Mapping and Modeling Auto Theft Trends in Toronto for Strategic Urban Crime Prevention*

An Analysis of 2014-2023 Toronto Auto Theft Open Data

Heyucheng Zhang

April 14, 2024

This study analyzes auto theft incidents in Toronto using publicly available data, mapping out the prevalence of thefts across neighborhoods and identifying temporal trends. Our findings indicate that certain areas experience consistently higher rates of theft, with predictive modeling suggesting an upward trend in incidents. These findings are pivotal for shaping effective policies and law enforcement tactics to curb urban auto theft, thereby bolstering community safety and awareness.

Table of contents

*Code and data are available at: https://github.com/heyuchengzhang/Toronto-Auto-Theft. Shiny Web App is available at: https://79muet-heyucheng-zhang.shinyapps.io/Toronto-Auto-Theft-Data-Explorer/.

1 Introduction

A subtle undercurrent of crime continues to exist in Toronto's bustling metropolis, taking the form of auto theft, a scourge that plagues neighborhoods and poses a challenge to law enforcement. Our study provides a statistical and spatial exploration that spans several years to reveal the ebbs and flows of automobile theft. It is motivated by the urgent need to comprehend and combat this illicit activity.

Our research goes beyond being an academic study because it provides insights and predictions into the trends of auto theft using extensive publicly accessible data. We have used advanced statistical techniques and geographic information systems within the versatile R framework to map an area with high-theft zones and temporal activity peaks. This detailed portrayal reveals the trends that obscure the communities that pose a threat and provides the ability to predict what will happen next.

The estimand in our study is dual-natured: we aim to estimate both the cumulative auto theft incidents in each of Toronto's neighborhoods over a decade and the city-wide annual totals. We were able to pinpoint the neighborhoods most affected by this crime wave through careful analysis, which made it clear that specific measures were required. The results went even further, illustrating not just the current situation but even the days ahead, suggesting that if action is not taken, auto thefts may increase in the future. This foresight serves as a clarion call for policymakers and law enforcement, urging the adoption of data-driven strategies and community-centric approaches to dampen the tides of theft and secure the peace of Toronto's streets. Our study has multiple benefits: it not only closes the empirical gap on Toronto's auto theft trends but also serves as a benchmark for the development of public policy and crime prevention strategies. We provide a directed narrative that places our approach, conclusions, and their ramifications in the larger framework of public welfare and urban safety.

This paper is structured to walk the reader through our study, starting with an in-depth look at the data sources, methodology and measurement in Section 2, followed by the statistical models and their justifications in Section 3. We then present our results, providing a granular analysis of neighborhood-specific trends and spatial distributions in Section 4. The discussion in Section 5 synthesizes our insights, contextualizes the study within the broader spectrum of crime analysis, acknowledges the limitations of our approach, and charts a course for future exploration.

2 Data

2.1 Methodology

In our analysis of Toronto auto theft incidents, we utilized the R programming language (R Core Team 2022) and essential packages from the tidyverse (Wickham et al. 2019), specifically dplyr (Wickham et al. 2023) and tibble (Müller and Wickham 2023), for efficient data manipulation. The janitor package (Firke 2023) facilitated data cleaning by tidying variable names and removing duplicates. For visualizations, we employed ggplot2 (Wickham 2016) to create insightful graphics, and kableExtra (Zhu 2021) for stylized data tables. The here package (Müller 2020) was important for managing file paths and ensuring reproducibility, particularly important in projects involving multiple data sources.

Spatial analyses were conducted using the sf package (Pebesma 2018), enhancing our maps with precise data from Toronto's Open Data portal via the opendatatoronto package (Gelfand 2022). Statistical results were efficiently formatted with modelsummary (Arel-Bundock 2022) and broom (Robinson, Hayes, and Couch 2022) packages. Our primary dataset was sourced from the Toronto Police Service's open data (Service 2024), processed using the arrow package (Richardson et al. 2023) for its rapid data import capabilities.

2.2 Dataset Sources

2.2.1 Auto Theft Open Data

Auto Theft Open Data (Service 2024): Obtained from the Toronto Police Service, this dataset contains detailed records of auto theft incidents, including the date of the incident (REPORT_DATE), the neighborhood identifier (NEIGHBOURHOOD_158), and precise geographical coordinates (LONG_WGS84, LAT_WGS84). This dataset facilitates a detailed examination of auto theft patterns both temporally and spatially.

2.2.2 Toronto Ward Profiles

Toronto Ward Profiles (Planning 2024): Published by City Planning through Open Data Toronto, the Ward Profiles provide the mapping data for Toronto. This resource was instrumental in obtaining high-resolution maps of Toronto, which served as a backdrop for plotting the geographical coordinates of auto theft incidents. The availability of such detailed and open geographic data enables an enriched spatial analysis, allowing for the precise marking of theft locations on the map.

2.3 Data Preparation and Data Selection Justification

Following the acquisition, the auto theft incident data underwent a process of cleaning and preparation. A focused dataset, cleaned_data, was created by selecting essential variables and excluding unnecessary ones. Table 1 forms the foundation for our analysis.

