Predicting Fire Incident Impact in Toronto: Financial Losses and Casualties*

Hechen Zhang

April 18, 2024

The present study examined the Fire incident data for the City of Toronto in order to elucidate any associations between fire incidents and numerous independent variables. The research revealed that these variables exhibit a significant predictive capability in determining the probability and consequences of fire occurrences, encompassing estimated financial damage and casualties. The study identified building status, business impact, fire alarm system operation, method of fire control, and presence of smoke alarms at the fire origin as noteworthy factors for predicting fire incidents. In addition, the arrival time to fire control and the duration from alarm to arrival were found to be significant predictors of fire incidents, emphasizing the criticality of a prompt response time. The results of the study make a valuable contribution to the existing body of knowledge on fire incidents and bear significant implications for public safety and strategies aimed at prevention. Subsequent research endeavors may concentrate on the development of a prognostic framework for fire occurrences derived from the noteworthy determinants unveiled in this investigation, as well as an assessment of the efficacy of various fire mitigation tactics.

Table of contents

*Code and data are available at: https://github.com/HechenZ123/Analyzing-Fire-Incident-Risks-in-Toronto.git

1 Introduction

The occurrence of fires represents a considerable risk to civic welfare, as they have the potential to result in loss of life and property damage, with consequent implications for the economic stability of affected regions. Annually in Toronto, the fire department is confronted with a substantial number of fire incidents, underscoring the significance of understanding the root causes of these occurrences. Prior research has explored different elements of fire incidents, including their origins, consequences, and the time it takes for a response. The current literature lacks information on how fire incidents are related to different factors like building status, business impact, fire alarm system operation, method of fire control, presence of a smoke alarm at fire origin, presence of a sprinkler system, time taken to control the fire, and time taken for rescue teams to arrive after the alarm. This research aims to analyze fire incident data for Toronto and predict potential financial losses and injuries based on certain factors.

The document includes Data Section with an overview of the data and methods used in the study, including data cleansing, reformatting, and statistical analysis. The Results Section includes analysis results such as descriptive statistics, chi-square tests, ANOVA tests, and logistic regression. In the Discussion Section, I summarized the research findings, discussed study limitations, and recommended future research directions. This research aims to identify the correlation between fire incidents and independent factors and forecast financial loss and casualties. The findings from this study shows that the status of building, impact on business after fire, the method used to control the fire, if there is smoking alarm at the sight of incident, the time taken by the fire fighters to response and the time taken to control the fires are the major variables that increase the risk of fatalities and money loss in the fire incidents. The results from this study can help to reduce the fire incidents in future by taking the measurable steps by both the civilians and the fire department. The following sections will further cover the data and methods use in the analysis, results of the study comprising tables and graphs and the discussion.

2 Data

The data was extracted using the package Gelfand (2020). The fire incidents in the dataset was reported between the year 2011 to 2022 by Toronto fire services.The data was transferred to Ontario fire marshal for reporting purposes and used for analysis.

The dataset contains 29425 observations and 43 variables. However, for this analysis, we will focus on the following variables as shown in Table 1 below.

Table 1: Description of variables used in the analysis

Variable Name	Description
Civilian Casualties	Civilian casualties observed at the scene

Estimated Dollar Loss	Estimated dollar loss
Building Status	OFM Building status code and description
Business Impact	OFM Business Impact code and description
Fire Alarm System Operation	OFM Fire Alarm System Operation code and description
Method Of Fire Control	OFM Method Of Fire Control code and description
Smoke Alarm at Fire Origin	OFM Smoke Alarm at Fire Origin code and description
Sprinkler System Presence	OFM Sprinkler System Presence code and description
TFS Alarm Time	Timestamp of when TFS was notified of the incident
TFS Arrival Time	Timestamp of first arriving unit to incident
Fire Under Control Time	Timestamp of fire under control
Number of responding apparatus	Number of TFS responding apparatus
Number of responding personnel	Number of TFS responding personnel

The chosen variables in this analysis offer understanding into the consequences of fire incidents, how the TFS reacts, and the elements influencing the spread and management of fires. It's worth mentioning that no personal information is included, and specific addresses have been combined to the nearest major or minor intersection to protect privacy. This paper employs the R programming language (R Core Team 2023) for its analysis. Various libraries such as tidyverse (Wickham et al. 2019), dplyr (Wickham et al. 2023), readr (Wickham, Hester, and Bryan 2024), ggplot2 (Wickham 2016), knitr (Xie 2017), kableExtra (Zhu 2024), grid (R Core Team 2024), gridExtra (Auguie 2017), psych (William Revelle 2024), arrow (Richardson et al. 2024), patchwork (Pedersen 2024) and broom (Robinson, Hayes, and Couch 2023) were utilized. The information is brought in from a CSV file, then processed by eliminating any empty values and generating new variables. The information is then examined using statistical techniques such as summary statistics, chi-square tests, ANOVA tests, logistic regression, and linear regression. The findings are displayed through tables and figures generated with ggplot2, knit, and kableExtra. The ultimate model is chosen through step-by-step removal of variables.

The missing observations were identified and then removed from the data. The final sample size for the data was 14264.

