



COSC 480S1 PROGRAM USE & DESIGN DOCUMENT

Dhruv Sharma
dsx11@uclive.ac.nz
76749332

INTRODUCTION

Security is everybody's responsibility. With the same idea in mind, we decided to pursue with the project of performing demographic analysis of crimes and victimizations in different regions of New Zealand. To achieve this objective, we shifted our focus on analysing two types kinds of data. One is the data related to Offenders; people who have committed the crime while the other is data for victimizations time and place.

The features which we analysed related to offenders were age, sex and ethnicity. We restricted our study to these offender demographic factors only, because analysing factors including their traits, personal behaviour, name etc would have violated their privacy. We collected data from the NZ police department from 2014 to February 2018 for the both the analysis.

The features which we analysed related to victimizations time and place were the type of crime committed, the territorial authority, area unit, mesh block, hour day month and year and the location type where the crime took place. The results of this analyses are discussed in the Results section of the report.

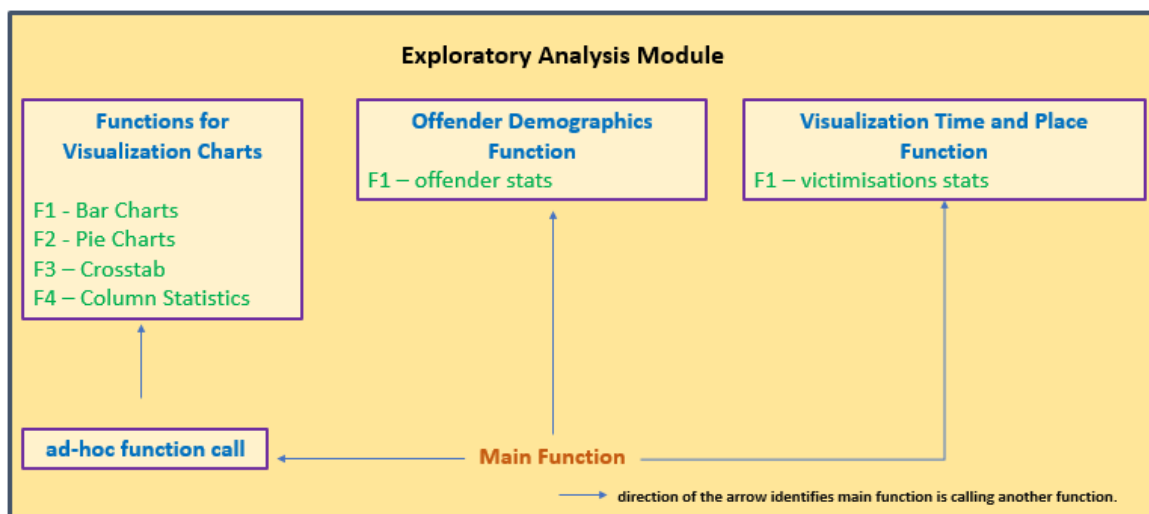
METHODOLOGY

We used Python 3.6.0 as the programming language and Wing as the integrated development environment to analyse the data. We made use of Python Pandas and Numpy libraries to perform exploratory analysis of the data.

Data preparation: We had two CSV files which contained offender data, and victimizations time and place data. A total of 0.4 Million records were extracted and analysed using the Python Pandas library. Data quality was good, except a few cases where data entry issues were encountered and recorded as bad data. However, we did not remove these cases as we wanted to show the True data. Most of the data was categorical in nature which restricted us to perform any meaningful regression analyses.

Program Design: The program was designed in a piece mill approach. The program is divided into functions and each function is created to perform a unique task. The various functions defined in the module are represented below in Figure 3. The Main function calls other functions and displays the results.

Figure 1



User Interface: The user interface for the program is very interactive. The Figure 2 below shows the options displayed when the program is run. While running the program, user can choose Option 1 or 2 to see canned analysis for offender demographics and victimizations time and place. By canned analysis we mean the analysis which had been done to solve the objective of the project.