Table 1: Sample of Cleaned Data

REPORT_DATE	NEIGHBOURHOOD_158	LONG_WGS84	LAT_WGS84
2014/01/01 05:00:00+00	Cliffcrest	-79.23612	43.72183
2014/01/01 05:00:00+00	Victoria Village	-79.30675	43.73465
2014/01/01 05:00:00+00	Etobicoke City Centre	-79.52969	43.61899
2014/01/02 05:00:00+00	Edenbridge-Humber Valley	-79.51245	43.68552
2014/01/02 05:00:00+00	Mount Olive-Silverstone-Jamestown	-79.59534	43.74430
2014/01/02 05:00:00+00	Agincourt North	-79.27393	43.81356
2014/01/02 05:00:00+00	Woodbine-Lumsden	-79.31380	43.68810
2014/01/03 05:00:00+00	Bendale-Glen Andrew	-79.25424	43.77664
2014/01/03 05:00:00+00	Bendale South	-79.24854	43.74843
2014/01/03 05:00:00+00	O'Connor-Parkview	-79.29707	43.69825

Through cleaned_data, we explore the variables of REPORT_DATE, NEIGHBOURHOOD_158, LONG_WGS84, and LAT_WGS84. Each plays a important role in understanding the dynamics of auto theft in Toronto, allowing for temporal trends and spatial distribution analysis. The incorporation of Toronto's ward maps enhances this spatial analysis by providing a detailed geographic context. Initial exploration of the dataset reveals key insights into the pattern and distribution of auto theft across Toronto's neighborhoods. Summary statistics and visualizations derived from the data provide a preliminary understanding of the auto theft landscape, identifying areas with higher incidents and potential temporal patterns.

The chosen datasets provide a granular view of auto theft incidents, offering both the specificity of individual events and the broader geographic context necessary for a detailed spatial analysis. The choice of Open Data Toronto's ward profiles as a mapping source ensures that our spatial visualizations are grounded in accurate and city-sanctioned geographic data. This study intends to identify patterns and trends in auto theft in Toronto by utilizing precise mapping tools and detailed historical information. The findings could provide insights that could influence policy and aid in the creation of focused efforts to reduce such crimes.

2.4 Measurement

In our study on auto theft incidents in Toronto, we meticulously navigated the complex terrain of measurement, akin to the nuanced challenges highlighted across various domains of research. The dataset, sourced from the Toronto Police Service, provides a detailed account of auto theft incidents, encoding each event within specific spatial and temporal dimensions through variables such as REPORT_DATE, NEIGHBOURHOOD_158, LONG_WGS84, and LAT_WGS84. Similar to how historical developments in measurement instruments, such as microscopes and atomic clocks, have transformed our understanding in domains ranging from biology to physics, this granularity allows us to analyze temporal trends and spatial distributions. However, there were inherent difficulties in using these variables to capture the essence of auto theft. These difficulties reflect larger measurement problems, such as the need to balance the richness of reality in quantitative research with the accuracy of identifying phenomena and operational simplicity.

To ensure the integrity and analytical robustness of our dataset, rigorous cleaning and preparation were undertaken, focusing on variables critical to our study's objectives. This process mirrors the meticulous consideration essential in measurement, aiming to distill complex information into a coherent form for analysis. Much like the careful selection of units and the evolution of measurement definitions that have underpinned scientific advancements, our methodical approach to dataset preparation seeks to provide a reliable foundation for exploring auto theft patterns. Through this, our study not only contributes to the quantitative analysis of crime but also engages with the broader discourse on the challenges and considerations inherent in the measurement, underlining our commitment to capturing the multifaceted nature of auto theft incidents in Toronto.

3 Model

3.1 Model Set-up

Through our data analysis, we identified a significant relationship between time and the number of auto theft incidents in Toronto. To explore this relationship further and predict future trends in auto thefts, we constructed a linear regression model.

We formalize the model as follows:

$$Y_i = \beta_0 + \beta_1 Time\_Index_i \tag{1}$$

In Model:

- Y_i represents the number of auto theft incidents at time index i.
- β_0 is the intercept, indicating the estimated number of thefts at the starting point of our time index.
- β_1 is the coefficient for the time index, capturing the rate of change in the number of thefts over time.

The independent variable in our model is the Time_Index, which is a numerical representation of time, enabling us to examine how theft incidents evolve. The dependent variable is the Thefts, denoting the number of auto thefts recorded.

3.1.1 Model Justification

The rationale behind using a linear regression model stems from the initial analysis, suggesting a linear trend in the increase or decrease of auto theft incidents over time. Given the straightforward relationship between time and the occurrence of these incidents, a linear model provides a clear and interpretable way to quantify and predict changes in auto theft rates. What's more, the linear model's simplicity makes it an attractive choice for initial investigations into time series data, allowing for straightforward interpretation and application of results. By examining the coefficient associated with the Time_Index, we can assess whether there's a general increase or decrease in theft incidents and at what rate these changes are happening.

Before constructing the model, we analyzed the distribution and trend of auto theft incidents using historical data. Our preliminary findings indicated a potential linear relationship between time and the number of incidents, justifying the selection of a linear regression model for further analysis. Our goal is to use the linear regression model to acquire a better understanding of the patterns of auto theft in Toronto. This understanding will serve as a foundation for future research, the creation of policies, and the implementation of anti-auto theft measures. The use of a linear model is further justified by its simplicity of interpretation and the correlation it establishes between time and auto theft incidents, which makes it a useful instrument for identifying and forecasting patterns in urban crime.

