Interim Report

Capstone Project

Toronto Crime Analysis

Submitted By

Dikeshkumar Patel- 0794600

Parthkumar Dhameliya - 0784114

Kajal Arora – 0774848

Amandeep Kaur - 0793780

Kanishk Tripathi - 0789972

Under the guidance of

Manoop paliot

Contents:

Toronto Crime Analysis

History:

Canada's largest city, Toronto, with a population of about 2.7 million. The 1980s and 1990s saw a spike in gang-related violence in Toronto, especially in the housing projects of the city. A rise in organized crime activities was also observed in the city, particularly in the areas of drug trafficking and money laundering.

Beyond the year 2000 - Gun violence has increased recently in Toronto, especially in the less affluent areas of the city. In order to address the underlying causes of gun violence and develop strategies to avoid it, law enforcement and community organizations are now putting in greater effort.

Executive Summary:

Crime is a major issue in Toronto, and understanding its pattern and trends is essential to effectively address it. Toronto crime analysis is the process of analysing data related to crime in order to discover patterns and trends.

In this presentation, we will explore the types of crime, the trends, and the strategies to reduce crime in Toronto.

Toronto has many types of crime, ranging from violent crimes such as homicide and assault, to property crimes such as theft and vandalism.

The most common types of crime in Toronto are property crimes, followed by violent crimes and drug-related crimes.

The scope of the project would likely involve collecting and analysing crime data, including information on the types of offenses, locations, time and date, and demographics of offenders and victims.

The analysis may also involve identifying hotspots or areas with high crime rates, analysing the effectiveness of existing crime prevention programs, and exploring the relationship between crime and other factors such as social and economic conditions.

The overall objective of a Toronto crime analysis project would be to provide politicians, law enforcement organizations, and community organizations with insights and suggestions for developing evidence-based measures to lower crime and enhance safety for city residents and visitors.

Progress Update:

We initially chose to focus on US crime analysis, but it didn't work out, so we switched to Toronto crime analysis. We pick a dataset that includes information on crime in Toronto. We verify that the dataset is trustworthy, pertinent, and current. We look for information from sources including Statistics Canada, the Data Library at the University of Toronto, and the Open Data portal of the Toronto Police Service.

We make sure the dataset is correct and consistent and clean it. This could entail getting rid of duplicates, adding values where there are blanks, and fixing mistakes or discrepancies.

After the data has been cleansed, We begin to investigate the dataset. Analysing the data's variables and distributions is necessary for this. We analyze some patterns or trends you think might be interesting.

We make graphs and charts based on your exploration to help you see the facts. You can use this to spot patterns and trends that might be hard to spot in the raw data. Bar charts, line graphs, scatterplots, and heat maps are a few types of graphs we made using our dataset to ensure our project is successful.

We are now working on creating and hosting the website, and after that is done, we will work on the dataset for the online presentation.

Key Progress:

We worked on datasets and information gathering on our project. Our key findings for this project are listed below:

Dataset selection and gathering of data on Toronto crime. We make sure your data is organized and clear, and study the dataset to find any patterns or trends. This may help in the accuracy and refinement of the machine learning model.

Model evaluation and improvement: Assess the effectiveness of our machine learning model and make some necessary improvements. To fine-tune the model's parameters and raise its accuracy, we use employ strategies like Random Forest, Classifier grid search or cross-validation.

We face some challenges in the beginning on our dataset and topic like we didn’t get sponsors for our project and didn’t get enough information which make our project successful. So we decided to change the topic to Toronto crime analysis which overcome our project challenges.

We use user testing to get feedback and enhance the user experience on our Figma prototype. By doing this, we are better able to spot any usability problems or user-experience problems and make any necessary design changes.

