Project Report

Research Topic:

The following question is the objective of our investigation:

How do the covid lockdowns influence rates of various crime types across areas of distinct socio-economic status in Victoria?

This research question concerns the liveability aspect of Victoria, as a broad range of factors can enhance the liveability of a community and among them are low crime rates. Furthermore, inclusivity is yet another feature to be examined. Racially profiled crimes as a consequence of covid runs rampant in numerous communities, and a reduction in such crimes will be yet another step towards a truly inclusive living environment.

Datasets Used:

- Socio-economic level by area from 2016 (554Rx13C): https://stat.data.abs.gov.au/Index.aspx?DataSetCode=ABS_SEIFA_LGA#
- Crime statistics in Victoria (Data Tables - LGA Recorded Offences Visualisation)
  - (Year ending Dec, 2020, 343949Rx9C): https://www.crimestatistics.vic.gov.au/crime-statistics/latest-victorian-crime-data/download-data-0
  - (Year ending Sep, 2020, 341529Rx9C): https://www.crimestatistics.vic.gov.au/crime-statistics/historical-crime-data/download-data-3
  - (Year ending June, 2020, 339091Rx9C): https://www.crimestatistics.vic.gov.au/crime-statistics/historical-crime-data/download-data-2
  - (Year ending March, 2020, 336719Rx9C): https://www.crimestatistics.vic.gov.au/crime-statistics/historical-crime-data/download-data-1
- Lockdown Timeline: https://www.abc.net.au/news/2021-08-19/melbourne-200-days-of-covid-lockdowns-victoria/100386078

Background Information:

- The crime statistics are extracted from the *Crime Statistics Agency (CSA),* it consists of **six** major divisions, denoted by alphabets ***A - F***.
    - Crime type A is **Crimes against the person**. E.g. Robbery and sexual offences
    - Crime type B is **Property and deception offences.** E.g. Theft and arson
    - Crime type C is **Drug offences.** E.g. Manufacture and trafficking of drugs
    - Crime type D is **Public order and security offences.** E.g. Public nuisance and security offences
    - Crime type E is **Justice procedures offences.** E.g. Disruption of court and resisting arrest
    - Crime type F is **Other offences.** E.g. Driving and other miscellaneous offences
    - The full definition of each sub-division is outlined in *figures 1* and *2* below
- Time periods of analysis are also split into two, **"Pre/Post"** and **"During"** lockdown. "Pre/Post" is over the months of **January - March** and **October - December**, whereas "During" is from **April - September**.
- The socio-economic-percentiles are discretized into **four** "SES" classes:
    - [1, 25] Low, [26, 50] Medium-Low, [51, 75] Medium and [76, 99] High.
- Due to the scale of the y-axis, some fluctuations in data may not be visually observable.

CSA Offence Division	CSA Offence Sub-Division
A Crimes against the person	A10 Homicide and related offences
	A20 Assault and related offences
	A30 Sexual offences
	A40 Abduction and related offences
	A50 Robbery
	A60 Blackmail and extortion
	A70 Stalking, harassment and threatening behaviour
	A80 Dangerous and negligent acts endangering people
B Property and deception offences	B10 Arson
	B20 Property damage
	B30 Burglary/Break and enter
	B40 Theft
	B50 Deception
	B60 Bribery
C Drug offences	C10 Drug dealing and trafficking
	C20 Cultivate or manufacture drugs
	C30 Drug use and possession
	C90 Other drug offences

Fig. 1 Table of Crime Classification (A ~ C)

CSA Offence Division	CSA Offence Sub-Division
D Public order and security offences	D10 Weapons and explosives offences
	D20 Disorderly and offensive conduct
	D30 Public nuisance offences
	D40 Public security offences
E Justice procedures offences	E10 Justice procedures offences
	E20 Breaches of orders
F Other offences	F10 Regulatory driving offences
	F20 Transport regulation offences
	F30 Other government regulatory offences
	F90 Miscellaneous offences

Fig. 2 Table of Crime Classification (D ~ F)

Data Linkage & Wrangling Methods:

We used six datasets for our investigation which include the socio-economic level (SES) by LGAs, crime statistics by LGAs of 2020 divided into quarters, as well as the lockdown timeline during 2020. All datasets used are formatted into csv files except the lockdown timeline is in texts.