3.2 Model Prediction

Following the construction and validation of our linear regression model, which estimates the number of auto theft incidents in Toronto as a function of time, we proceeded to forecast auto theft incidents for the upcoming 12 months. This predictive exercise aims to provide insights into future trends in auto theft, potentially guiding policy and resource allocation for crime prevention.

Using the established linear model:

$$Y_i = \beta_0 + \beta_1 Time\_Index_i \tag{2}$$

we extended our Time_Index beyond the current dataset to predict theft incidents for the next year. The Time_Index was incremented by 12 months, representing our forecast horizon.

3.2.1 Predictive Insights

The resulting predictions, juxtaposed against observed data, offer a compelling visualization of the expected trend in auto theft incidents. In Figure 1, observed theft incidents up to the present are marked in black, while our model's predictions for the future 12 months are highlighted in red.

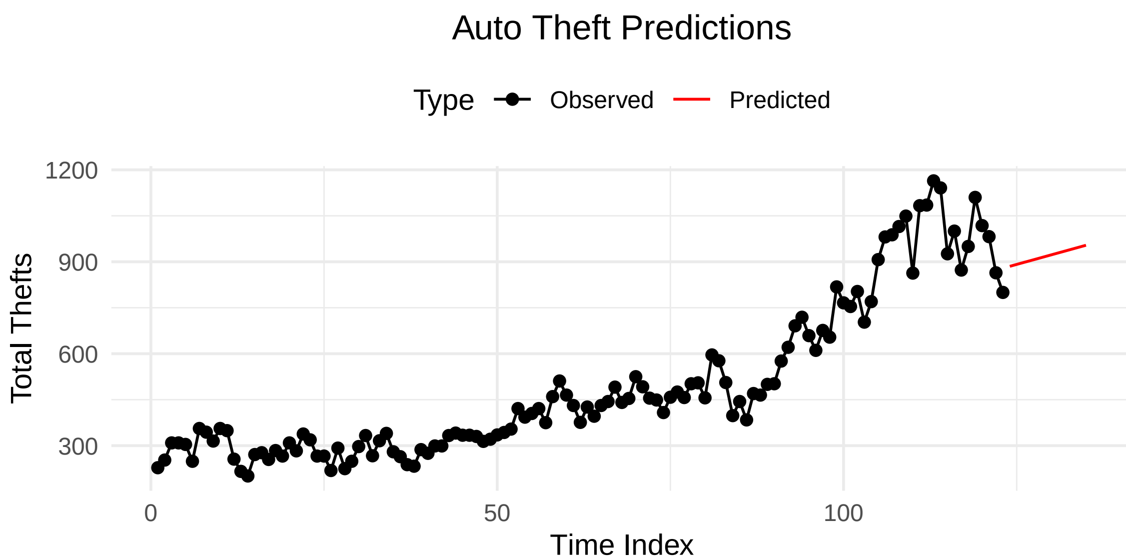

Figure 1: Auto Theft Predictions

The linear model predicts that the current trend will continue into the near future as is shown. Even though the model is predicated on a linear sequence of events, it is to recognize the possible influence of additional factors that the model does not account for and which might affect actual future theft rates.

3.2.2 Implications for Policy and Prevention

The predictive model underscores the necessity for proactive measures in addressing auto theft. By anticipating future incidents, policymakers and law enforcement agencies can better allocate resources, design targeted interventions, and implement preventive strategies aimed at curbing auto theft rates.

It is also to continuously refine the predictive model with updated data and potentially integrate more variables that could affect auto theft trends. Such iterative improvements will enhance the accuracy of future forecasts and, consequently, the effectiveness of preventive measures. This predictive analysis represents a foundational step towards a data-driven approach in combatting auto theft, highlighting the value of statistical models in informing and guiding public safety strategies.

4 Results

Our investigation into auto theft incidents throughout Toronto produced a wealth of information, which is presented in a series of tables and figures to give an in-depth understanding of the patterns and trends in auto theft in the city.

4.1 Neighborhood-Specific Trends

Table 2 showcases the distribution of auto thefts across various neighborhoods, indicating a significant variance in incident counts. West Humber-Clairville stands out with the highest number of recorded thefts, suggesting a targeted area for policy intervention and community safety initiatives.

Table 2: Sample of Neighbourhood Counts

NEIGHBOURHOOD_158	Count
West Humber-Clairville	4897
York University Heights	1567
Etobicoke City Centre	1332
Humber Summit	1074
Milliken	965
Wexford/Maryvale	956
Yorkdale-Glen Park	956
Oakdale-Beverley Heights	920
Glenfield-Jane Heights	899
Bedford Park-Nortown	882

In Figure 2, we present a bar chart detailing the top 10 neighborhoods with the highest counts of auto thefts. This visualization not only prioritizes areas for crime prevention strategies but also may prompt further investigation into the socio-economic and environmental factors prevalent in these neighborhoods.

