ISYS5050 Knowledge Management

Final Report

Date: 29/05/2019

Group 3

NSW CRIME REPORT (1995-2017)

Team Members:

Dai, Wei - 470248191

Hu, Xinyu - 470475340

Gao, Fan - 480099589

Ning, Xin - 460346559

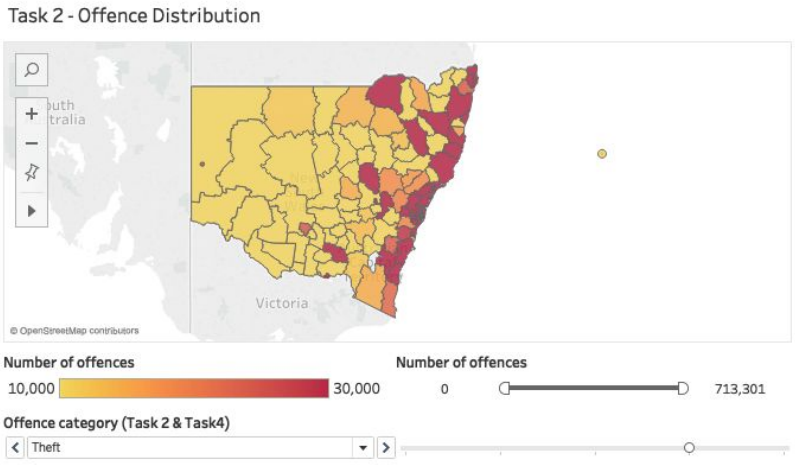
Wang, Yili - 470106260

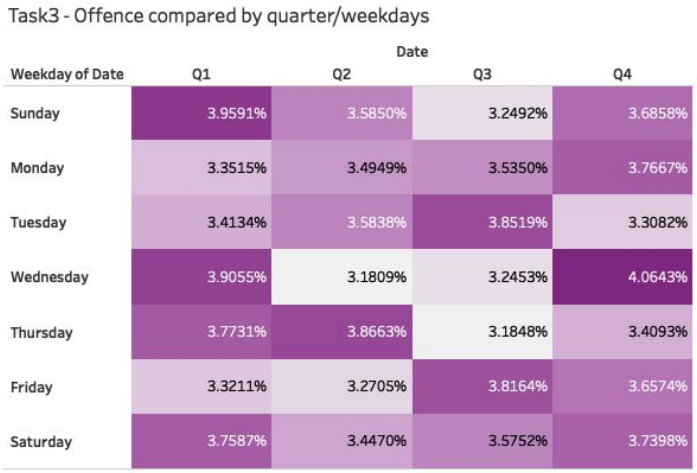
TABLE OF CONTENTS

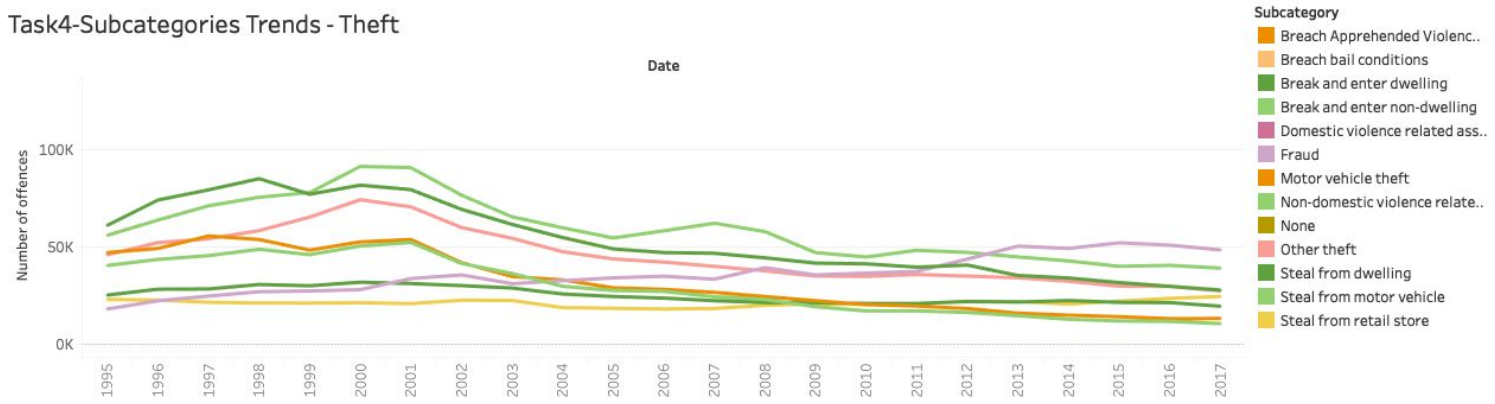
1. INTRODUCTION

This project is aiming to analyzing data from several aspects based on the monthly criminal incidents recorded by the NSW police of the period between 1995 and 2017. In general, there are four major parts in the project. Firstly, to better manipulate the data, data pre-processing transforms the previous data format for OLAP operations. For the next part, summary data is presented from a regional perspective according to the total number of different types of offences committed. Moreover, in the third part, the relationship of offence patterns with time and season is explored. Lastly, further research is applied to predict the crime situation tendency.

In addition, there is an interactive dashboard generated for the project. As the dashboard screenshot shown below, the heatmap of task two is on the top left, which can be filtered by the range of offences numbers and the offence category options. The upper right of the dashboard shows the task three's crime patterns from the perspective of weekdays and seasons. At the bottom, the tendency of offences number for different subcategories over time based on a chosen offence category is presented for task four. It is worth mentioning that for different offence category options, the dashboard of task two and task four will change synchronously.







Based on the introduction above, the dashboard can be applied by the following link: https://public.tableau.com/profile/david3945#!/vizhome/Task3_15574060828310/Final-Dashboard?publish=yes

2. DATA PRE-PROCESSING

The original format of raw data consists of three dimensions ("LGA", "Offence category", "Subcategory"). In order to analyze the "Number of offences" based on the "Date" dimension, the ETL process requires to implement in the first step to create the snowflake schema for adding the date dimension shown as figure 2.1.

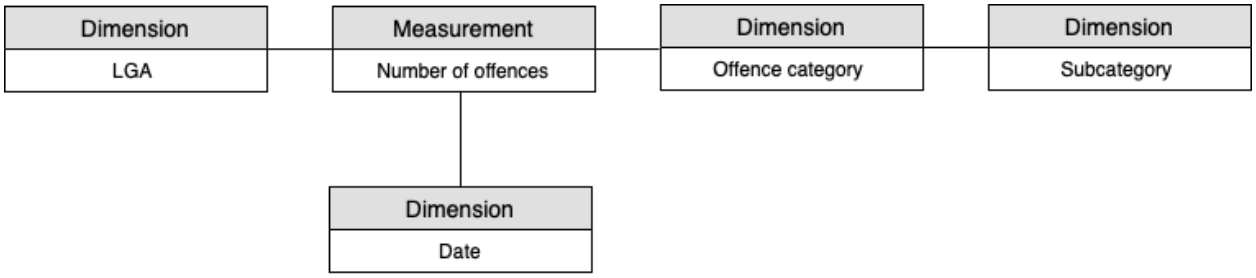
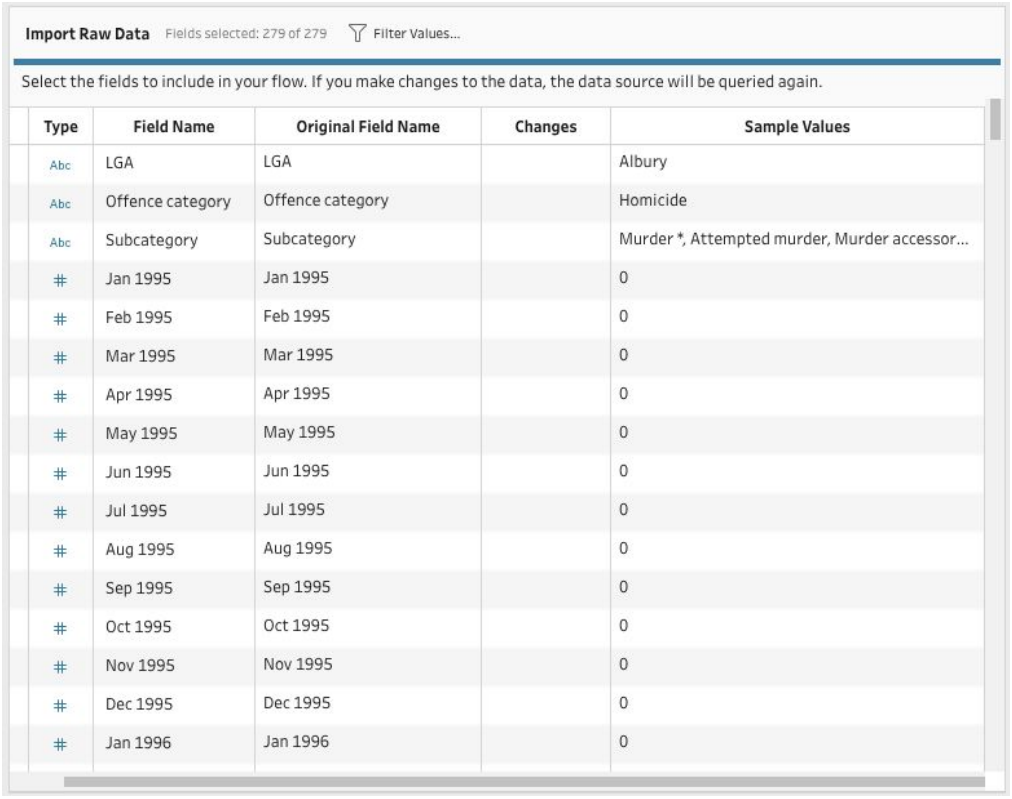


Figure 2.1 Snowflake Schema of NSW Crime

The tool utilized to preprocess data is the Tableau Prep Builder which empowering the ability to combine, shape and clean data with the non-technical staff. (Tableau Prep, 2019) There are six steps within the ETL process as below (Figure 2.2):

(1) Extract/download the raw data from the Canvas
(2) Import the raw data into the Tableau Prep Builder (Figure 2.3)
(3) Check and clean the null value within the different dimensions, only the subcategory has the null value, replaced the null value by the “Others”.
(4) Pivot the separated date columns into the rows below the *Date* dimension, the value corresponding to date is Measurement (*Number of offences*).
(5) Export the data as the CSV file (Figure 2.4)
(6) Import the CSV file into the Tableau Public to analyze

Figure 2.2 Workflow in Tableau Prep Builder

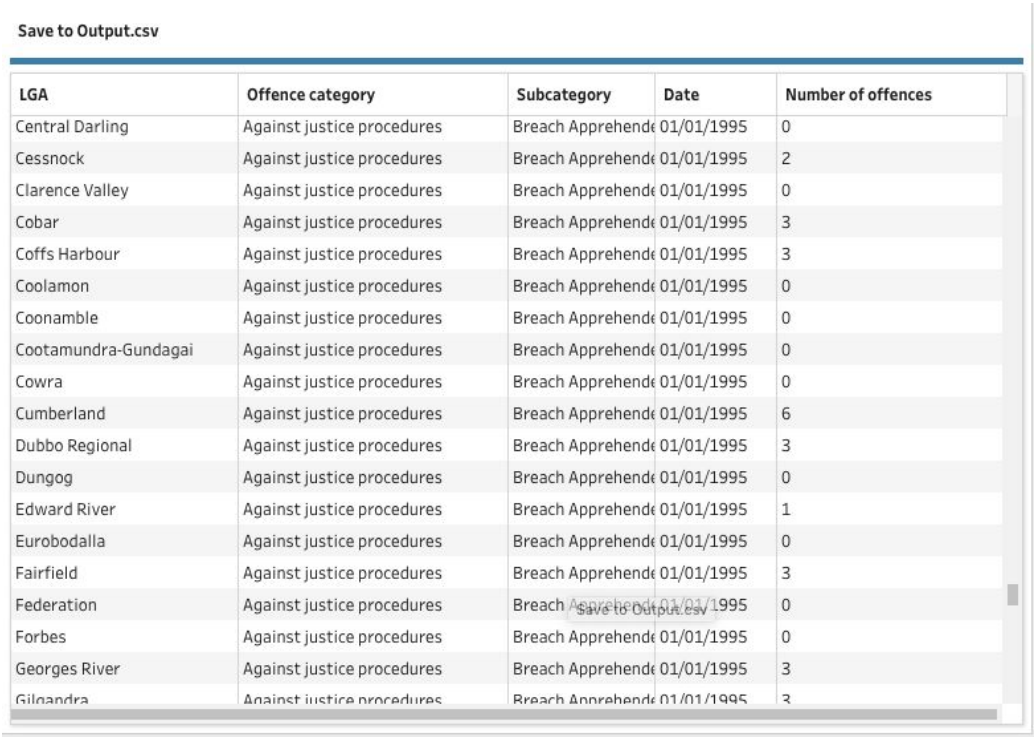


Import Raw Data Fields selected: 279 of 279 Filter Values...

