COVID-19 and Canadian Crime Rates

Prerna Babbar

TABLE OF CONTENTS

Objectives
Problem Statement

Impacts of the Coronavirus outbreak on Canadian Crime in 2020 and future trajectories.

In this report, I reviewed and analyzed data surrounding Canadian crime in 2020 and cross-examine it with SARS-CoV-2 data sources. In doing this, I was looking for any relationship and to use this to predict future outcomes in crime related to the novel Coronavirus *(SARS-CoV-2)*.

The future outcome I was looking to predict specifically was the trajectory of crime or crimes depending on a rise or fall of Coronavirus cases in Canada. This may include the possible release of a vaccine in early 2021.

Hypothesis and Early Assumptions

Hypothesis:

With lockdowns going into effect, it was easy to say it may lead to a decrease in crime due to a decrease in opportunity. However, with the lockdown comes financial hardship. The gap between the wealthy and poor was growing, which may lead to an increase in crime as the needs of the people are not being met.

Initial Assumptions and Notes:

- With the data I are using, I are not incorporating outside factors that may be affecting crime. This may lead to a bias in the analysis, certainly creating a weakness.
- I are assuming a causational relationship initially. This may prove not to be true, or if it was true it may not be properly represented in the analysis outcome.
- Results may differ based on how I group and organize the initial data.
- Over the past few years, according to StatsCan, crime in Canada has been in decline over the past two decades. This points to the fact that outside factors may have a larger impact on the data than thought. A causal-relationship will be difficult to establish.

Data Preparation

Data Sources

a **Crime Statistics in Canada 2019-2020**
b **Crime Statistics Province-wise 2020**
c **COVID cases in Canada**   *(to 05.2020)*
d **COVID cases in Canada**   *(up-to-date)*
e **COVID cases in Ontario**

Assumptions and Data Cleaning

A   SOURCE: CRIME STATISTICS IN CANADA 2019-2020

Column: All

This section contains details of total criminal incidents in March-June of both 2019 and 2020 each in their own column. While the information was not broken down by city or province, I can still use it for time-series analysis of Criminal Incidents and Calls for Service in the early stages of the Covid-19 pandemic to verify our assumptions.

LIMITATION:

The comparison of June 2020 to the present date cannot be captured as the related database currently cannot be found. This would be a task for future analysis once the online database was updated.

SOLUTION:

I used the data source to perform our time series analysis, in which I compare the selected Criminal Incidents and Calls for Service data between both March-June of 2019 and March-June of 2020.

ASSUMPTIONS:

The Covid-19 pandemic lockdowns began in March of 2020. People's outdoor social activities have been limited so outdoor activity-related criminal incidents are assumed to reduce in number when compared to the same period of the previous year. This would include crimes such as (non-family related) assaults, sexual assaults, and robbery.

With many people working and studying from home, and most social outings impossible, it was assumed that domestic abuse and assault, domestic disturbances, and mental health services would be on the rise.

B   SOURCE: CRIME STATISTICS PROVINCE-WISE 2020

Column: Geo

Regions cover cities, provinces, and in some instances, Royal Mounted Police constabulary statistics.

LIMITATION:

Regions initially appeared to be provinces, however, upon further inspection, the regions seem to be

quite sporadic. Some information was included for Ontario regions as Ill but not all provinces are covered in the data set.

SOLUTION:

To begin, I decided to look at Ontario related reports. This was done by separating the data by city in order to select Ontario data. I also created a separate data frame to capture crime information for other provinces. The aforementioned was done by splitting the GEO data to find the province and city.

Column: Violations and Calls for Service

LIMITATION:

This included over 30 types of crime with totals mentioned as individual rows which skewed the data if not removed.

SOLUTION:

Types of crime Ire condensed and added into a new column: "Crime_Category". I also dropped all rows containing "total" in the column name.

Column: REF Date

LIMITATION:

This included over 30 types of crime with totals mentioned as individual rows which skewed the data if not removed.

SOLUTION:

Types of crime Ire condensed and added into a new column: "Crime_Category". I also dropped all rows containing "total" in the column name.

C   SOURCE: COVID CASES IN CANADA

Column: Date LIMITATION:

The data in this table was collected only up to May 2020, and the data on crime starts in May.

SOLUTION:

As a result, I decided to not use this data source. In our analysis, I use a more up-to-date version from Canadian statistics sources such as Data Ontario.

D   SOURCE: COVID CASES IN ONTARIO

Column: Reporting_PHU_City LIMITATION:

The Covid cases Ire reported by region, with many more regions than categories of crime in the data set.

SOLUTION:

As ultimately I want to merge this table with the Crime Statistics table, I created a new column which would rename the cities as per the regions mentioned in the crime data set for easier analysis.

At this point, as there are four regions in Ontario for which I have crime data, I may drop multiple cities or regions from the Covid table if they don't contain relevant crime information.

Column: Accurate Episode Date LIMITATION:

The data here was recorded daily, ie. each case was individually recorded. However, the crime table was sampled monthly.

SOLUTION:

I created a count column and set the value for each row as 1. I then resampled the data into monthly and added the count to get total cases per month.

Column: Accurate Episode Date LIMITATION:

While counting total COVID cases, I realized I Ire getting unrealistic numbers (ex. 200'000 cases in Alberta for the month of May).

SOLUTION:

On further examination, I noticed that the column closed from the file was the accumulative total number of cases on any given date. I then Int back to the file and found a column "Covid Region" which I used to resample the data for further analysis.

Column: Province

LIMITATION:

Apart from provinces, there was data recorded separately for Canada (total) and repatriated travellers.

SOLUTION:

By dropping these rows, I ensure that the data I use was province-specific.

Feature Creation

Downloading Crime Statistics for Canada (Provincially) and the analysis of various columns and entries. This data was collected betIen the months of May and August.

