Toronto's Fire Incidents and Contributing Factors: An Explorative Analysis*

Jimmy Luc

April 18, 2024

This paper uses publicly available data from Toronto Open Data to identify contributing factors to fire incidents and their resulting financial losses. 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

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

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 damage of fire incidents is crucial for implementing effective fire prevention measures to minimize the impact of such incidents. In this analysis, I explore the dataset provided by Toronto Open Data (Toronto Open Data 2024) which captures fire incidents from 2011 to 2022 within the Toronto area as defined by the Ontario Fire Marshal (OFM). Through this dataset, I aim to identify contributing factors to fire incidents and their resulting financial losses.

The motivation behind this analysis stems from the importance of fire safety. Beyond the protection of assets, it's about keeping people safe. To ensure the safety of fellow Canadians, it is imperative to comprehend the underlying patterns and determinants of fire incident. Analyzing factors such as the fire's 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 analyzing the factors that contributing to the estimated dollar loss per fire incident in the Toronto area such as sociological factors. To address this gap, I employ a data-driven approach, utilizing explorative data analysis techniques to reveal patterns and relationships within the dataset. My 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 fire incidents and their associated financial losses, 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 Section 2. Subsequently, I present my findings, discussing the relationships uncovered and their implications for fire safety in Section 3. Finally, I conclude by highlighting the importance of the results by discussing the current laws and regulations in place and suggest avenues for future research in the domain of fire safety in Section 4.

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), `arrow` (Richardson et al. 2024), `lubridate` (Grolemund and Wickham 2011), and `scales` (Wickham, Pedersen, and Seidel 2023). Further details regarding data cleaning, variables of interest, and measurement methodologies will be discussed in the following subsections.

2.1 Data Source and Collection

The data for this paper was acquired through the R package `opendatatoronto` (Gelfand 2022), which provides publicly available datasets to be freely used, reused, and redistributed. Specifically, the "Fire Incidents" dataset is used as the basis of this analysis as it contains information regarding fire incidents in the area of Toronto provided by the Toronto Fire Services (TFS). The dataset captures incidents from the beginning of the year 2011 to the end of 2022. It is important to note that the dataset only includes incidents as defined by the Ontario Fire Marshal (Government of Ontario 2023b).

2.2 Variables of Interest

The dataset acquired contains various data points for analysis, but I only concern myself with a select few. In particular, the incident's area of origin, the estimated dollars loss in damages, the longitude and latitude of the incident, the number of responding personnel from the TFS, the possible causes of the incident, the determined property use, the presence, functionality and configuration of fire alarms, and finally the TFS notification and arrival time to the incidents.

2.3 Measurement

The methodologies behind the dataset's measurement is meticulously organized to account for scalability. Throughout the dataset, there are classifications methods used in scenarios where many instances require categorization. This is a decision that made with the concept of scalability in mind as there are far too many descriptors in the real world to account such that a singular description sufficiently and accurately describes the events of each incident. For example, regarding the area of origin, the raw data directly from Toronto's Open Data contains 74 unique classifications for an area of origin. The authors of the dataset hold the responsibility to properly assign the classification of each incident. This may result in inaccurate classifications without a clearly defined outline of what differentiates one from another.

One difficulty with measuring data concerning fire incidents is the nature of fire; fire is destructive and has the capacity to reduce tangible data to ash. The dataset has many instances

of undetermined data which is likely related to nature of the incident. Taking a look at the types of smoke alarms at the locations of these incidents, nearly a fifth of all incidents recorded from 2011 to 2022 in Toronto are undetermined. While there may be numerous justifications for such values, it is difficult to analyze the data without disregarding a large portion of data, potentially leading to inconsistencies in the analysis and measurement errors.

The dataset's concern for privacy is a bit alarming in regards to the provided latitude and longitude data points. The dataset contains information regarding the incident's nearest intersection which is generally comprised of two streets. While this alone does not breach the privacy of the victims of the fire incident, the latitude and longitude measurements bring you directly to the location of the incident. This may not be a concern for those two reside within high-rise apartments, but when considering detached homes with unique home addresses, this does raise some privacy concerns.

3 Results

In Figure 1, we can see the total number of fire incidents that occur each year from 2011 to 2022 in Toronto. While there is not much to decipher, we find that the number is relatively similar each year for the past decade. While some may argue that fire incidents have not reduced or increased substantially each year, another way of interpreting these numbers is that fire incidents have maintained a stable and infrequent occurrence in Toronto.

