

|  |
| --- |
| Data Analysis - Victoria Crime  Statistics 2021 |
|  |
| January 3  Monash Data Analytics Bootcamp  Authored by: Kelvin Hung Nguyen  Anh Huong, Jacqueline Xia, Antoinette Boyle  Your Name |

# Contents

# 1 Introduction 3

# 2 Key definitions 4

# 3 Types of Offences 5

# 4 Offences by Police Region 8

|  |
| --- |
| **Introduction** This report aims to analyze the trend of Victoria crimes by using 5 different excel data sets from Crime Statistics Agency, including the number of recorded offences, criminal incidents, alleged offender incidents, victim reports and family incidents. The data we use is cleaned by using advanced excel techniques including filters, sorting and pivot tables. In addition, python and a number of python modules including matplotlib, scipy, numpy, sklearn and pandas are being used to generate graphs to further support our findings.  Our main focuses are types of offences, number of offences in different police regions   1. **Key Definitions**   2.1 Criminal Incidents  A recorded criminal incident is a criminal event that may include one or more offences, alleged offenders and/or victims, and that is recorded on a single date and at one location.  The criminal incident count is designed to represent the broad event that occurs as a more representative measure of instances of crime, rather than using the bulk number of offences that Victoria Police members record within one incident. Consequently, there's could be many offences in one criminal incident.  2.2 Police Region  Local government areas (LGAs) in Victoria, are sorted into five wider police regions. LGA’s are also referred to as municipalities, the 79 Victorian LGAs are classified as cities (34), shires (38), rural cities (6) and boroughs (1). In general, an urban or suburban LGA is called a city and is governed by a City Council, while a rural LGA covering a larger rural area is usually called a shire and is governed by a Shire Council.  Local councils have the same administrative functions and similar political structures, regardless of their classification.  The sorting of LGAs into five wider police regions we used for our project has no legal or administrative significance.   1. **Types of Offences**   We have found that of the incidents recorded in Victoria the most common incidents relate to Property and deception offences. Please refer to figure 1 below:  Figure 1:  ../Desktop/Screen%20Shot%202022-01-02%20at%2011.25.14%20pm.png  As per table above, the number of incidents has remained fairly stable in the last 10 years except for a peak in 2016 and a significant drop in 2020. The peak in 2016 relates to Theft crimes. Covid 19 lockdowns resulted in a drop of crime in 2020.  An analysis of the breakdown of “Property and deception offences” over the 10 years has revealed that the crimes in this category are mostly “Theft” related crimes. The most significant of these crimes relate to stealing from a motor vehicle (37%). Followed by other crimes are “Burglary or break and enter” and “Deception crimes”. Incidents of Arson and Bribery show the least among others. Please see figure 2 below:  Figure 2:  ../Desktop/Screen%20Shot%202022-01-02%20at%2011.27.34%20pm.png  Apart from that, other more serious crimes like “Crimes against the person” have also occurred over last 10 years, but with lower incidents. “Assault and related offences” have dominated this category with 64%. Please refer to figure 3 below:  Figure 3:    “Drug Crimes” over the last 10 years were mostly for drug possession at 69%. Drug trafficking only 19%.  The category “Public Order and Security offences” were dominated by “Disorderly and offensive conduct” at 57% and “Weapons and explosive offenses” were close behind at 37%.  “Justice procedures offences” included “breaching family violence order” 40% and “breaching bail conditions” 37%.   1. **Offences by Police Region**   Offences per Police Region over the last 10 years were mostly stable and once again 2016 had the increase in theft offences and 2020 had a decrease in crime due to the COVID Lockdowns. The rate per 100,000 population echo’s this trend. Please refer to figure 3:  **Figure 3:**  ../Desktop/Screen%20Shot%202022-01-02%20at%2011.59.15%20pm.png  Noteworthy, North West Region has the highest incidents of crime, which includes the city centre. Followed by The Eastern and Southern Regions. Please refer to figure 4:  **Figure 4:**  ../Desktop/Screen%20Shot%202022-01-02%20at%2011.52.50%20pm.png  As per figure 5 below, the rate per 100,000 population shows a decrease for Melbourne LGA since 2012. This is possibly because of increased policing and higher population density in the city.  Yarra LGA has a relatively higher crime rate per 100k population. Most crimes are occurring in Melbourne LGA (depicted by Green line) in North West police Region.  Figure 5:  ../Desktop/Screen%20Shot%202022-01-02%20at%2011.58.04%20pm.png   1. **Trend of Victim Reports**   Over last 10 years, we can see a surge in victim reports filed by people in 2016. This is consistent to our previous analysis as there were more crime offences happened during 2016. Then the number of victim reports dropped significantly from 2019 to 2021 which is due to Covid-19 lockdown (Please see figure 6).  Figure 6:  ../Desktop/Screen%20Shot%202022-01-01%20at%206.50.03%20pm.png  We also identified that the most victim reports were filed by the age group between 25 and 34 years old and followed by the age group between 35 and 44 years old, which are 25.2% and 20.5% respectively (Please see figure 7). Possibly because that these two groups of people have most professionals who are highly well educated and knows how to stand by themselves.  Figure 7:  ../Desktop/Screen%20Shot%202022-01-01%20at%206.50.15%20pm.png   1. **Trend of Affected Family Members** |
|  |
|  |
|  |

