Relationship between Crimes and the Number of Liquor Stores

By: William Liandri - 728710

Research Questions

- Where are the top 5 LGAs with the highest crime rates?
- What type of crimes that are mostly happened in Victoria ?
- Is there any relationship between the number of liquor stores and the certain type of crimes happening in Victoria? If so, what type of crimes is that?

Why is It Worth Tackling?

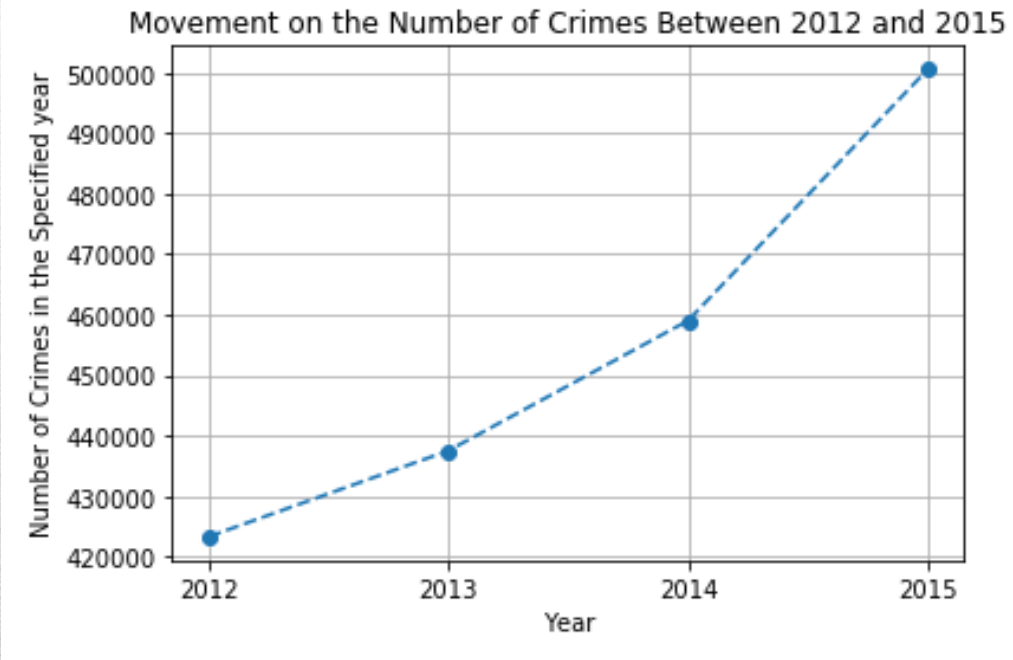

Figure 1

- Number of crimes have been increasing in Victoria by around 20% between 2012 and 2015
- This statistic should draw government's attention to reduce the number of crimes in Victoria

What are the Datasets? Why?

- *"Crime_location.csv"*:
  - **Information:** the number of crimes between 2012 and 2016 by Crime Statistics Agency in Victoria.
  - **Reason:** To get information about the total number of crimes and the total number of crimes for each type of crimes in each LGA in Victoria in 2015.
- *"Liquor_Metro_2015.csv" and "Liquor_Region_2015.csv"* :
  - **Information:** list of liquor stores in metropolitan and region in Victoria in 2015 by Victorian Commission for Gambling and Liquor Regulation.
  - **Reason:** To know the total number of liquor stores in each LGA in Victoria in 2015.
- *"Population_by_LGA_2015.csv"*:
  - **Information:** the total number of populations based on gender and by age in 2015 in Australia by Australian Bureau of Statistics
  - **Reason:** To get information about the total number of populations in each LGA in Victoria, which will be used to normalise the data

Data Wrangling Methodologies

- Detects some missing values:
  - 4 LGAs (*"Hepburn", "Hinmarsh", "West Wimmera", and "Indigo"*) with type crime *"F Other offences"* in *"Crime_location.csv"*
    - **Solution:** Replaced the value with 0
  - 1 LGA ("*Queenscliffe*") in *"Population_by_LGA.csv"*
    - ***Solution:*** Replaced the value with the average population in Victoria

Figure 2

- Detects some outliers:
  - 16 LGAs as outliers ("*Boroondara", "Brimbank", "Casey", "Darebin", "Greater Dandenong", "Greater Geelong", " Horsham", " Hume", "Latrobe", "Melbourne", "Monash", "Mornington Peninsula", "Port Phillip", "Queenscliffe", "Stonnington", "Yarra*")