Select the fields to include in your flow. If you make changes to the data, the data source will be queried again.

Type	Field Name	Original Field Name	Changes	Sample Values
Abc	LGA	LGA		Albury
Abc	Offence category	Offence category		Homicide
Abc	Subcategory	Subcategory		Murder *, Attempted murder, Murder accessor...
#	Jan 1995	Jan 1995		0
#	Feb 1995	Feb 1995		0
#	Mar 1995	Mar 1995		0
#	Apr 1995	Apr 1995		0
#	May 1995	May 1995		0
#	Jun 1995	Jun 1995		0
#	Jul 1995	Jul 1995		0
#	Aug 1995	Aug 1995		0
#	Sep 1995	Sep 1995		0
#	Oct 1995	Oct 1995		0
#	Nov 1995	Nov 1995		0
#	Dec 1995	Dec 1995		0
#	Jan 1996	Jan 1996		0

Figure 2.3 NSW Crime Data before Transformation



Save to Output.csv

LGA	Offence category	Subcategory	Date	Number of offences
Central Darling	Against justice procedures	Breach Apprehende	01/01/1995	0
Cessnock	Against justice procedures	Breach Apprehende	01/01/1995	2
Clarence Valley	Against justice procedures	Breach Apprehende	01/01/1995	0
Cobar	Against justice procedures	Breach Apprehende	01/01/1995	3
Coffs Harbour	Against justice procedures	Breach Apprehende	01/01/1995	3
Coolamon	Against justice procedures	Breach Apprehende	01/01/1995	0
Coonamble	Against justice procedures	Breach Apprehende	01/01/1995	0
Cootamundra-Gundagai	Against justice procedures	Breach Apprehende	01/01/1995	0
Cowra	Against justice procedures	Breach Apprehende	01/01/1995	0
Cumberland	Against justice procedures	Breach Apprehende	01/01/1995	6
Dubbo Regional	Against justice procedures	Breach Apprehende	01/01/1995	3
Dungog	Against justice procedures	Breach Apprehende	01/01/1995	0
Edward River	Against justice procedures	Breach Apprehende	01/01/1995	1
Eurobodalla	Against justice procedures	Breach Apprehende	01/01/1995	0
Fairfield	Against justice procedures	Breach Apprehende	01/01/1995	3
Federation	Against justice procedures	Breach Apprehende	01/01/1995	0
Forbes	Against justice procedures	Breach Apprehende	01/01/1995	0
Georges River	Against justice procedures	Breach Apprehende	01/01/1995	3
Gilgandra	Against justice procedures	Breach Apprehende	01/01/1995	3

Figure 2.4 NSW Crime Data after Transformation

3. SUMMARY OF SUBURBS CRIME OFFENCES

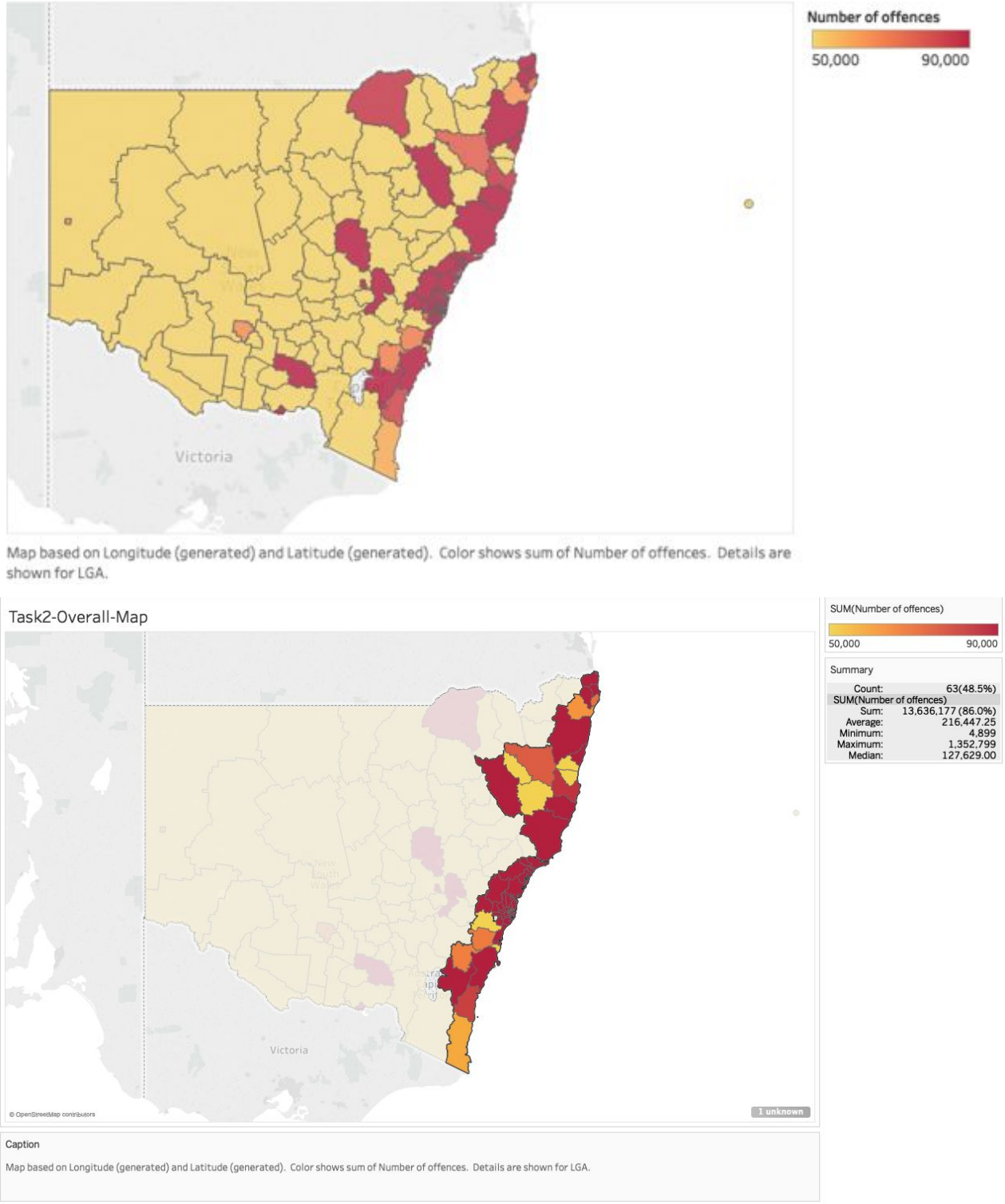


Figure 3.1 Offences Distribution Map

The high crime offences areas mainly aggregate at the eastern coastal areas of NSW. According to Figure 3.1, ***48.5%*** of NSW suburbs located in the eastern coastal areas contributes ***86.0%*** of the total number of offences.

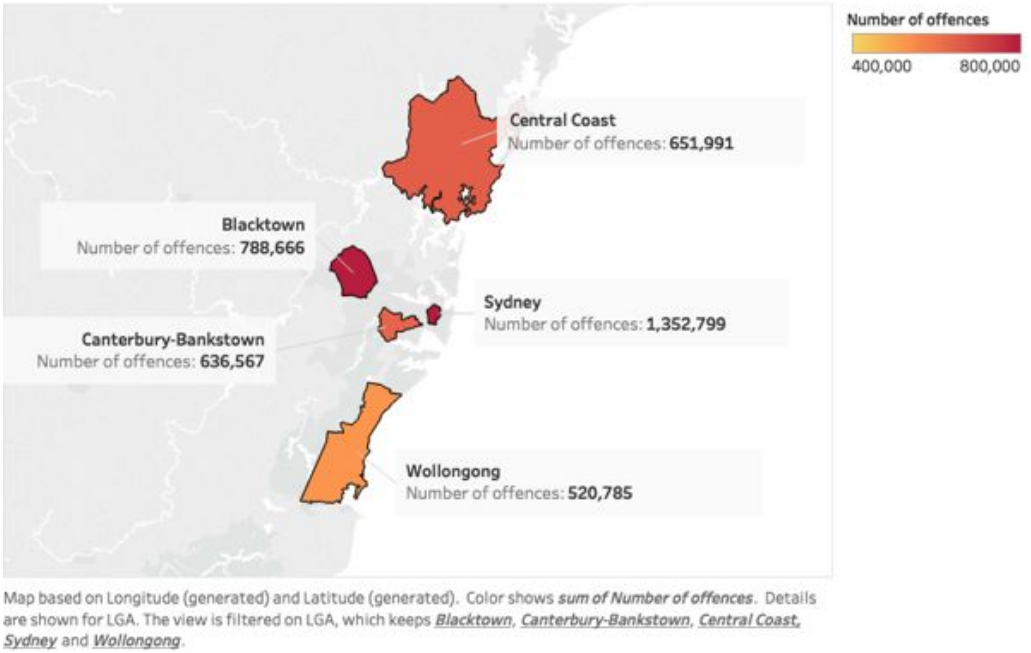


Figure 3.2 Map of Total Number of Top Five Suburbs

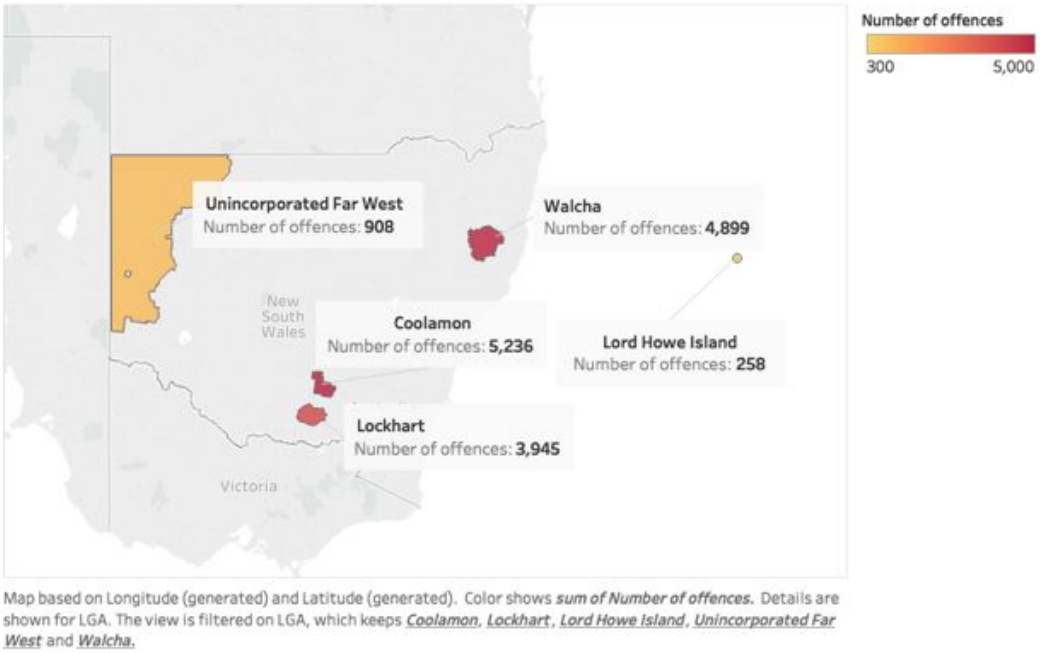


Figure 3.3 Map of Total Number of Bottom Five Suburbs

On the one hand, dig out the top five and bottom five suburbs in terms of the total number of offences. Figure 3.2 and 3.3 outlines the suburb with the highest crime number is ***Sydney (1,352,799)*** and the suburb with the lowest crime number is ***Lord Howe Island (258)***. The other four top crime suburbs are ***Blacktown (788,666), Central Coast (651,991), Canterbury-Bankstown (636,567), Wollongong (520,785)*** in descending order. The other four bottom crime suburbs are ***Unincorporated Far West (908), Lockhart (3945), Walcha (4,899), Coolamon (5,236)*** in ascending order.

