[11] and the load time for it was horrible and the dashboard was too cluttered. This made us realize the importance of our minimalistic approach and the KISS principle.

6.1 Future Work and Extensions

We identified a few limitations of our mini-project in the previous section and are determined on resolving them. To keep the data up to date, a pipeline can be developed that automatically connects to the open data catalogue and updates our database. We also need to collect user feedback on the trends that we have visualized to customize the graphs to user specific needs to make it dynamic and appealing.

This portal has been developed for visualizing the crime statistics of Vancouver, however, we would love to see it extended to other major cities of the world. Our code is easy to manage and understand and the only hurdle in this extension would be to format the initial data in a comma separated values files to provide as an input.

References

[1] Chicago Crime Mapping, 2019. https://www.kaggle.com/uds5501/chicago-crime-mapping.

[2] Open Data catalogue — City of Vancouver, 2019. https://vancouver.ca/your-government/open-data-catalogue.aspx.

[3] Shiny, 2019. https://shiny.rstudio.com/.

[4] Open source and enterprise-ready professional software for data science - RStudio, 2019. https://www.rstudio.com/.

[5] Leaflet for R - Introduction, 2019. https://rstudio.github.io/leaflet/.

[6] Create elegant data visualisations using the grammar of graphics ggplot2, 2019. https://ggplot2.tidyverse.org/.

[7] Read Rectangular Text Data readr, 2019. https://readr.tidyverse.org/.

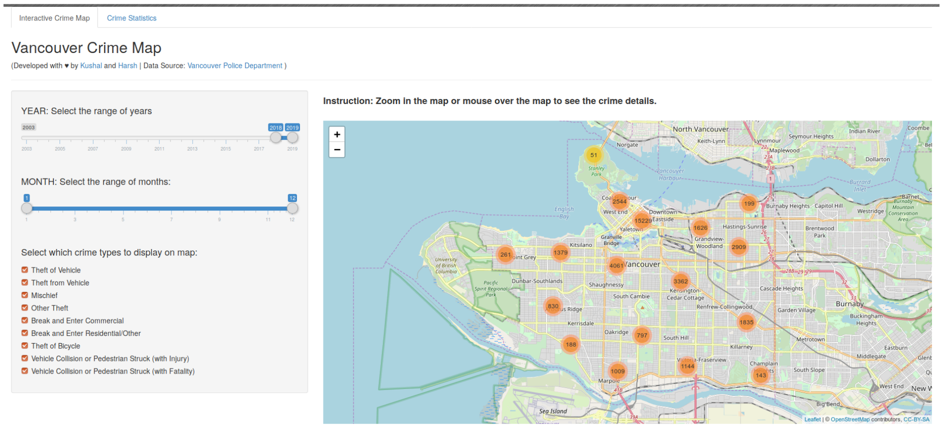
[8] A Grammar of Data Manipulation dplyr, 2019. https://dplyr.tidyverse.org.

[9] CRAN - Package rgdal, 2019. https://cran.r-project.org/web/packages/rgdal/index.html.

[10] CRAN - Package rsconnect, 2019. https://cran.r-project.org/web/packages/rsconnect/index.html.

[11] VPD GeoDASH, 2019.

7 APPENDIX A - Output Images



Vancouver Crime Map

(Developed with ♥ by Kushal and Harsh | Data Source: Vancouver Police Department)

YEAR: Select the range of years

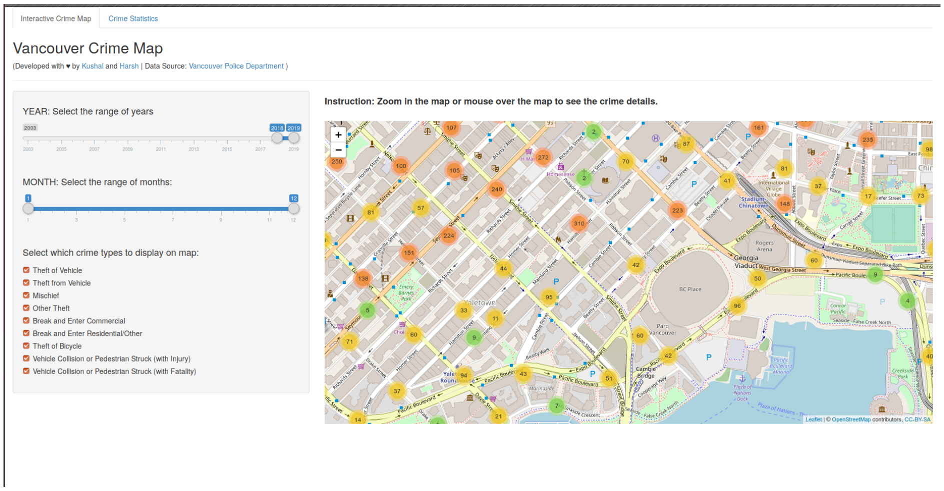
MONTH: Select the range of months:

Select which crime types to display on map:

- ☑ Theft of Vehicle
- ☑ Theft from Vehicle
- ☑ Mischief
- ☑ Other Theft
- ☑ Break and Enter Commercial
- ☑ Break and Enter Residential/Other
- ☑ Theft of Bicycle
- ☑ Vehicle Collision or Pedestrian Struck (with Injury)
- ☑ Vehicle Collision or Pedestrian Struck (with Fatality)

Instruction: Zoom in the map or mouse over the map to see the crime details.

Figure 1: Screen after first load



Vancouver Crime Map

(Developed with ♥ by Kushal and Harsh | Data Source: Vancouver Police Department)

YEAR: Select the range of years

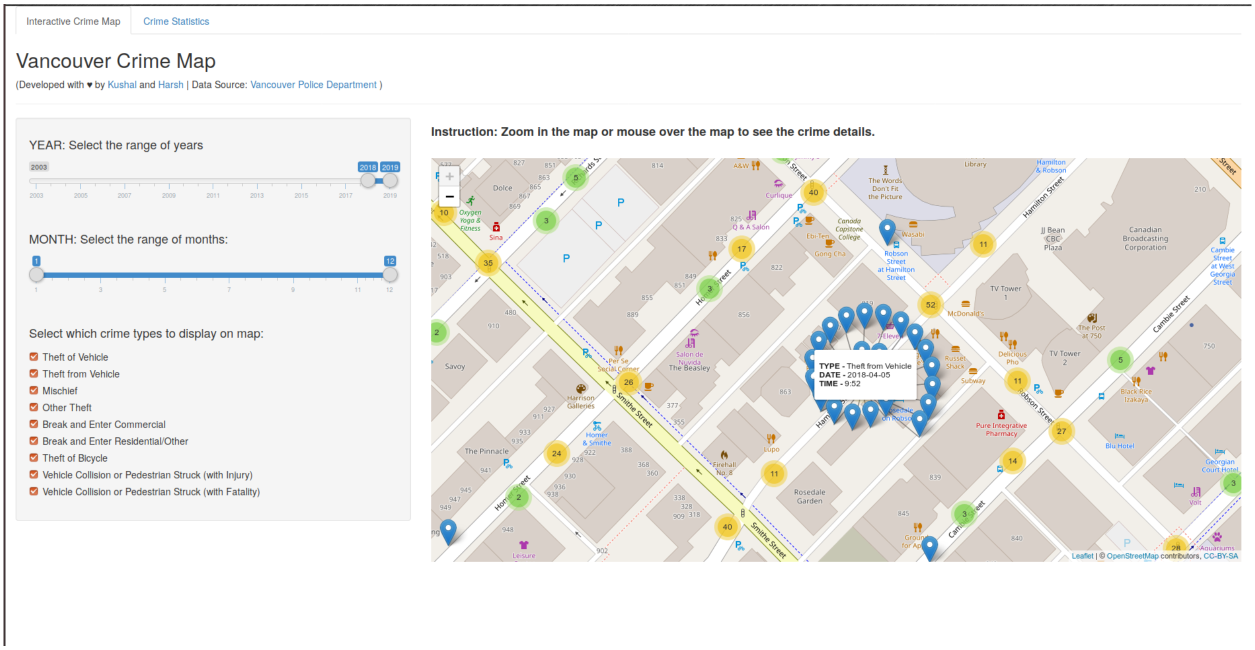
MONTH: Select the range of months:

Select which crime types to display on map:

- ☑ Theft of Vehicle
- ☑ Theft from Vehicle
- ☑ Mischief
- ☑ Other Theft
- ☑ Break and Enter Commercial
- ☑ Break and Enter Residential/Other
- ☑ Theft of Bicycle
- ☑ Vehicle Collision or Pedestrian Struck (with Injury)
- ☑ Vehicle Collision or Pedestrian Struck (with Fatality)

Instruction: Zoom in the map or mouse over the map to see the crime details.