2.1 Measurements:

The measurement of this study involves creating new variables and recoding existing ones to better understand their relationship with fire incidents. We created a new variable for risk type based on the criteria of civilian casualties, and estimated dollar loss. The risk type was divided into five categories Low, very low, moderate, high, and very high. We also recorded existing variables such as building status, business impact, fire alarm system operation, method of fire control, smoke alarm at fire origin, and sprinkler system presence into categorical variables. We alos calculated the difference between TFS arrival time and TFS alarm time to determine the alarm to arrival time and the difference between fire under control time and TFS arrival time to determine the arrival to fire control time.

Below are some descriptive statistics and graphs for the variables selected:

Table 2 presents an overview of statistical data pertaining to four variables associated with occurrences of fire incidents. The average time it takes for fire control teams to arrive at the scene is 0.25 hours, with a standard deviation of 0.79. The time it takes to control fire from arrival is 5 hours, with a standard deviation of 1,39 hours. Additionally, the average number of fire trucks or other vehicles dispatched to a fire incident is approximately 9, while the average number of firefighters or other personnel dispatched to a fire incident is approximately 31. The average estimated dollar loss is 44,743.18, with the standard deviation of 505,450.18 dollars which shows that the estimated dollar loss is highly skewed.

Table 2: Summary Statistics Table

	Mean	SD	Median	Min	Max
Arrival to fire control time	0.25	0.79	0.11	0.0	3.347e+01
Alarm to arrival time	4.99	1.39	4.87	0.3	2.443e+01
Number of responding appartaus	9.46	8.49	6.00	1.0	4.360e+02
Number of responding personnel	31.12	26.06	22.00	2.0	1.275e+03
Estimated loss of dollar	44,743.18	502,450.18	3,000.00	0.0	5.000e+07

The dependent variable risk type is shows in Figure 1, indicates that 1823 with high risk, 1444 with low risk, 2476 with moderate risk, 1291 with very high risk and 7230 most of with very low risk.

Figure 2 shows the histogram for alarm to arrival time. It was calculated by taking the difference between the time TFS arrived at the incident and the time when TFS was notified. The histogram shows the minutes range from less than 5 to more than 200 minutes. The data shows the approximately normal distribution.

Figure 3 shows the histogram for arrival to fire control time. It was calculated by taking the difference between the time the fire was completely controlled and the time TFS arrived at the incident. The histogram shows the hours range from 0.1 to 30 hours. The data shows the non normal distribution.

2.2 Models:

We performed logistic regression to determine the relationship between fire incidents and various independent variables. We will also perform backward elimination to select the most significant variables.

