

Evaluating Toronto's Fire Incidents to Extrapolate Future Occurrence Data*

Jimmy Luc

April 17, 2024

This paper uses publicly available data from [Toronto Open Data](#) to predict the estimated dollar loss per fire incident in the Toronto area by property use. Further analysis of Toronto Fire Services' response time, area of origin, and various fire prevention measures as considering factors in the estimation. The results reveal a positive correlation between estimated dollar loss and possible causes of the incident, namely, mechanical/electrical failures yielding the highest loss. These findings bring urgency to fire safety as they highlight the importance of proper education, maintenance, and safety inspections. This analysis encourages homeowners, businesses, and property managers to take initiative in order to ensure the well-being of others.

Table of contents

1	Introduction	2
2	Data	3
3	Model	3
3.1	Model set-up	3
3.1.1	Model justification	3
4	Results	4
5	Discussion	6
5.1	First discussion point	6
5.2	Second discussion point	6

*Code and data are available at: <https://github.com/rex009x/fire-incidents-toronto>.

5.3	Third discussion point	6
5.4	Weaknesses and next steps	7
Appendix		8
A Additional data details		8
B Model details		8
B.1	Posterior predictive check	8
B.2	Diagnostics	8
References		9

1 Introduction

Fires pose a significant threat to both residential and commercial properties, often resulting in substantial financial losses and, more importantly, endangering lives. Developing a better understanding of the factors that influence the amount in damages per fire incident is crucial for implementing effective fire prevention measures to minimize the impact of such incidents. In this analysis, I delve into the extensive dataset provided by Toronto Open Data (Toronto Open Data 2024) which captures fire incidents within the Toronto area as defined by Ontario Fire Marshal (OFM) from 2011 to 2022. Through this dataset, I aim to predict the estimated dollar loss per fire incident across different property uses and identify the factors contributing to these losses.

The motivation behind this analysis is that fire safety matters and is important. Beyond the protection of assets, it’s about keeping people safe. In order to ensure the safety of fellow Canadians, it is imperative to comprehend the underlying patterns and determinants of fire incident. By dissecting the data provided by Toronto Open Data, I seek to not only predict the estimated amount in damages but also to identify significant factors influencing these losses. Analyzing factors such as Toronto Fire Services’ response time, area of origin, and the presence, functionality, and configuration of fire alarms. These additional variables are essential in understanding the dynamics of fire incidents and formulating informed strategies for fire prevention and mitigation.

Despite the abundance of data available, there remains a gap in comprehensively analyzing the factors contributing to the estimated dollar loss per fire incident in the Toronto area. To address this gap, I employ a data-driven approach, utilizing statistical modeling techniques to uncover patterns and relationships within the dataset. Our methodology involves exploring various factors such as property use, possible causes of the incident (e.g., mechanical/electrical failures), and the presence of functioning fire alarms. Additionally, I incorporate the number of responding personnel from the Toronto Fire Services (TFS) as a variable to assess its impact on the estimated dollar loss.

The findings of this analysis hold significant implications for fire safety practices and policies in the Toronto area. By understanding the factors influencing the estimated dollar loss per fire incident, stakeholders can prioritize resources and allocate them accordingly to mitigate fire risks effectively. This paper is structured as follows: I first provide an overview of the dataset and variables of interest, followed by the methodology employed for analysis in sec-data. Subsequently, I present our findings, discussing the relationships uncovered and their implications for fire safety in sec-results. Finally, I conclude by highlighting the importance of our findings and suggesting avenues for future research in this domain concerning fire safety in sec-discussion.

2 Data

All data collection and analysis was completed using statistical tools such as R (R Core Team 2023), and Rstudio IDE (RStudio Team 2020) to better streamline workflow. The analysis was done using the R program and the following supporting packages: `tidyverse` (Wickham et al. 2019), `dplyr` (Wickham et al. 2023), `here` (Müller 2020), `opendatatoronto` (Gelfand 2022), `ggplot2` (Wickham 2016), `ggmap` (Kahle and Wickham 2013), `lubridate` (Grolemund and Wickham 2011), and `scales` (Wickham, Pedersen, and Seidel 2023). Further details regarding data cleaning and analysis will be discussed in the following subsections.

