

My title*

My subtitle if needed

David Qi

September 27, 2024

This paper discusses some basic findings around the monetary and lives loss in Toronto fire incidents. By using data from open data Toronto and choosing the relevant entries, we have identified possible factors that influences monetary and lives loss in fire incidents.

1 Introduction

Fire incidents are a major threat to urban lifestyles, causing substantial loss of life and property. In recent years, there has been an increase in number of fire incidents reported in Toronto (Figure 1), which highlights the importance of identifying trends and patterns in historical data. These trends can provide helpful insights on causation of casualties and monetary loss, which may further enable us to find areas of improvement for our fire services.

In this paper, we make use R (R Core Team (2023)) and the tidyverse package ((tidyverse?)). The data set is from Open Data Toronto ((opendatatoronto?)). The figures in our study are generated by ggplot ((ggplot2?)) . We also have used ((sigmoid?)) when simulating and testing with our data.

The remainder of this paper is structured as follows: We first give a brief overview of the dataset under **Data** section. Providing some basic information. Then, we use graphs and models to present the results we find from our data in the **results** section. We then discusses the possible interpretations of our find further in the **discussion** section. While also pointing our the weaknesses and areas for further investigations in the **Weaknesses and next steps** section.

*Code and data are available at: https://github.com/UTDQi/starter_folder

2 Data

2.1 Overview of dataset

The data set includes 32000 fire incidents from year 2011 to 2023. As description provided by open data Toronto([opendatatoronto?](#)) tells us, “This dataset includes only fire incidents as defined by the Ontario Fire Marshal (OFM) up to December 31, 2023.”

The raw data set contains 43 columns, we have chooses 5 of them for further analysis. For each of the variables, some of the entries may be missing, each time we refer to the variables, we will removed the missing entries. The actual number of incidents presented in each graph will be mentioned before each figure.

The selected variables are: estimated dollar loss, number of responding personnel, possible cause, TFS alarm time, and civilian casualties. We believe these variables have high correlations, and may suggest the reason for life and property loss.

2.2 Basic informations

This section provides basic information about some of the variables independently, without referencing other variables. This gives as an overview of the structure of our data, which will help us understand the results we find in analysis.

To highlight the increased importance of our study, we first present the number of reported fire incidents each year in Toronto(Figure 1). There are no missing entries in alarm time, so this figure includes 32000 incidents.

It is important to note the great increase of fire incidents from 2683 incidents in 2022 to 3457 in 2023. This may implies that as we reach the end of the end of covid pandemic, the threat of fire has increased due to increased human activities.

We then present the distribution of estimated dollar loss on 23449 incidents(Figure 2).

The graph is similar to an exponential distribution. With a few peaks at 50, 100 and 150 thousand, due to the nature of this data is being estimated.

We then present possible causes of fire on 23425 incidents.(Figure 3)

We see the three main causes of fire identified are: Electrical failure, Improperly discarded ignition source and Unattended ignition source. We will focus on these causes when we analysis the data.