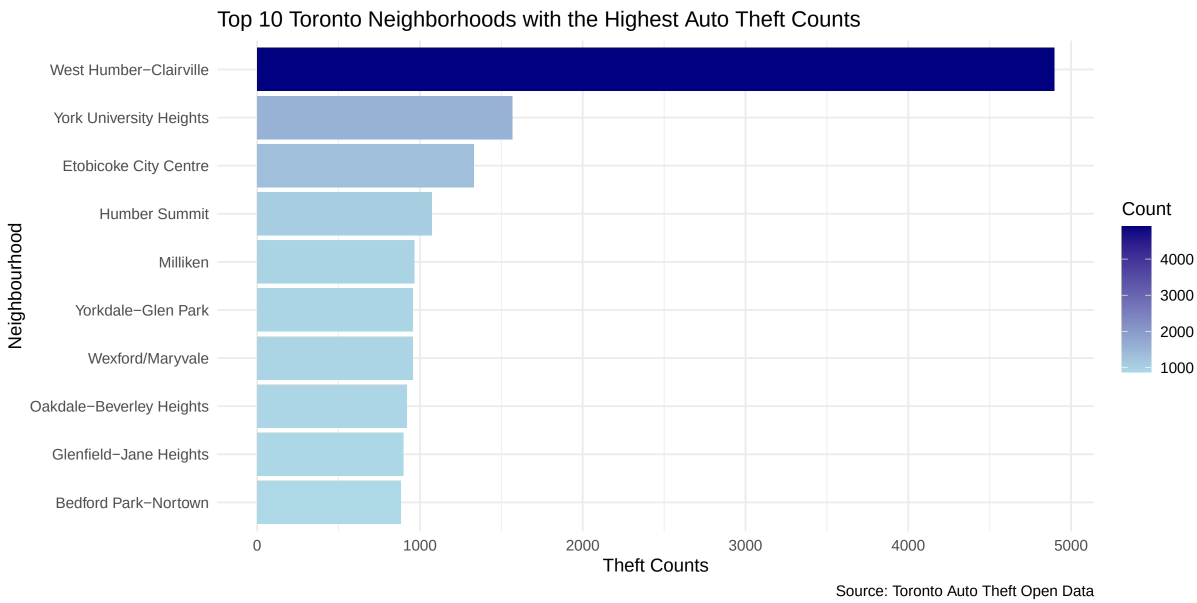

Figure 2: Top 10 Toronto Neighborhoods with the Highest Auto Theft Counts

4.2 Annual Theft Evolution

In Table 3, we reveal the annual progression of auto thefts from 2014 through 2023. There is an evident increase in incidents over the years, with a notable surge in recent years. This uptick raises concerns about the factors contributing to the rise and underscores the need for enhanced security measures.

Table 3: Sample of Theft Counts Per Year

REPORT_YEAR	Total_Thefts
2014	3628
2015	3285
2016	3348
2017	3617
2018	4804
2019	5362
2020	5787
2021	6642
2022	9835
2023	12262

Figure 3 displays the total auto thefts per year, complemented by a smooth trend line that underscores the steady increase in thefts over the examined period. The predictive model applied to this trend suggests that if current patterns hold, the upward trajectory of auto thefts is likely to continue.

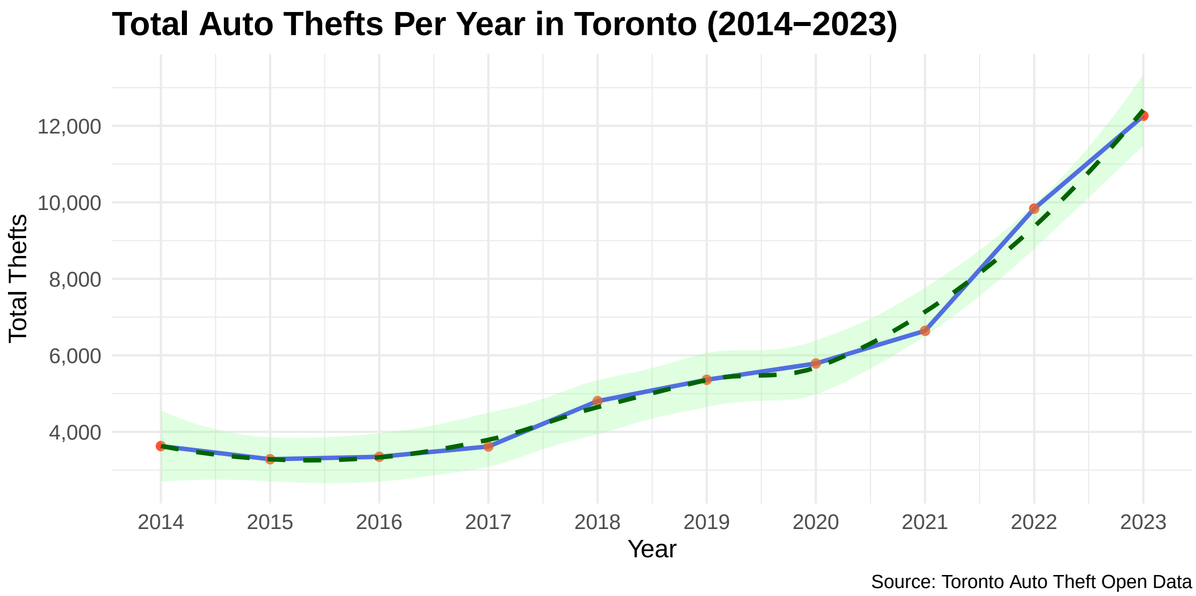

Figure 3: Total Auto Thefts Per Year in Toronto (2014-2023)