Data Gathering and Metrics:

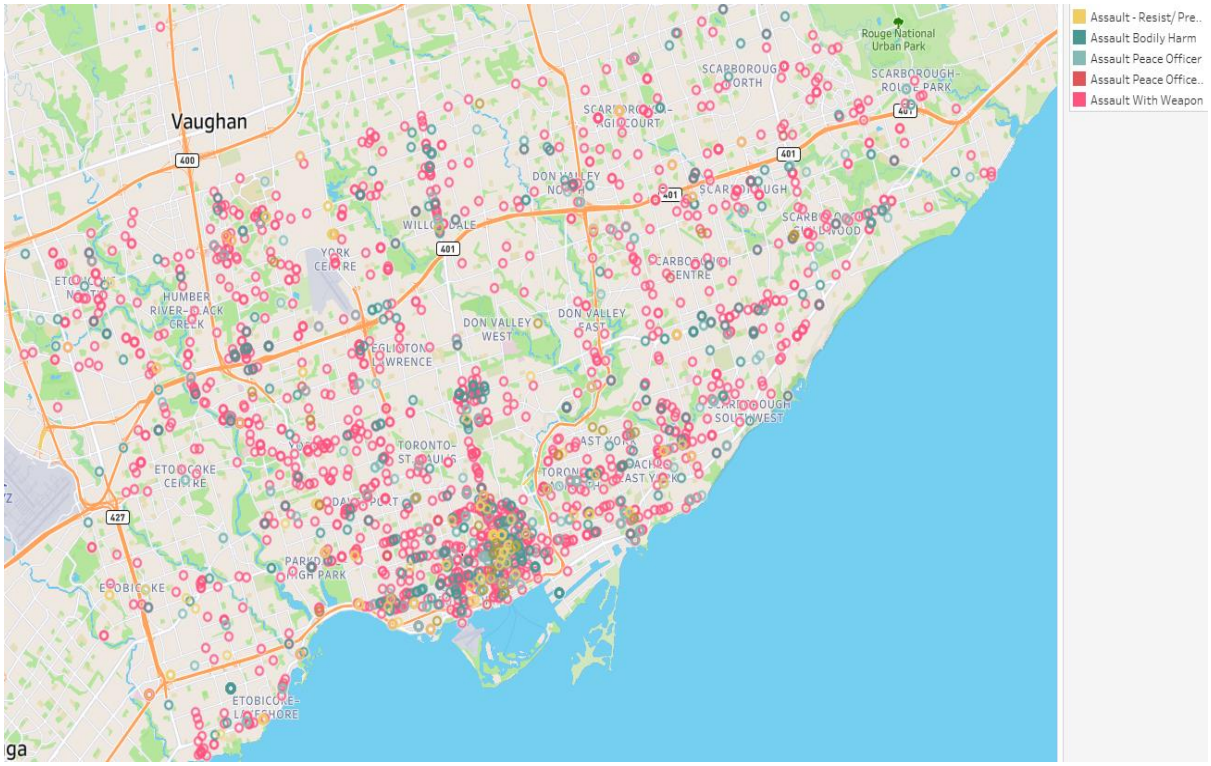
Index_	event_uni	Division	occurrenc	reporteddi	location_t	premises_	ucr_code	ucr_ext	offence	reportedy	reportedm	reportedd	reportedd	reportedd	reportedh	occurrenc	occurrenc	occurrenc	occurrenc	occurrenc	occurrenc	mci_categ
201	GO-20141	D31	2014/01/0	2014/01/0	Apartment	Apartment	1430	100	Assault	2014	January	3	3	Friday	11	2014	January	3	3	Friday	11	Assault
202	GO-20141	D42	2014/01/0	2014/01/0	Single Hon	House	2120	200	B&E	2014	January	3	3	Friday	14	2014	January	3	3	Friday	14	Break and
203	GO-20141	D22	2014/01/0	2014/01/0	Open Area	Outside	1430	100	Assault	2014	January	3	3	Friday	13	2014	January	3	3	Friday	13	Assault
204	GO-20141	D53	2014/01/0	2014/01/0	Other Com	Commerci	2130	210	Theft Over	2014	January	3	3	Friday	13	2014	January	3	3	Friday	12	Theft Over
205	GO-20141	D22	2014/01/0	2014/01/0	Convenien	Commerci	1610	210	Robbery -	2014	January	3	3	Friday	16	2014	January	3	3	Friday	14	Robbery
206	GO-20141	D51	2014/01/0	2014/01/0	Convenien	Commerci	2120	200	B&E	2014	January	3	3	Friday	11	2014	January	2	2	Thursday	23	Break and
207	GO-20141	D33	2014/01/0	2014/01/0	Other Com	Commerci	1430	100	Assault	2014	January	3	3	Friday	15	2014	January	2	2	Thursday	22	Assault
208	GO-20141	D14	2013/12/3	2014/01/0	Parking Lo	Outside	2132	200	Theft From	2014	January	3	3	Friday	18	2013	December	31	365	Tuesday	11	Theft Over
209	GO-20141	D13	2014/01/0	2014/01/0	Apartment	Apartment	1430	100	Assault	2014	January	3	3	Friday	17	2014	January	3	3	Friday	17	Assault
210	GO-20149	D11	2014/01/0	2014/01/0	Apartment	Apartment	2120	220	B&E W'Int	2014	January	3	3	Friday	20	2014	January	3	3	Friday	20	Break and