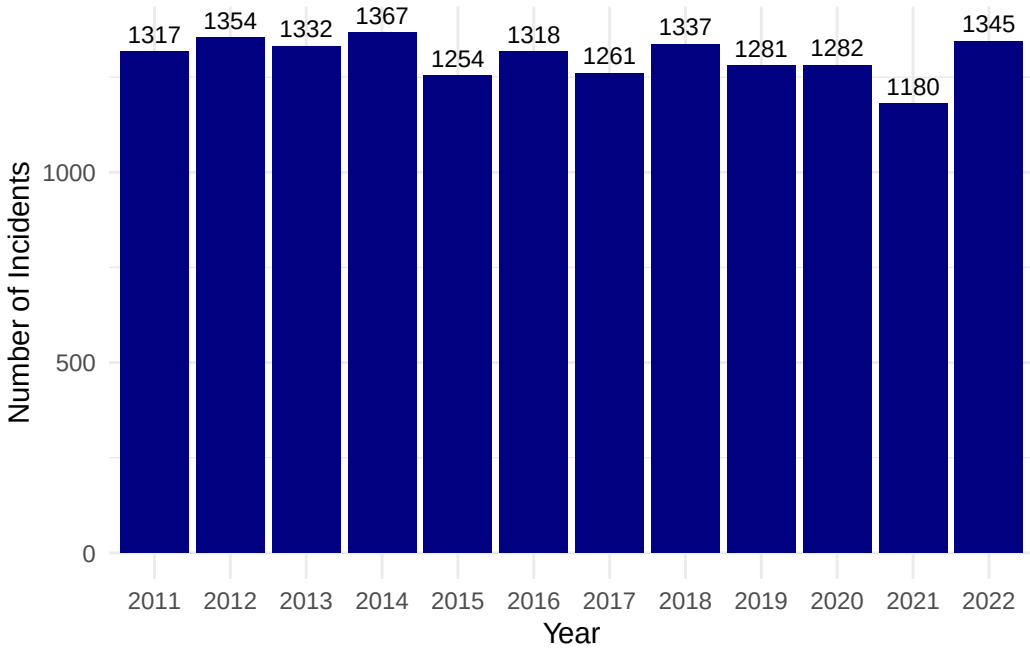

Figure 1: Number of fire incidents in Toronto by Year (2011-2022)

The dataset captures fire incidents that occur and are reported for the Toronto Fire Services and Figure 2 depicts where these incidents are located geographically. We can observe a greater number of incidents taking place in the heart of Toronto. This is likely due to the higher concentration of people with a population density of 3,088 persons per square kilometer which is the densest urban area or population centre in North America (Statistics Canada 2021). Looking further away from the city center, we find a higher concentration of incidents in the form of vertical lines that follow major roadways. Simply looking at the raw fire incident data, we find vehicles to be the third highest contributing factor to the source of ignition when considering electrical and mechanical incidents.