Year	Year ending	Local Government Area	Postcode	Suburb/Town Name	Offence Division	Offence Subdivision	Offence Subgroup	Offence Count
2020	September	Alpine	3691	Dederang	rty and deception offences	B40 Theft	B41 Motor vehicle theft	1
2020	September	Alpine	3691	Glen Creek	C Drug offences	Drug use and possession	C32 Drug possession	1
2020	September	Alpine	3691	Glen Creek	F Other offences	90 Miscellaneous offences	F93 Cruelty to animals	1

Table 1. Raw crime data

As shown in Table 1, the raw data containing various offence counts, separated by crime divisions, of certain LGA's required filtering in order to separate the relevant columns. First and foremost, the rows that were dated before the year 2020 were removed. Since the original table was sorted in descending order by year, this was achieved by finding the index of the last recorded crime incident in 2020. The next step was to extract the useful columns in the table, which incidentally were, Local Government Area, Offence Division and Offence Count. Upon discovering certain formatting differences in the previously mentioned columns, text preprocessing methods were utilized in order to homogenize the data, namely both replace() and strip() functions. A number of entries in Offence Count were found to be of type string instead of int and contained commas, presumably for better readability. Once the replace() function was used to remove occurrences of commas, the string values were converted to int. Many of the values in

Offence Division had discrepancies in the form of leading and trailing spaces characters, which were removed using the strip() function.

LGA	offence_div	offence_count
Alpine	A Crimes against the person	127
Alpine	B Property and deception offences	191
Alpine	C Drug offences	33

Table 2. Quarterly crime count

LGA	offence_div	Score	total_div_count
Alpine	A Crimes against the person	970	256
Alpine	B Property and deception offences	970	492
Alpine	C Drug offences	970	60

Table 3. During-period crime count

The resulting data can be seen in Table 2, displaying the relevant columns, grouped by both LGA and offence_div, for each quarter of the year 2020. For the purposes of analysis, the four tables would be further joined to produce only two tables that would represent periods in and out of lockdown. Upon reaching the desired table format, an additional column was added, called Score, representing the SES score, thus producing Table 3.

Measure		Score	RANK WITHIN AUSTRALIA			RANK WITHIN STATE AND TERRITORY			Minimum score for SA1s in area	Maximum score for SA1s in area	Usual resident population
			Rank within Australia	Rank within Australia - Decile	Rank within Australia - Percentile	Rank within State or Territory	Rank within State or Territory - Decile	Rank within State or Territory - Percentile			
Local Government Areas - 2016											
New South Wales	Albury (C)	956	254	5	47	64	5	49	642	1151	51076
	Armidale Regional (A)	976	339	7	63	87	7	67	747	1119	29449
	Ballina (A)	987	383	8	71	92	8	71	673	1117	41790
	Balranald (A)	927	136	3	25	30	3	23	874	1031	2287

Table 4. Raw SES data

Table 4 shows the raw socio-economic level data from 2016, which includes the ranks, scores, percentiles within Australia and State by LGA. Since our scale of investigation is within Victoria, we first have to filter in the rows for Victorian LGAs. As for the next step, the text-processing of LGA names with the use of regular expressions function, re.sub(), was done to discard irrelevant texts and characters such as “(C)” and “(A)” written after each LGA name. The other columns we extracted were “Score” and “Rank within State or Territory - Percentile” since we were only interested in creating tables that clearly show the SES levels of each LGA within Victoria. Values of those two columns were converted into integers from strings. The resulting tables are the following, Table 5 and 6.

Low	Medium low	Medium	High
Campaspe	Baw Baw	Darebin	Boroondara
Colac-Otway	Towong	Hobsons Bay	Bayside
Greater Shepparton	Murrindindi	Maribyrnong	Stonnington

Table 5. SES classes

LGA	Percentile	Score	Class
Boroondara	99	1128	high
Bayside	98	1125	high
Stonnington	97	1120	high

Table 6. Rankings of SES levels

In Table 5, the SES levels are separated into four classes, ‘Low’ include LGAs with percentiles between 1 - 25 and ‘Medium low’ include LGAs with percentiles between 26 - 50, and so on. Table 6 contains the percentiles and scores of SES levels in the descending order and the assigned SES divisions.

Crime Type	Low	Medium low	Medium	High
A Crimes against the person	32225	27567	54523	49882
B Property and deception offences	88500	74806	189464	233139
C Drug offences	13293	10499	23950	27887

Table 7. During-period total crime count by SES level divisions

Table 2 and 5 were then utilised to create tables that show the total counts of each crime type based on the four SES level divisions, which were used for creating bar charts that will be analysed in the latter part of the report.