Figure 2: Zoomed in map showing detailed statistics for the neighborhood



Interactive Crime Map | Crime Statistics

Vancouver Crime Map

(Developed with ♥ by Kushal and Harsh | Data Source: Vancouver Police Department)

YEAR: Select the range of years

2003 – 2018 2019

MONTH: Select the range of months:

1 – 12

Select which crime types to display on map:

- ☑ Theft of Vehicle
- ☑ Theft from Vehicle
- ☑ Mischief
- ☑ Other Theft
- ☑ Break and Enter Commercial
- ☑ Break and Enter Residential/Other
- ☑ Theft of Bicycle
- ☑ Vehicle Collision or Pedestrian Struck (with Injury)
- ☑ Vehicle Collision or Pedestrian Struck (with Fatality)

Instruction: Zoom in the map or mouse over the map to see the crime details.

TYPE - Theft from Vehicle
DATE - 2018-04-05
TIME - 9:52

Figure 3: Map at max zoom level for street level information

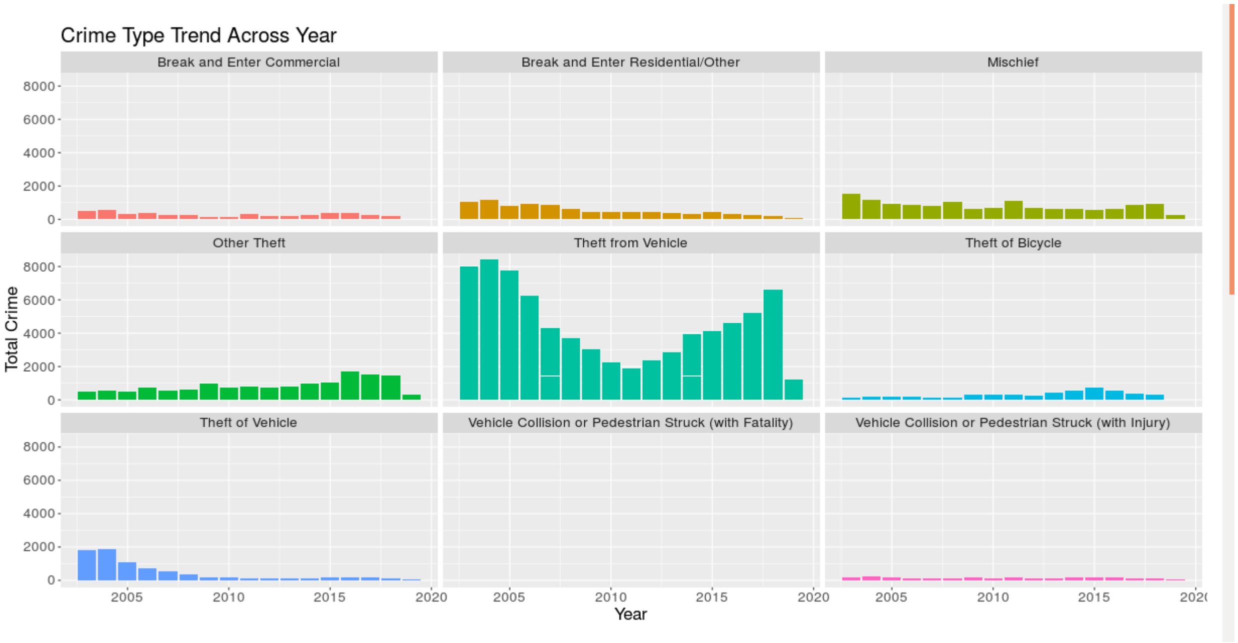


Figure 4: Yearly trends for all the crime types