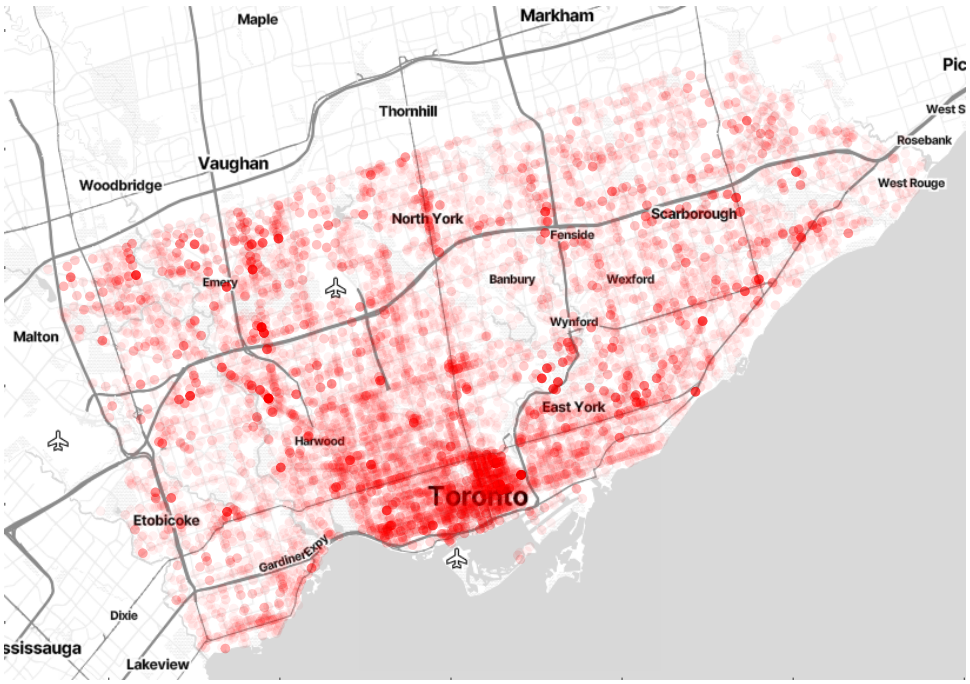

Figure 2: Heatmap of Toronto's fire incidents (2011-2022)

Looking at Figure 3, the plot illustrates the average dollar loss in damages from fire incidents across various categorized areas of origin. Each bar representing a distinct area of origin, with the height indicating the average amount in damages. The categories with the highest average amount in damages all share the fact they are all highly populated areas or require the utmost fire safety precaution. While the leading area of origin is "other/undetermined", it is largely because of the dataset was unable to classify the incident to one particular area hence the high estimated loss due to damages. The largest estimated dollar loss recorded in the dataset which is estimated at $50 million occurred on May 7, 2019 at the York Memorial Collegiate Institute (Freeman 2021). This is one example of an incident that is categorized under "other/undetermined" which resulted in a high dollar loss. As there are far too many classifications for large buildings that may serve many purposes, it is likely the reason such area of origin is leading.