211	GO-20141	D42	2014/01/0	2014/01/0	Single Hon	House	2120	200	B&E	2014	January	3	3	Friday	17	2014	January	3	3	Friday	15	Break and
212	GO-20141	D42	2014/01/0	2014/01/0	Single Hon	House	2120	200	B&E	2014	January	3	3	Friday	19	2014	January	3	3	Friday	13	Break and
213	GO-20141	D12	2014/01/0	2014/01/0	Open Area	Outside	1430	110	Assault - F	2014	January	3	3	Friday	14	2014	January	3	3	Friday	14	Assault
214	GO-20141	D11	2014/01/0	2014/01/0	Apartment	Apartment	1460	100	Assault Pe	2014	January	3	3	Friday	20	2014	January	3	3	Friday	20	Assault
215	GO-20141	D43	2014/01/0	2014/01/0	Single Hon	House	2120	200	B&E	2014	January	3	3	Friday	18	2014	January	3	3	Friday	18	Break and
216	GO-20141	D14	2013/12/1	2014/01/0	Apartment	Apartment	2120	200	B&E	2014	January	3	3	Friday	17	2013	December	18	352	Wednesda	19	Break and
217	GO-20141	D53	2014/01/0	2014/01/0	Single Hon	House	2120	200	B&E	2014	January	3	3	Friday	20	2014	January	3	3	Friday	20	Break and
218	GO-20141	D32	2014/01/0	2014/01/0	Apartment	Apartment	2120	200	B&E	2014	January	3	3	Friday	21	2014	January	3	3	Friday	8	Break and
219	GO-20141	D42	2014/01/0	2014/01/0	Unknown	Other	1430	100	Assault	2014	January	3	3	Friday	23	2014	January	2	2	Thursday		Assault
220	GO-20141	D51	2014/01/0	2014/01/0	Commerci	Commerci	1430	100	Assault	2014	January	3	3	Friday	13	2014	January	3	3	Friday	11	Assault
221	GO-20141	D52	2014/01/0	2014/01/0	Commerci	Commerci	2130	210	Theft Over	2014	January	3	3	Friday	15	2014	January	3	3	Friday	14	Theft Over
222	GO-20141	D52	2014/01/0	2014/01/0	Streets, Rc	Outside	1430	100	Assault	2014	January	3	3	Friday	23	2014	January	3	3	Friday	22	Assault
223	GO-20141	D42	2014/01/0	2014/01/0	Bar / Rest	Commerci	1420	100	Assault Wi	2014	January	3	3	Friday	5	2014	January	3	3	Friday	5	Assault