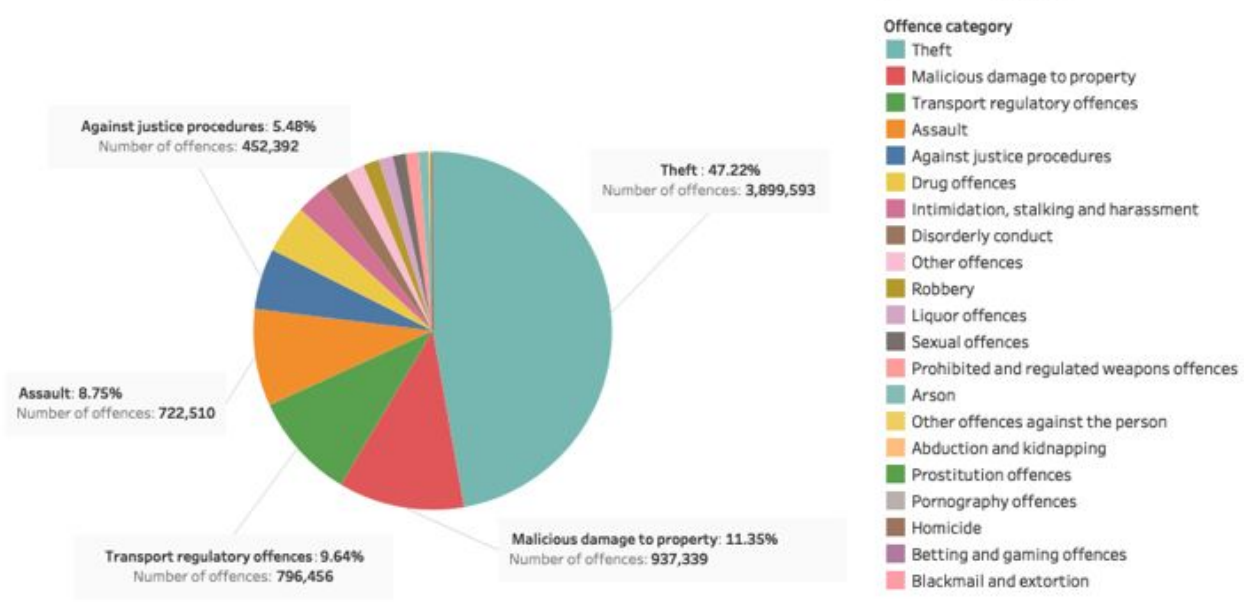


Figure 3.4 Pie of Offence Category Distribution

On the other hand, figure 3.4 illustrates the top five offence categories which are ***Theft*** (***47.22%***), ***Malicious damage to property (11.35%)***, ***Transport regulatory offences (9.64%)***, ***Assault (8.75%), and Against justice procedures (5.48%)*** respectively.

The detail of the top five suburbs in terms of the total number of top five offense categories shown as Figure 3.5.

In terms of LGA, the largest number of offence category is ***Theft*** which contributes over ***49.15%*** of crime number of total offence number for both individual suburbs based on Figure 3.6.

In terms of different offence category, with the perspective of Theft, Assault, Against justice procedures, the area with the largest number of offence is ***Sydney*** (***37.83% - Theft***, ***31.08% - Assault, 31.36% - Against justice procedures)***. In terms of Transport regulatory offences, the area with the largest crime number is ***Blacktown (29.85%)***. For the Malicious damage to

property, the suburb with the largest offence number is ***Central Coast (24.01%).*** Shown in Figure 3.7.

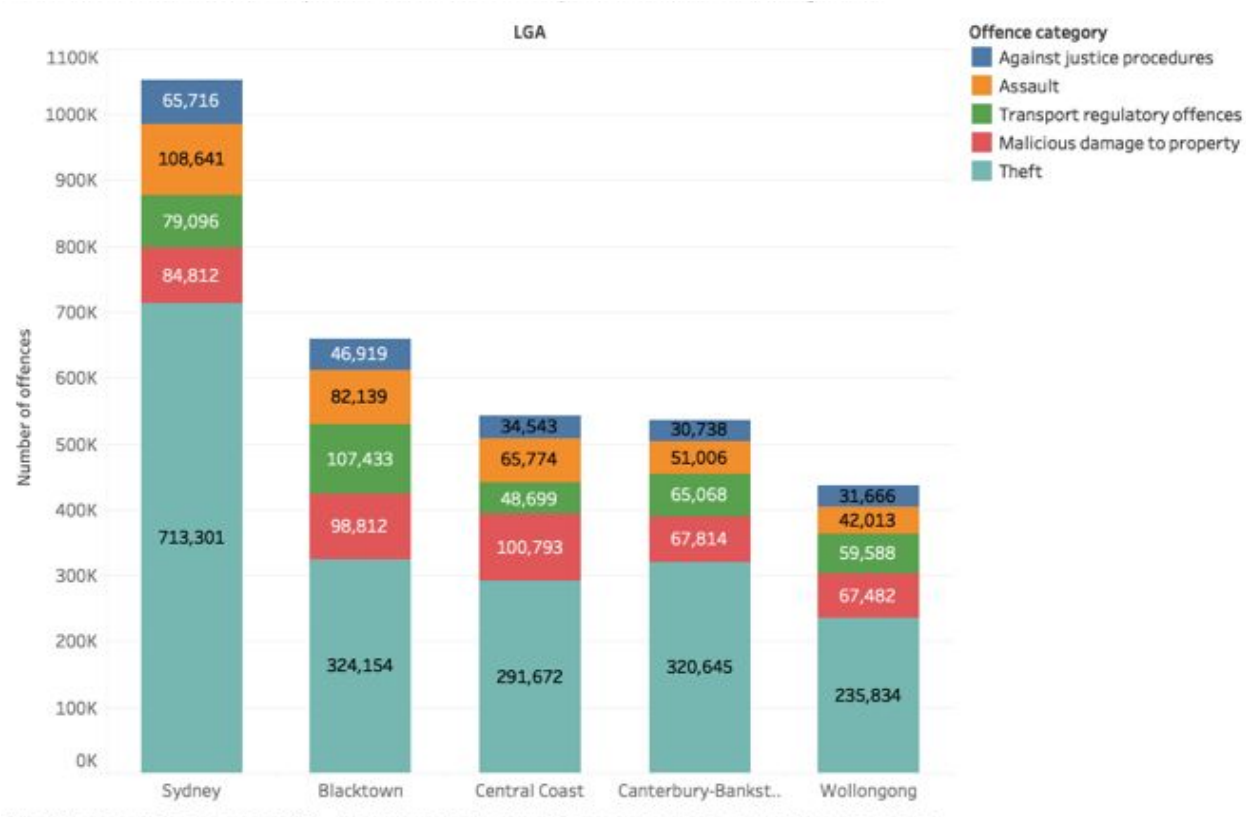


Figure 3.5 Bar of Total Number of Top Five Suburbs of Top Five Offence Categories

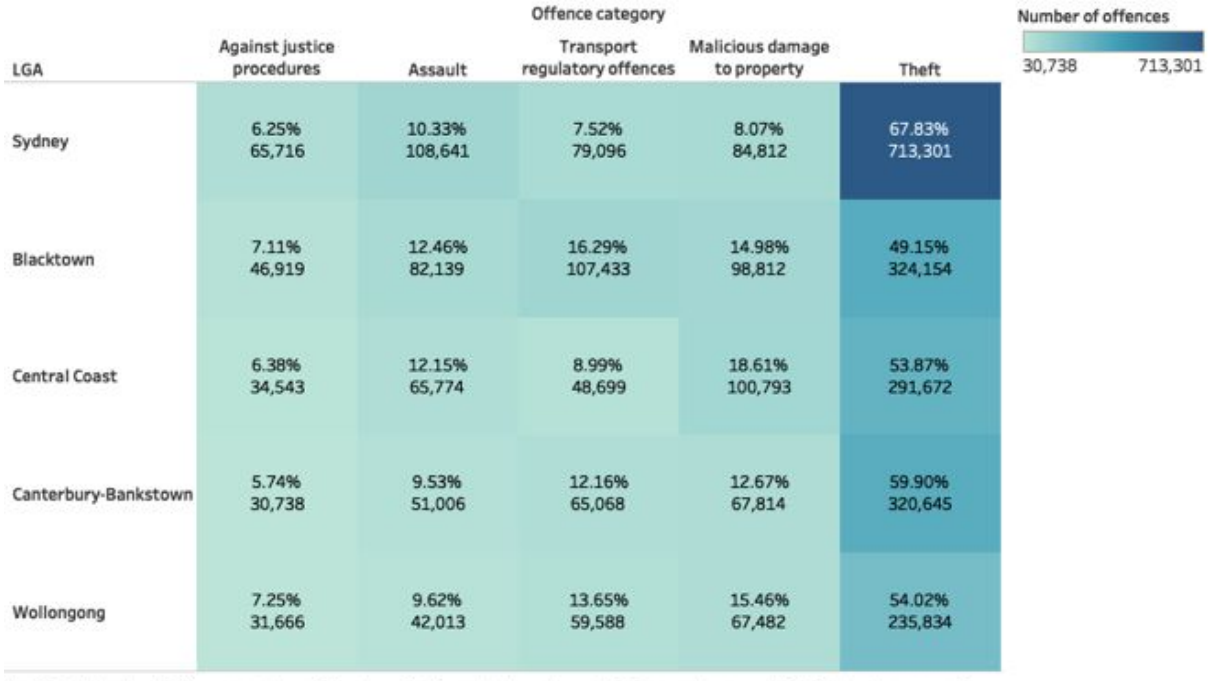
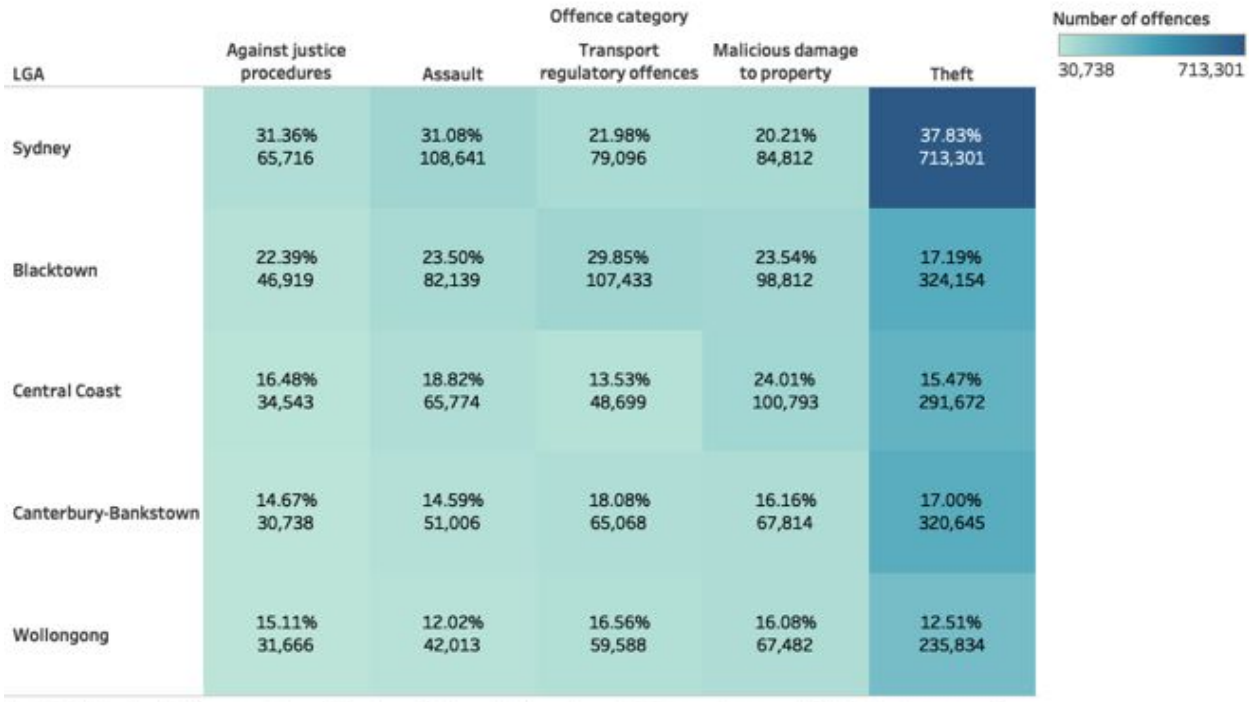


Figure 3.6 Offence Category Distribution in LGA for Top Five



LGA	Against justice procedures	Assault	Transport regulatory offences	Malicious damage to property	Theft
Sydney	31.36% 65,716	31.08% 108,641	21.98% 79,096	20.21% 84,812	37.83% 713,301
Blacktown	22.39% 46,919	23.50% 82,139	29.85% 107,433	23.54% 98,812	17.19% 324,154
Central Coast	16.48% 34,543	18.82% 65,774	13.53% 48,699	24.01% 100,793	15.47% 291,672
Canterbury-Bankstown	14.67% 30,738	14.59% 51,006	18.08% 65,068	16.16% 67,814	17.00% 320,645
Wollongong	15.11% 31,666	12.02% 42,013	16.56% 59,588	16.08% 67,482	12.51% 235,834

Figure 3.7 LGA Distribution in Offence Category for Top Five

The detail of the bottom five suburbs in terms of the total number of top five offense categories shown as Figure 3.8.