*A picture containing diagram

Description automatically generated*

* *Most Residential crimes are happening inside Homes.*
* *Streets are where most of the community crimes are occurring.*
* *Crimes in other locations are mostly in the Retail setting*

# Victoria Crimes Trends

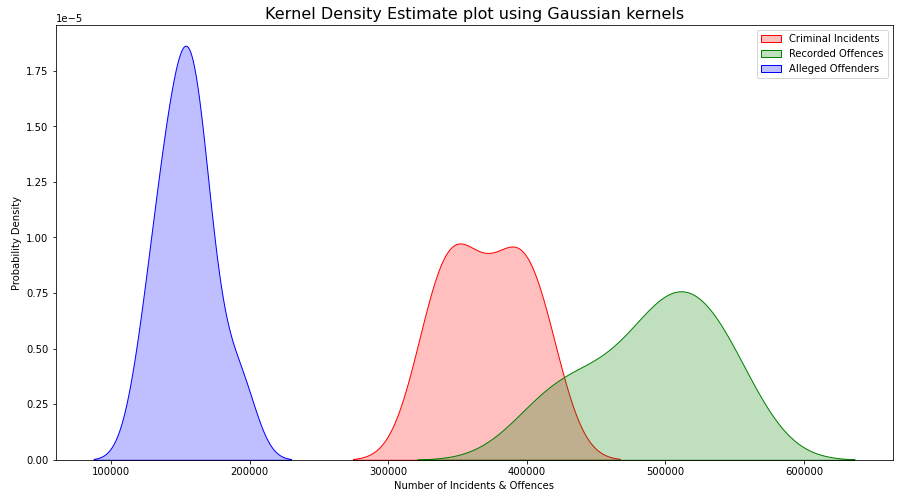
The study is for the period of September 2012 to September 2021. Various techniques and statistical calculations had been applied to ensure the findings are consistent and best interpreted the data.

* 1. **Kernel Density Estimation (KDE)**

In statistics, Kernel Density Estimation (KDE) is a non-parametric way to estimate the Probability Density Function (PDF) of a random variable.

This study applied Gaussian kernels with automatic bandwidth determination based on “scotts\_factor” formula: *n\*\*(-1./(d+4))*.

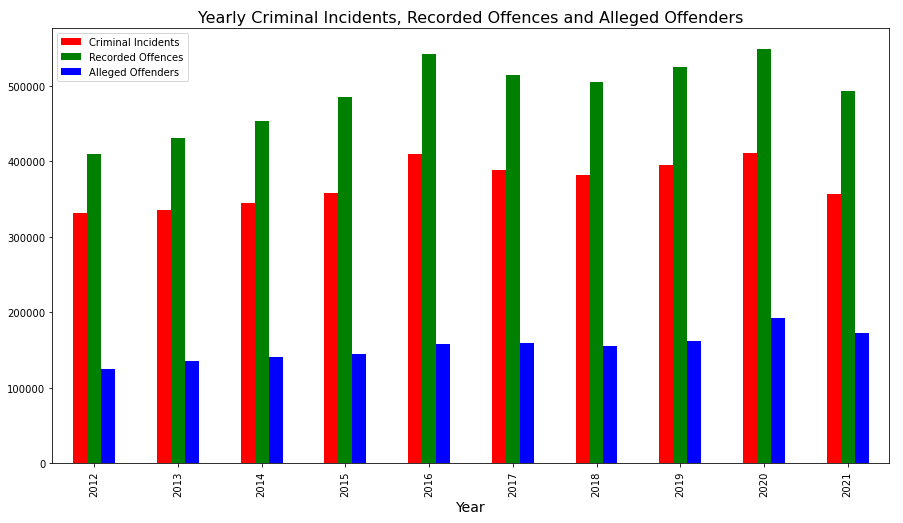
(D.W. Scott, “Multivariate Density Estimation: Theory, Practice, and Visualization”, John Wiley & Sons, New York, Chicester, 1992.)