Figure 2

```
=====
Enter 1 to see Offender Demographics Analysis
Enter 2 to see Victimizations Time and Place Analysis
Enter 3 to do Adhoc Analysis
Enter q to exit
=====
Your Option -
```

However, from the Figure 2, we may also choose option 3 to ad-hoc analysis. The Figure 3 below displays various options, we get after we select Option 3 from Figure 1. You may also exit the loop anytime while running the program by pressing 'q' from the keyboard. To cater to various exceptions in the code, we have used Exception Handling and While loops

Through this program, Option 1 produces 3 analyses for offender's data.

- Number of Crimes by Ethnic Group
- Number of Crimes by Age Group
- Number of Crimes by Year and Sex.

Whereas, for Option 2 for victimizations time and place, the program produces 6 different analyses.

- Top 10 victimizations by Territorial authority.
- Victimisations by Location type division
- Victimisations by Occurrence hour of the day
- Victimisations by Occurrence day of the Week
- Victimisations by Year Month
- Top by Victimisations by Location Type

Figure 3

```
=====
Your Option - 3
Adhoc Analysis
=====
Enter the File Name Number to Read from the below options:
1. Victimisation Time and Place.csv
2. Offender_Demographics.csv
Your Option: 2

For analysis of Offender Demographics - X variables could be any one of the following -
Age Group, Ethnic Group, Mop Division, Mop Group, SEX, ANZSOC Division, Year Month
Y Variables could be any one of the following -
Number of Crime, Proceedings

For analysis of Victimisations Time and Place - X variables could be any one of the following -
Anzsoc Division, Area Unit, Locn Type Division, Occurrence Day Of Week, Occurrence Hour Of Day, Territorial
Authority, Year Month
Y Variables could be- Victimisations

Enter X variable: Area Unit
Enter Y variable: Victimisations
=====
Column Statistics for : Area Unit
=====
Please enter valid X and Y variable entry
=====
Enter 1 to see Offender Demographics Analysis
Enter 2 to see Victimizations Time and Place Analysis
Enter 3 to do Adhoc Analysis
Enter q to exit
=====
Your Option -
```

RESULTS

From the data we performed Year on year trend and various exploratory analysis. Below are some of the outcomes-

- Males have been committing almost 2 times more crimes as compared to females.
- Out of the total, half of the total crimes are committed by people aged between 15 to 35.
- There is a significant Year on Year increase in crimes committed in 2014 to 2015 as compared other years.
- We can identify that the proportion of crimes committed by the people belonging to an ethnic group was proportional to the respective ethnic population in the country in 2017.
- More people are victimized in the day time, opposite to the belief of there being more victimizations during night time.
- We see that there are more crimes during the weekend as compared to the week days.
- Auckland and Christchurch lead the way in highest number of victimizations. As the area of the city increased, trend for number of victimizations also increased proportionately.
- Administrative and Professional locations were the locations where the maximum number of crimes took place.
- The maximum proportion of the crimes committed were related to 'Acts intended to cause injury', 'Public order offences', 'Theft & Related Offences' and 'Traffic & vehicle Regulatory Offences'.

The analyses charts can be found in Appendix section of the report.

DISCUSSION

Using the program, many interesting outcomes came out including the time of the day and the day of the week when victimizations take place. We can share our analyses with the police department to validate our findings, and further appropriate strategies can be devised to curb the crimes.

As such, there are many more use cases which can be analysed using the ad-hoc functionality of the program. However, to mention a few, using the data and the platform, further studies can be done to predict possibility of crimes to be committed in near future in a region. Also, analysis can be done to check if there is any pattern to crimes committed across a region, if yes what could be the possible reasons.

REFERENCES

<https://pandas.pydata.org/pandas-docs/stable/visualization.html>

<http://jonathansoma.com/lede/algorithms-2017/classes/fuzziness-matplotlib/understand-df-plot-in-pandas/>

<https://www.police.govt.nz/about-us/publications-statistics/data-and-statistics/policedatanz?reportName=Offender%20Demographics%20%28Proceedings%29&reportPath=/Live/Reports/&viewerMode=Classic&reportViewOnly=true>

APPENDIX

The Appendix section contains the charts which we produced using the Python program. Many more can be produced and analysed using the ad-hoc functionality of the program.