4.3 Geospatial Analysis

The geographical dispersion of auto thefts is visually represented in Figure 4, where we plot the latitude and longitude of reported incidents onto a map of Toronto. This map is critical for understanding the spatial context of auto theft, highlighting neighborhoods with dense clusters of incidents and potentially guiding the allocation of policing resources.

4.4 Shiny Web App: Toronto Auto Theft Data Explorer

Shiny Web App is available at: https://79muet-heyucheng-zhang.shinyapps.io/Toronto-Auto-Theft-Data-Explorer/

In our study, we developed a Shiny web application to enable dynamic exploration of auto theft data across Toronto. This application allows users to interactively select specific years—from 2014 to 2023—or view aggregated data for all years. Upon selection, the user interface updates to display relevant data in two main formats:

- Table View: This tab displays a table of the auto theft incidents corresponding to the selected year or the entire period. It presents raw data including detailed attributes such as date, location, and other pertinent information, allowing users to gain an in-depth look at the specifics of each incident.
- Map View: The second tab features an interactive map rendered using Leaflet(Cheng, Karambelkar, and Xie 2021), which highlights the geographical distribution of auto theft

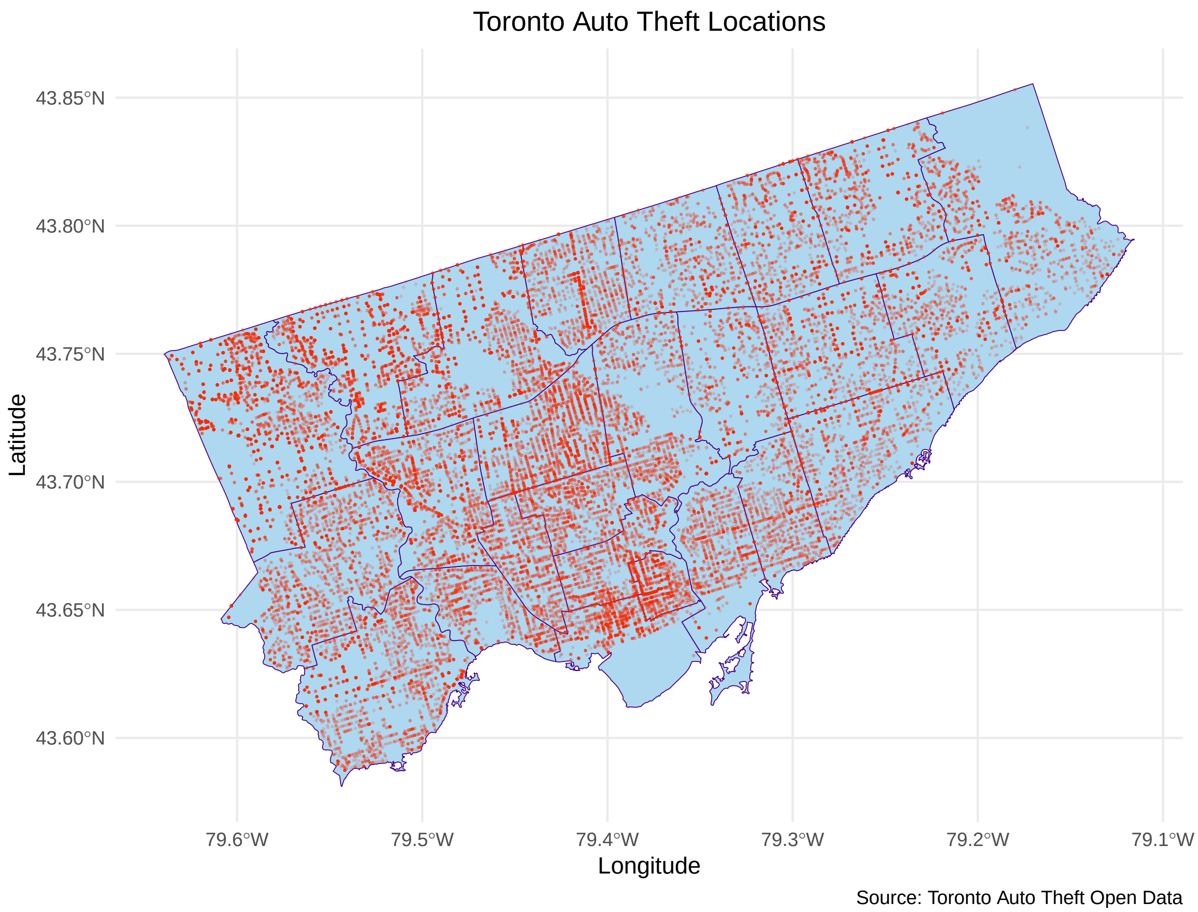

Figure 4: Toronto Auto Theft Locations

incidents. Markers on the map indicate the locations of thefts, with their placement based on the longitude and latitude coordinates from the data. This map provides a visual representation of theft hotspots and trends, enhancing understanding of the spatial dynamics of auto theft in Toronto.

