

Melting snow: Exploring Toronto Fire Incident Data*

Jacob Gilbert

January 25, 2024

Fire incidents have and will pose a significant problem to the wellbeing of public health and the infrastructure of urban areas. Understanding the patterns and characteristics is essential for policy makers to implement preventative measures. In the ever growing city of Toronto, where population is set to grow 45% across the Greater Toronto Area by 2046 (ontario.ca), the importance of this is ever prevalent.

Table of contents

1	Introduction	2
2	Data	2
3	Analysis	3
3.1	Geographical	3
3.2	Hour of day	4
3.3	Ignition sources	5
4	Conclusion	5

*Code and data are available at: [LINK](#)

1 Introduction

Prevention of fires is a multifaceted challenge with many unpredictable factors, thus understanding the information we have is essential. Building infrastructure, common causes and environmental factors can all play a part. In a densely populated city the concentration and constant evolution of structures create additional challenges. New buildings and skyscrapers are on the rise in Toronto for residential and industrial purposes and as such a proactive approach is essential for minimising losses.

With such a large geographical area to cover, the GTA divides its emergency services geographically. Different stations respond to incidents within their surrounding area as to not overencomber each location. With the changes mentioned to population size and density in many areas of Toronto the demands for different areas are likely to change. As a result policy makers will need to understand were this demand is required most to make efficient responses.

The following analysis aims to provide valuable insights for policy makers, emphasising areas of increased frequency and bringing attention to hotspots for areas of increased fiscal damage. Doing this both by geographic location as well as by location use. Finally, look at other features that may contribute to policy makers such as the source of fires, in order to implement educational measures, and time of day for resource allocation.

2 Data

The following report uses fire incident data, and fire service station data opendatatoronto (**opendatatoronto?**) obtained through City of Toronto's open data portal. Further map information used in (Figure 1)1 is also from opendatatoronto. this resource is updated regularly by the city of toronto, all rsources are as of 2023. Fire incident data is that of events responded to by the Toronto Fire service, and thus it is assumed that they are the data source. the only data that is incomplete is that which is still under investigation or classified, as the data used is from 2018 - 2022 this is deemed insignificant. The most recent 4 years of data were used in an aim for relevancy giving information on 10,246 incidents.

Key data points were taken from the vast amount of information to analyze patterns and behaviours of which any incomplete information was removed or set to appropriate inputs. One of which datapoints taken was estimated dollar lost as a result of the fire which was used as a measure for extent of damage in this paper. This was the most reliable data thruought the dataset aswell as set an unbiased measure for conscequences of the incidents.

The analysis will be done using R (R Core Team (2022)) using packages dplyr (Wickham et al. (2023)), sf (Pebesma (2018)), janitor ((**janitor?**)) and tidyverse (Wickham et al. (2019)). All graphs were generated by ggplot2 (Wickham (2016)).

3 Analysis

3.1 Geographical

From the fire incidents dataset provided each incident was allotted coordinate location of the nearest intersection allowing for observations into geographically analysis of the data. The geographical graph (Figure 1) contains a graphed map of toronto with red points indicating fires within buildings and blue for other, outside locations. The size of the points reflects the estimated dollar loss of the incident.

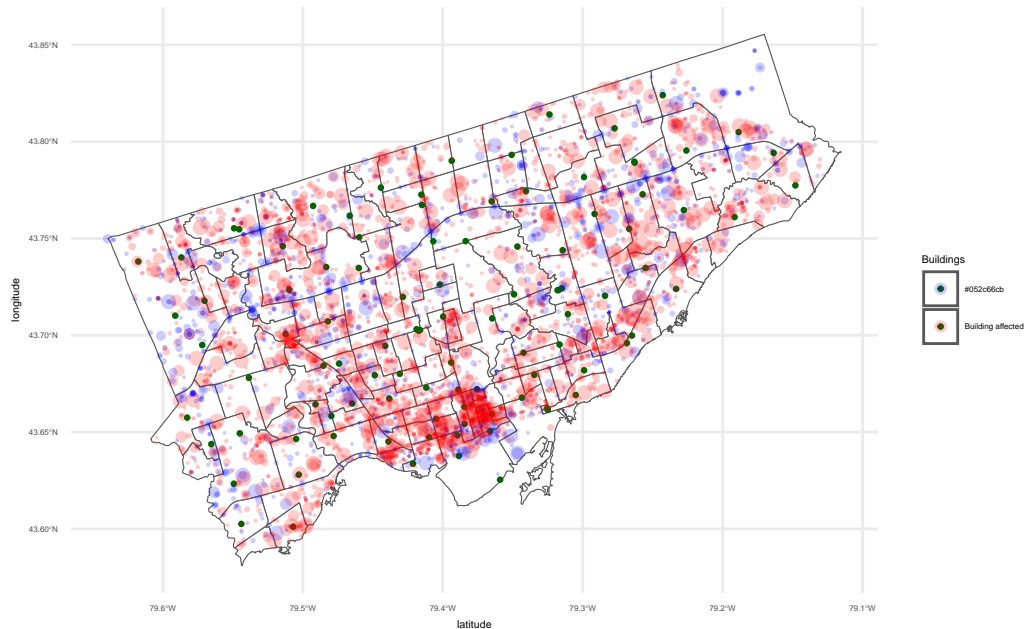


Figure 1: Fires and fire stations in toronto, size of point represents estimated dollar lost of fire, displays higher density and costs towards downtown