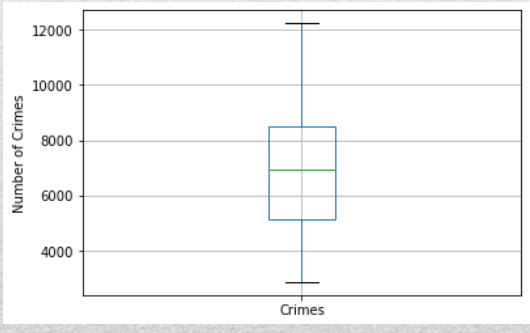

Figure 3

- Normalizing the data using this formula:

$$\textbf{Number of crimes per 100,000 people} = \frac{\textit{Total number of crimes in that LGA}}{\textit{Total number of populations in that LGA}} * \textbf{100,000 people}$$

Findings

- Top 5 LGAs with the Highest Crime Rates in 2015

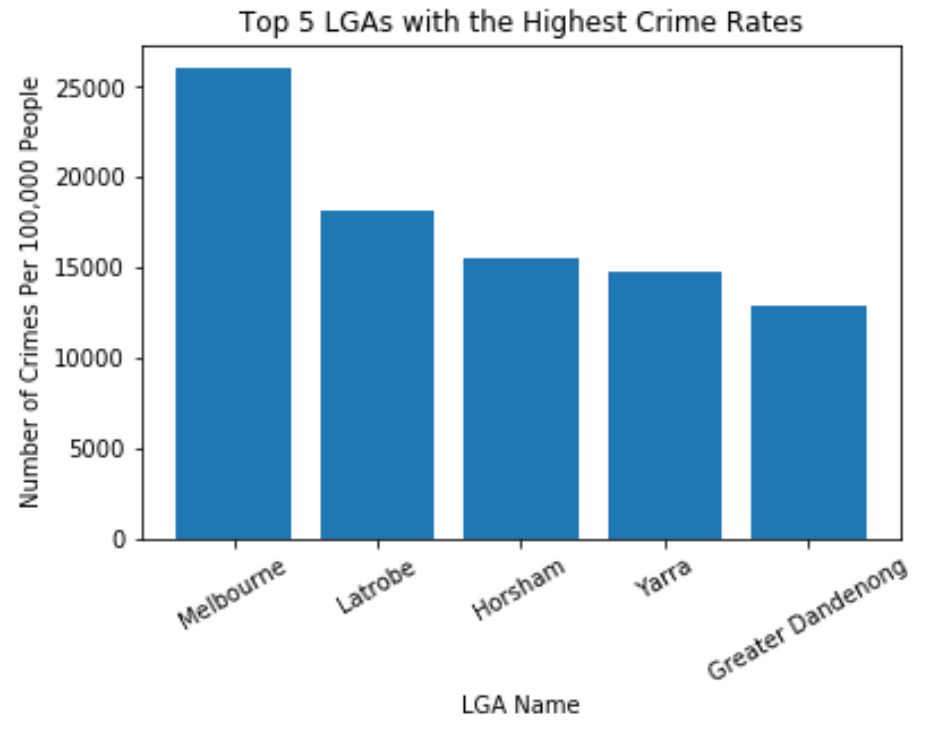

Figure 4

- Top 5 LGAs with the highest crime rates were Melbourne, Latrobe, Horsham, Yarra, and Greater Dandenong.
- Amongst them, Melbourne had the highest crime rates with around **26,000** crimes per 100,000 people.
- Emphasizes Government to reduce crimes in those LGAs