It was immediately realized there Ire too many unique values for the columns "Violations and Calls for Service". To work around this I converted them into the following types:

- Domestic
- Outside
- Mental Health
- Theft
- COVID
- Miscellaneous

There Ire two reasons for this:

- I. Ease of plotting/graphing
- II. Potential to check relationship with certain types of crimes with COVID cases

I also created a new column to capture Province information from the GEO column. Based on this I ended up with the following:

- Ontario
- British Columbia
- Manitoba
- Alberta
- Saskatchewan
- New Brunswick

Ontario had about half as many entries by itself as the rest of the provinces combined and was split further into 6 cities/regions. As a result, I separated this into two DataFrames:

- CrimeOntario

- CrimeOther
    I decided to run the analysis separately as I wondered if the high population in Ontario would skew the findings for the other provinces. I can always use the combined data frame at a later stage if required.

I created the following new columns:
- Cases
    - Copied from VALUE column with NaN filled with 0's. I wanted to keep the VALUE column in case of any errors at a later stage to keep the initial data unchanged.
- City
    - This was mainly designed for Ontario as there Ire many cities for which the criminal data Ire individually recorded.
- Province
    - To clean the GEO data
- Crime_Category
    - These Ire the initial buckets used to perform analysis. However, I realized the type of crime that may be directly impacted could be "domestic" and theft, while the one that could go down are the "outside" type due to lockdown, etc.
- Type
    - As explained above

Once I Ire happy with the data preparation on the Crime side, I continued to COVID data. For Ontario, I broke down the COVID data by CityRegion and renamed a few of the regions to ensure they matched the names on the CrimeOntario DataFrame. COVID Canada information was recorded by Province which I used to compare against crime in Other provinces (outside Ontario).

Conversion of all date columns into DateTime format and choosing them to be the index thus making the data frame TimeSeries. This was done to help us check the progress over time.

The regions in Ontario that I finally ran an analysis on are as follows:
- Waterloo Region (Kitchener)
- Halton Region (Oakville/Burlington)
- York Region
- Ottawa
- London
- Toronto

The provinces (outside Ontario) that I finally ran an analysis on are as follows:
- British Columbia
- Manitoba
- Alberta
- Saskatchewan
- New Brunswick

Analysis

1. **Reviewing the data on the following four tables and making sure they are ready for analysis. The "Data Preparation" section goes over this process.**

2. **Checking individual trends in each of the data frames.**

   a. *Selected police-reported crime and calls for service during the COVID-19 pandemic, March to June 2020 (correction)*

- The scatterplot below shows a strong positive-linear association between the amount of crime and calls for service in March to June of 2019 and the numbers in March to June of 2020. That means the crime and calls for service have the same seasonal trends during the reported period.
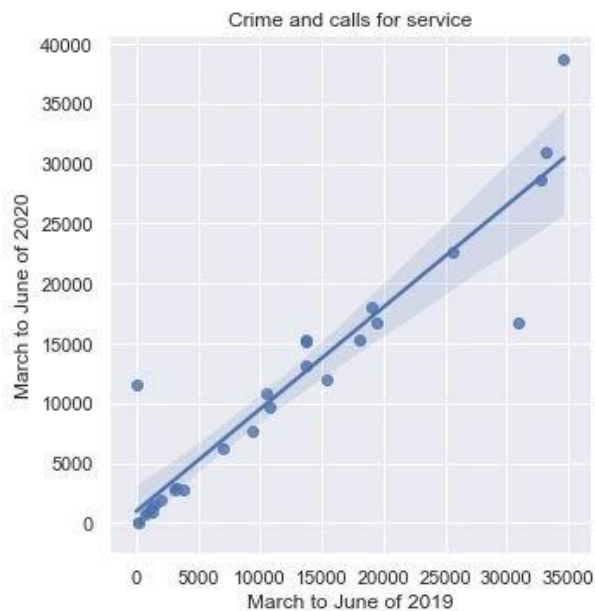
- The graph below shows the numbers of selected violations and calls for service by each category, in which the percentage of differences was over 10% compared to the number of March to June of 2019.

Violations and Calls for services of March-June 2019 vs March-June 2020
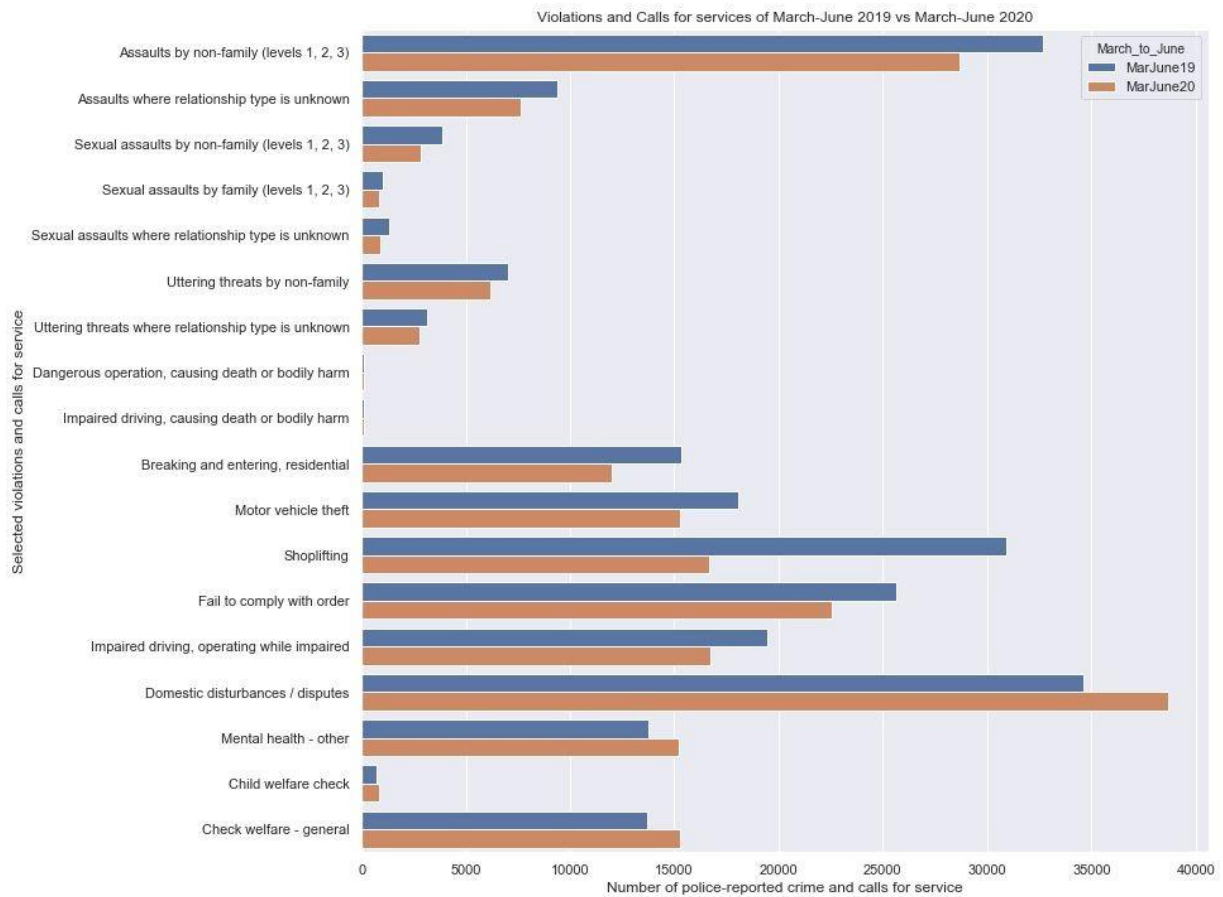
- The graph below shows the percentage that indicates the differences in selected violations and calls for service from March to June 2020 compared to March to June 2019.