Key Findings and Analysis:

- **Graphical Analysis: Bar Charts**

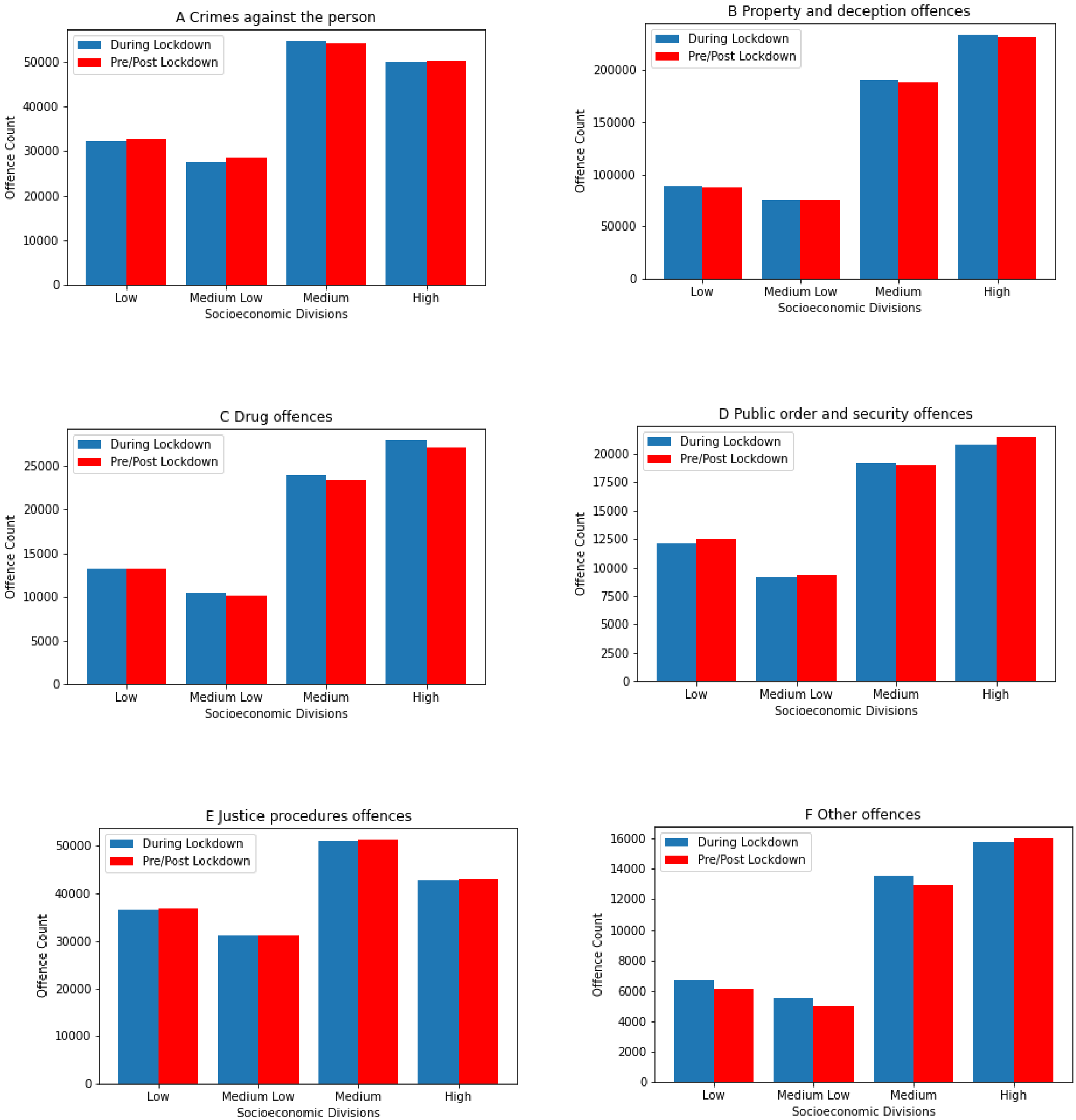


Fig. 3-8 Bar charts of different crime types in different socio-economic divisions

Across the six crime types examined, the general trend observed is that medium to high SES levels have the greatest crime rates, while medium-low are consistently the lowest. Furthermore, the fluctuations in crime during and out of lockdowns are insignificant on this scale, with crime type “justice and procedure offences” having nearly identical rates regardless of lockdown. However, crime rates for “other offences” are noticeably higher during lockdown when compared to out of it, with the exception being the high SES level, where an opposite effect is present.