Here is our Clean dataset of Toronto crime. We choose this dataset from the official website of the Toronto Police Service. Here I am providing the link to the dataset.

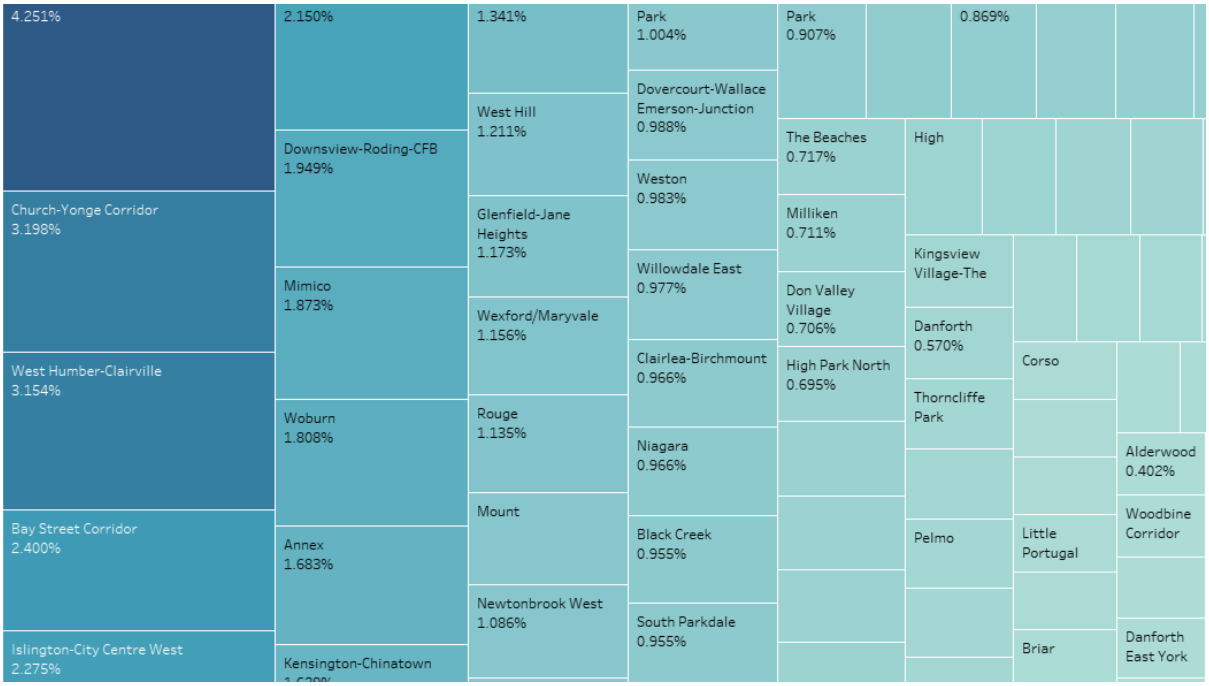
Link: https://data.torontopolice.on.ca/datasets/TorontoPS::major-crime-indicators-open-data/explore

- Toronto Major crime from 2014-06-01 to 2022-06-30
- It has 27 Column
- 0.3 Million Records in the dataset

We use Tableau to check if the dataset records work with the visualization so we made some graph, here I am providing some graph that shows the crime rate in Toronto city with what weapon used during the crime.



This map shows the type of crime that happens in some places.



This chart shows the crime that happens in street with the crime rate.

Changes Made During Project:

We started our project on US crime, but it didn't work well we didn't get more information and we didn't get a dataset on US crime, So we started searching for other datasets and topics then we finalize our project topic on Toronto crime analysis and our respected professor Manoop Paliot continues to guide us on how we can create this project on the effective way with Canadian location and analysis.

We choose Toronto crime analysis after researching crime in Canada. Give the feasibility testing, academic portion, and documentation portion new work. Therefore, it takes time, and our entire team is working on it.