The graph illustrates the probably density for Criminal Incidents, Recorded Offences and Alleged Offenders based on the ten years period.

The numbers of Alleged Offenders Incidents were sharply increased within the bell curve, while Criminal Incidents were steady increased and Recorded Offences were slowly increased with some outliers.

* 1. **Criminal Incidents, Recorded Offences and Alleged Offenders**



The graph above illustrates the yearly crime data for Criminal Incidents, Alleged Offenders and Recorded Offences.

The yearly data volume changes have been calculated in percentage and represent in the table below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Criminal Incidents | Recorded Offences | Alleged Offenders | Population |
| 2012 | (Baseline) | (Baseline) | (Baseline) | (Baseline) |
| 2013 | 0.012099 | 0.054790 | 0.090524 | 0.017587 |
| 2014 | 0.029842 | 0.052259 | 0.040208 | 0.024388 |
| 2015 | 0.039363 | 0.068144 | 0.028308 | 0.018456 |
| 2016 | 0.141908 | 0.116527 | 0.084495 | 0.021263 |
| 2017 | -0.051259 | -0.050605 | 0.008406 | 0.018048 |
| 2018 | -0.018528 | -0.018061 | -0.020907 | 0.029002 |
| 2019 | 0.035609 | 0.040722 | 0.043626 | 0.015371 |
| 2020 | 0.040523 | 0.044216 | 0.189915 | 0.021673 |
| 2021 | -0.131986 | -0.100593 | -0.105169 | -0.007842 |

According to the data, there is a significant increase in Criminal Incidents (0.14%) and Recorded Offences (0.12%) for year 2016. There is also a big jump in the number of Alleged Offenders incidents in year 2020 with 0.19%.

* 1. **Ten Years Overview Statistics**

The data had been calculate for its mean, standard deviation with minimum, maximum, first quartile, second quartile (median) and third quartile.

* The first quartile (*Q*1) is defined as the middle number between the smallest number (minimum) and the median of the data set. It is also known as the *lower* or *25th empirical* quartile, as 25% of the data is below this point.
* The second quartile (*Q*2) is the median of a data set; thus 50% of the data lies below this point.
* The third quartile (*Q*3) is the middle value between the median and the highest value (maximum) of the data set. It is known as the *upper* or *75th empirical* quartile, as 75% of the data lies below this point

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Criminal Incidents | Criminal Incidents  Rate | Recorded Offences | Recorded Offences  Rate | Alleged Offenders | Alleged Offenders  Rate | Total |
| count  (number of years) | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 |
| mean | 371,030.6 | 23,614.45 | 490,620.5 | 31,240.43 | 154,560.7 | 9,955.12 | 1,016,211.8 |
| standard deviation | 29,772.09 | 1,304.28 | 46,454.24 | 1,856.8 | 19,499.59 | 713.94 | 92,297.04 |
| min | 330,923.0 | 21,248.43 | 408,983.0 | 28,472.58 | 124,409.0 | 8,828.47 | 864,315.0 |
| 1st Quartile - 25% | 347,828.75 | 23,052.66 | 461,668.25 | 29,749.42 | 142,124.75 | 9,619.61 | 952,109.25 |
| 2nd Quartile - 50% | 369,846.0 | 23,453.34 | 498,991.0 | 31,448.79 | 156,385.5 | 9,827.37 | 1,031,891.5 |
| 3rd Quartile - 75% | 393,172.5 | 24,177.39 | 522,423.75 | 32,315.31 | 161,301.75 | 10,256.3 | 1,076,898.0 |
| max | 410,764.0 | 26,304.42 | 548,465.0 | 34,796.25 | 192,965.0 | 11,485.34 | 1,152,194.0 |

The findings based on the observed data demonstrate that:

* Year 2016 has the highest Crime Incidents (26,304 incidents) and Recorded Offences (34,796 offences) rate per 100,000 population.
* Year 2021 has lowest Crime Incidents (21,248.43) rate per 100,000 population.
  1. **Sdafafd**
  2. **dafdf**