This Shiny application not only provides a hands-on approach for users to interact with the data but also serves as a practical tool for enhancing public understanding and engagement with the data. By allowing users to select specific years and observe the corresponding data and maps, the application encourages deeper exploration and understanding of the patterns and trends in auto theft across Toronto.

4.5 Model Summary and 90% Credibility Interval

Figure 5 encapsulates the results of our model, displaying the correlation between time and the number of thefts. The blue data points represent actual theft counts, while the red line indicates the fitted model's predictions. The close alignment of the model with the observed data points indicates a strong linear relationship, reinforcing the model's applicability in predicting future trends.

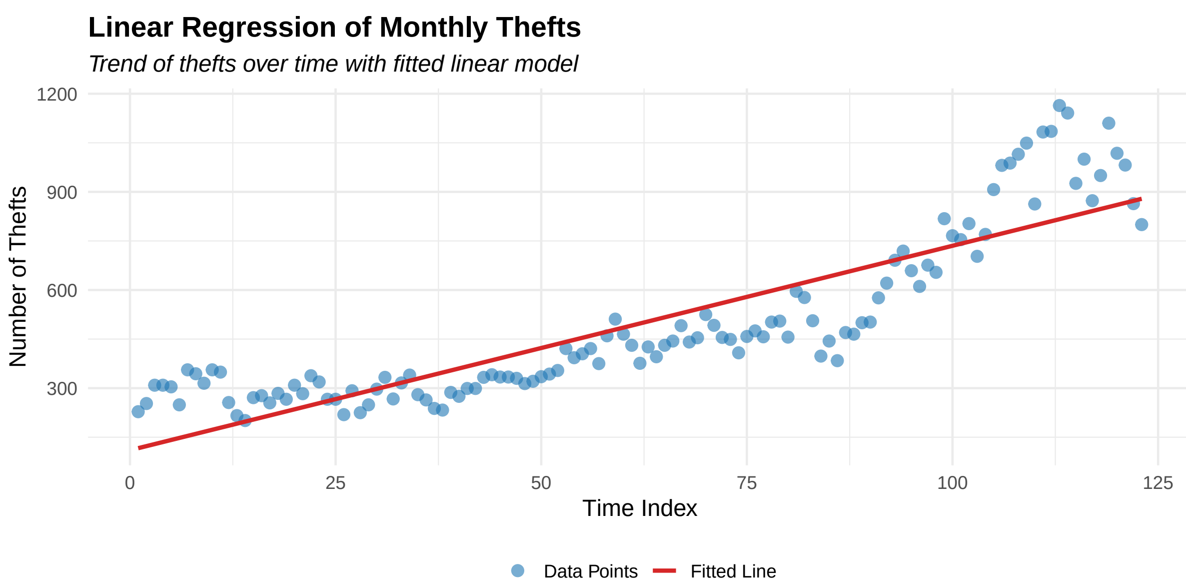

Figure 5: Linear Regression of Monthly Thefts

In addition to the spatial and temporal trends observed, our quantitative analysis includes a linear regression model, depicted in Table 4, which delineates the relationship between time and auto theft counts. The model estimates an intercept of 110.24, indicating the starting point of our time index, with a slope coefficient of 6.25. This suggests that, on average, the number

Table 4: Explanatory models of Auto Thefts Counts based on Time Index

	Auto Thefts Counts
(Intercept)	110.24
	(22.15)
Time_Index	6.25
	(0.31)
Num.Obs.	123
R2	0.771
R2 Adj.	0.769
AIC	1535.0
BIC	1543.5
Log.Lik.	−764.510
RMSE	121.10

of auto thefts increased by approximately 6.25 for each additional time index increment, a proxy for time passage in our dataset.

The robustness of the model is reflected in the R-squared value of 0.771, indicating that approximately 77.1% of the variability in monthly theft counts is explained by the time index. The adjusted R-squared of 0.769 accounts for the number of predictors in the model, maintaining a similar level of explained variability, which indicates a good fit. The model's predictive capability is supported by a reasonable Root Mean Square Error (RMSE) of 121.10, suggesting that the typical prediction error of the number of thefts is within this range. The model's coefficients, including the standard errors in parentheses, provide insight into the precision of the estimates, with smaller standard errors indicating more precise estimates.

Figure 6 offers a detailed look at the explanatory power of our model through a graphical representation of the 90% credibility interval around the coefficients for the time index and the intercept. This interval visualization aids in understanding the degree of certainty in our model's estimates and its reliability in extrapolating future incidents.