- **Graphical Analysis: Scatter Plot**

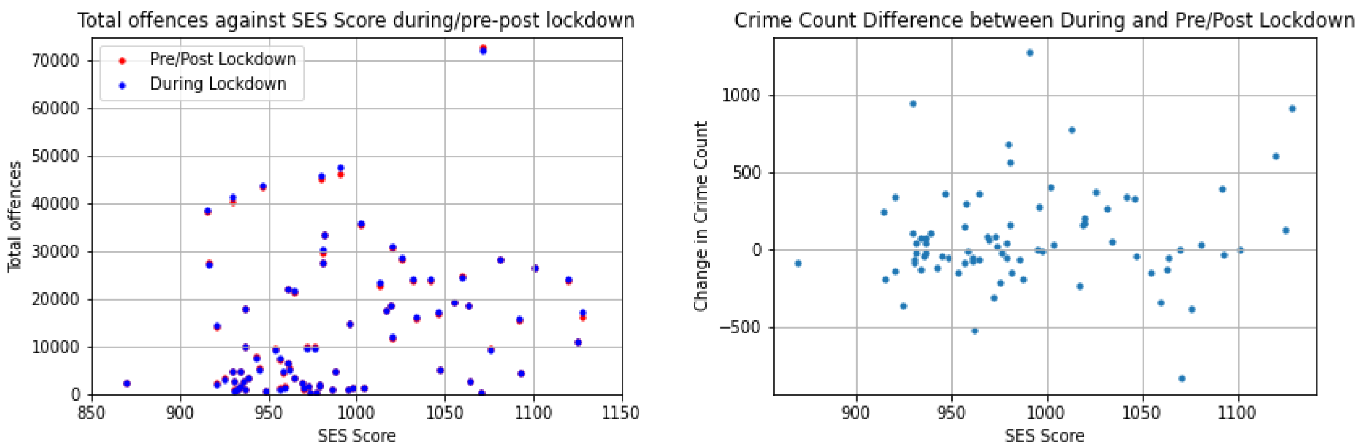
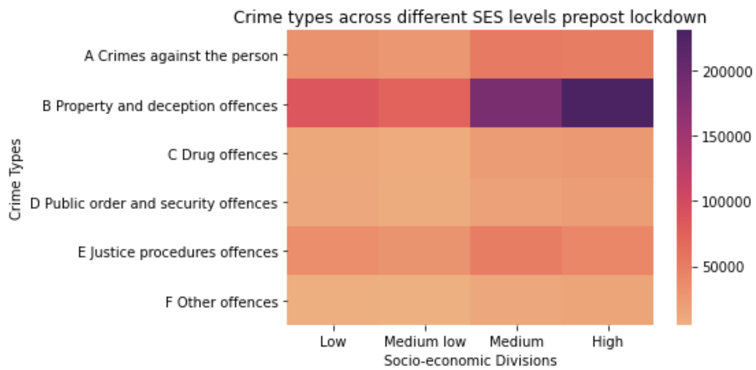


Fig. 9-10 Scatter plots of total offences and crime difference against SES score

By overlapping the total offences that occurred during and pre/post lockdown, minimal change was observed between the total offences during lockdown compared to when there was no lockdown throughout most SES levels, rather some areas can be seen having a higher number of offences during lockdown. This is further supported after using “Pearson Correlation " between both total offences happening with and without lockdown with a strong positive correlation of 0.9997755. Furthermore, observed crime rates in low to medium SES levels were greater compared to that of the higher SES levels, with Melbourne being the outlier at around 70,000 total offences. A scatter plot where the crime counts during lockdown was subtracted with the crime count without lockdown was plotted and the LGAs whose observed difference were the lowest are mostly low SES levels, whereas the difference in higher SES levels are more significant.