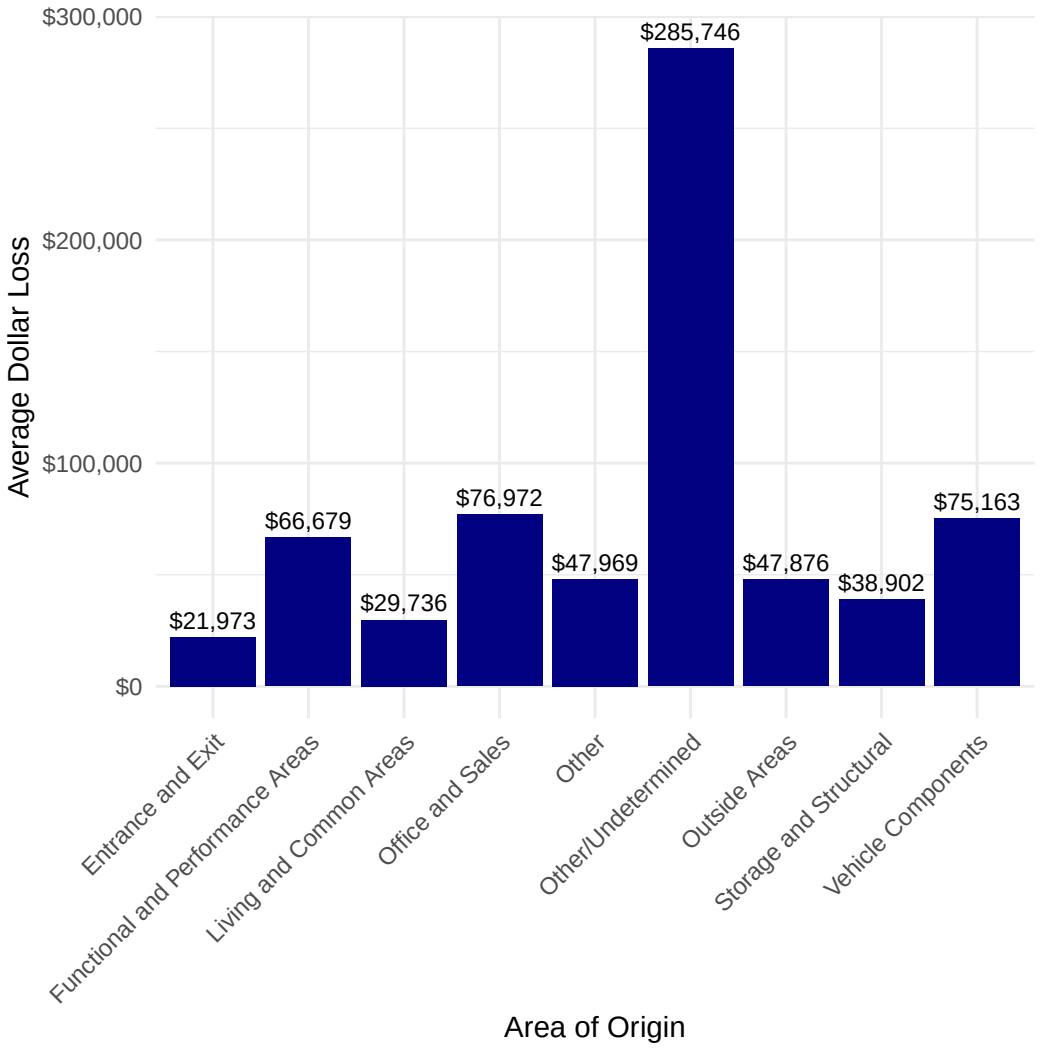

Figure 3: Average estimated dollar loss by grouped area of origin

Figure 4 shows the number of each type of smoke alarms found present during fire incidents in Toronto from 2011 to 2022. Each bar representing a specific type or configuration of smoke alarm, while the height corresponds to the number of fire incidents associated with that particular fire alarm. We can see wireless smoke alarms to be the lowest number of smoke alarms found at the location of the incidents with the remaining smoke alarm types being within a reasonable range of each other. No alarm/undetermined is the most frequent which is likely related to the nature of the emergency as fire alarms may have been destroyed as a result.