Difference (%) of crime and calls for service during Covid-19 Pandemic
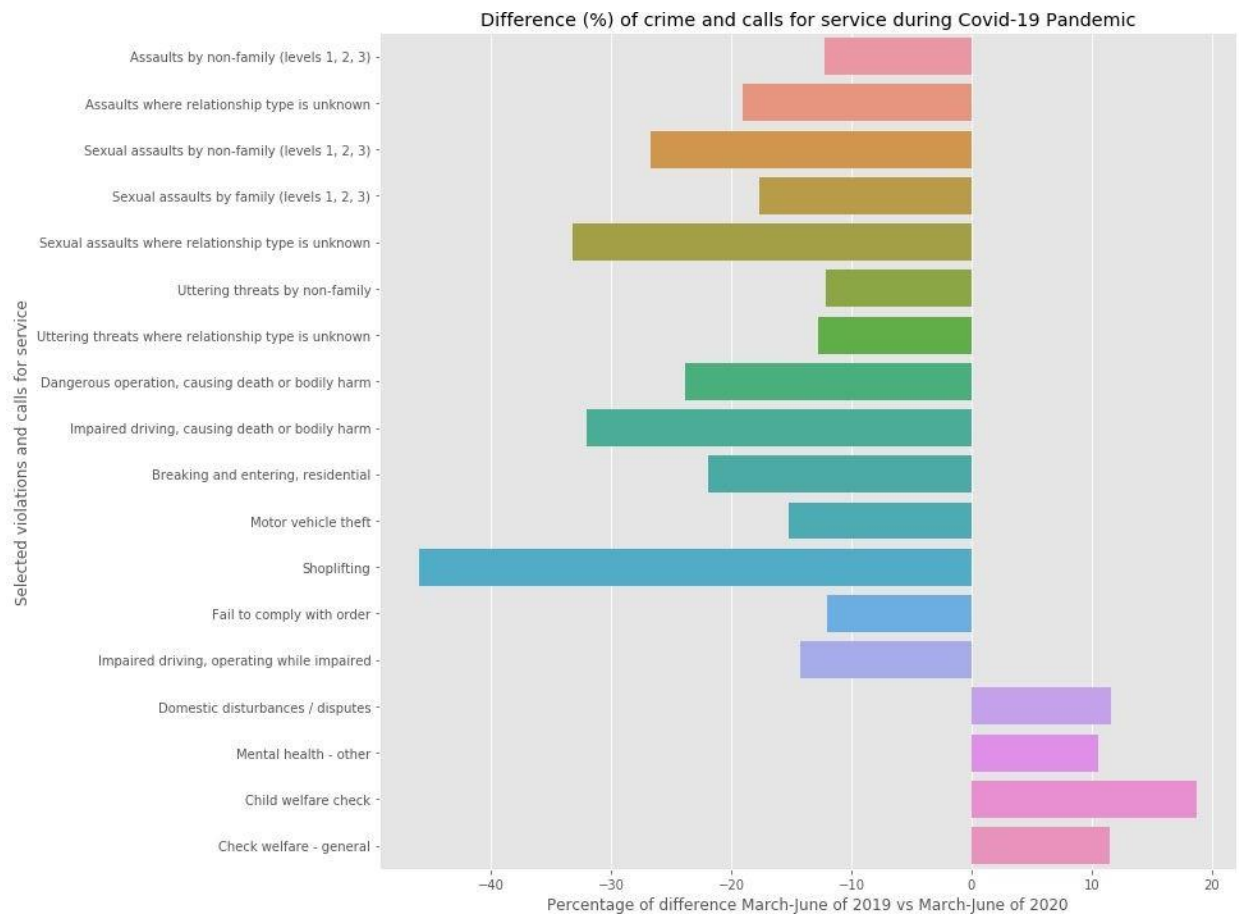
- The negative percentage on the left means the rates are going down over 10%, while the right side means the rates are going up compared to the data of March to June 2019. For example, the criminal incidents rates of assaults, sexual assaults, robbery, and other similar instances decreased.
- This verifies our assumption that outdoor activity-related criminal incidents would be going down during the pandemic.
- In addition, the rates of domestic disturbance/disputes within families, and mental health-related services Ire going up since people Ire working at home remotely and students Ire taking classes remotely, the majority of people's daily activities switched to indoor. This also matches our assumptions.
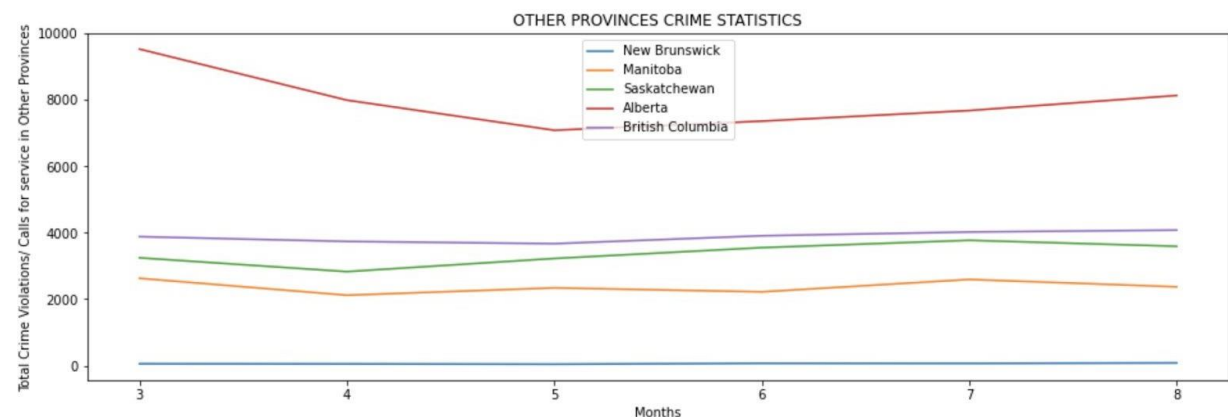
- The subtotal of categorized criminal incidents and calls for service are listed in the graph below, which shows the obvious decrease in assaults and robbery, and increase in Domestic disturbances/disputes and Mental health during the pandemic.
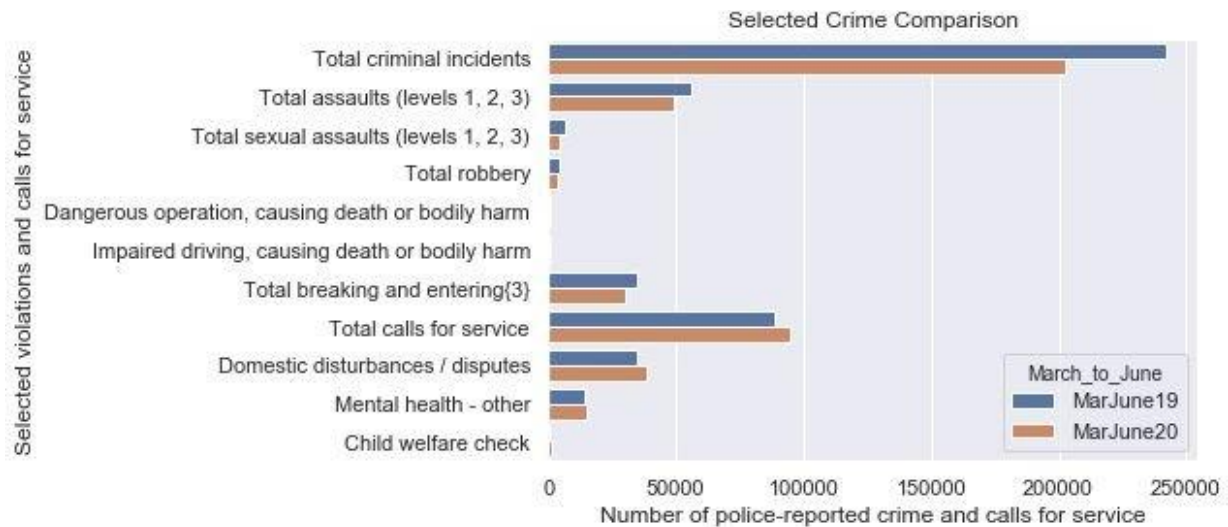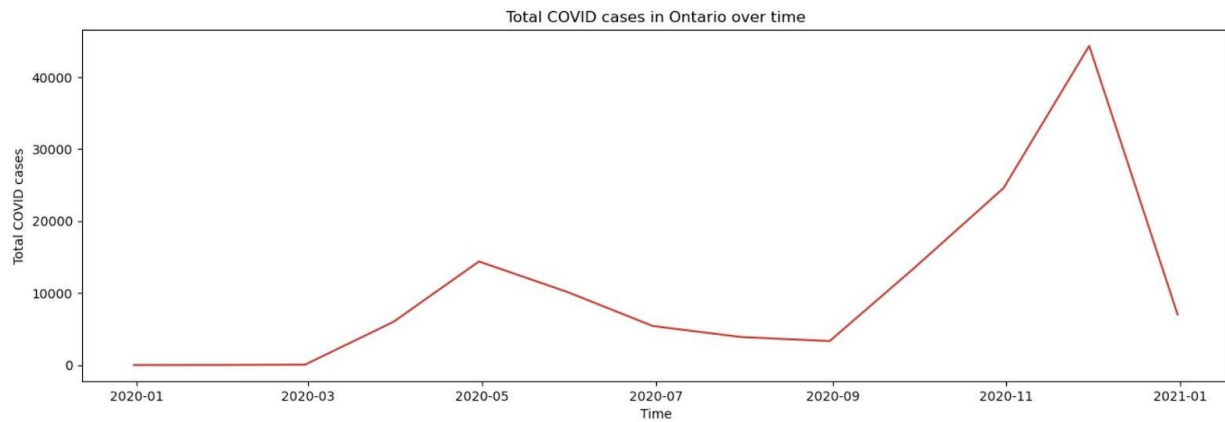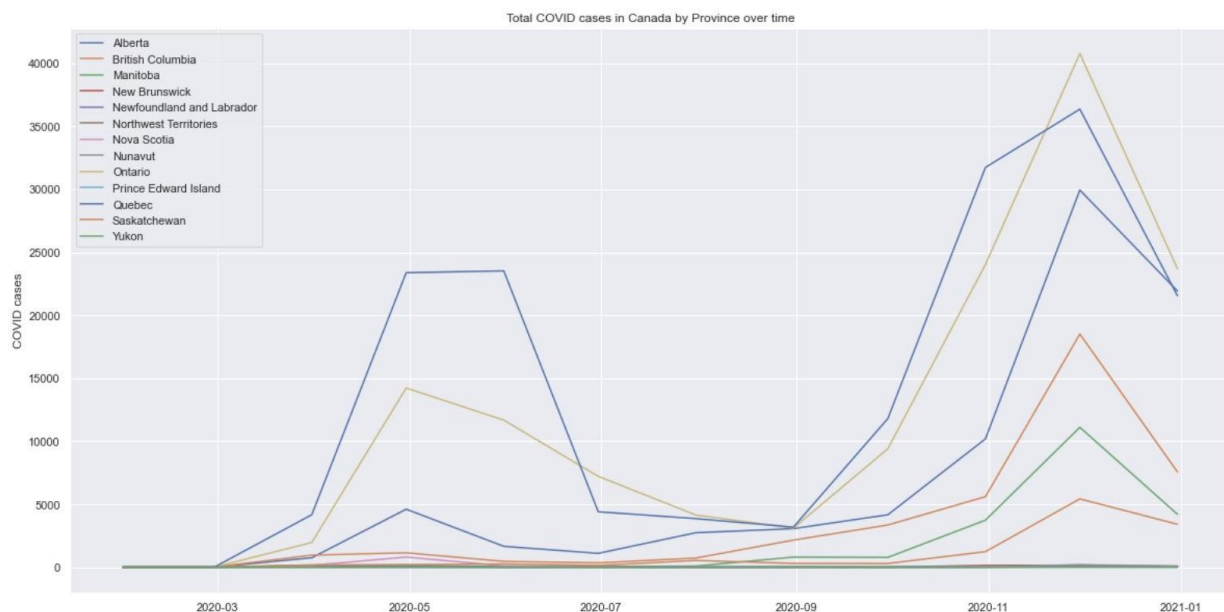