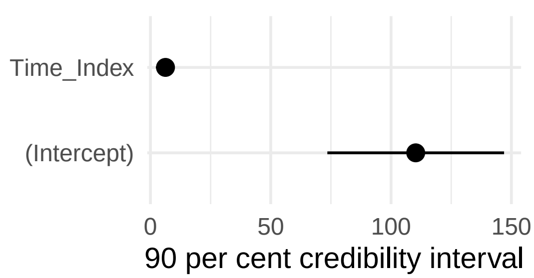

Figure 6: Explanatory models of Auto Theft Counts based on Time Index: 90% Credibility Interval

Together with the geographical and temporal analysis previously discussed, this regression model is an essential tool for understanding the dynamics of auto theft incidents and anticipating future trends. It provides a statistical foundation for estimating the occurrence of thefts over time, offering insightful information to law enforcement and politicians entrusted with reducing these crimes.

5 Discussion

5.1 Analyze Another Surge

In an extension of our analysis on the auto theft trends in Toronto, we analyzed the most recent and pronounced increase in incidents during the years 2022 and 2023. Figure 7 labeled "Monthly Auto Thefts Counts for 2022 and 2023" provides a focused examination of these two years, highlighting the monthly theft counts and revealing a comparative perspective between them. The visualization captures the ebb and flow of auto theft incidents throughout each month for both years, presented side by side for direct comparison. In 2022, represented by the color green, we see a pattern that suggests a certain seasonality to the thefts, with some months peaking higher than others. The year 2023, denoted by the color orange, shows not only a continuation of this pattern but also an elevated count of thefts across all months.

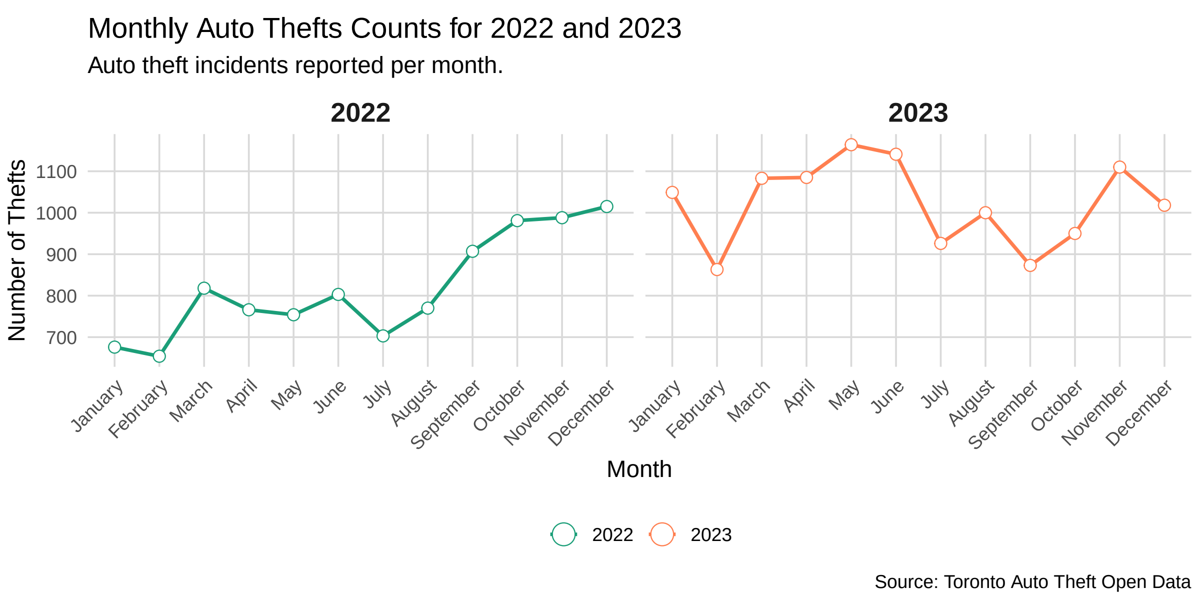

Figure 7: Monthly Auto Thefts Counts for 2022 and 2023

By employing line and point geoms, the graph offers a clear view of the theft counts' trajectory, with lines connecting the monthly data points that accentuate the trend over time. The points, with their white fill, draw attention to the actual data, allowing for each month to stand out within the annual trend. The facet wrap segregates the years, facilitating a focused analysis of each year's pattern without the visual complexity that might arise from overlapping data series. This detailed month-by-month breakdown underscores the acceleration in auto theft incidents during 2022 and 2023 and signals potential underlying factors that may be contributing to this uptrend. Such granular insights are for devising timely and effective preventative strategies, allowing stakeholders to adapt to and address the evolving landscape of auto theft in Toronto.

5.2 Insights on Auto Theft Trends

Through spatial analysis, we learned that urban crime, particularly auto theft, is not randomly distributed but is concentrated in specific neighborhoods. One of the key findings of our study is the identification of hotspots for auto thefts, such as West Humber-Clairville, which could signify areas where preventive measures could be intensified. This suggests a spatial component to crime that may be influenced by factors such as population density, urban design, and socio-economic variables. Our study's spatial visualizations can help guide targeted law enforcement and community interventions. The upward trend in theft incidents over the years indicates not only the growing concern of auto theft in the city but also possibly reflects broader socioeconomic trends that could be influencing crime rates.

The linear regression model's prediction of thefts increasing over time calls attention to the potential for these trends to continue if interventions are not effectively implemented. This trend underscores the need for continuous monitoring and adaptive strategies in policing and community programs to address and prevent auto thefts effectively.