Findings

- Crime Rates for Each Type of Crimes

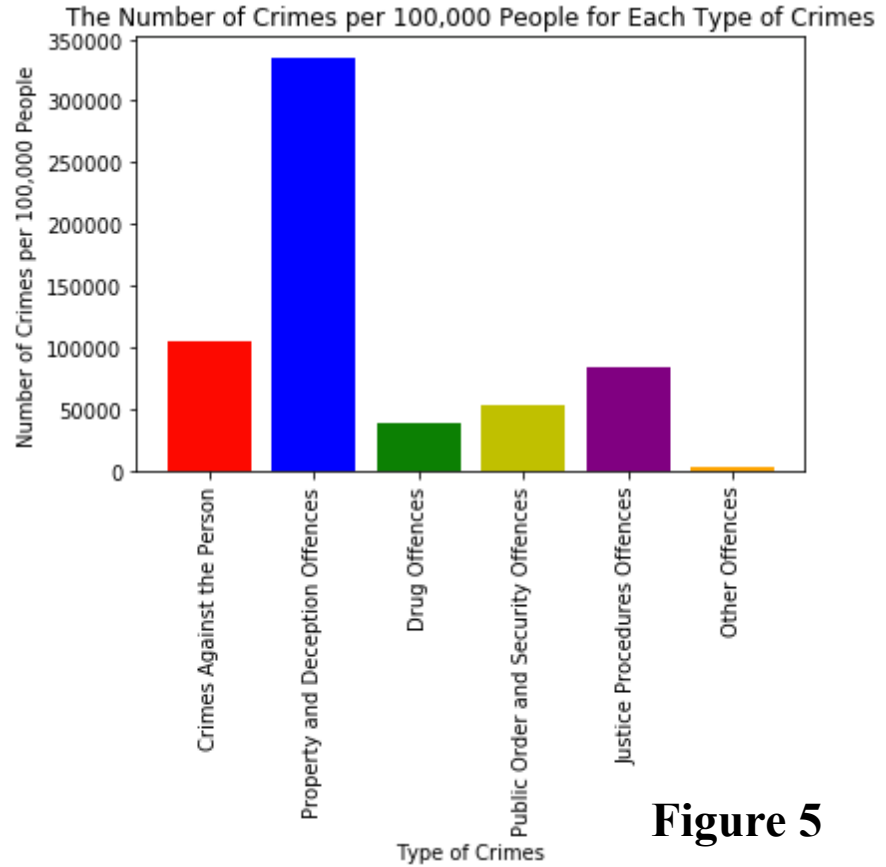

Figure 5

- **Highest:** Property and Deception offences with **335,509** crimes per 100,000 people
- **Lowest:** Other offences with **3,017** crimes per 100,000 people

Findings

- Correlation between Certain Type of Crimes and the Number of Liquor Stores

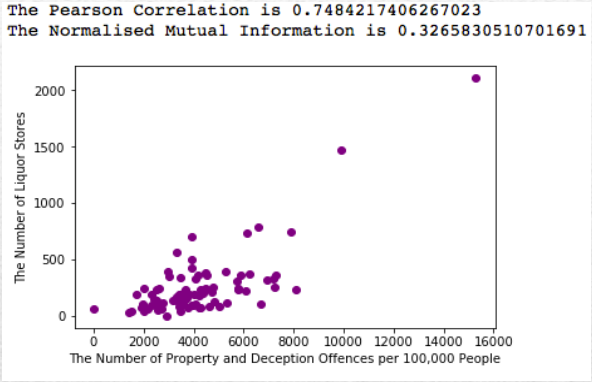

Figure 6

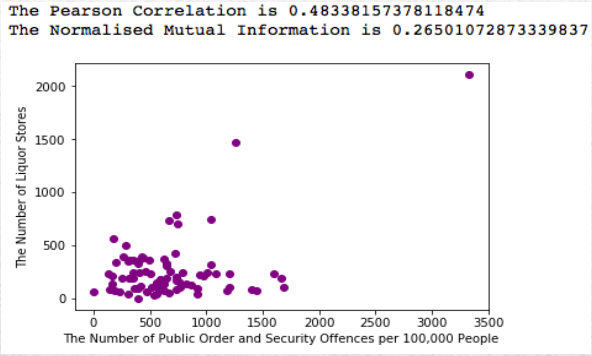

Figure 7

Analyze Results for *"Property and Deception Offences":*

- **Pearson Correlation:** strong relationship with the number of liquor stores (since above 0.5)
- **Normalised Mutual Information:** quite high correlation (more dependent) with the number of liquor stores

Analyze Results for *"Public Order and Security Offences":*

- **Pearson Correlation:** moderate relationship with the number of liquor stores (since below 0.5)
- **Normalised Mutual Information:** quite high correlation (more dependent) with the number of liquor stores

From these findings, it can be concluded there is a relationship between crimes, especially crime with those 2 categories above, and the number of liquor stores. Thus, the government can reduce the number of crimes by reducing the number of liquor stores.

Challenges

- Spent around 5 days and changed the topic 3 times before choosing the right topic and dataset.
- Changing the dataset from 2016 to 2015, since the population data in 2016 could not be found to normalize the data. Thus, some codes were also changed and there were some new codes to be written to suit the new dataset.

THANKS FOR YOUR ATTENTION !