b. *Crime in Ontario over time*
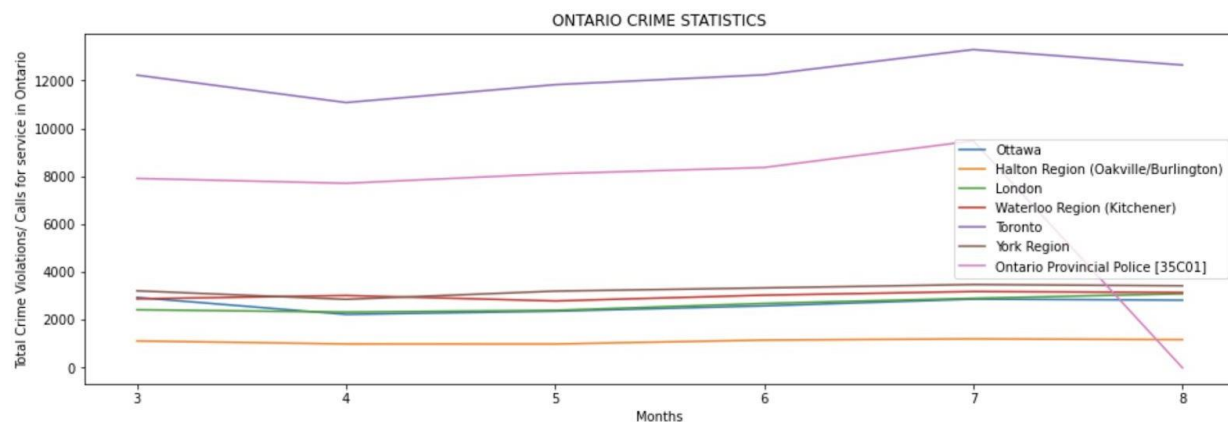
c. *Crime in other Provinces over time*