5.3 Overview and Review

In our extensive study, we explored the intricate landscape of auto theft incidents in Toronto, meticulously analyzing data collected over several years. By employing statistical methodologies and geospatial analysis, we have been able to provide a nuanced depiction of the patterns and fluctuations in auto theft across various neighborhoods and time periods. Utilizing R's robust statistical environment and applying the accessibility of public data sources, our research underscores the powerful role that data analysis plays in shaping informed public policy and enhancing community safety measures.

Our findings have led us to identify neighborhoods with the highest rates of auto theft, a step in understanding the geographical distribution of this criminal activity. Furthermore, by forecasting future auto theft occurrences through our predictive models, we have laid the groundwork for law enforcement and policymakers to develop targeted strategies aimed at preventing and reducing crime. These predictive insights offer a proactive approach, allowing for the allocation of resources and the implementation of security measures in anticipation of potential theft hotspots.

By bridging data-driven insights with public safety objectives, our study acts as a testament to the potential of open data in facilitating impactful societal changes. As a result of our analyses, authorities are equipped with actionable intelligence to institute measures that not only prevent auto theft but also address its root causes, fostering safer communities throughout Toronto.

5.4 Limitations and Weaknesses

Our study has provided valuable insights into the patterns of auto theft within Toronto, applying robust statistical methods and data analysis. However, we must acknowledge the inherent limitations that accompany our study. Our reliance on officially reported data inherently presents a challenge, as it is well-established in criminology that not all crimes are captured in police datasets. While our model provides a statistically significant fit for the data, it does not encapsulate the complexity of external influences that may drive variations in auto theft rates. Potential confounding factors such as economic fluctuations, unemployment rates, and demographic changes are not accounted for, which could offer additional explanatory power if included. For instance, economic downturns may increase criminal activity due to heightened financial pressures, or shifts in police resource allocation could inadvertently impact theft rates.

Another limitation is our model's RMSE, which, although reasonable, highlights the potential for improved accuracy. This measure of prediction error suggests that our model, while useful, could benefit from further refinement. Future research endeavors could look to incorporate broader datasets, potentially including real-time economic indicators, sociodemographic profiles, vehicle security features, and even weather patterns, which could all have a bearing on auto theft occurrences. By integrating these additional layers of data, future models could provide a more nuanced and precise understanding of the predictors of auto theft, thus offering stronger predictive capabilities. Enhanced models could then serve as a more reliable tool for policy recommendations, allowing for proactive and targeted crime prevention strategies that are tailored to address the multifactorial nature of auto theft.

5.5 Future Exploration

As we look to the future, our research serves as a foundation for a more nuanced understanding of auto theft in urban environments. We plan to expand our analysis to include the methods used in these thefts, which will involve cataloging the various techniques thieves employ to steal vehicles. This will enable us to identify patterns in the methodologies of auto theft, which could be critical in developing preventative measures. The exploration of auto theft methods will include the collection and statistical analysis of incident reports, with particular attention to the details of each crime. This analysis is expected to yield valuable insights that could inform law enforcement strategies, leading to targeted interventions that address the most prevalent theft techniques.

Through our continued research, we aim to contribute actionable recommendations to the police and policymakers. By understanding how cars are stolen, strategies can be refined to anticipate and mitigate these methods. For example, if a specific technology is found to be frequently exploited, our research could lead to recommendations for vehicle security enhancements or public awareness campaigns focused on the vulnerabilities. Furthermore, we anticipate that our future studies will combine these quantitative analyses with qualitative research. Interviews with law enforcement officials, crime prevention experts, and victims of auto theft could provide additional context to the statistical patterns we observe. These narratives will enrich our understanding of the data and could uncover aspects of auto theft that are not immediately apparent through numbers alone.

In closing, while our current study has illuminated the patterns and trends of auto theft incidents in Toronto, it is just the beginning of a more extended approach to combating this type of urban crime. Our commitment to ongoing data collection, innovative analysis, and the fusion of quantitative and qualitative research methodologies is poised to deliver impactful insights. These efforts will be instrumental in crafting evidence-based crime prevention strategies that can significantly reduce the incidence of auto theft.

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Cheng, Joe, Bhaskar Karambelkar, and Yihui Xie. 2021. *leaflet: Create Interactive Web Maps with the JavaScript "Leaflet" Library*. https://CRAN.R-project.org/package=leaflet.

Firke, Sam. 2023. *janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gelfand, Sharla. 2022. *opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Müller, Kirill. 2020. *here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439–46. https://doi.org/10.32614/RJ-2018-009.

Planning, City. 2024. *Ward Profiles*. Open Data Toronto. https://open.toronto.ca/dataset/ward-profiles-25-ward-model/.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, and Apache Arrow. 2023. *arrow: Integration to Apache Arrow*. https://CRAN.R-project.org/package=arrow.

Robinson, David, Alex Hayes, and Simon Couch. 2022. *broom: Convert Statistical Objects into Tidy Tibbles*. https://CRAN.R-project.org/package=broom.

Service, Toronto Police. 2024. *Auto Theft Open Data*. https://data.torontopolice.on.ca/datasets/TorontoPS::auto-theft-open-data/about.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jenny Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with "kable" and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.