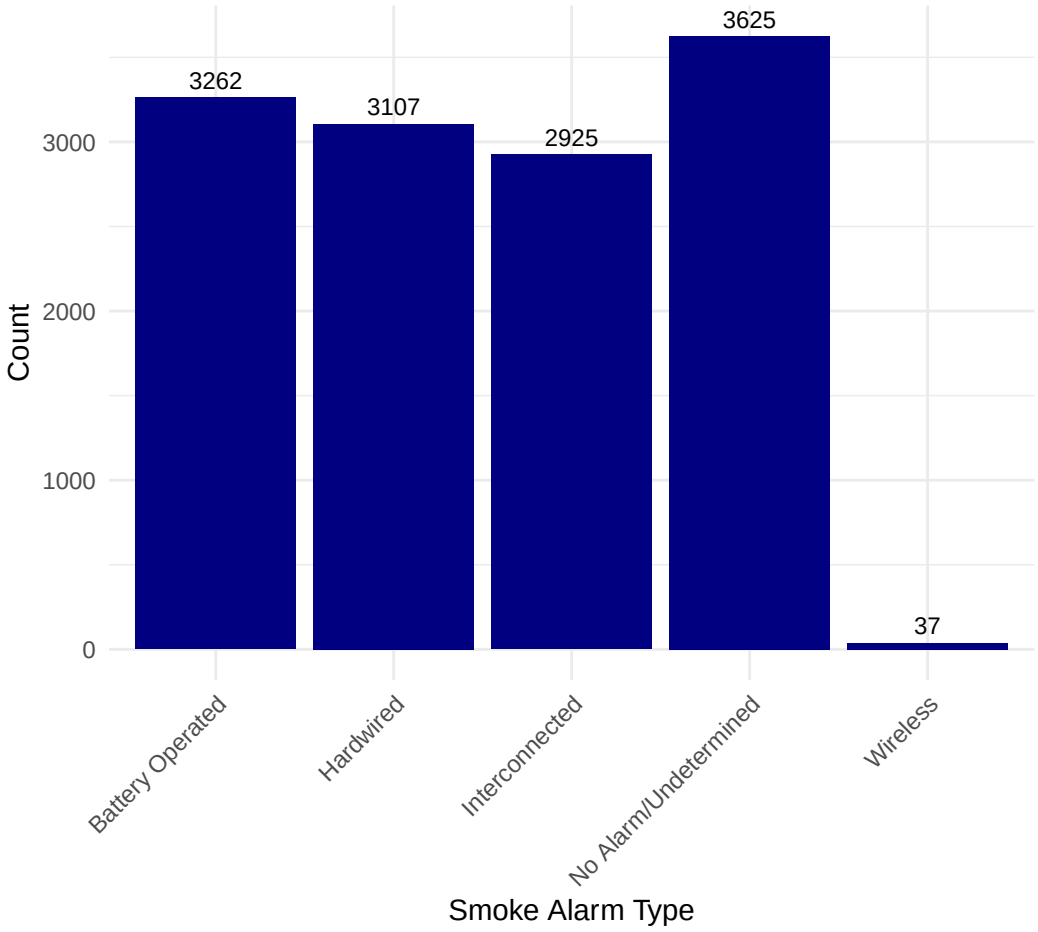

Figure 4: Number of each type of smoke alarm present during fire incidents in Toronto

4 Discussion

4.1 First discussion point

This analysis shows that while there are many contributing factors and places that fire may occur causing potentially life-altering financial loss or loss of life, there are only a few safety measures to be made that can mitigate such events from happening. The Government of Ontario outlines home fire safety information including precautionary measures, laws and regulations regarding fire safety that is imperative for any homeowner, landlord, and tenant to ensure the safety of those around you (Government of Ontario 2023a). While it is required for every home to have a smoke alarm on every story, the City of Toronto recommends replacing the batteries of your smoke alarms every two years (City of Toronto, n.d.). As most fatal

fires occur at night when people are asleep, after seeing the types of fire alarms present at fire incidents in Figure 4, it is advised that fire alarms be interconnected which would trigger all connected alarms if one where to ever activate. Making such change would be an incentive for public safety and also promote fire safety for all of those who reside within multiple unit dwellings especially in the busy city of Toronto. There are many systems that are already in place as preventative measures for the threat of fires almost everywhere today. From our schools to the underground subways, fire safety is present, but it is everyone's responsibility to be aware and understand the necessary procedures in the event of a fire. While Toronto is renowned for its diversity, one thing most would agree with is that fire has the capacity to bring happiness, despair, or destruction. It is important to acknowledge fire safety as it saves lives, protects our belongings, and ensures the safety of our communities from the devastating impacts of fires.

4.2 Weaknesses and next steps

A weakness of this analysis is the integrity of the dataset after cleaning and processing which is outlined in Section A. This analysis could be improved in regards to the consideration of sociological perspectives concerning peoples feelings towards fire safety. I believe discussing the awareness and sentiments towards fire safety and the steps they have taken to ensure the well-being for themselves and those around them. Including some survey or statistical evidence through a sociological lens would greatly benefit the overarching message of fire safety. Additionally, the analysis of the Toronto Fire Services response time to emergency calls could have been analyzed to determine whether or not it has changed in recent years as Toronto continues to increase in population. In Figure 2 we see that the most densely populated locations of Toronto also have the most concentrated number of fire incidents and incorporating some data model to predict the estimated dollar loss in populated areas would inform those within these areas of the dangers of fire.

A Appendix A

A.1 Additional data details

The dataset acquired through Toronto's Open Data (Gelfand 2022) was processed using `R` (R Core Team 2023) which omitted all "NA" values as once one was present in any instance in the data, it was common to find many "NA" values in data points of the respective instance. This decision reduced the dataset from a total of 29,425 instances of data to 15,628 instances. Additionally, the data point `area_of_origin` was reorganized into broader categories for analysis as there were initially 74 unique areas of origin in the dataset. Furthermore, in Figure 4 the variable of interest is `smoke_alarm_at_fire_origin_alarm_type` which looked at the types of fire alarms without accounting for "undetermined" types as it provided little insight towards the goal of the paper.

A.2 Datasheet

Extract of the questions from Gebru et al. (2021).

Motivation

1. *For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.*
    - The dataset provides more detailed information compared to the basic dataset that the Toronto Fire Service (TFS) responds to. The dataset was acquired through Toronto's Open Data which is an open source delivery tool that aims to connects people and data. These publicly available datasets are targeted towards various demographics such as app developers, students, or citizens.
2. *Who created the dataset (for example, which team, research group) and on behalf of which entity (for example, company, institution, organization)?*
    - Toronto's Open Data, an open data initiative by the City of Toronto government.
3. *Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.*
    - No funding was received.
4. *Any other comments?*
    - No.