d. *COVID spread in Ontario over time*

Total COVID cases in Ontario over time

e. COVID spread in other Provinces over time

Total COVID cases in Canada by Province over time

3. I then decided to merge the CrimeOntario and CovidOntario using an inner join which would mean that only matching records would show up.

This meant that, while the COVID Ontario data frame had cases recorded until the end of the year, I Ire analyzing data only betIen May and August as this was the time period that was included in the Crime Ontario data frame. Additionally, I only analyzed the cities that Ire mentioned in both data frames.
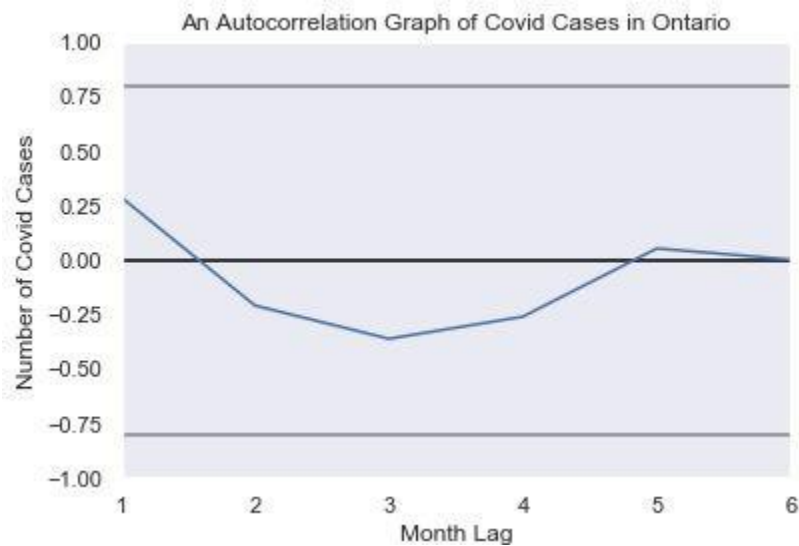
a. Analysis correlation betIen Covid cases and Crime numbers in Ontario