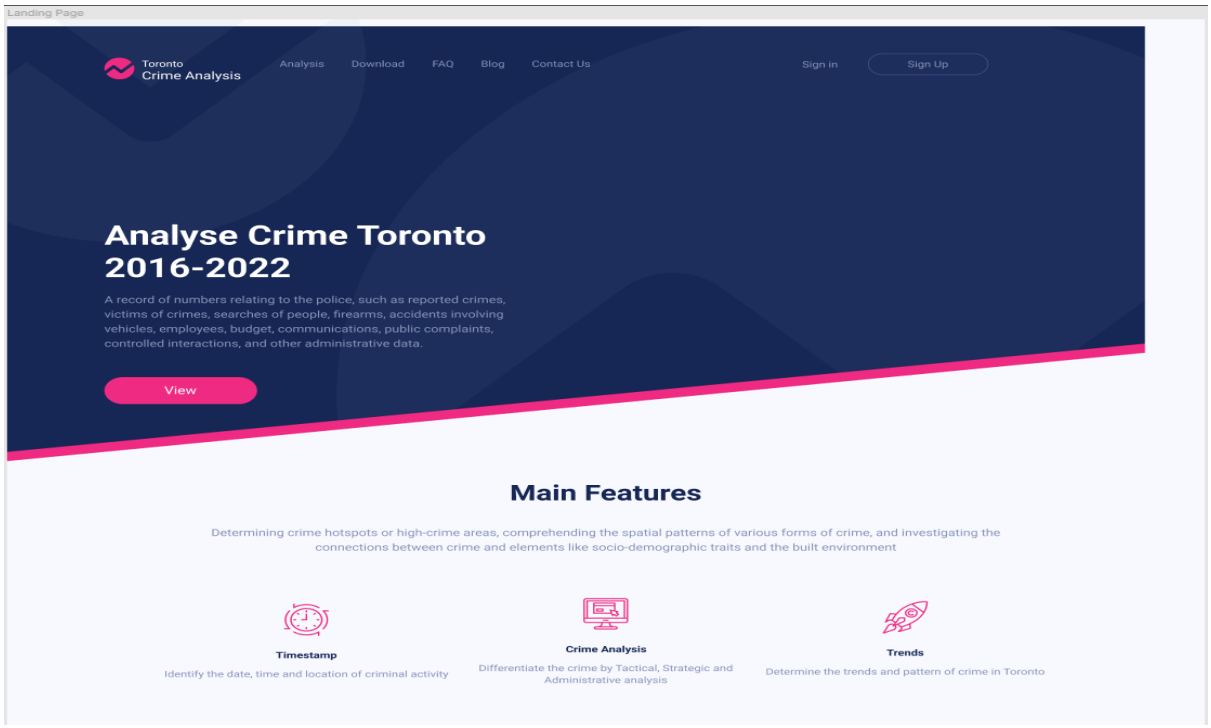
Methodology:

In the first step in building the Toronto crime analysis project, We started using some Tools and Technology to prepare a graphical presentation and working on the tableau to create the visualization.

We use Figma to create the prototype of the first page of our website. We train the model in Jupiter notebook, we try coding part in it, and tested the dataset which gives accuracy to use this dataset for further use in the future. We will create our project in WordPress which will give an analysis of Toronto crime.

We selected the website-creating platform and also got the hosting site from which we can access our dataset from this will connect our dataset from the hosting site to WordPress. Here I am providing the Figma prototype which shows the first page of our project how it will look like.

Figma Prototype:



ML model (Train and Test):

```
#Balanced Class Weight doesn't make a big difference for results:

classifier = RandomForestClassifier(n_estimators = 100, criterion = 'entropy', random_state = 42, class_weight='balanced')
classifier.fit(X_train, y_train)
y_pred = classifier.predict(X_test)
print(accuracy_score(y_test, y_pred)) #accuracy at 0.63
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test,y_pred, target_names=definition_list_MCI))

#------------------------------------------------#


grad_class = GradientBoostingClassifier(learning_rate=0.1,n_estimators = 10, random_state = 42)
grad_class.fit(X_train_OH, y_train_OH)
y_pred_OH = grad_class.predict(X_test_OH) # Predicting the Test set results

print(accuracy_score(y_test_OH, y_pred_OH)) #modest improvement to 0.648
print(confusion_matrix(y_test_OH, y_pred_OH))
print(classification_report(y_test_OH,y_pred_OH, target_names=definition_list_MCI))
```

```
0.64636882077894
[[34058  2848    79   805  1859]
 [ 7221  6816   102    90   576]
 [ 1604   489    30    55   244]
 [ 3737   374    19  2452   522]
 [ 4330   877    38   299  4474]]
                precision    recall  f1-score   support

       Assault       0.67      0.86      0.75     39649
Break and Enter      0.60      0.46      0.52     14805
     Theft Over      0.11      0.01      0.02      2422
        Robbery      0.66      0.35      0.45      7104
     Auto Theft      0.58      0.45      0.51     10018

       accuracy                           0.65     73998
      macro avg      0.52      0.42      0.45     73998
   weighted avg      0.62      0.65      0.62     73998
```

Result:

In Toronto, some areas have greater rates of violent crime than others, with the downtown area seeing the majority of these incidents.