Composition

1. *What do the instances that comprise the dataset represent (for example, documents, photos, people, countries)? Are there multiple types of instances (for example, movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.*

    - Each row of the dataset is an individual instance of a fire incident that is recorded. Each column represents a specific point of information for each incident.

2. *How many instances are there in total (of each type, if appropriate)?*

    - 29425

3. *Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (for example, geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (for example, to cover a more diverse range of instances, because instances were withheld or unavailable).*

    - All individuals included in the dataset are reported by the Ontario Fire Marshal (OFM) up to December 31, 2022. Incidents in the dataset that include incomplete data may be under investigation or is classified as a no loss outdoor fire. Some incidents have been excluded pursuant to exemptions under Section 8 of Municipal Freedom of Information and Protection of Privacy Act (MFIPPA). For privacy purposes, all personal information regarding the fire incidents have not been included in the dataset.

4. *What data does each instance consist of? "Raw" data (for example, unprocessed text or images) or features? In either case, please provide a description.*

    - Each instance consists of information regarding the fire incident such as the time that the TFS received an emergency call to the time the respondents have arrived at the scene, estimated loss in damages, the classified source of the incident, etc. Various classifications for each incident was used throughout the dataset

5. *Is there a label or target associated with each instance? If so, please provide a description.*

    - Yes, there is a unique id associated with each instance. This id is a unique number that represents the incident in the dataset.

6. *Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (for example, because it was unavailable). This does not include intentionally removed information, but might include, for example, redacted text.*

    - Yes, there is missing information throughout the raw dataset. Information such as area of origin, building status, building impact, estimated dollars loss due to damages, and more. This information is likely unavailable due to ongoing investigations or information that was never recorded or received.

7. *Are relationships between individual instances made explicit (for example, users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.*

    - No, ther eare no relationships between instances.

8. *Are there recommended data splits (for example, training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.*

    - No.

9. *Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.*

    - No.

10. *Is the dataset self-contained, or does it link to or otherwise rely on external resources (for example, websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (that is, including the external resources as they existed at the time the dataset was created); c) are there any restrictions (for example, licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.*

    - Self-contained.

11. *Does the dataset contain data that might be considered confidential (for example, data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? If so, please provide a description.*

    - No.

12. *Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.*

    - No.

13. *Does the dataset identify any sub-populations (for example, by age, gender)? If so, please describe how these subpopulations are identified and provide a description of their respective distributions within the dataset.*

    - No.

14. *Is it possible to identify individuals (that is, one or more natural persons), either directly or indirectly (that is, in combination with other data) from the dataset? If so, please describe how.*

    - No.

15. *Does the dataset contain data that might be considered sensitive in any way (for example, data that reveals race or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)? If so, please provide a description.*

    - No.

16. *Any other comments?*

    - No.

Collection process

1. *How was the data associated with each instance acquired? Was the data directly observable (for example, raw text, movie ratings), reported by subjects (for example, survey responses), or indirectly inferred/derived from other data (for example, part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.*

    - The data was gathered by the TFS throughout the occurrence of the incident, first beginning with a emergency phone call or report of a fire incident. Information such as estimated loss in damages is likely reported by the victims of the incident. A collaborative effort to collect information throughout the process.

2. *What mechanisms or procedures were used to collect the data (for example, hardware apparatuses or sensors, manual human curation, software programs, software APIs)? How were these mechanisms or procedures validated?*

    - Telephone. Citizen reports and trained operators at the TFS.

3. *If the dataset is a sample from a larger set, what was the sampling strategy (for example, deterministic, probabilistic with specific sampling probabilities)?*

    - The dataset is not a sample from a larger set.

4. *Who was involved in the data collection process (for example, students, crowdworkers, contractors) and how were they compensated (for example, how much were crowdworkers paid)?*

    - Emergency call operators, firefighters, and citizens. Emergency dispatchers and firefighters classify as government employees and have their salaries publicly available by the Canadian government. Citizens that provide information are voluntary and do not receive monetary gain.

5. *Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (for example, recent crawl of old*

news articles)? If not, please describe the timeframe in which the data associated with the instances was created.

- The dataset contains fire incidents from 2011 to 2022 which accurately represents the time frame of the data.

6. *Were any ethical review processes conducted (for example, by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.*

   - No.

7. *Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (for example, websites)?*

   - The data was obtained from Toronto's Open Data which was published by Toronto Fire Services.