- The correlation coefficient (-0.9498778810318722) has been calculated using the monthly number of covid cases and the number of criminal incidents and calls for service. The scatterplot shows a strong, negative, linear association betIen them, where the R-squared value was 0.902, P value was 0.004, which means the relationship
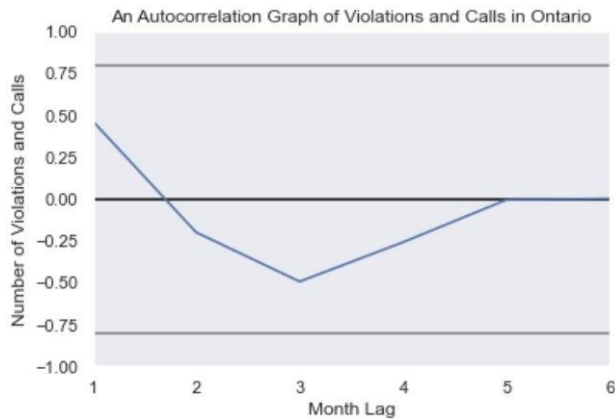
11

betlen covid cases and crime numbers was statistically significant, as covid cases increase, the number of crimes decreases in Ontario.

Covid Cases and Criminal Incidents - Ontario March to August of 2020
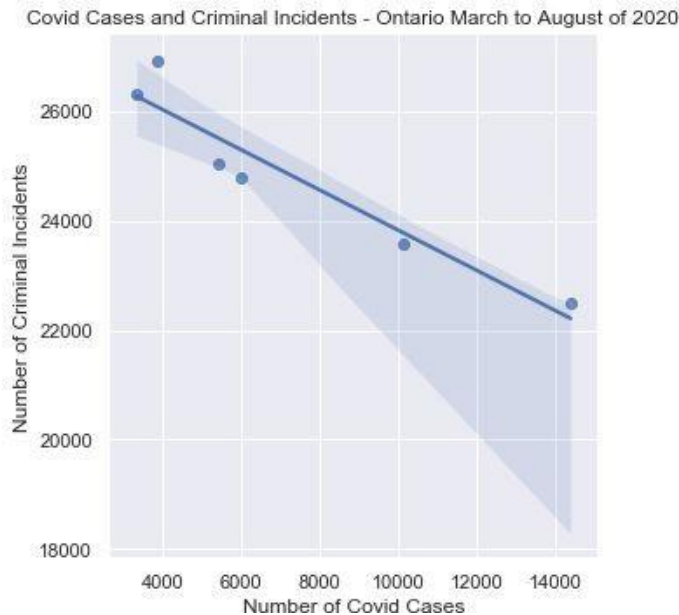
- A positive autocorrelation present and future values of the time series move in the same direction, whereas negative values mean that present and future values move in the opposite direction. The function value range was [-1,+1]. The horizontal lines displayed in the plot correspond to 95% and 99% confidence bands.

- The graph of autocorrelation below measures the linear relationship between the number of Covid cases by month (lagged values of a time series). They are autocorrelated, but not significant. The positive values are shown at the beginning, which means the present and future values of this time series move in the same direction. After that, the negative values are shown in the curve, which means the present and future values of the time series move in the opposite direction.

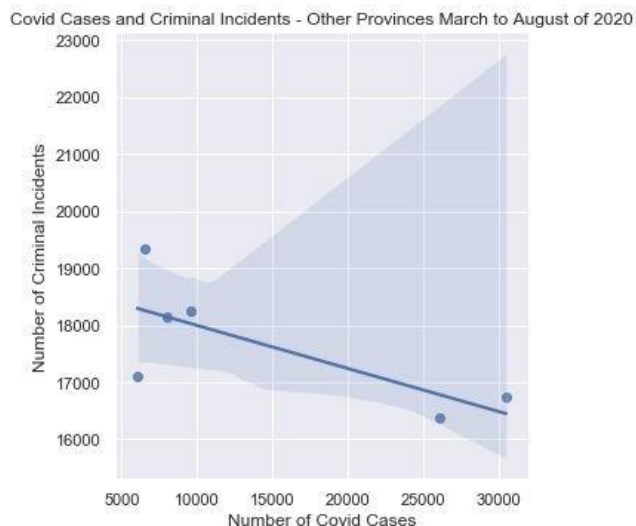
An Autocorrelation Graph of Covid Cases in Ontario

- The autocorrelation below measures the linear relationship between the number of Violations and Calls by month (lagged values of a time series). They are autocorrelated, but not significant. The positive values are shown at the beginning, which means the present and future values of this time series move in the same direction. After that, the negative values are shown in the curve, which means the present and future values of the time series move in the opposite direction.
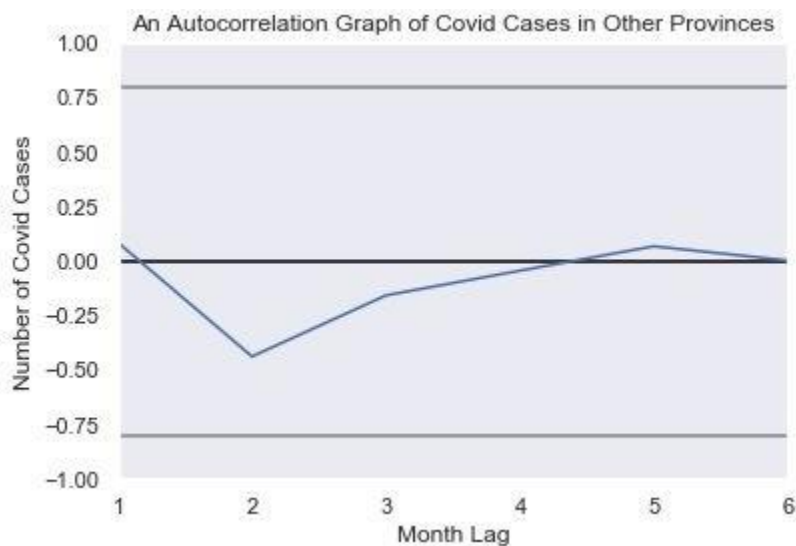