Towards the southern end of the GTA a hotspot appears in the data. This area is compact with larger and more densely populated red points. This area is downtown Toronto, filled with highrises and home to 16,000 people per square (Ontario.ca (2022)) kilometre, it clearly comes with fire risk. The increase in density may be attributed to higher condensation of population. Further it is worth noting that the estimated dollar lost here may be positively skewed as a result of more expensive property. However the amount of fires outside of the buildings is proportionally minimal, that is, compared with the rest of the GTA.

Outside of downtown we see plenty of incidents not occurring in residences or businesses but outside. As the estimated cost histogram (Figure 2) emphasises these events are less damaging on average. Thus policy makers may consider this when allocating resources. However the less

urban areas aren't without some large incidents, more sporadic as well. Making them further from relative fire stations.

The graph below (Figure 2) does highlight how the proportion of events that are more damaging are on properties. Thus giving an even clearer image of the likelihood of costs for red data (Figure 1) points being high. Showing avoiding these types of incidents is a way to mitigate damages.

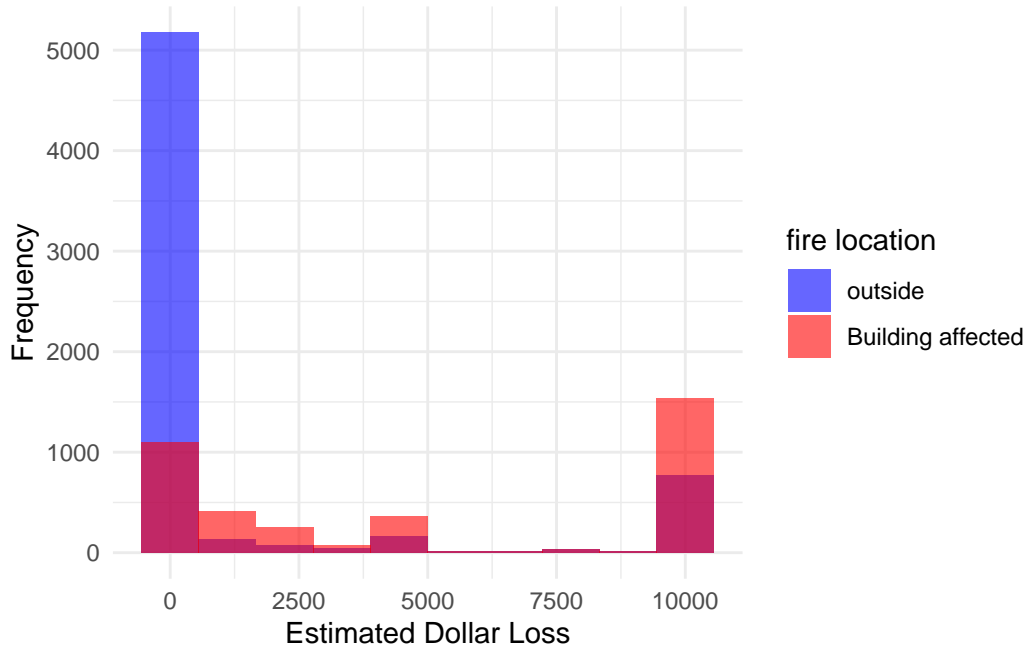


Figure 2: Histogram of fire incident cost separated by indoor and outdoor events, with large left skew. final bin indicates 10,000+

3.2 Hour of day

Time and space go together hand in hand and the hourly data (Figure 3) looks at the former. Hourly data was available for all observations of the data and thus tells a significant story. Based on the data, fire incidents are less likely to occur in the early hours, from 2 to 9am. Based on time of day, (Figure 3) the Toronto fire department is most likely to be called to an incident in the late hours of the night. Ending around 2am, it is worth noting that the majority of establishments close at this time in Toronto. This could be seen in a positive or negative way due to the potential vacancy of buildings in the early hours. What certainly is an assurance is these hours will coincide with times residents are asleep.