8. *Were the individuals in question notified about the data collection? If so, please describe (or show with screenshots or other information) how notice was provided, and provide a link or other access point to, or otherwise reproduce, the exact language of the notification itself.*

   - Information is recorded by the government and associated government services. All 911 emergency calls are recorded for government records and all information protected through the MFIPPA (Municipal Freedom of Information and Protection of Privacy Act) will not be disclosed.

9. *Did the individuals in question consent to the collection and use of their data? If so, please describe (or show with screenshots or other information) how consent was requested and provided, and provide a link or other access point to, or otherwise reproduce, the exact language to which the individuals consented.*

   - All individuals are subject to the MFIPPA and have the information concerning the incident recorded for government records.

10. *If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses? If so, please provide a description, as well as a link or other access point to the mechanism (if appropriate).*

    - Information collected has revoked personal information for privacy purposes. Government records are subject to the MFIPPA.

11. *Has an analysis of the potential impact of the dataset and its use on data subjects (for example, a data protection impact analysis) been conducted? If so, please provide a description of this analysis, including the outcomes, as well as a link or other access point to any supporting documentation.*

    - No.

12. *Any other comments?*
    - No.

Preprocessing/cleaning/labeling

1. *Was any preprocessing/cleaning/labeling of the data done (for example, discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remaining questions in this section.*
    - No, no cleaning has been stated.
2. *Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (for example, to support unanticipated future uses)? If so, please provide a link or other access point to the "raw" data.*
    - No.
3. *Is the software that was used to preprocess/clean/label the data available? If so, please provide a link or other access point.*
    - None that is disclosed.
4. *Any other comments?*
    - No.

Uses

1. *Has the dataset been used for any tasks already? If so, please provide a description.*
    - No.
2. *Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point.*
    - No.
3. *What (other) tasks could the dataset be used for?*
    - General fire prevention measures.
4. *Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? For example, is there anything that a dataset consumer might need to know to avoid uses that could result in unfair treatment of individuals or groups (for example, stereotyping, quality of service issues) or other risks or harms (for example, legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?*

- No.

5. *Are there tasks for which the dataset should not be used? If so, please provide a description.*

   - No.

6. *Any other comments?*

   - No.

Distribution

1. *Will the dataset be distributed to third parties outside of the entity (for example, company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.*

   - No.

2. *How will the dataset be distributed (for example, tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?*

   - The dataset is publicly available at the Open Data portal. It can be downloaded directly, through API request, or the `opendatatoronto` R package.

3. *When will the dataset be distributed?*

   - The dataset is available now.

4. *Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/ or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.*

   - The dataset is subject to Ontario's Open Government Licence outlined in the Open Data website.

5. *Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.*

   - No.

6. *Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.*

   - No.

7. *Any other comments?*

- No.

Maintenance

1. *Who will be supporting/hosting/maintaining the dataset?*
    - The dataset is maintained by Toronto's Open Data and the TFS.
2. *How can the owner/curator/manager of the dataset be contacted (for example, email address)?*
    - The publisher of the dataset can be contacted at the provided email address.
3. *Is there an erratum? If so, please provide a link or other access point.*
    - No.
4. *Will the dataset be updated (for example, to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (for example, mailing list, GitHub)?*
    - Yes, the dataset is stated to refresh annually. It appears as through the data is refreshed daily automatically.
5. *If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (for example, were the individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced.*
    - No.
6. *Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers.*
    - No.
7. *If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.*
    - No.
8. *Any other comments?*
    - No.

References

City of Toronto. n.d. "Smoke Alarms." https://www.toronto.ca/community-people/public-safety-alerts/safety-tips-prevention/safety-equipment-devices/smoke-alarms/.

Freeman, Joshua. 2021. https://toronto.ctvnews.ca/tdsb-files-90m-lawsuit-against-city-police-and-province-over-2019-fire-that-gutted-west-end-toronto-high-school-1.5415600.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92.

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

Government of Ontario. 2023a. "Fire Safety at Home." https://www.ontario.ca/page/fire-safety-home.

———. 2023b. "The Office of the Fire Marshal." https://www.ontario.ca/page/office-fire-marshal.

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/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

RStudio Team. 2020. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC. http://www.rstudio.com/.

Statistics Canada. 2021. "2021 Census of Population." Government of Canada. https://www12.statcan.gc.ca/.

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.