An Autocorrelation Graph of Violations and Calls in Ontario

b. _I repeated the same process for CrimeOther and CovidOther._

- I resampled both CrimeOntario and CrimeOther "monthly" and generated a plot to check the total number of criminal cases over time just to have an initial look. I did this for both Ontario and other provinces separately and saw different trends indicating it was right for us to separate analysis as the population could have had an impact on crime.
- The correlation coefficient (-0.7339597270635329) has been calculated using the monthly number of covid cases and the number of criminal incidents and calls for service. The scatterplot shows a moderately strong, negative, linear association between them, where the R-squared value was 0.539, P value was 0.097, which means the relationship covid cases and crime numbers was significant at 0.1 level, as covid cases increase, the number of crimes decreases in other provinces, but it was not significant as it in Ontario.

Covid Cases and Criminal Incidents - Other Provinces March to August of 2020
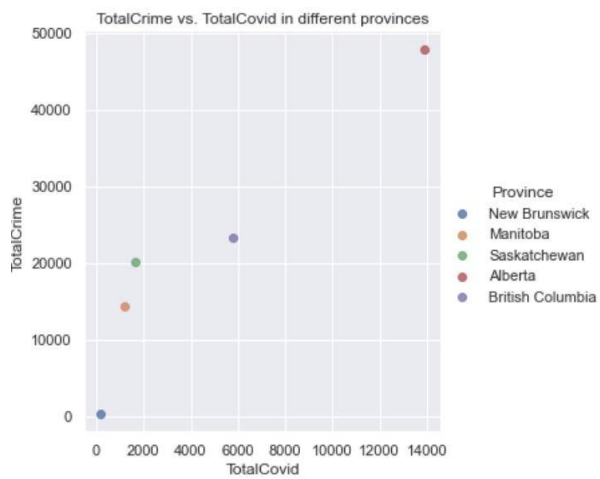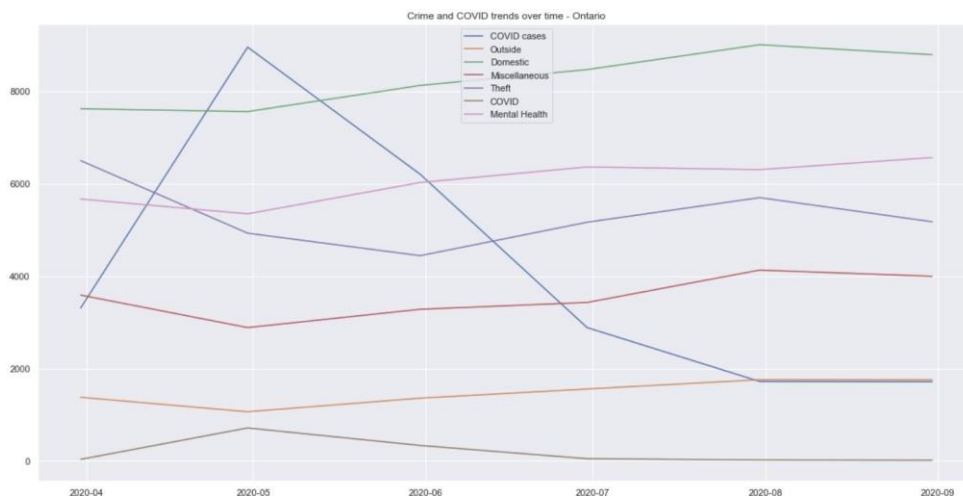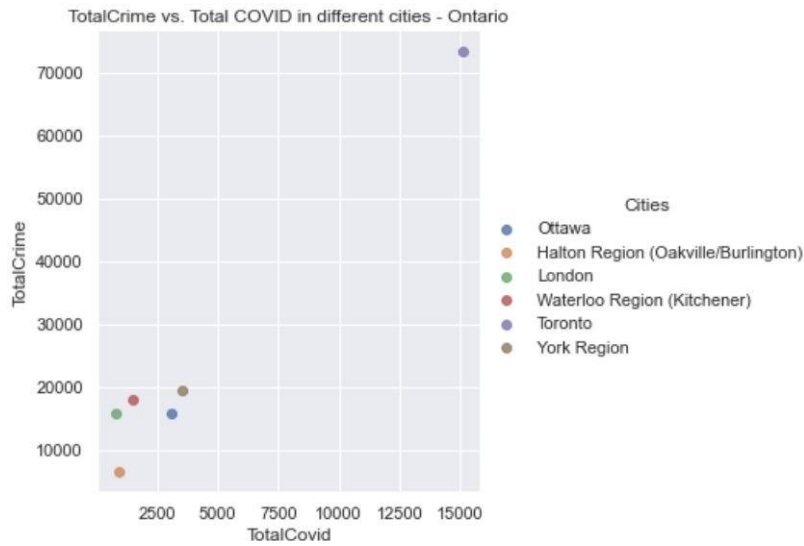
13

- The graph of autocorrelation below measures the linear relationship between the number of Covid cases by month (lagged values of a time series) in other provinces. They are autocorrelated, but not significant. The negative values are shown in the curve, which means the present and future values of the time series move in the opposite direction. While the positive values are shown in the curve after that, which means the present and future values of the time series move in the same direction.