- **Graphical Analysis: Heatmaps**



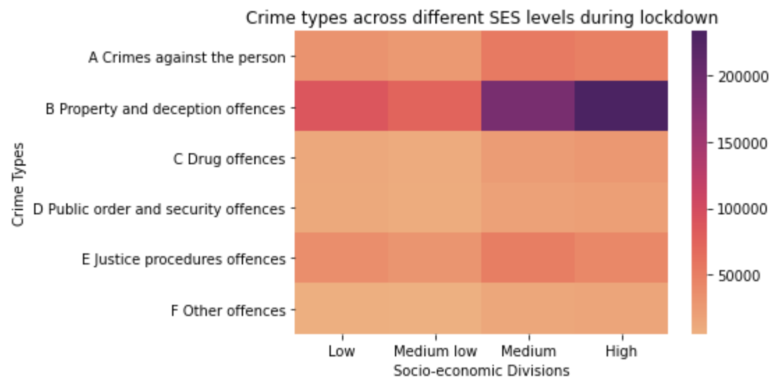


Fig. 11-12 Heatmaps of different crime types against socio-economic division

Among the six crime types examined, it was observed that the crimes committed during and out of lockdown remained the same across all SES levels as there is a negligible change in colour across the heatmap. Furthermore, it was observed that the crime type “property and deception offences” were most prominent among the four SES levels with the medium to high SES levels having upwards of 100,000 reported offences during both periods of lockdown and without lockdown.

Value and Significance of Results:

The expectation for our investigation was that lockdown measures would directly contribute to the recession of all crime types with the exception of online criminal activity, reason being that individuals were prohibited by law to leave their premises unless for essential purposes. Hence, there should theoretically be fewer counts for both individuals subject to crime, and also those with criminal intent out amidst the public. However, our findings have failed to support such a claim, where according to various visualisations, the fluctuations in crime are erratic and minuscule on the scale of our crime count, which could simply be the consequence of seasonal variation in crime. Furthermore, yet another explanation for this phenomenon could be that due to certain leeways in place which allow individuals to exit their homes, criminals are able to abuse such freedoms to continue illegal activities. Those individuals are also predisposed to break the law, thus various lockdown measures are merely ignored to fulfill their nefarious motives. Moreover, visible in the bar charts, type C and F crimes are one of the few crime types where there is a notable increase during lockdown in comparison to out of it. As aforementioned, due to covid restrictions, a large proportion of citizens must stay indoors for prolonged periods of time, meaning any online misconduct cases such as hacking, stalking or even drug dealing will be expected to climb. Consequently, during a pandemic or any similar crisis in the future, the authorities must allocate more resources towards combating cyber crime.

Following the expectations, we look at the significance of the fact that the state of lockdown does not help in deterring the crime rate. This is significant as the change in crime rates observed between periods of lockdown and no lockdown show that the increase or decrease in crime could well be due to other underlying variables as stated previously rather than the lockdown. Furthermore, it was observed that medium to low SES areas tend to have a higher crime rate as observed from the scatter plots above, with the exception of Melbourne being an outlier, where it has a high SES level as well as the highest crime rate in Victoria. Evidently authorities should allocate more personnel to these specific areas so that crime rate could be lowered overall. This holds true especially for Melbourne with it being a very busy as well as the most populated city in Victoria.

Limitations and Improvements:

The raw data obtained, particularly tables containing crime statistics, did not come without their limitations. Since the Crime Statistics Agency only made their data available on a quarterly basis, this hindered the ability to observe potential fluctuations that could have only been viewed on either a monthly or weekly basis. An example could be that during a given week, Victoria may have had a lockdown protest, which in turn would generate a spike in crime. However, given the periodicity of the data, such a spike would be lost once visualised. Another problem that arises from quarterly data is how it was applied to the lockdown timeline. Both lockdown start and end dates were misaligned, which would cause issues relating to the accuracy of analysis and results. This can be seen with the treatment of the December offence statistics, as it was included in the "Pre/Post" lockdown time period, yet lockdown extended into the month of October, meaning some reports would have occurred whilst Victoria was in lockdown. The nature of crime data relies on incidents to be recorded, and as such, may prove to be incomplete, which in turn, may lead to slight errors in the results produced. Additionally, factors which can affect changes in crime extend beyond both SES and the state of lockdown. This could ultimately have an effect on the correlation and causation of these factors. One such factor is seasonal variation, which could have been studied further by employing crime statistics from past years.

As for the socio-economic level statistics, the latest data we were able to obtain was from 2016. However, our investigation is based on 2020 when the first lockdown began. As a result, it hindered the accuracy of scatter plots and bar charts analysis. Additionally, due to the nature of data, we were not able to take other measures into consideration although they could have potential impacts on the crime rates. Those measures include mask requirements, social distancing, movement restrictions, and even economic measures.