In Toronto, there is a significant link between poverty and crime, with greater crime rates in areas with higher poverty rates.

These findings are important because they can guide targeted crime prevention strategies and policies that are tailored to the particular characteristics of Toronto. For example, robbery and auto theft are crimes that are more likely to occur in particular locations and at particular times of the day.

These findings are significant because they help guide focused crime prevention programs and policies that are made to fit Toronto's particular needs.

A crime analysis project in Toronto can contribute to the field by generating fresh ideas and information regarding the nature of crime in cities, as well as by pointing up potential areas for further investigation and action.

The output of a machine learning model includes numerous sorts of crime as well as precision, recall, an F-score, and support level. Overall models yield a score of 0.65, whereas individual MCI yield accuracy ranging from 0.78 to 0.66. We use test size 0.25 in our model.

Right now, our team is working on developing a website. This is our final product for the capstone project.

Additionally, a crime analysis project can support the development of trust and involvement between policymakers, law enforcement organizations, and the community by employing rigorous data

analysis techniques and disseminating the findings in a transparent and approachable manner.

Budget and Timeline:

In our project, we have a timeline of how and when we do complete our work. So here is the timeline of our project.

Week 1- 4:

We make a group with selected members. We started searching for project topics and datasets on a project. We finalize the topic of US crime and gather some information. All members work on data gathering for the topic.

Week 5- 9:

We didn't get enough information on the US dataset and didn't get information, so change the project topic to Toronto crime analysis. Our team started searching for project information, where we finalize a new dataset on Toronto crime. This week we distributed work to all members, where we did a feasibility study and documentation work.

Week 9- 13:

This week we completed half of the work, and we started working on a prototype. For the prototype, we use Figma which shows what our first page looks like. Then we started working on the coding part to create an ML model in Jupiter notebook to check the dataset train and test to work in our final project on our website. We finalized the website platform and hosting site.

For our Project Budget, we calculate the cost of the website creation part and hosting platform. For our Website, it will cost around 400 Dollars to build our project.

Future Work (Next step):

While crime statistics are significant, it may also be helpful to take into account social and economic issues that influence crime rates, such as poverty, unemployment, and inequality. These elements could contribute to a more detailed understanding of the underlying causes of crime.

Evaluation of interventions: Following their implementation, interventions should be assessed for their efficacy in lowering crime rates. This could involve comparing the crime rates in areas with and without the intervention using a quasi-experimental design or a randomised controlled trial.

Predictive analysis: By using predictive analysis techniques to identify places or times where crime is expected to occur, law enforcement organisations can more efficiently allocate their resources.

Conclusion:

In order to inform evidence-based measures to lower crime and enhance public safety, a crime analysis project in Toronto would seek to understand the patterns, trends, and causes of crime in the city.

The project would require gathering and examining crime data from a variety of sources, using statistical and geographical analysis methods to find patterns and trends, and collaborating with stakeholders all along the way.

The accuracy and applicability of the data, the rigor and openness of the analysis, and the value and accessibility of the findings to stakeholders are just a few of the variables that will determine if the project is a success.

Providing information for targeted crime prevention strategies: The project's findings may be used to pinpoint the neighbourhoods, crime kinds, and hours of the day where actions are most likely to lower crime rates.

Improving community-police relations: Participating in community organizations and other stakeholders' activities throughout the project may contribute to the development of trust and better community-police relations.

Overall, a crime analysis initiative in Toronto has the potential to significantly advance the discipline of crime analysis and raise community safety. However, the caliber of the research techniques, data sources, and stakeholder involvement would determine its effectiveness.