3 Model

Work in progress.

3.1 Model set-up

Work in progress.

3.1.1 Model justification

Work in progress.

4 Results

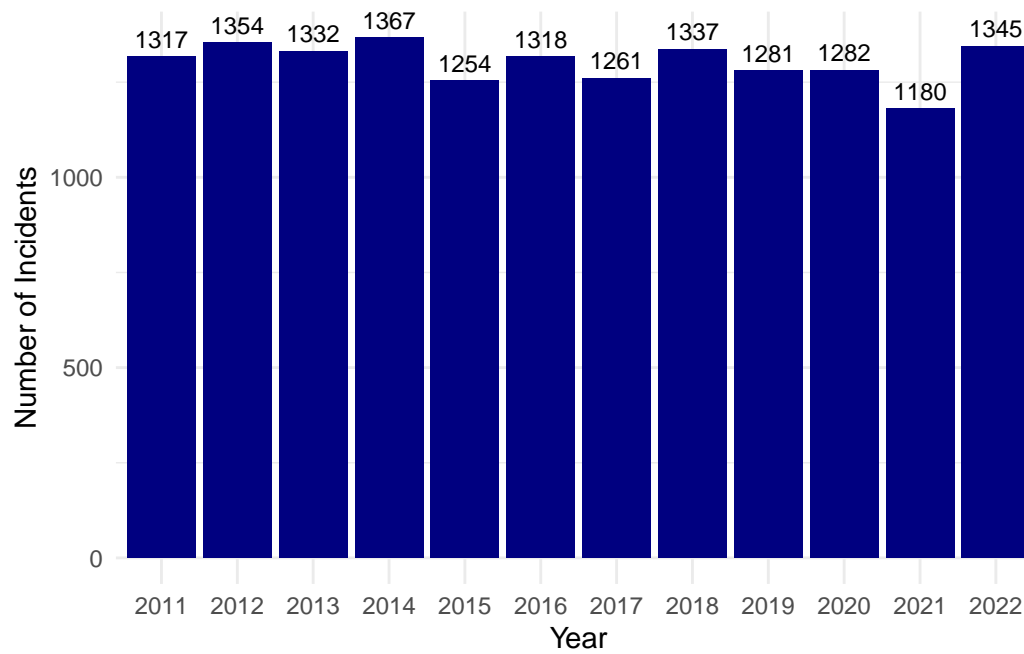


Figure 1: Number of Fire Incidents in Toronto by Year (2011-2022)

Figure 1 testing testing hello.