In terms of LGA, the largest number of offence category is also ***Theft*** which contributes over ***49.49%*** of crime number of total offence number for both individual suburbs based on Figure 3.9.

In terms of different offence category, with the perspective of Theft, Malicious damage to property, the area with the largest number of offence is ***Coolamon (35.75% - Theft, 38.49% - Malicious damage to property)***.with the perspective of Assault and Against justice procedure, the suburb with the largest crime number is ***Walcha (36.71 - Assault, 48.58% - Against justice procedure).*** As for the transport regulatory offences, the areas with the largest offence number are ***Walcha*** and ***Lockhart (38.46%)***. Shown in Figure 3.10.

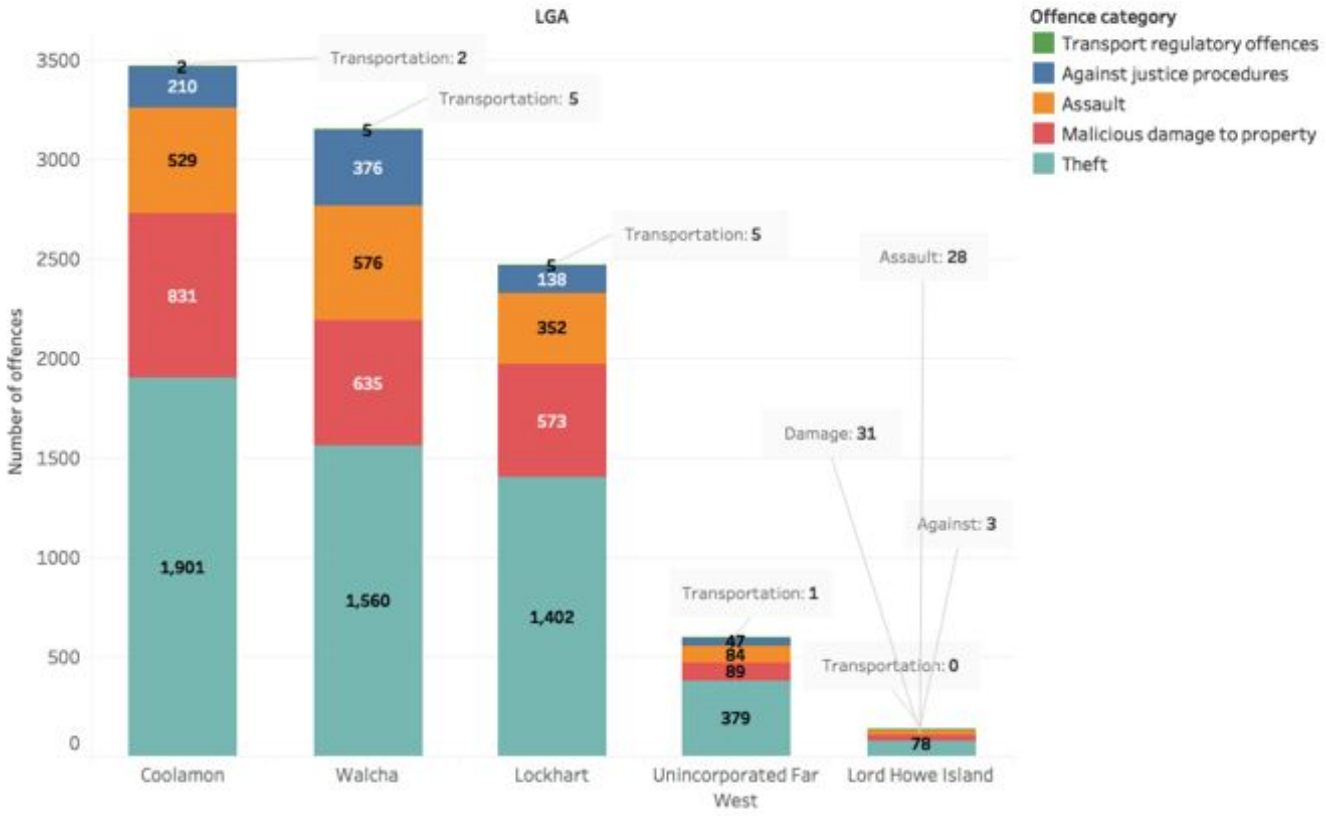


Figure 3.8 Bar of Total Number of Bottom Five Suburbs of Top Five Offence Categories

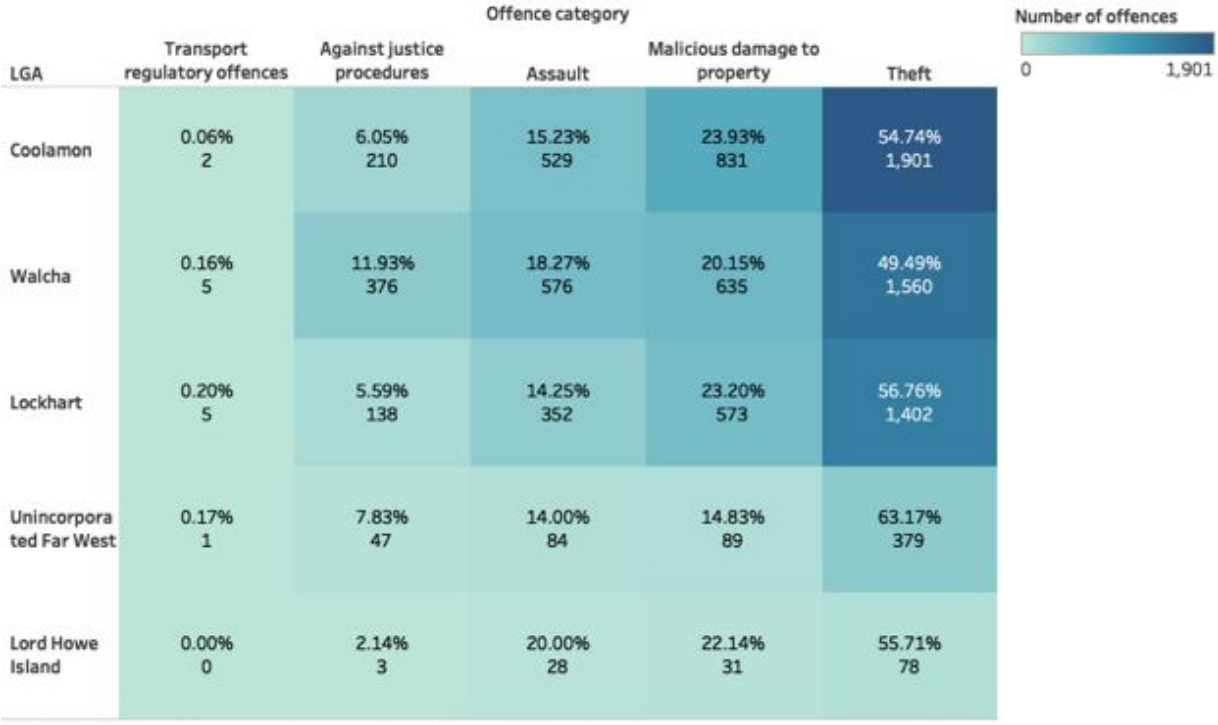


Figure 3.9 Offence Category Distribution in LGA for Bottom Five

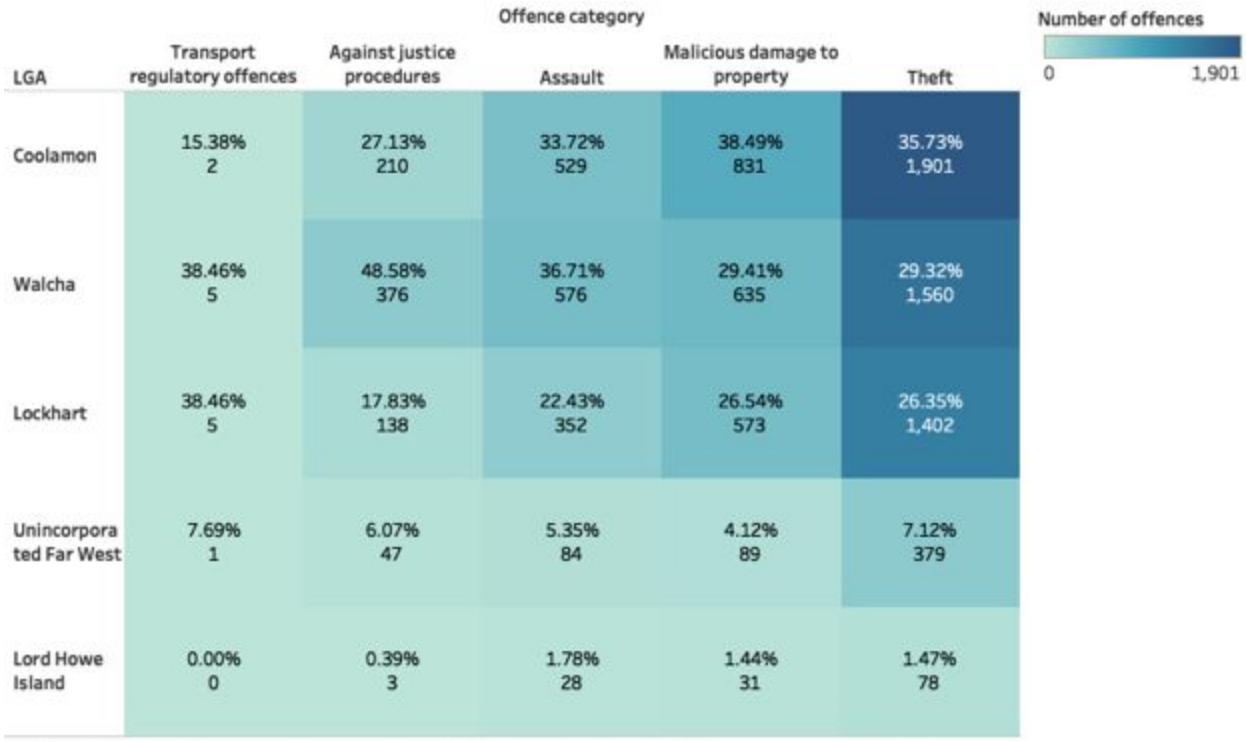


Figure 3.10 LGA Distribution in Offence Category for Bottom Five

To sum up, comparing the top five suburbs with the bottom five suburbs, the theft is the most serious crime for both two groups. In terms of other offense categories, the number and percentage of offense for the two groups has a significant difference.

With perspective of offence category, In terms of the top five suburbs, the suburb of the largest offence number of *Against justice procedure, Assault,* and *Theft* is the *Sydney.* The suburbs of the largest offence number of *Transport regulatory offences* and *Malicious damage to property* are *Blacktown* and *Central Coast* respectively.