Offender Demographics Analysis:

Figure 4

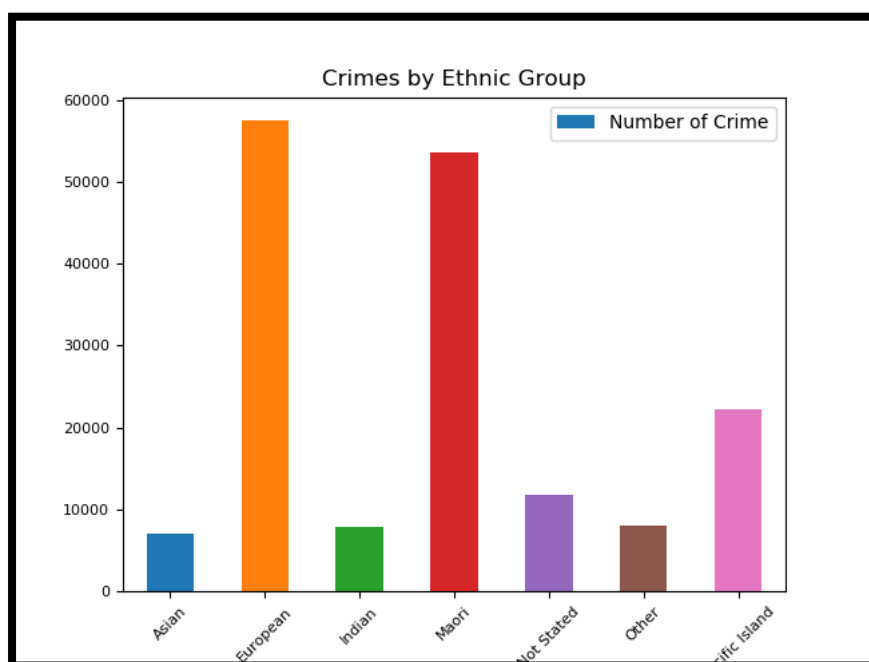


Figure 5

