

|  |
| --- |
| Data Analysis - Victoria Crimes  Statistics 2021 |
|  |
| January 4  Monash Data Analytics Bootcamp  Authored by: Kelvin Hung Nguyen  Anh Huong, Jacqueline Xia, Antoinette Boyle  Your Name |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **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.   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. **Analysis 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. **Region of Offences**   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 4:  **Figure 4:**  ../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 5:  **Figure 5:**  ../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. **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 criminal 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 well educated and knows how to stand by themselves.  Figure 7:  ../Desktop/Screen%20Shot%202022-01-01%20at%206.50.15%20pm.png   1. **Family Incidents**   The family members got affected from the family incidents have been increasing steadily over the past 5 years, and they’re mainly concentrated on the age groups of 25 – 34 years and 35 – 44 years (please see figure 8 and figure 9). According to ABC news, family violence offences has been increasing and the reason the number of victim reports have been increasing could result a greater awareness of family violence nowadays.  **Figure 8:**  ../Desktop/Screen%20Shot%202022-01-01%20at%206.50.30%20pm.png  **Figure 9:**  **../Desktop/Screen%20Shot%202022-01-01%20at%206.50.39%20pm.png**  Moreover, the number of both affected females and males are increasing over last 5 years, with the females outweigh the males significantly. If we divide it by region, the family incidents mostly happened in North West compared to other region.  **Figure 10:**  ../Desktop/Screen%20Shot%202022-01-01%20at%206.50.30%20pm.png  **Figure 11:**  **../Desktop/Screen%20Shot%202022-01-04%20at%202.44.41%20pm.png**   1. **Police service area - incident charge status**   Local government areas (LGAs) in Victoria, are sorted into four 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.  There are too many LGA’s, those must be allocated into seven administrative regions for ease of referencing.    For more simplicity and better service, the police further group LGA into only four regions     * 1. **The dataset**   There were three categories for charge status:   * Charges Laid, * No Charges Laid, * Unsolved   For purpose of understanding progress of a criminal case, criminal charges statuses are grouped into two main categories, solved or unsolved.  We took a snapshot of criminal incident unsolved charge status in 2021 and concluded the data is normal distributed and there are no outliners.  Figure \*:    On average, a LGA has unsolved charges rate of 41.5% with a standard deviation of 10.4%.  68% of LGA has unsolved charge percentage between 31.1% and 51.9%  95% of LGA has unsolved charge percentage between 20.7% and 62.3%  99.7% of LGA has unsolved charge percentage between 10.3% and 72.7%  Top vs bottom 5 LGA’s in 2021:   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Ranking | Local Government Area | Incidents Recorded | Total Incidents | Unsolved Status Percentage | | 1 | Swan Hill | 362 | 1810 | 20.00 | | 2 | Southern Grampians | 172 | 780 | 22.05 | | 3 | Ararat | 259 | 1134 | 22.84 | | 4 | Wangaratta | 395 | 1631 | 24.22 | | 5 | Northern Grampians | 233 | 879 | 26.51 | | ... | ... | ... | ... | ... | | 75 | Bayside | 2007 | 3435 | 58.43 | | 76 | Glen Eira | 3276 | 5440 | 60.22 | | 77 | Maribyrnong | 4180 | 6903 | 60.55 | | 78 | Boroondara | 3473 | 5535 | 62.75 | | 79 | Queenscliffe | 47 | 72 | 65.28 |  * 1. **Number of police stations and charge status**   In 2021, there’s a medium negative correlation between number of police station in each LGA and percentage of unsolved charge (r-value is -0.45). The more police station a LGA has, the less is charge status percentage. The relationship is depicted with equation y = -2.17 x + 50.56, whereas y is status percentage, x is number of police station in LGA    However, number of police stations only explain the how high is the charge status percentage by only 20.25% (r-square)  As per figure below, it shows that building more polices closer to the center metropolitan might improve the charge status.     * 1. **Population and charge status**   Population is another factor that influence the unsolved charge status. The more population a LGA has, the higher the percentage of unsolved charges (r value is 0.86)     * 1. **Time series analysis on unsolved charge**   LGAs dataset are grouped into four PSA regions for this analysis.  Population for North West metropolitan region was highest in term of growth rate, followed by Southern metropolitan region. Although Eastern region has higher population compared to Southern metropolitan region, the growth rate is more steadily    In or each region, unsolved charge percentage, however, showed decreased in trend despite the increase in population.  Especially for North West metropolitan region, the population growth rate was higher, but the unsolved charge percentage reduced at a little higher rate compared to other regions.  On the other hand, Southern metropolitan with lower population growth rate compared to North Western metropolitan, but the unsolved charge percentage reached to equivalent level in last three years.   Victoria Crimes Trends The study is for the period of September 2012 to September 2021. Various data mining, scrapping, cleansing, data visualisation techniques and statistical calculations had been applied to ensure the findings are consistent and best interpreted the government crime 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 visualised 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 crimes data for Criminal Incidents, Alleged Offenders and Recorded Offences.  