In terms of the bottom five suburbs, the suburb of the largest offence number of *against justice procedure,* and *Assault* is *Walcha.* The suburbs of the largest offence number of *Transport regulatory offences* are *Walcha* and *Lockhart.* Moreover, the largest offence number of *Malicious damage to property* and *Theft* is *Coolamon.*

4. SUMMARY OF OFFENCE SEASONAL PATTERNS

SUM of offences compares by quarter/weekdays

Weekday of Date	Date: Q1	Q2	Q3	Q4
Sunday	3.9591%	3.5850%	3.2492%	3.6858%
Monday	3.3515%	3.4949%	3.5350%	3.7667%
Tuesday	3.4134%	3.5838%	3.8519%	3.3082%
Wednesday	3.9055%	3.1809%	3.2453%	4.0643%
Thursday	3.7731%	3.8663%	3.1848%	3.4093%
Friday	3.3211%	3.2705%	3.8164%	3.6574%
Saturday	3.7587%	3.4470%	3.5752%	3.7398%

Summary	
Count:	28
% of Total SUM(Number of offences)	
Sum:	100.0000%
Average:	3.5714%
Minimum:	3.1809%
Maximum:	4.0643%
Median:	3.5795%
SUM(Number of offences)	
Sum:	15,853,736
Average:	566,204.86
Minimum:	504,298
Maximum:	644,350
Median:	567,487.00

Table 4.1 Number of offences breakdown by quarter and weekdays

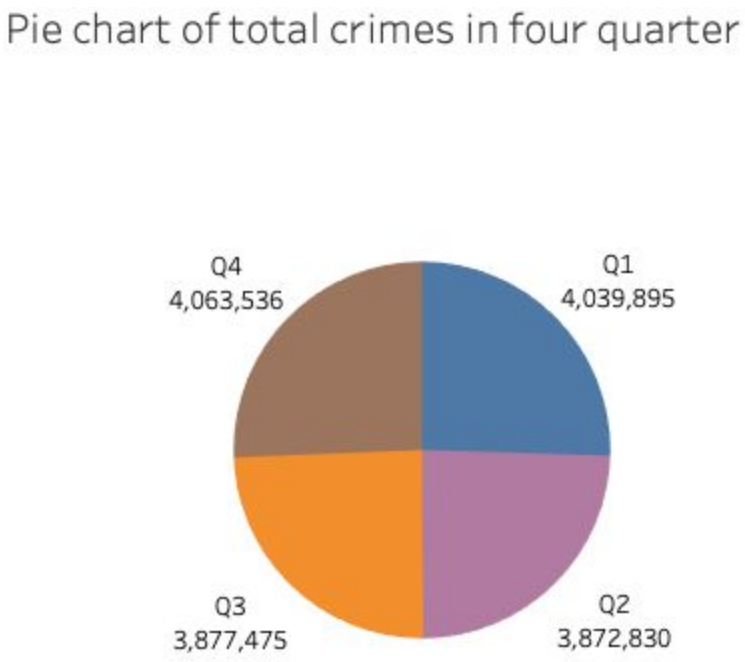


Figure 4.2 quarter breakdown of offence amount of total

Total Number of offences broken down by Date **Quarter** vs. Date **Weekday**. Color shows sum of number of offences from deep to shallow representing the quantitative. The maximum percentage (4.0643% of total) and the minimum (3.1809% of total) both shows on the **wednesday**. The crime quantity has relatively **no relationship with weekdays or weekends**.

The similar trend could be seen on the third maximum and minimum percentage(3.9055% and 3.2453%) . The top 3 and bottom 3 distribute in the Q1/Q4 and Q2/Q3 respectively, not hard to conclude, **more crime happens in Q1/Q4 than Q2/Q3** .

In addition, saturday and sunday show **equally division** of the week, taking up (28.9998% of total). It appears no relationship of crime between holidays or weekdays.

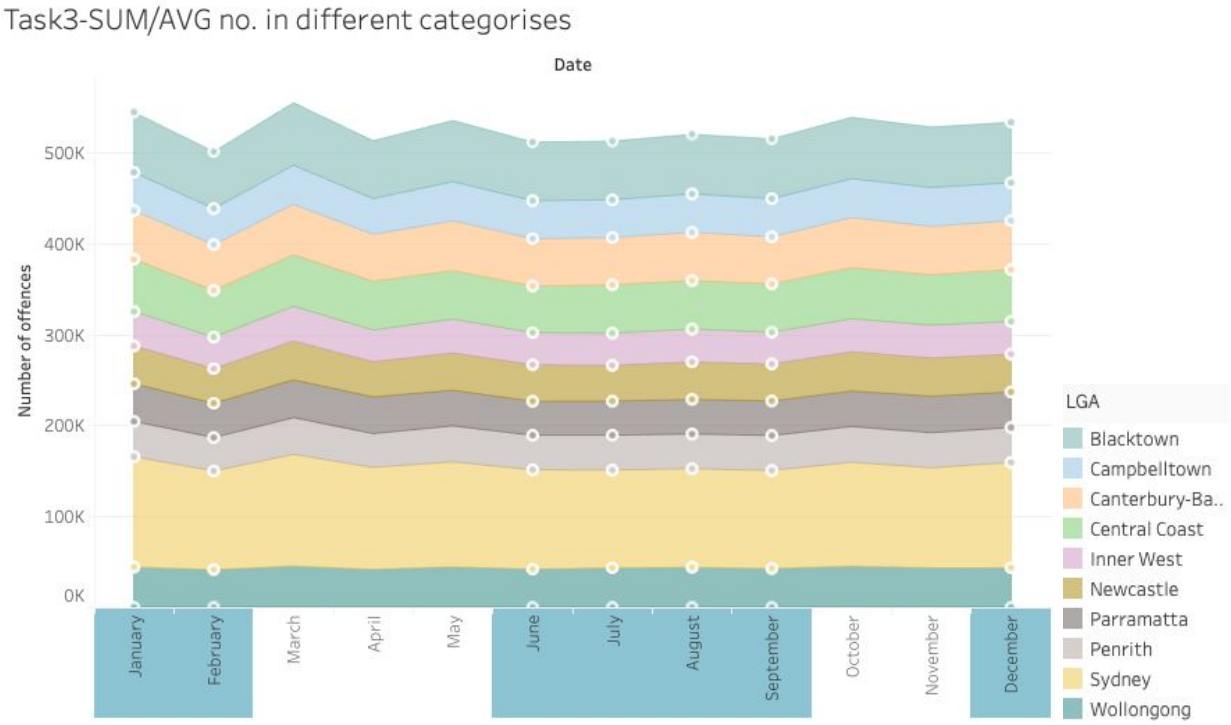


Figure 4.3 Top 10 areas breakdown by month of offence number

Sum of Number of offences for each month. Color shows details about LGA. The view is filtered on LGA, which keeps top 10 areas.

From the figure, June to September, the **winter** of Australia, showing the relatively low crime possibility of the year. The **lowest point** of the line chart could be found in winter, keeping **steady continuously** till September.
In the **summer**, December to march, the **biggest fluctuation** appears consisted by both highest point and second lowest point.

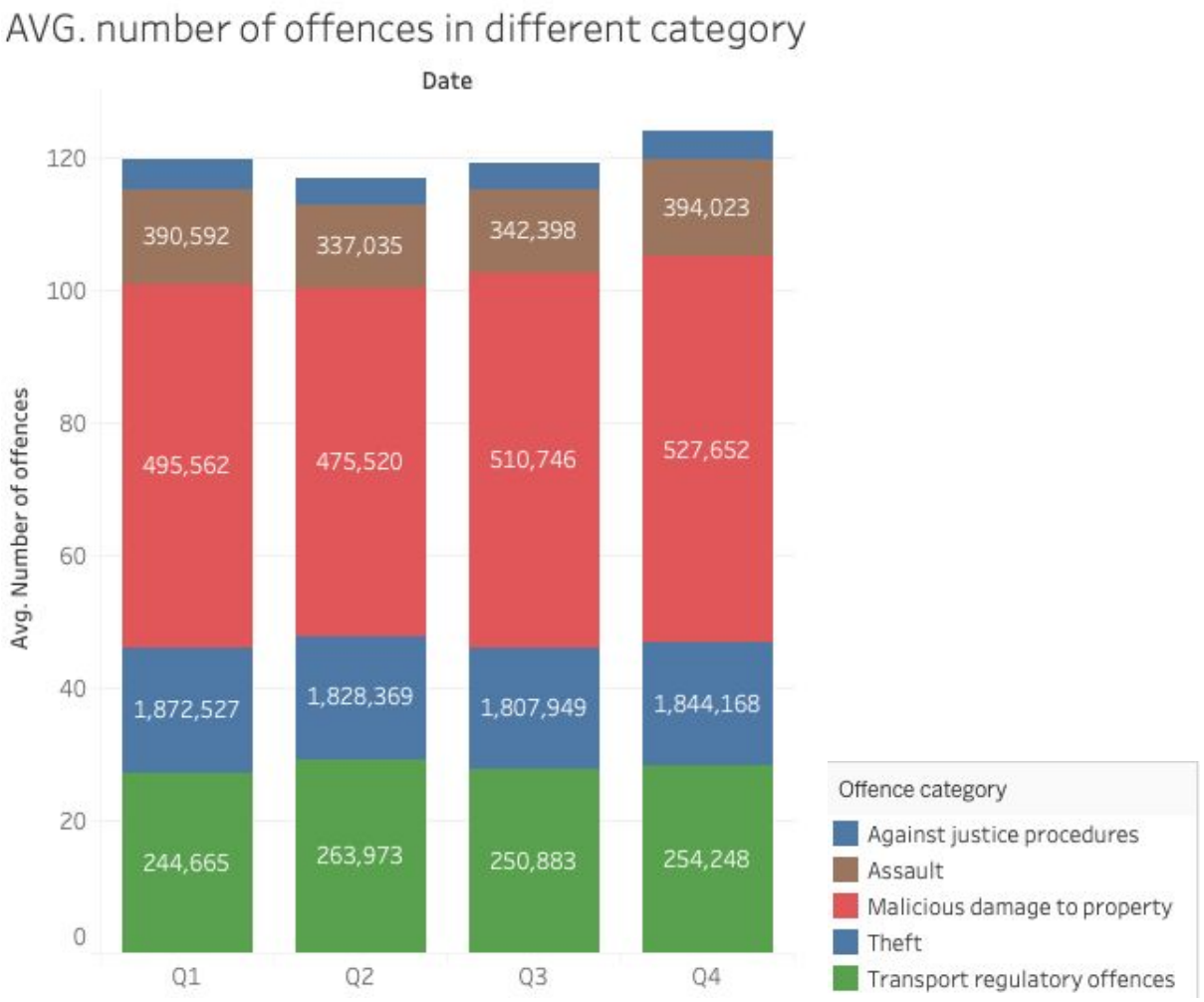


Figure 4.4 Top 5 offence categories breakdown by quarter

Average of number of offences for each quarter. Color shows details about offence category. The view is filtered to top 5 on Offence category, which keeps Assault, Intimidation, stalking and harassment, Malicious damage to property, Theft and Transport regulatory offences.

Malicious damage to property takes up the largest percentage of top 5 crime classifications, following the same distribution of yearly division in TASK 2. Not hard getting the offence category has little relationship with quarter.

To summary, cold days are safer than others. Crimes could happens any days of the week no matter weekend or weekdays. There is not specific type crime happens in one quarter more than others.

5. SUMMARY OF CRIME TRENDS

According to the requirement of Dr.Ian Oppermann, which is mining the data tendency of crime in the NSW and reasons of that, we choose crime dimension as Y-axis and time dimension as X-axis to show the overview of crime changing through time. The tendency line is showing that crime in NSW is increase before 2007 and then decrease. And prediction by Tableau shows that crime in the future would also keep this downtrend.