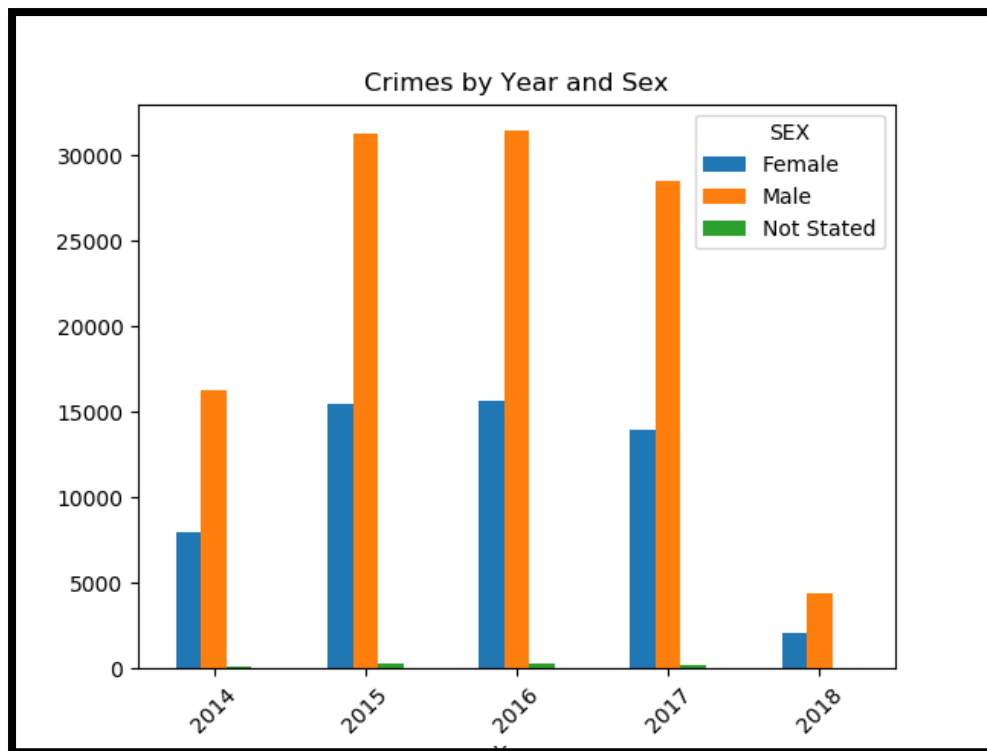
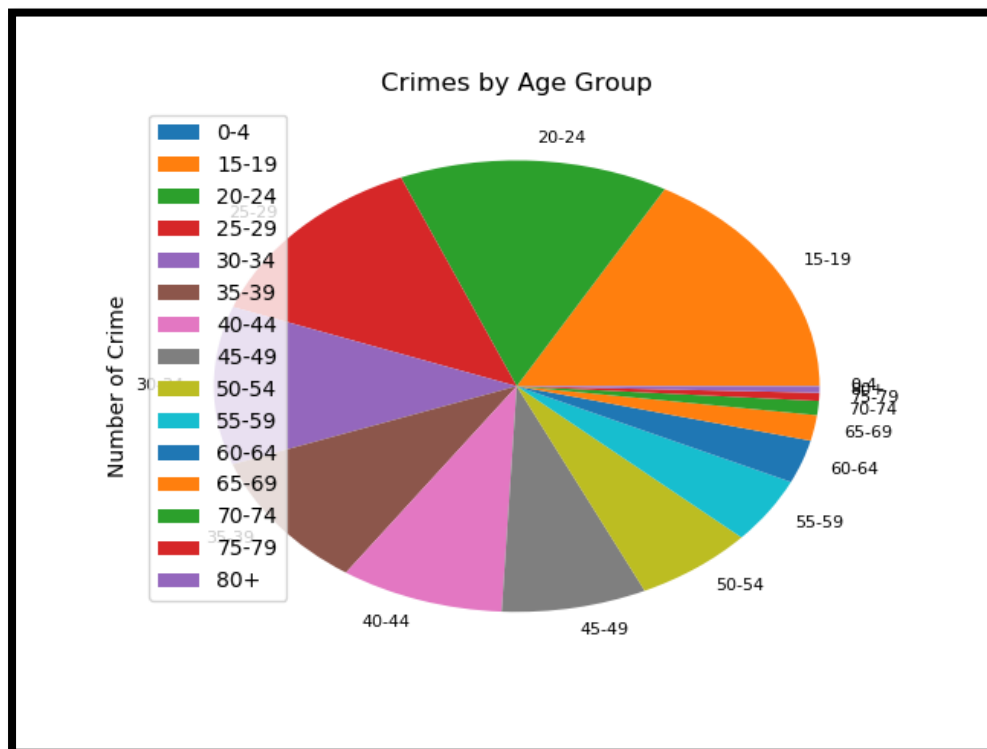


Figure 6



Victimization Time & Place Analysis:

Figure 7

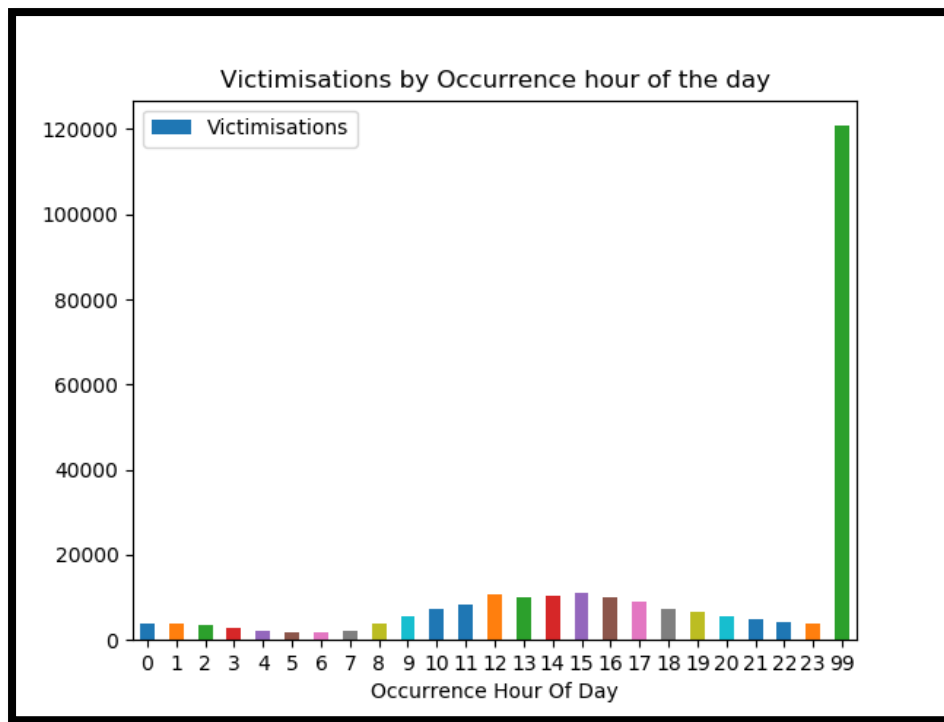


Figure 8

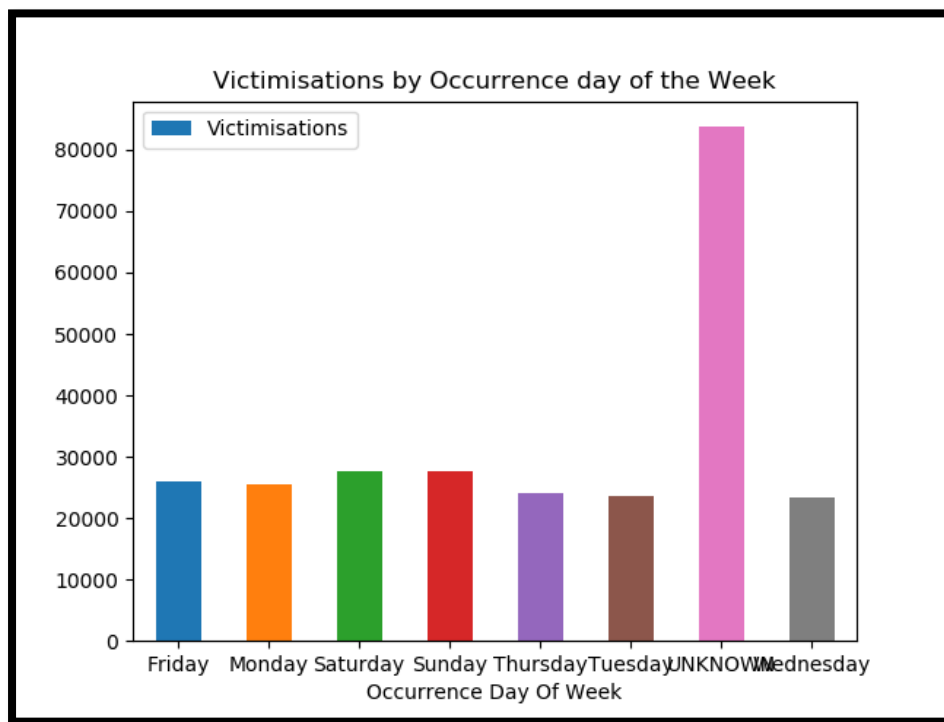