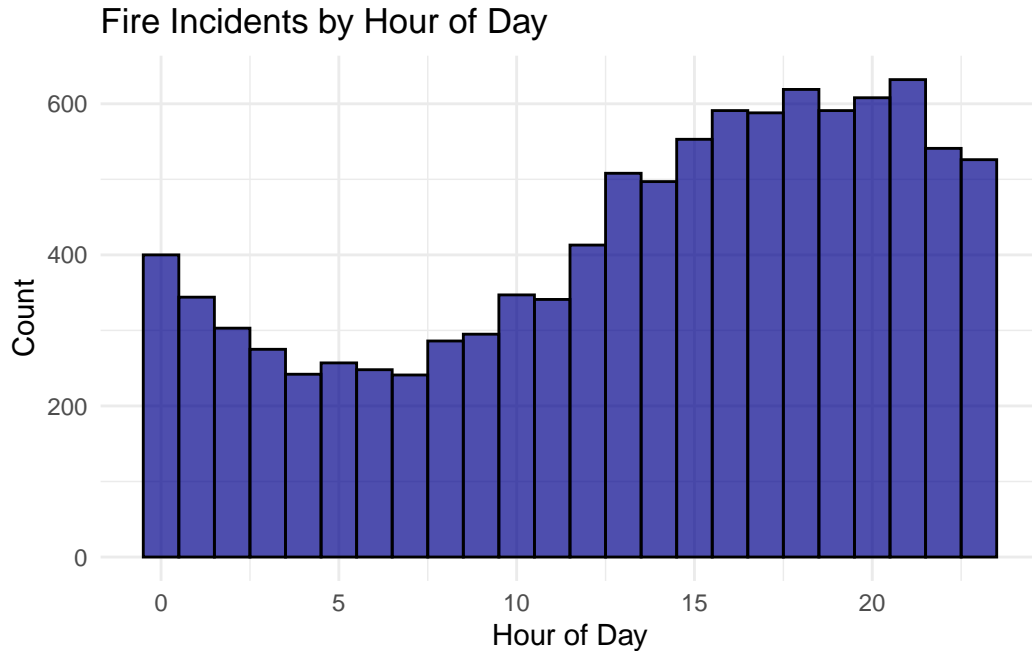


Figure 3: frequency of fire incidents by hour, highest frequency later in the evening

3.3 Ignition sources

- (1) gives information on the most common initial ignition sources for fires. The largest proportion of these top causes is “undetermined”, this is not a result of bad data. In fact this is a result of the type of data, as some fires are already past the point of collecting this information. However the remainder of the information does tell a story. The largest portion is due to kitchen appliances telling us the area where fires are likely to start. Next we see smokers’ articles are a large contributor. Implying that those who smoke are more likely to be involved in a fire. This along with candles, perhaps unsurprisingly, indicate that open flames are an incredibly common cause of fires. Finally we see that vehicle fires are clearly frequent in the data, it is worth noting that the cause of a vehicular fire may be easily assumed and thus its prevalence bias in the data.

4 Conclusion

This data has supplied information regarding fire incidents in Toronto. Looking into how the data changes regarding different circumstances such as location and time of day. Also looking at what caused the fires and the resulting damages related to these possible factors. Geographical data showed an increase in frequency around the downtown area especially for fires occurring

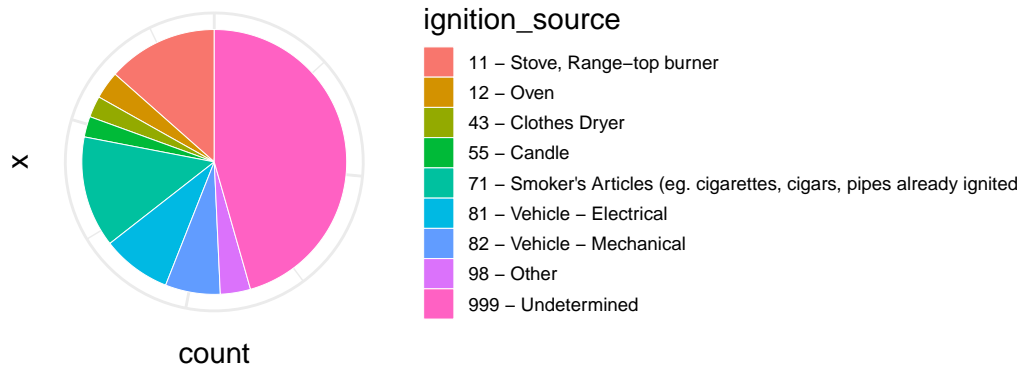


Figure 4: pie chart of the top 10 most common ignition sources

on properties. Outside of downtown we see a higher frequency of low damage off property incidences, with more sporadic higher damage events. Insight was also given to which times of the day result in a higher frequency of fires responded to by the Toronto fire department. Night was seen to have a much higher frequency, and far fewer in the day on average. Finally we observed the most common causes that lead to these causes seeing that smoker and kitchen appliances would be a good focus for education in the name of prevention. Policy makers can take the findings of this data to allocate resources, and understand what prevention measures are most effective.

- Ontario.ca. 2022. “Ontario Population Projections.” <https://www.ontario.ca/page/ontario-population-projections>.
- 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>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- 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://CRAN.R-project.org/package=dplyr>.