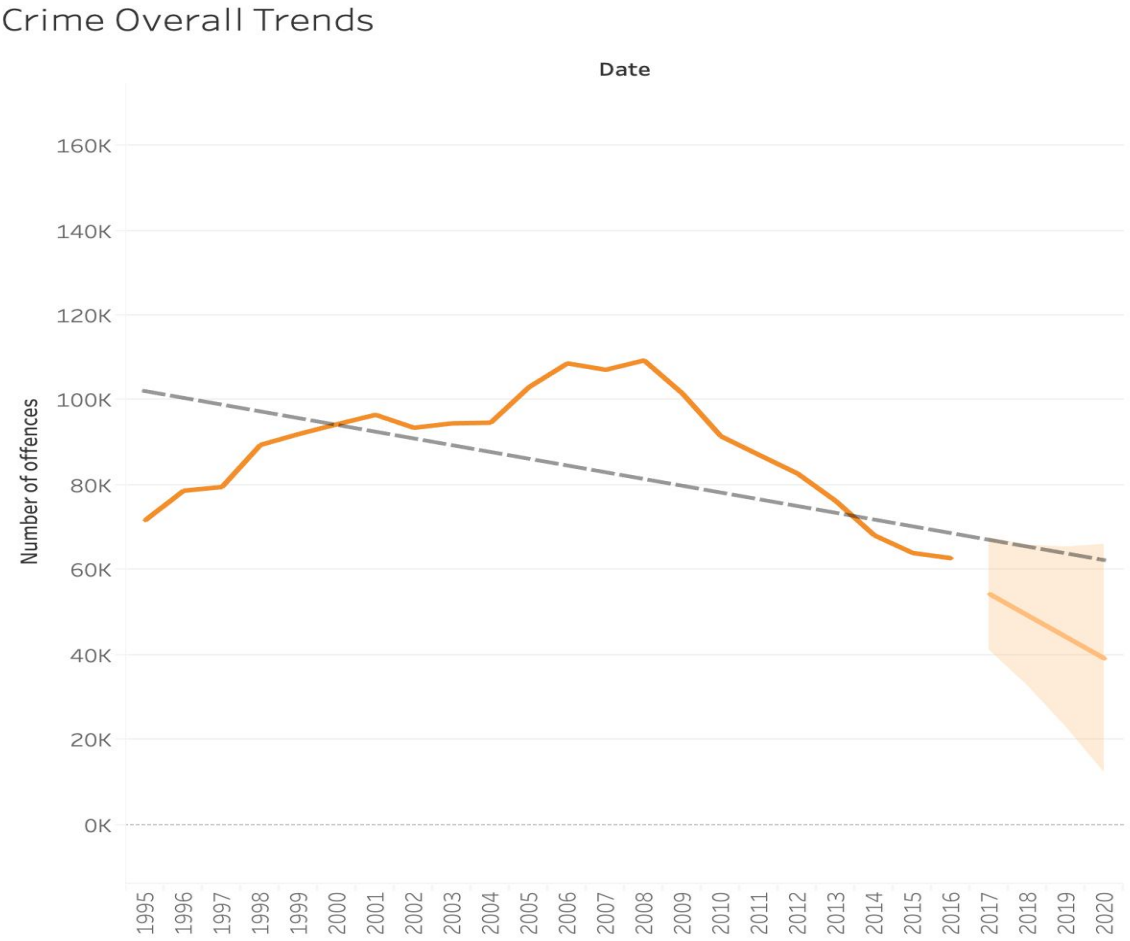


Figure 5-1 Overall of Crime in NSW

For finding the reason of the above conclusion, We drill down the data to the subcategories level and keep the top five crime types which cause the most influence on the overall tendency of crime: ***Theft, Malicious damage, Transportation, Assault,Transportation, Against justice***. And the following part would analysis how top four types of crime influence the overall tendency.

5.1 Theft

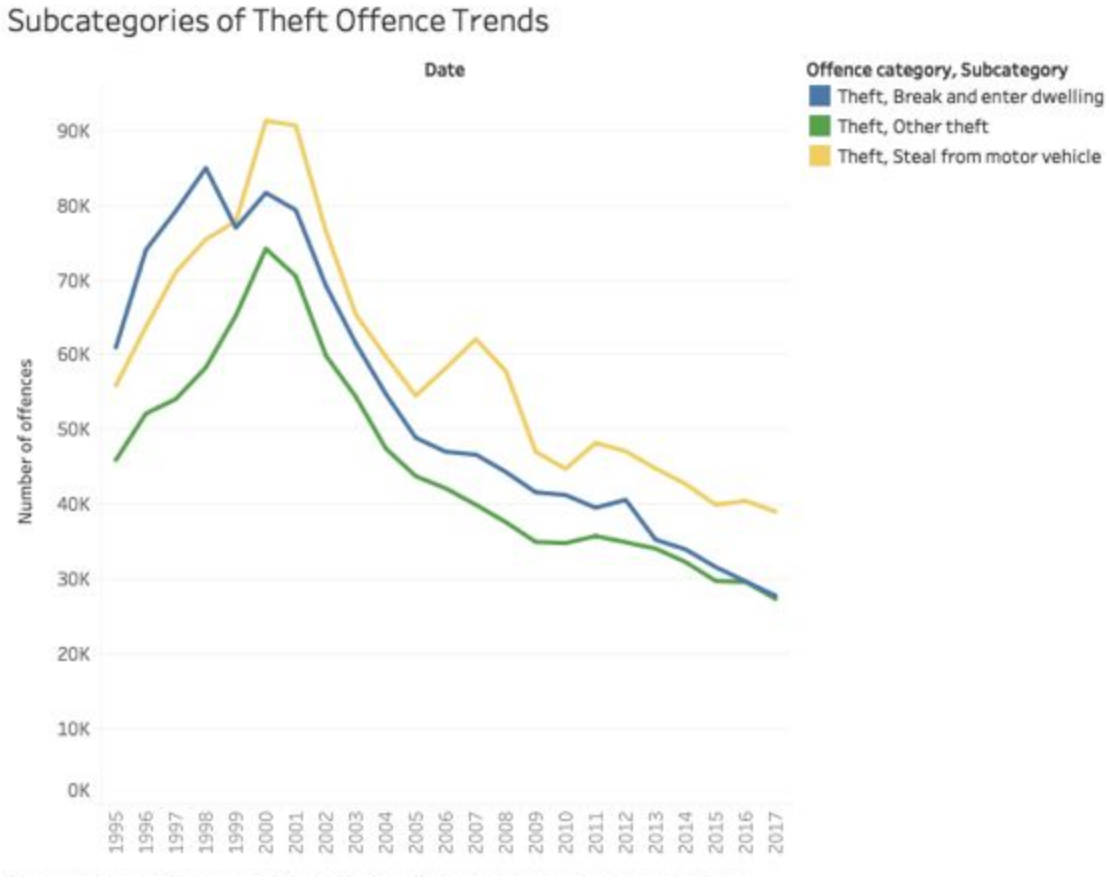


Figure 5-2 Theft crime in NSW

Theft in NSW is changing through time as the above-presented data shows, which we could find that the trend of top three sub-types of crime are showing a considerable decrease from 2001. To detect the reasons for this, we do lots of research and summary the possible reasons as following:

1.Australian heroin shortage

For supporting the expensive price of heroin, some depended Heroin addict would be a thief for obtaining extra money (Dobinson & Poletti, 1995; Dobinson & Ward, 1987; Dobinson & Poletti,1989). There is suitable proof shows that the demand for heroin would sudden drop when the price is rising. And according to Degenhardt (2004), Price of Heroin in Australia is increasing dramatically because of the shortage of drugs, which cause the decreasing the number of Heroin addicts in the NSW also decrease correspondingly.

2. Salary increasing

Another reason for the tendency of theft in NSW is keeping going down is that the salary of NSW citizen is increasing (Wan et al., 2012). With the improvement in disposable income, people would be satisfied with basic food and clothes, which would cause less theft crime. In this aspect, the Salary of NSW citizen would continually improve because the Australian economy keeps increasing (Wan et al., 2012).

3. Police and security is improving

According to the research of the NSW Bureau of Crime Statics and Research, with the improve of police number and configuration, the risk of arrest keeps increasing(Wan et al., 2012). In that case, people would be scared of crime and obey the law. In this aspect, the Police situation would improve continually, which would help to keep the theft crime from decreasing (Chilvers &Weatherburn, 2001).

4. Changes in the number of people in the peak offender-prone age bracket (16-24 years)

The age structure is a potential factor which would influence the crime because teenager would have a higher probability of crime (ACS Distance Education,2011). The Australian population is aging and that could cause down tendency of theft crime.

In that case, theft crime would have a tendency to keep in a low and stable level.

5.2 Malicious Damage

For Malicious Damage, the picture (Figure 5.3) below show its trend from 1995 to 2017. It shows overall slight decrease.

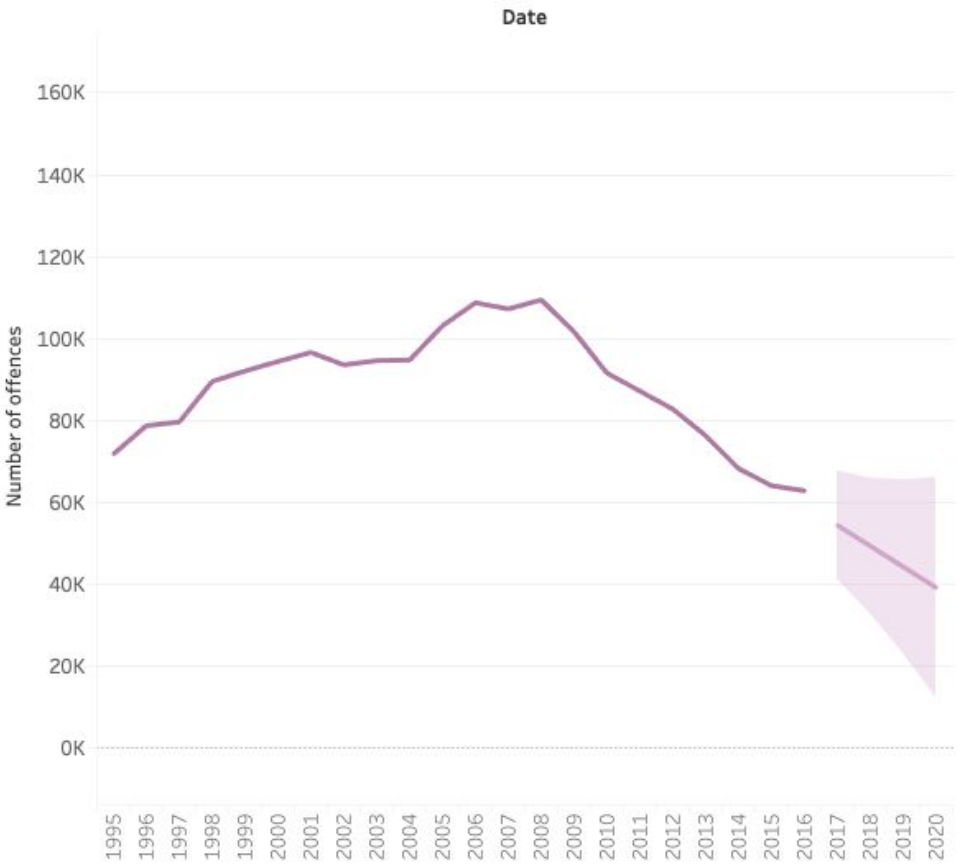


Figure 5.3 Malicious Damage Trend in NSW

There is a close connection between the crime rate and the improvement of the law. Thus, the rate of imprisonment and court sentences can be considered as an important reference for estimating the malicious damage trend. The picture below shows the variation of the court sentences & rate of imprisonment between 2002 and 2017(Gladstone, 2019). From the picture, we can see both the number of court sentences and imprisonment experience a gradual rise in NSW, in particular, they all significantly ascend after 2012. This phenomenon indicates the continuous improvement of the current legal system in NSW, which is extremely useful for suppressing the case of malicious damage. Hence, we are optimistic that the incidence of future crime about malicious damage will be lower due to the improvement of the law system.