Figure 9

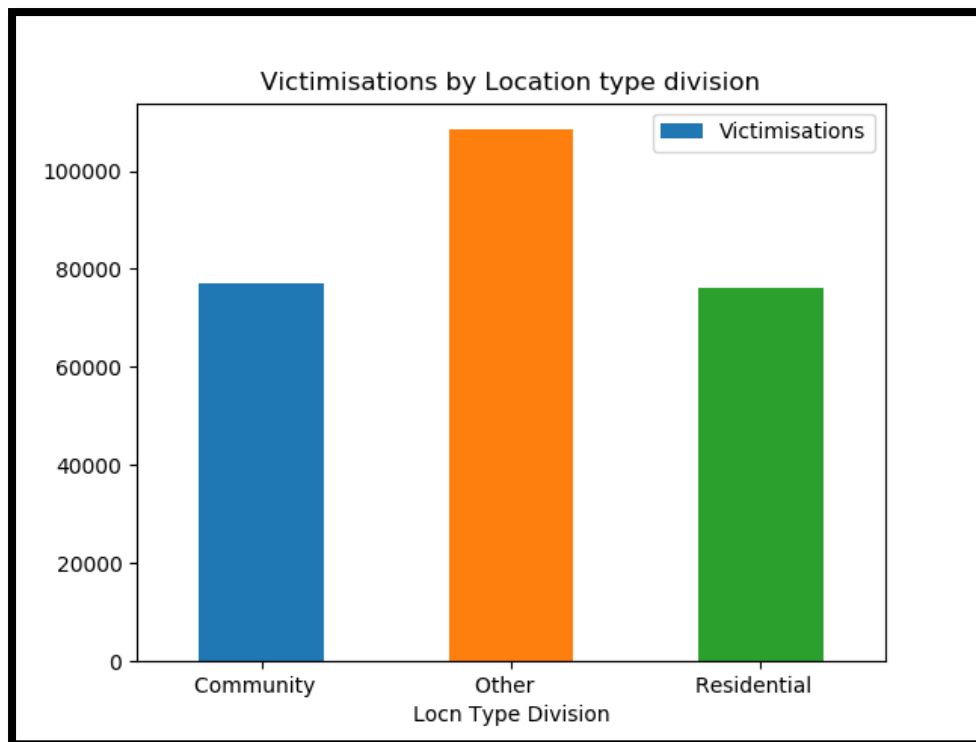


Figure 10

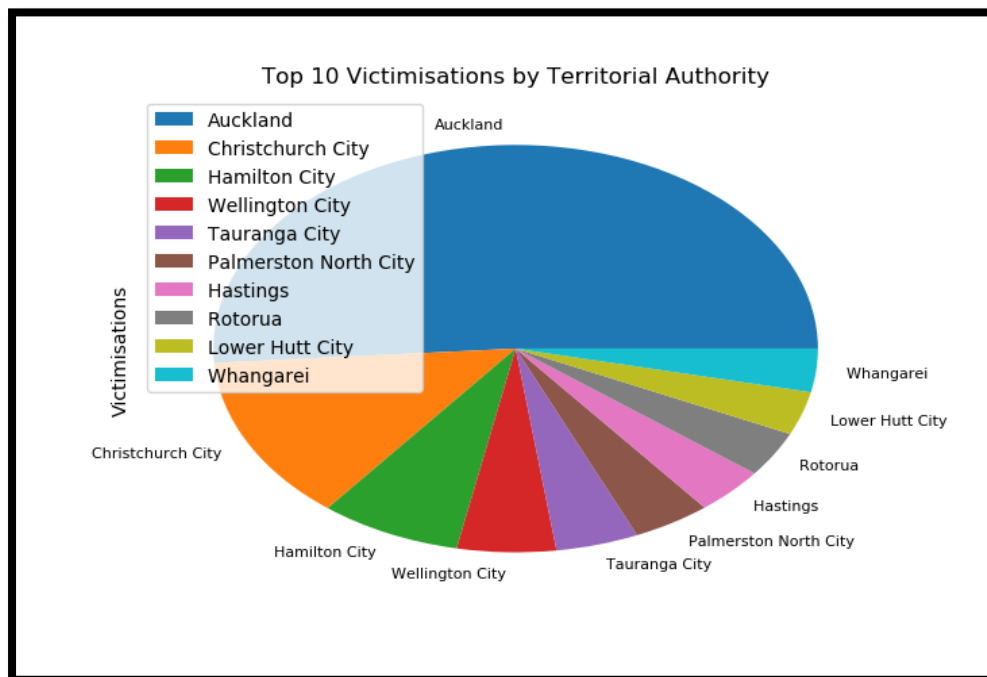