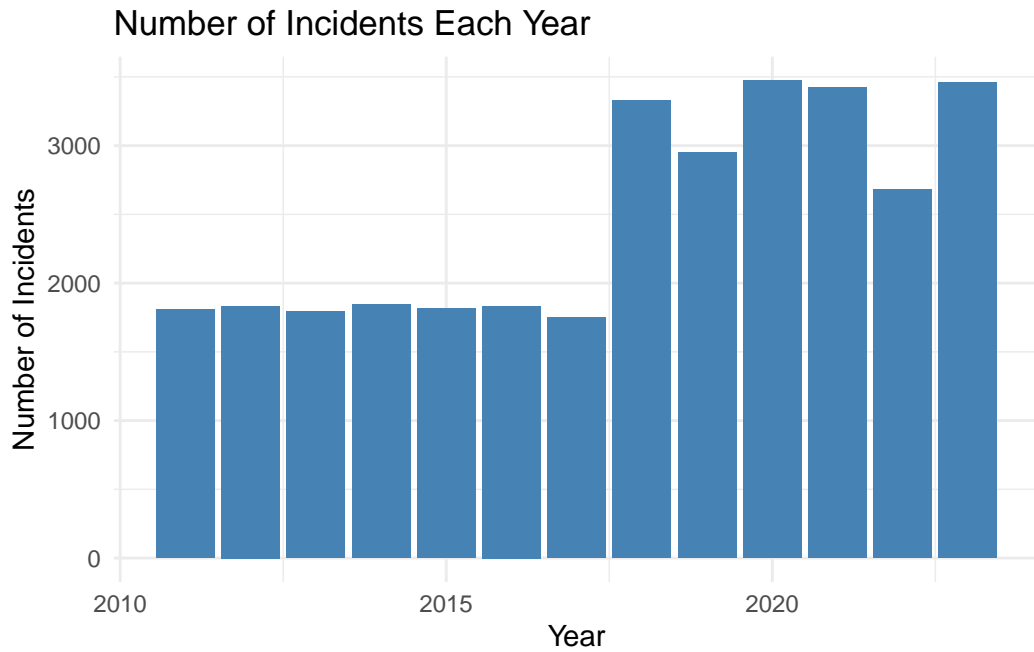


Figure 1: Number of incidents recorded each year

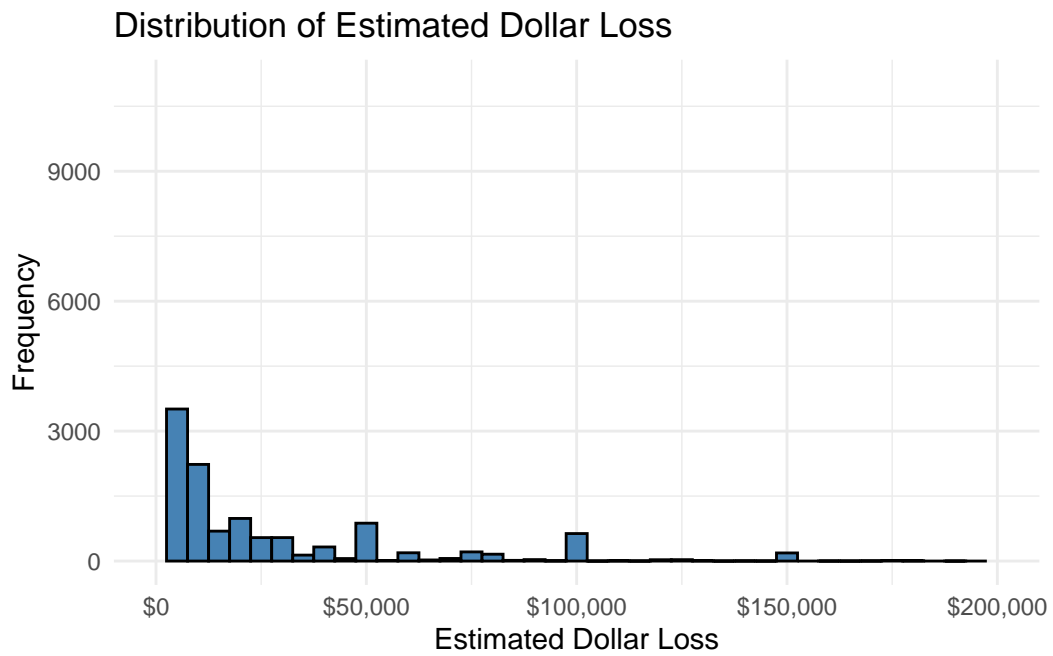


Figure 2: Distribution of Estimated Dollar Loss

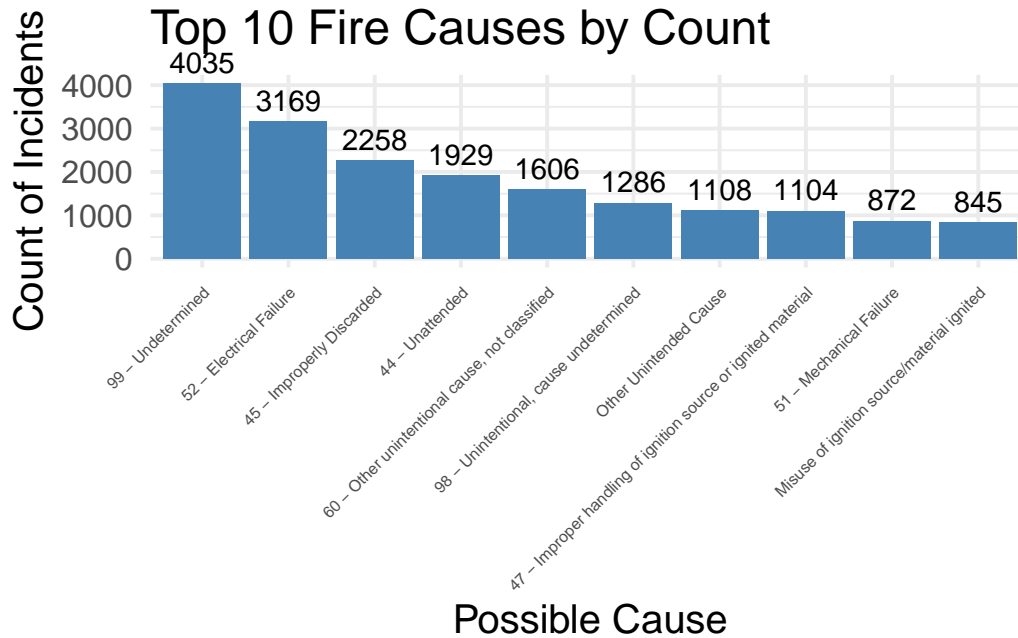


Figure 3: Count of incidents for each cause

3 Model

3.1 Model set-up

3.1.1 Model justification

4 Results

Our results are summarized in [?@tbl-modelresults](#).

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

B.2 Diagnostics

References

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.