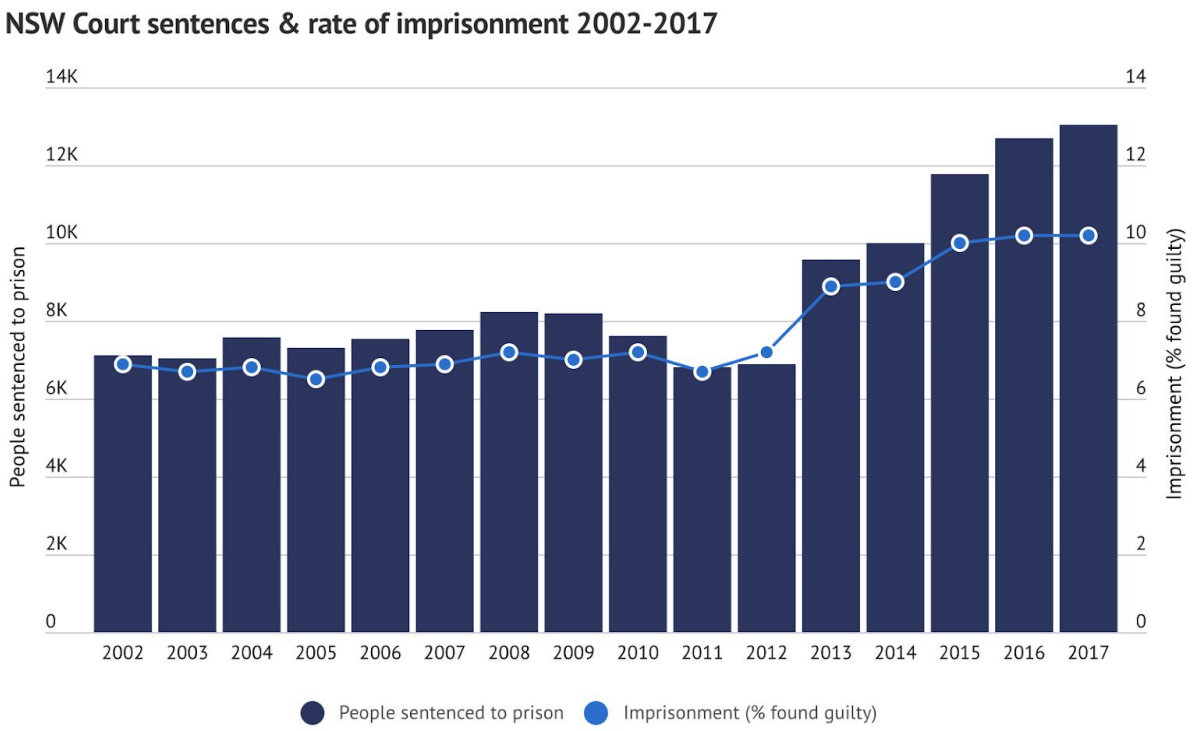


Figure 5.4 NSW Court sentences & rate of imprisonment 2002-2017(Gladstone, 2019)

5.3 Transportation

Due to the increase in population and the development of cities, the number of crime of transport regulatory has increased year by year. As can be seen from the below picture(figure 5.3), the incidence of transport regulatory offence experiences a sharp rise from 1995 to 2017.

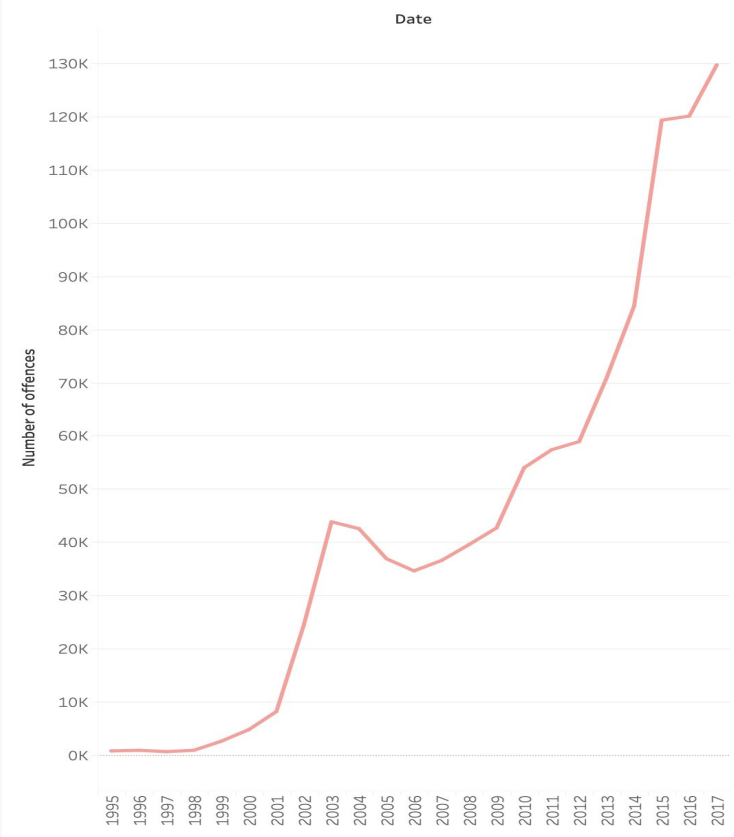


Figure 5.5 Transport Regulatory Offences Trend in NSW

Furthermore, it should be noted that the incidence of transport regulatory is extremely low before 2002. However, during 2002 and 2003, this offence experiences the dramatic increase from approximately **10,000** to **50,000** which shows a fivefold variation. The picture below(figure 5.4) displays the statistics of vehicle registration in NSW from 2000 to 2018(Roads Maritime & Services, 2019). The solid blue line represents the actual data, the dash blue line is its trend line. The dash orange line shows the forecast trend. The overall trend of vehicle registration in NSW is continuous growth which shows the similar trend to transport regulatory offence. Moreover, there is one interesting thing that cannot be ignored, the number of vehicle registration in NSW significantly grows up, especially showing a maximum growth rate between 2002 and 2003. This obvious variation coincides with the trend of transport regulatory offence in the same time

period. Thus, we propose that the incidence of transport regulatory offence can be impacted by the number of vehicle registration and it shows a positive correlation.

As can be seen from the predicted trend in the figure (dash orange line in figure 5.4), the number of vehicle registration will continue increasing in the next few years. Thus, we consider the transport regulatory offence will continue rising in the next few years due to the increase of the number of vehicle registration in NSW.

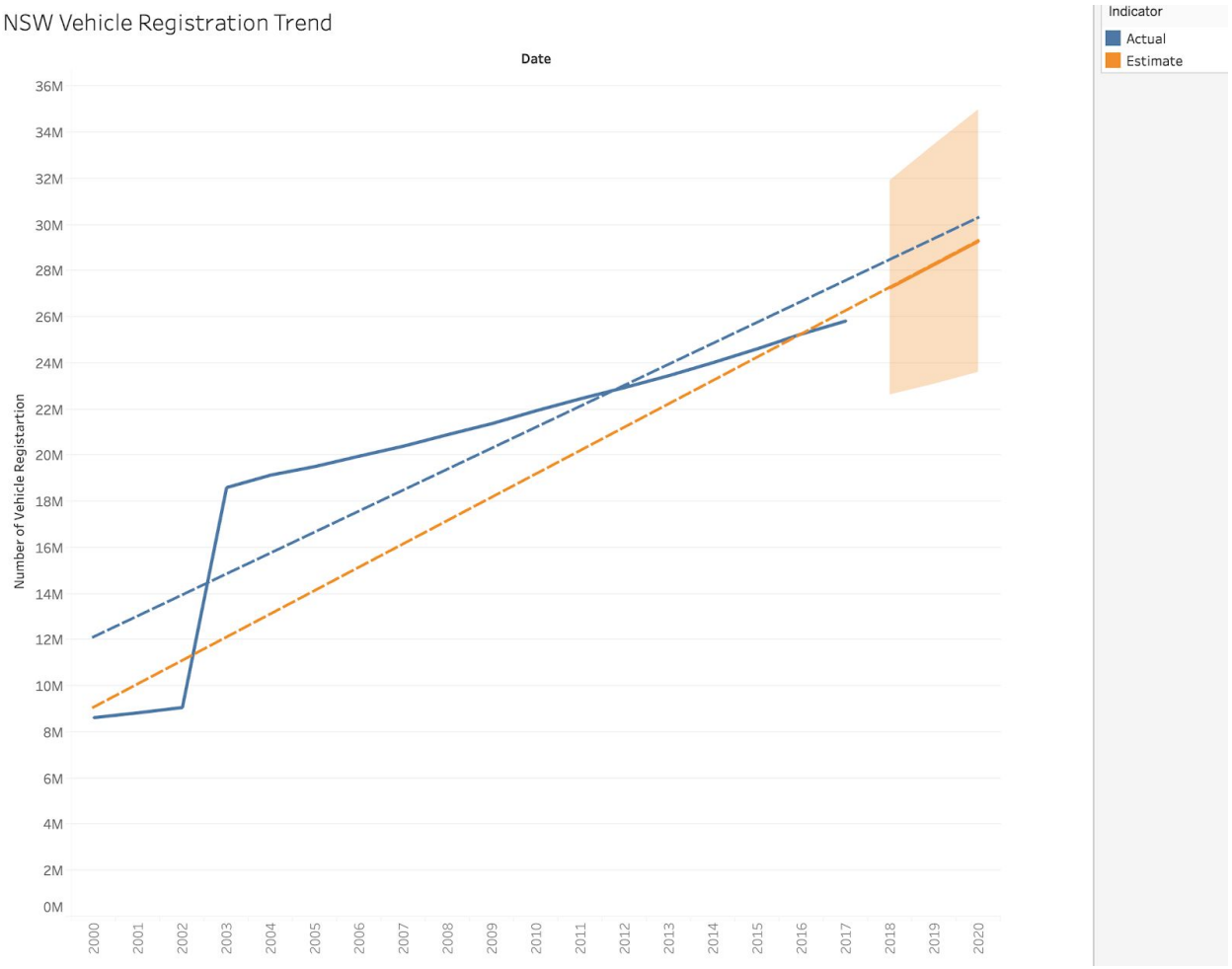


Figure 5.6 Vehicle Registration Trend in NSW

5.4 Assault

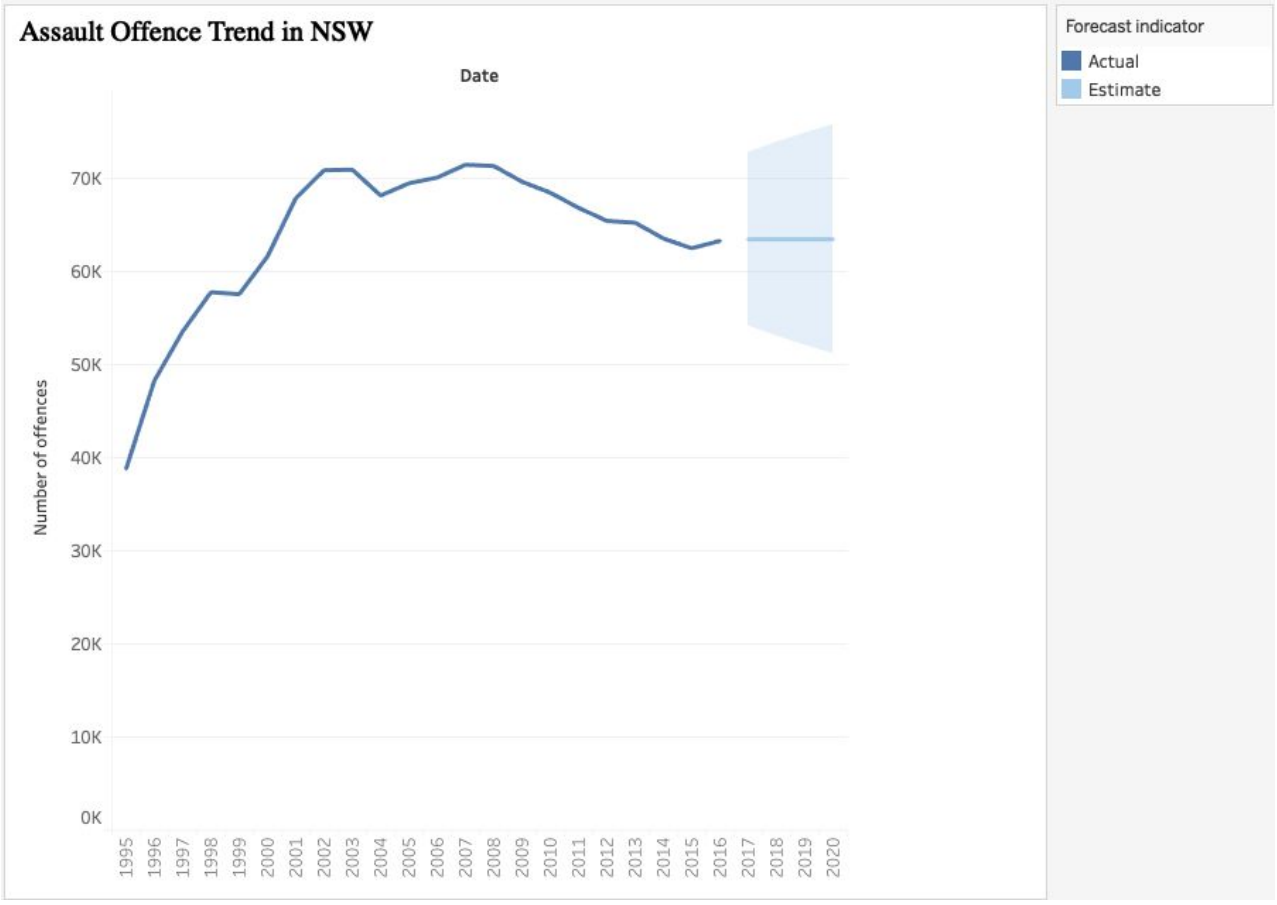


Figure 5.7 Assault Offence Trend in NSW

In terms of the assault offence category, the trend of assault is decreasing based on the prediction function of the tableau. After drill down into the subcategory level, it is not so hard to find the specific situation of assault crime from domestic and international violence related assault, which is shown as follows:

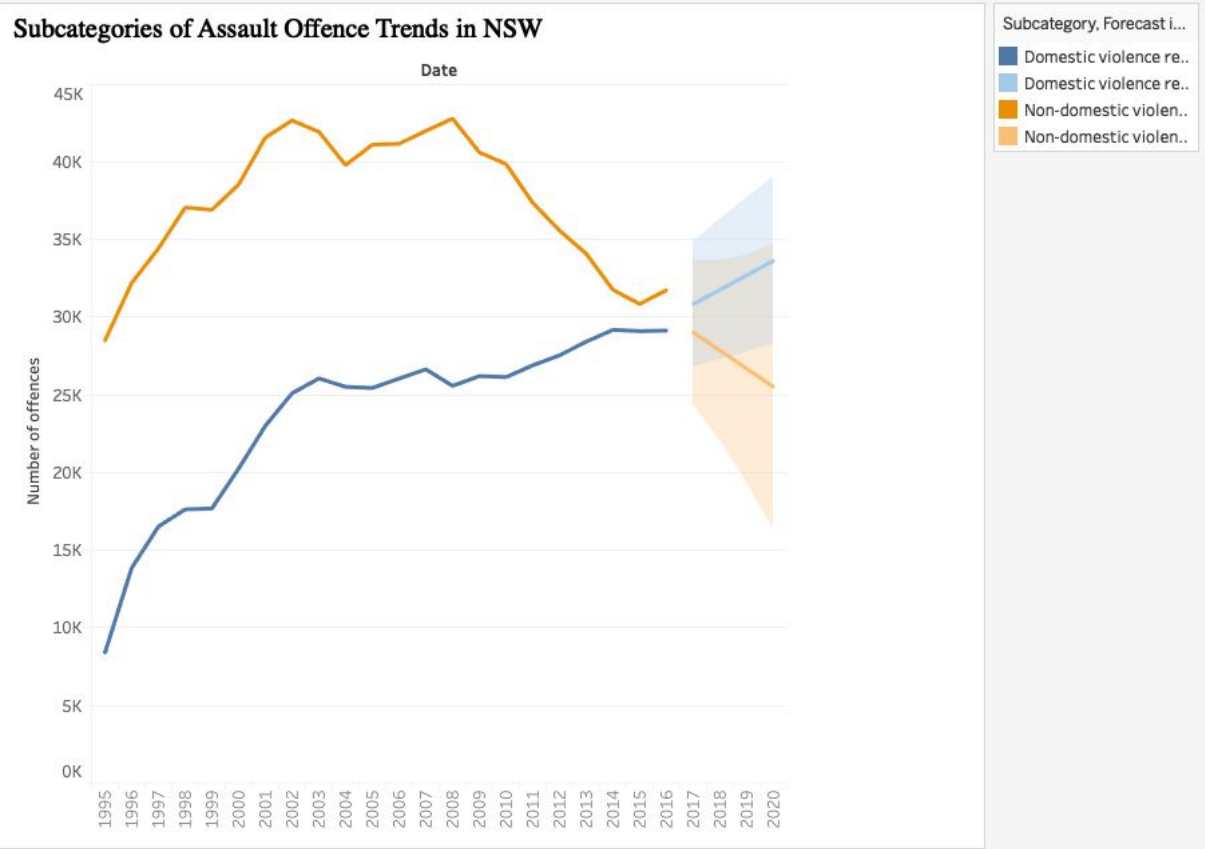


Figure 5.8 Subcategories of Assault Offence Trends in NSW

As the data above presented, non-domestic assault shows an overall down tendency and domestic assault shows an up tendency and then be smooth and stable. It is not hard to find the population is increasing year to year based on the population trends of NSW (Figure 5.9). We believe the population is a significant factor for the assault offence. The increasing population results in increasing trends of domestic assault. In terms of non-domestic assault, the population is also the most significant factor influence the change of the trend. The population of non-domestic is increasing and the biggest portion is the students based on Figure 5.10. It is not hard to find even the population is increasing, the reason why the non-domestic assault offense decrease is that the international population is restricted by the visa limitation and most of them are the student who above bachelor degree which means they receive the high-level education.

Even though the domestic assault offence increase, the total decreasing number of non-domestic assault offence is higher than the increasing number, the overall trend of the assault is decreasing.

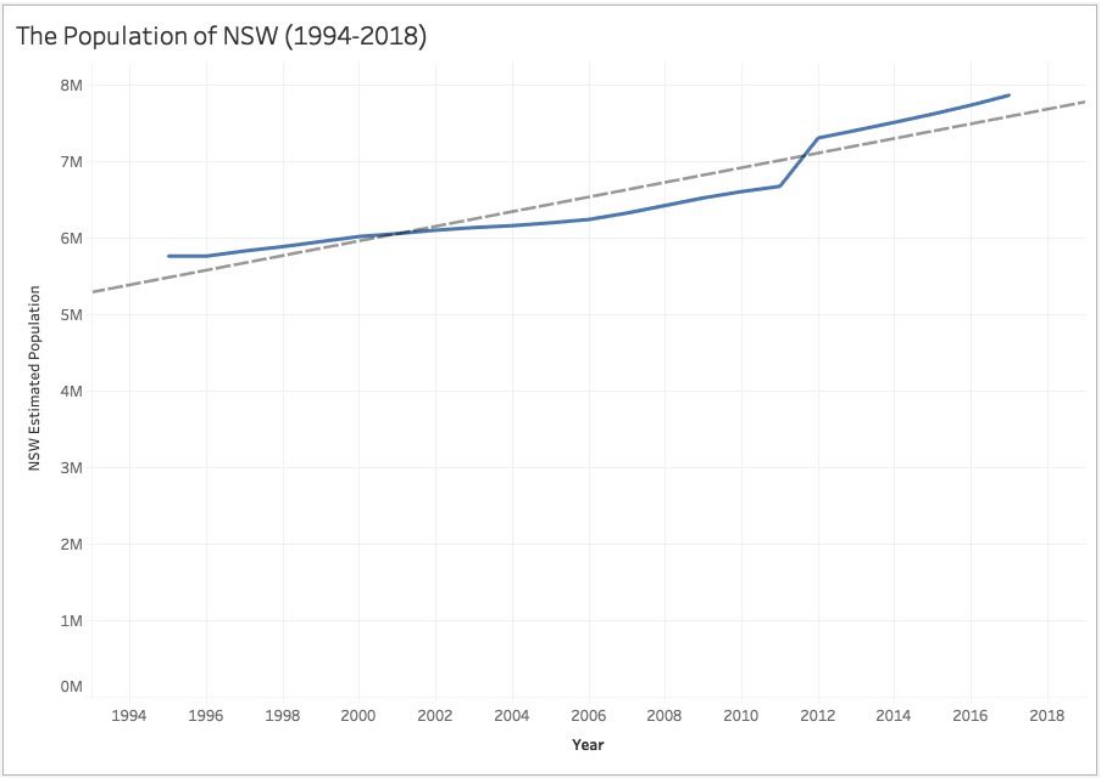


Figure 5.9 The Population Trends of NSW (1994-2018)

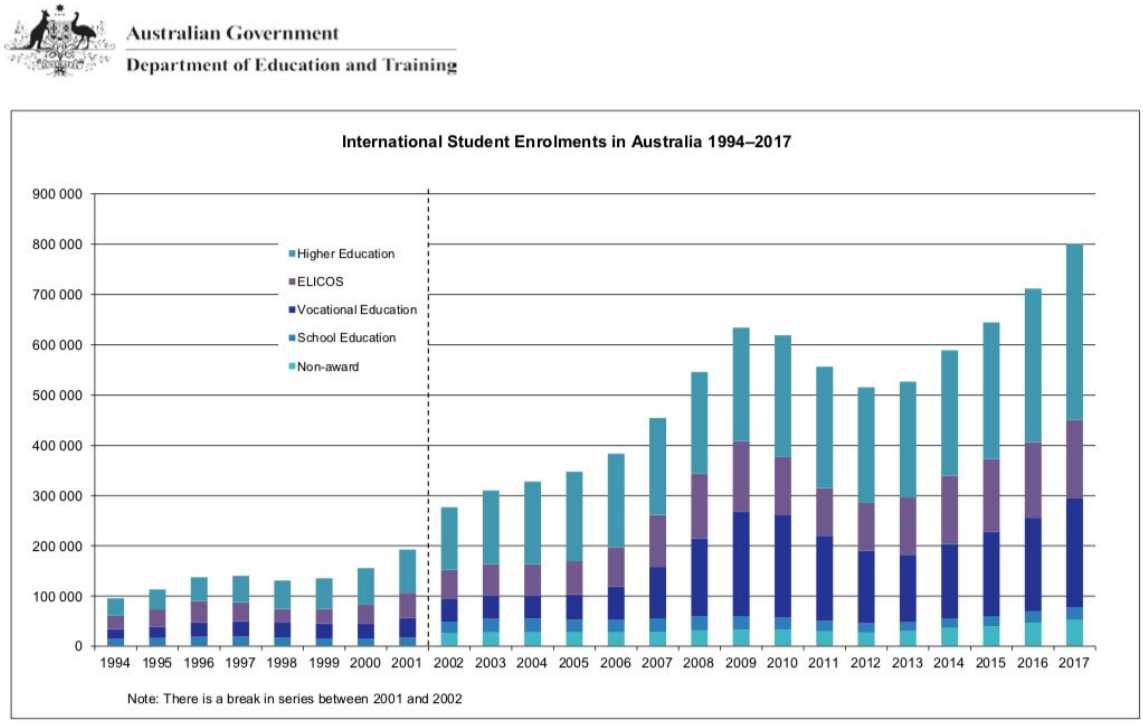


Figure 5.10 The International Student Enrollment in Australia 1994-2017

To sum up, according to the analysis of top four offence categories, only the fourth (*transportation*) is increasing, and top three offence categories are decreasing (*Theft, Malicious damage, Assault*). So the overall offence trend will decrease in the future.

6. REFERENCES

Acs Distance Education.(2011).*TEEN CRIME RISK FACTORS* cite from：https://www.acs.edu.au/info/psychology/child-development/crime-risk.aspx

Chilvers, M. & Weatherburn, D. (2001). *The New South Wales "Compstat" Process: Its Impact on Crime. Australian and New Zealand Journal of Criminology, 37(1), 22-48.*

Dobinson, I. & Ward, P. (1985). *Drugs and Crime: A survey of NSW prison property offenders. Sydney: NSW Bureau of Crime Statistics and Research.*

Dobinson, I. & Ward, P. (1987). *Drugs and Crime: Phase 11: A study of individuals seeking treatment. Sydney: NSW Bureau of Crime Statistics and Research.*

Dobinson, I. & Poletti, P. (1989). *Buying and Selling Heroin: A study of heroin user/dealers. Sydney: NSW Bureau of Crime Statistics and Research.*

Degenhardt, L., & Day, C. (2004). *The course and consequences of the heroin shortage in NSW. Sydney: National Drug and Alcohol Research Centre.*

Gladstone, N.(2019). *Crime in NSW has been dropping for more than a decade. This is why.* Retrieved from https://www.smh.com.au/national/nsw/crime-in-nsw-has-been-dropping-for-more-than-a-decade-this-is-why-20181108-p50eq4.html

International Student Data 2017. (n.d.). Retrieved from https://internationaleducation.gov.au/research/International-Student-Data/Pages/InternationalStudentData2017.aspx

Roads Maritime & Services (2018). *Quarterly trend in registered vehicles by vehicle type.*

Retrieved from https://www.rms.nsw.gov.au/about/corporate-publications/statistics/registrationandlicensing/tables/table111.html

Tableau Prep. (2019). Retrieved from https://www.tableau.com/products/prep

Wan, W., Moffatt, S., Jones, C., & Weatherburn, D. (2012). *The effect of arrest and imprisonment on crime (Crime and Justice Bulletin No. 158)*. Retrieved from NSW Bureau of Crime Statistics and Research website: http://www.bocsar.nsw.gov.au/lawlink/bocsar/ll_bocsar.nsf/vwFiles/CJB158.pdf/$file/CJB158.pdf