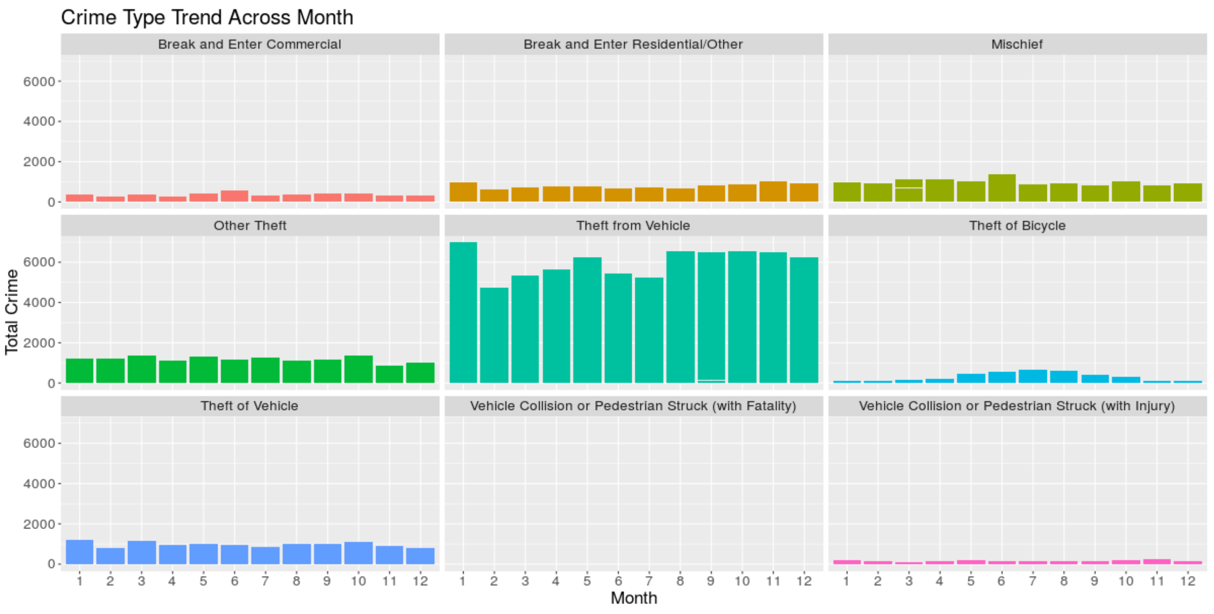


Figure 5: Monthly trends for all the crime types

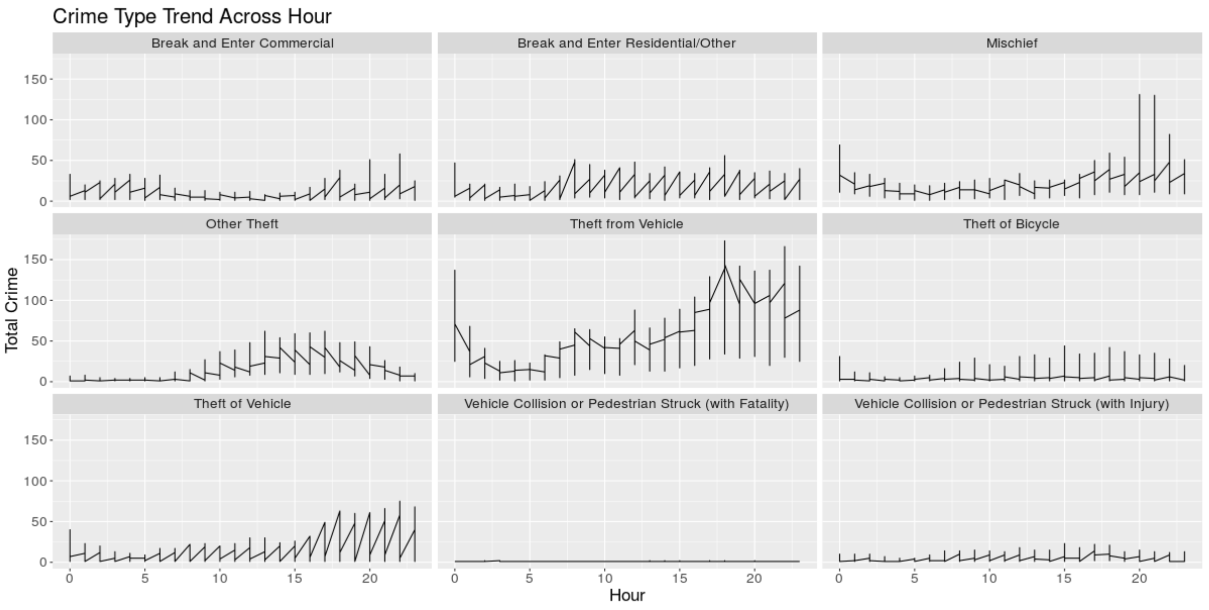


Figure 6: Hourly trends for all the crime types

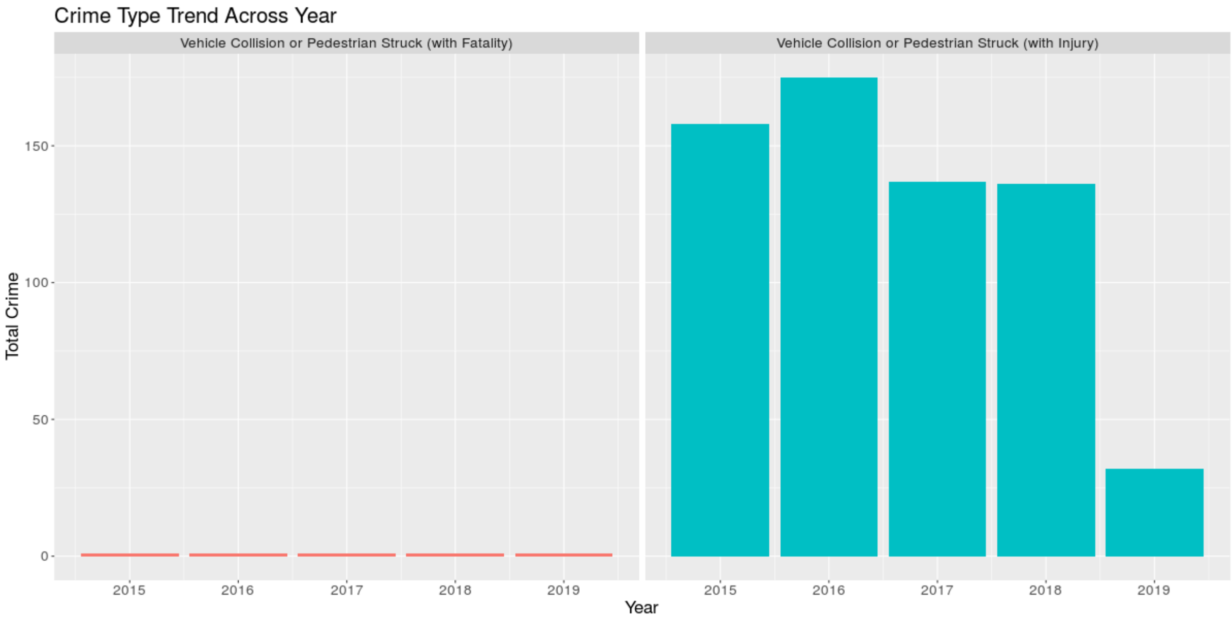


Figure 7: Yearly trends for vehicle crashes between 2015-2019

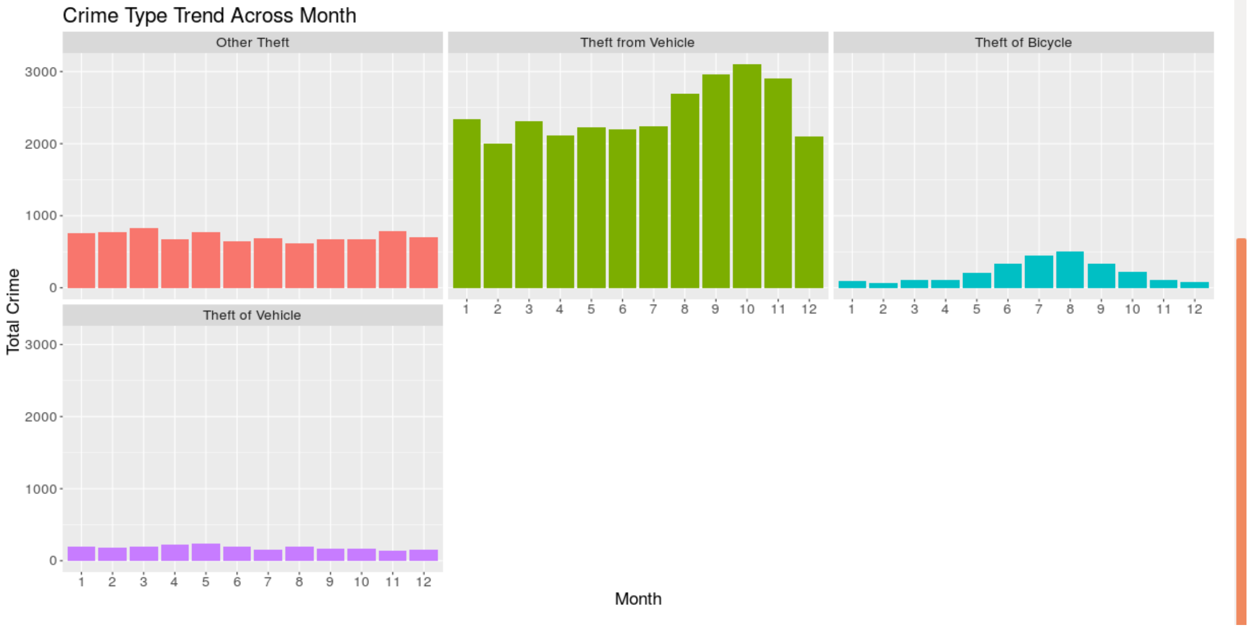
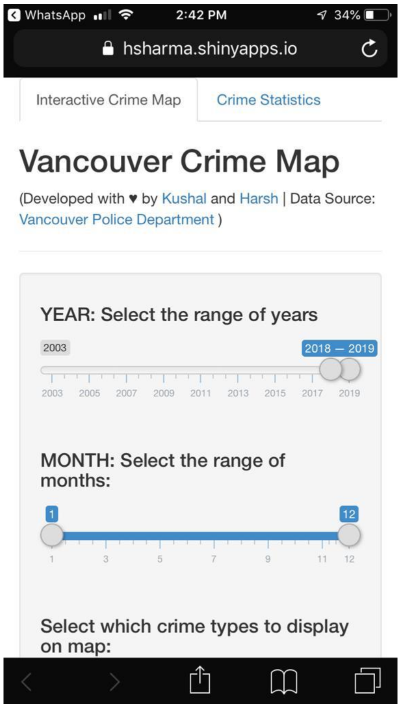


Figure 8: Monthly trends for thefts



Vancouver Crime Map