The yearly data volume changes had been calculated in percentage as shown 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 |   The finding depicts that:   * Year 2016 had a significant increase in Criminal Incidents with 0.14% and Recorded Offences with 0.12% (round up to the nearest two digits). * Year 2020 also had a big jump in the number of Alleged Offenders incidents with 0.19%.   1. **Ten Years Overview Crime Statistics**   Crimes data had been calculate for its mean, standard deviation with minimum, maximum, first quartile, second quartile (median) and third quartile.   * The first quartile (Q1) 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 (Q2) is the median of a data set; thus 50% of the data lies below this point. * The third quartile (Q3) 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 finding above demonstrates:   * 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. **Visualised Victoria Crime Trends by Police Service Region**   Several bar graphs had been drawn for each type of crime to visualise the number of crime incidents and offences occurred in the last ten years, with their associate rates per 100,000 population that recorded by the local’s police service region.  The four police service regions are: North West Metropolitan (blue), Eastern (green), Southern Metropolitan (orange) and Western (red).            The trend lines of the crime rate per 100,000 population for all crimes are decreasing between 2020 to 2021, except Family Incidents. This could be due to various local community events as highlighted in the previous sections of this report.   * 1. **Yearly Changes of Crime by Police Service Region**   The table below illustrates the percentage of changes of the number of incidents and offences that occurred in the four police regions.   |  |  |  |  | | --- | --- | --- | --- | | Types | Police Service Region | Highest In Year | Highest Yearly Percentage Change | | Criminal Incidents | Southern Metro | 2013, 2015, 2020 | 0.05, 0.07, 0.06 | | East | 2014, 2017, 2019 | 0.05, -0.01, 0.08 | | Recorded Offence | West | 2013, 2016, 2017, 2021 | 0.09, 0.13, -0.02, -0.06 | | Alleged Offender Incidents | Southern Metro | 2013, 2015, 2020 | 0.13, 0.42, 0.26 | | Victim Reports | East | 2013, 2014, 2017, 2019 | 0.02, 0.05, -0.04, 0.08 | | Family Incidents | North West Metro | 2018 | 0.08 | | East | 2019 | 0.09 | | Southern Metro | 2020 | 0.13 | | West | 2021 | 0.07 |   It is observed that:   * Year 2019, *East* region is taking the lead in increasing in Crime Incidents (0.06), Victim Reports (0.08) and Family Incidents (0.09) compares to other Police Service regions. * Year 2020, *Southern Metropolitan* region is having the highest increase in the Crime Incidents (0.06), Alleged Offender Incidents (0.26) and Family Incidents (0.13) compares to other Police Service regions.      * 1. **Victoria Crimes Data Maps**   Interactive maps had been produced to show the ten years average of Victoria crimes volume based on the local government boundaries with the utilization of the python geopanda library and geometry data from the government website. Summary average data of crimes will be displayed when a colored section of the map is selected. The maps had been created by using Jupyter notebooks, which can be located on [github](https://github.com/antoinetteboyle/Project-1).  Below are the snapshot samples of the map:    Each local government area (sometimes referred as Suburb or LGA) is coloured code. The ten year average volume of crimes is matching with the colour code, the higher crime rate is the lighter in colour. The colour is scaling from dark purple to green and then bright light yellow.  Another map was also produced for the crimes that grouped by police region. The map’s colour code is for the police service region rather than the density of crimes.   1. **Summary**   Overall, there is a spike in Criminal Incidents in 2016, 2017 because of multiple factors, including the population increase around 2015, Moomba festival incidents and Bourke Street rampage. Family incidents have been increasing which could due to a greater awareness of family violence. The findings also depict that there is a correlation between the number of police stations and solved crimes. Furthermore, if there is any open data that publicity available about of the victims relating the each type of crimes (in sub categories), then more data analysis can be performed and and further suggestions cane be made for the future resolution.  **References:**  Australian Government, 2021. *VIC DELWP – Vicmap Features of Interest – Police Stations (Points).* *.* Accessed 02 January 2022. available from <https://data.gov.au/dataset/ds-aurin-aurin%3Adatasource-VIC_Govt_DELWP-VIC_Govt_DELWP_datavic_VMFEAT_POLICE_STATION/distribution/dist-aurin-aurin%3Adatasource-VIC_Govt_DELWP-VIC_Govt_DELWP_datavic_VMFEAT_POLICE_STATION-0/details?q>  Australian Government, 2021. *VIC Local Government Areas - PSMA Administrative Boundaries GeoJSON.* Accessed 02 January 2022. available from <https://data.gov.au/dataset/ds-dga-bdf92691-c6fe-42b9-a0e2-a4cd716fa811/distribution/dist-dga-ce0a0ed3-6003-47fd-88ad-4b49d9337d47/details?q>  Australian Government, 2021. *VIC Local Government Areas - Geoscape Administrative Boundaries.* Accessed 02 January 2022. available from <https://data.gov.au/dataset/ds-dga-bdf92691-c6fe-42b9-a0e2-a4cd716fa811/details>  Crime Statistics Agency (2021). *Latest Victorian Crime Data.* Accessed 02 January 2022. available from <https://www.crimestatistics.vic.gov.au/crime-statistics/latest-victorian-crime-data/download-data>  Karen, P, 2016. *Victoria's crime rate spikes as Opposition warns of a 'crime tsunami'.*  ABC News. Accessed 02 January 2022. available from <https://www.abc.net.au/news/2016-09-29/victorian-crime-rate-spikes-as-opposition-warns-of-crime-tsunami/7890832> |
|  |
|  |
|  |