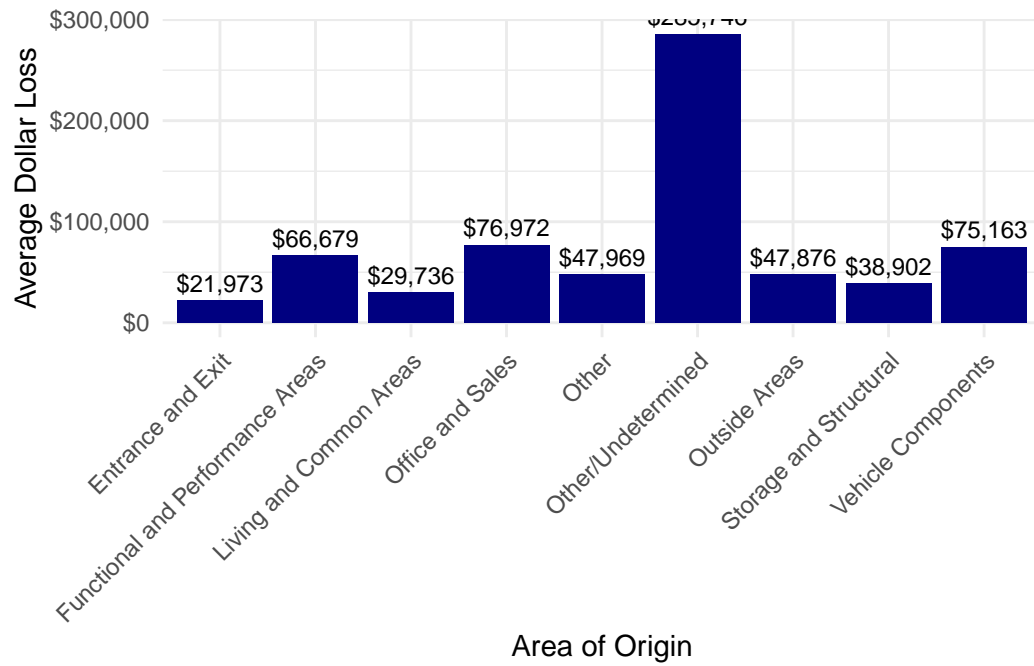


Figure 2: Average Estimated Dollar Loss by Group Area of Origin

Figure 2 testing testing hello

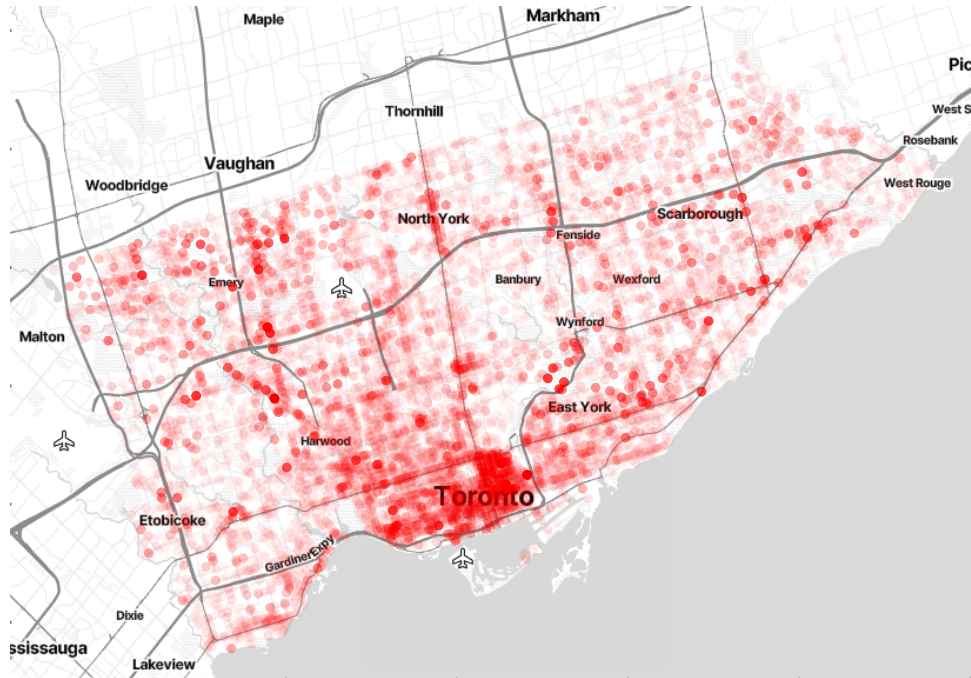


Figure 3: Fire Incidents Heatmap in Toronto

Figure 3 testing testing hello

5 Discussion

5.1 First discussion point

Work in progress.

5.2 Second discussion point

Work in progress.

5.3 Third discussion point

Work in progress.

5.4 Weaknesses and next steps

Work in progress.

Appendix

Work in progress.

A Additional data details

Work in progress.

B Model details

Work in progress.

B.1 Posterior predictive check

Work in progress.

B.2 Diagnostics

Work in progress.

References

- Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- Grolemund, Garrett, and Hadley Wickham. 2011. “Dates and Times Made Easy with lubridate.” *Journal of Statistical Software* 40 (3): 1–25. <https://www.jstatsoft.org/v40/i03/>.
- Kahle, David, and Hadley Wickham. 2013. “Ggmap: Spatial Visualization with Ggplot2.” *The R Journal* 5 (1): 144–61. <https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- RStudio Team. 2020. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC. <http://www.rstudio.com/>.
- Toronto Open Data. 2024. “Fire Incidents.” <https://open.toronto.ca/dataset/fire-incidents/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer 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, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org>.
- Wickham, Hadley, Thomas Lin Pedersen, and Dana Seidel. 2023. *Scales: Scale Functions for Visualization*. <https://CRAN.R-project.org/package=scales>.