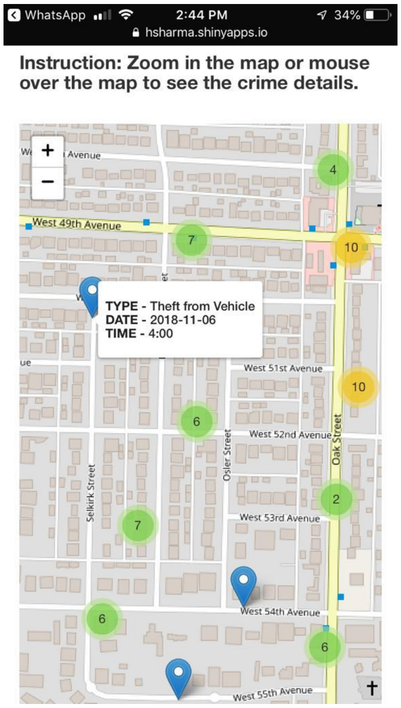
(Developed with ♥ by Kushal and Harsh | Data Source: Vancouver Police Department)

YEAR: Select the range of years

MONTH: Select the range of months:

Select which crime types to display on map:

Figure 9: Filter setting in a mobile phone



Instruction: Zoom in the map or mouse over the map to see the crime details.

TYPE - Theft from Vehicle
DATE - 2018-11-06
TIME - 4:00

Figure 10: Map visualization in an iPhone