Figure 11

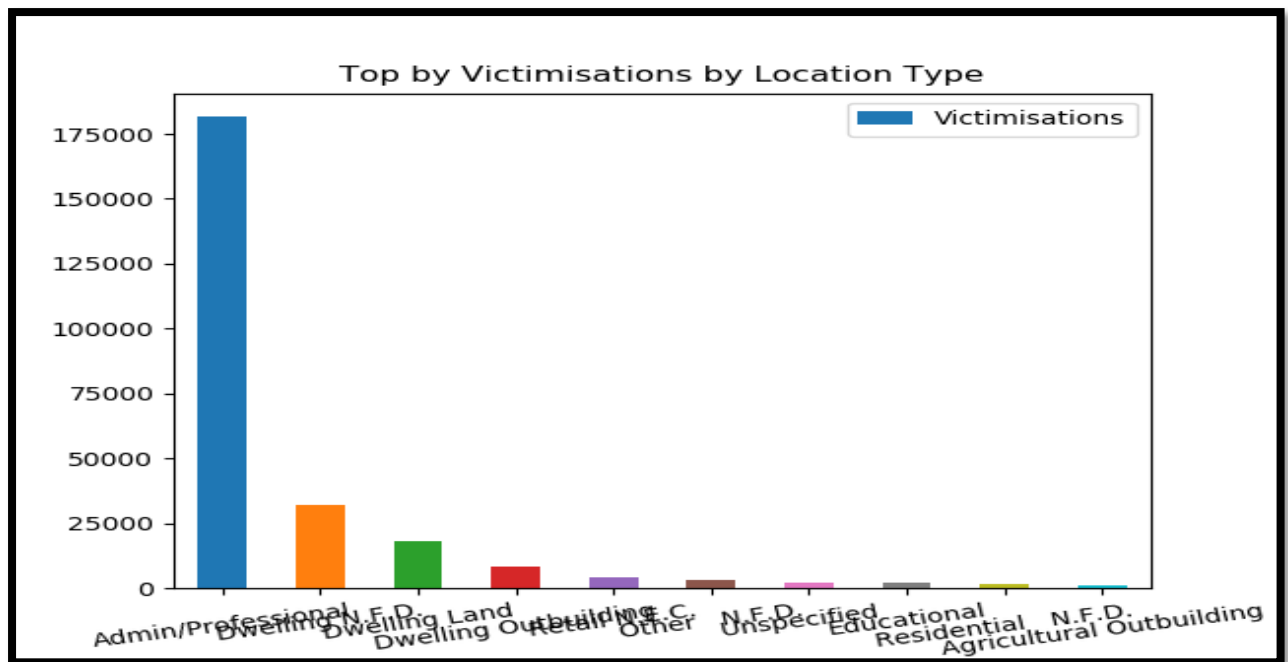
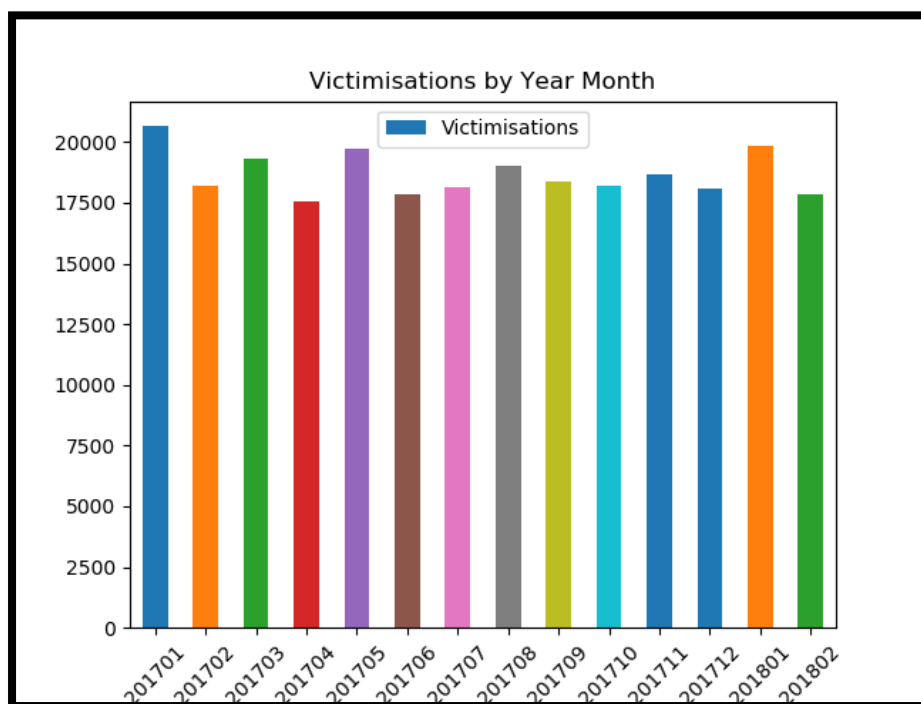