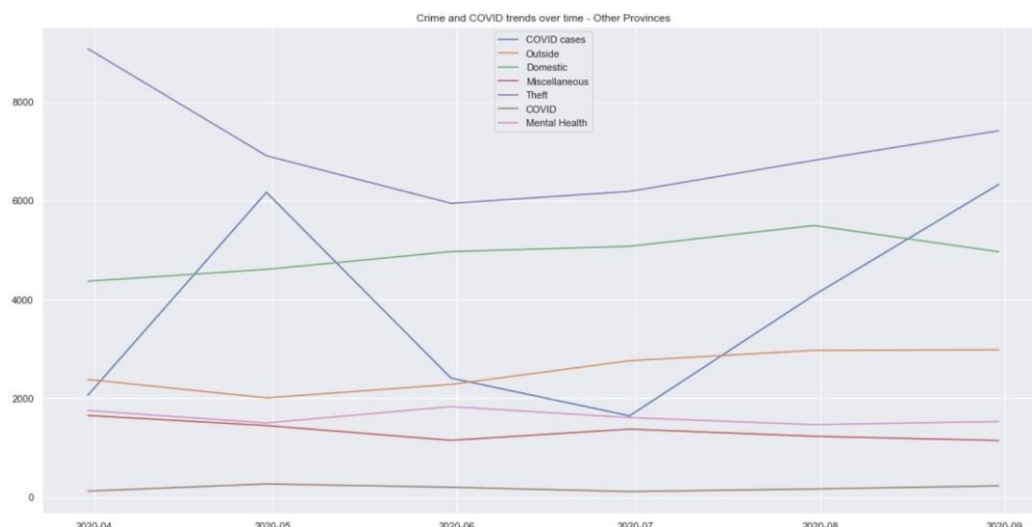An Autocorrelation Graph of Covid Cases in Other Provinces

- The autocorrelation below measures the linear relationship between the number of Violations and Calls by month (lagged values of a time series) in other provinces. They are autocorrelated, but not significant. The negative values are shown in the curve, which means the present and future values of the time series move in the opposite direction. While the positive values are shown in the curve, which means the present and future values of this time series move in the same direction.

An Autocorrelation Graph of Violations and Calls in Other Provinces

14
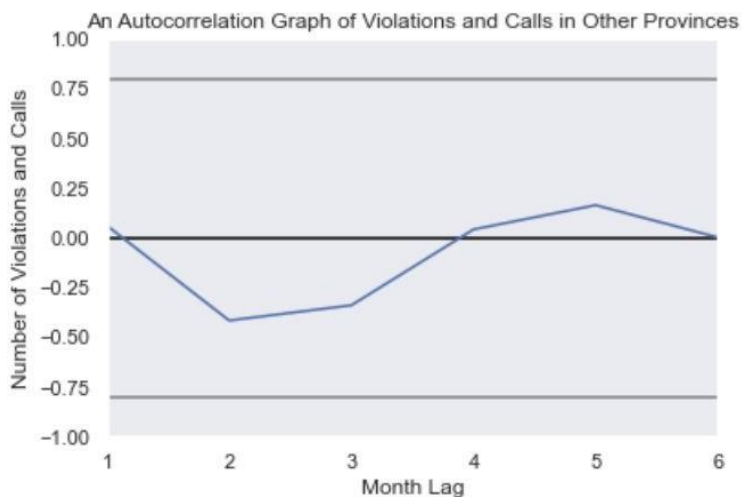
- I created graphs to visualize the trends in COVID and Crime over time and by region for both Ontario and other provinces.

TotalCrime vs. Total COVID in different cities - Ontario

Crime and COVID trends over time - Ontario

TotalCrime vs. TotalCovid in different provinces

Crime and COVID trends over time - Other Provinces

- I ran both OLS and SciKit regressions models to test the following:
  - Effect of time on COVID and Crime - Here I saw a slight negative correlation as mentioned above. Many types of crime Int down as a result of lockdown.
  - Effect of region and COVID on Crime - Here I saw a strong positive correlation. On deeper study, I realized that population and demographics played a large role in not just Crime, but COVID itself.

Conclusions

- There was a definite increase in the amount of crime in Canada in 2020 as compared to 2019.
- Crimes that relate to outdoor-activity were lower indicating a possible impact of COVID due to the lockdown.
- Crimes that relate to indoor activity, such as mental health calls and domestic disturbances, have increased during COVID. I can attribute this to the negative psychological impact of the lockdown, financial stress and also the weariness resulting from this pandemic.
- I ran linear regression models to find a relationship between Crime data and Covid data and found that criminal activity was definitely impacted by Covid.
- Based on the OLS model that I ran over time, I can see that there was a slight negative correlation between COVID and Crime. This could be due to the decrease in outdoor crimes and less calls for service.
- When I ran the regression model by region, I noticed that there was a strong positive correlation between COVID and Crime. On deeper study, I realized that this was due to other factors, including population. The demographics of each city/region affects both crime and COVID.

End Summary

Taking data from Canadia and Ontario sources, I compared the amount of crimes with covid cases. After cleaning the data I found there to be a definite increase in crime compared to 2019 before the pandemic hit. While crimes that Ireoutside
opportunity-based decreased, the increase came mainly from crimes and calls for service
that related to domestic disputes as Ill as mental health calls. This was a clear correlation between crime and Covid cases.

After running linear regression models, I found a definite relationship between the data on crime and the data on Covid.

Based on our OLS model that I ran over time, I saw a slight negative correlation between Covid and crime, possibly due to the decrease in outdoor crimes and calls for service.

After running our regression model by region, I noticed a strong positive relationship between Covid and crime as mentioned above. However, after deeper study I realized that this was due to other factors. This includes population as the demographics of each city or region affects both crime and Covid significantly.

As addressed in our initial notes and assumptions, there are many factors that can affect crime numbers. While I began working with it in mind that Covid would be our independent variable and crime our independant, it was important to recognize that both variables are dependent on a variety of other factors. Sometimes they are reliant upon the same factors such as population size or density.

It was easy to say that a decrease in Covid cases may ultimately lead to a decrease in overall crime due to the nature of the majority of crimes that are being committed. That was not to say that Covid would be the only, or even largest factor that would have an effect in the projection of crime.