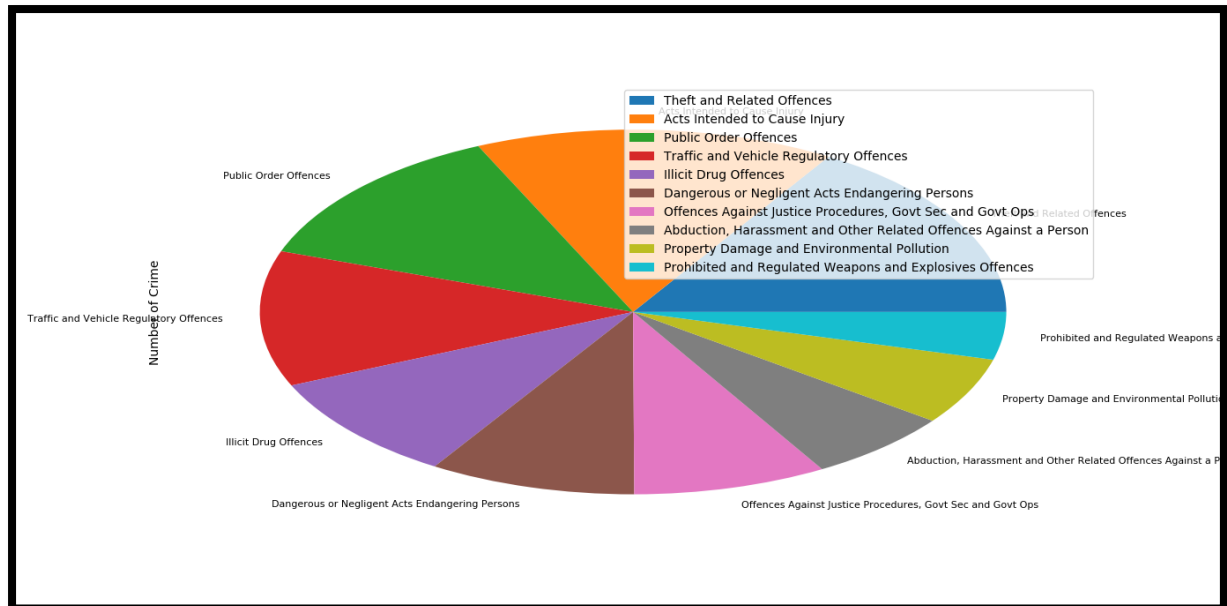
Figure 12



Ad-hoc analysis:

Number of Crimes by Crime Type

Figure 13



***** End of Report*****