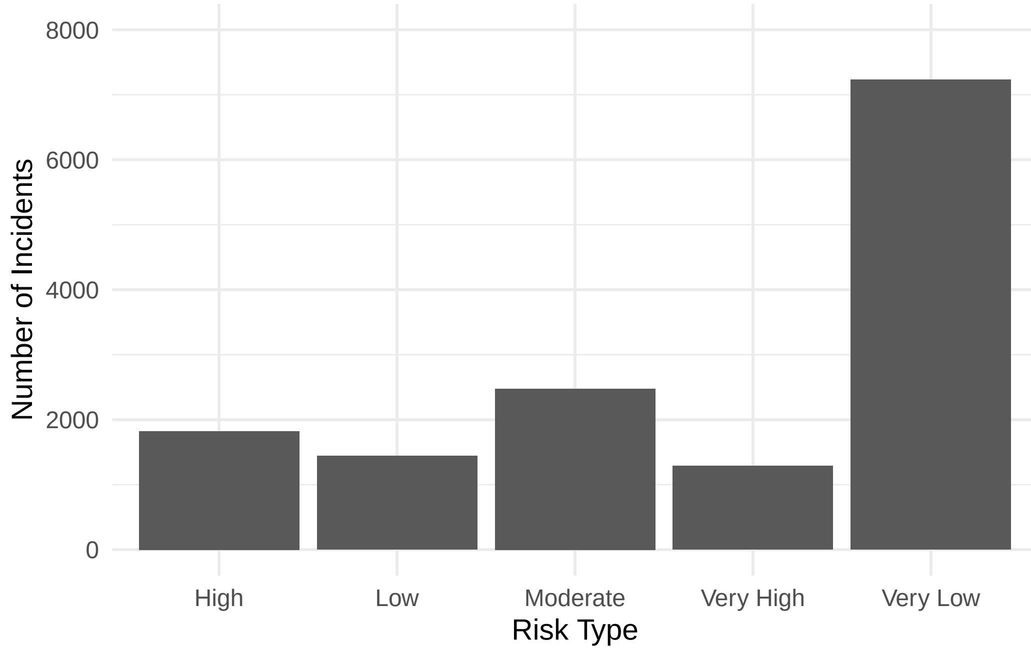

Figure 1: Bar graph for risk type

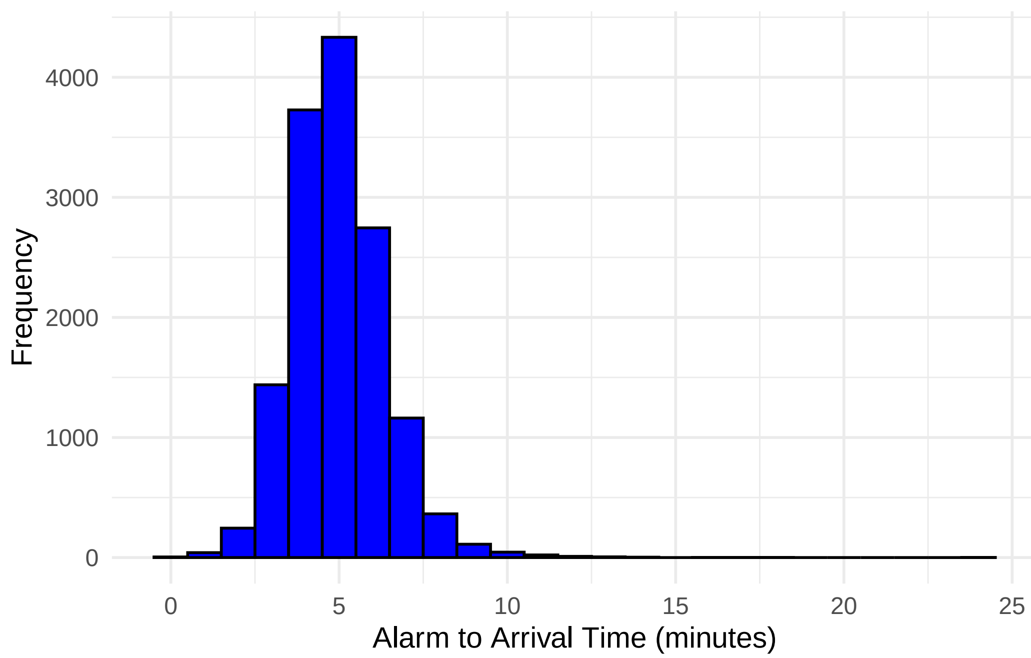

Figure 2: Histogram of Alarm to Arrival Time

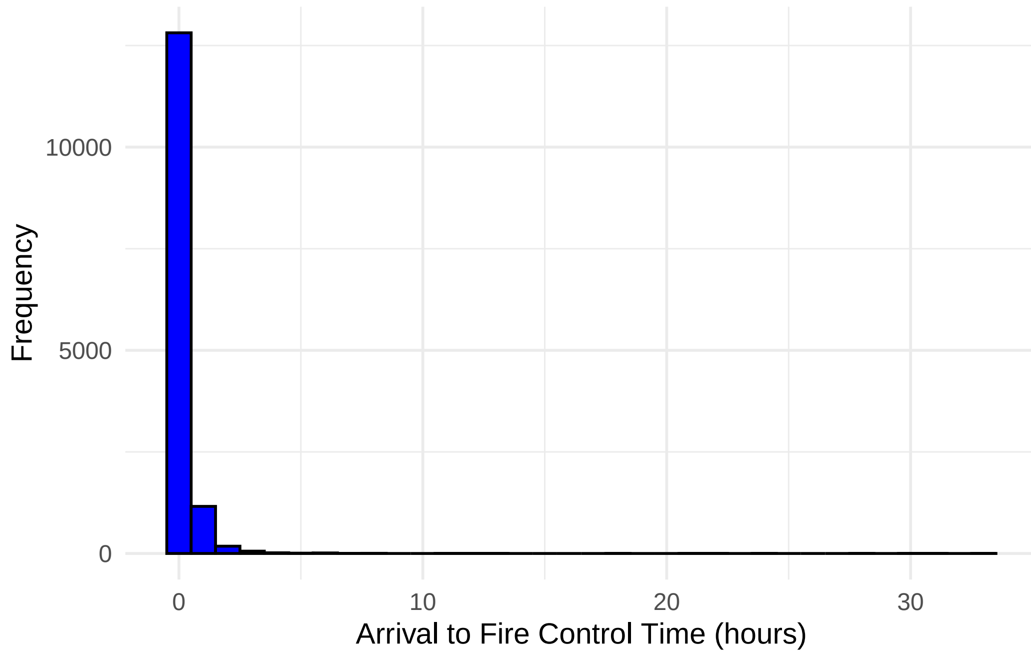

Figure 3: Histogram of Alarm to Arrival Time

2.2.1 Logistic Regression Model:

The binary response variable in question is representative of the categorization of fire incidents based on their risk level. In this context, the designation Y = 1 denotes a high-risk incident, whereas Y = 0 signifies a low to moderate-risk incident. The independent variables (X1, X2, . , Xp) utilized in the model comprise building status, business impact, fire alarm system functionality, method of fire control, presence of smoke alarm at fire origin, the existence of a sprinkler system, time taken to arrive at the scene of the fire, time elapsed from alarm activation to arrival, the number of responding apparatus, and the number of responding personnel. The logistic regression model can be written as:

$$\begin{aligned}\text{logit}(Risktype) = \beta_0 + \beta_1^{\text{BuildingStatus}} + \beta_2^{\text{BusinessImpact}} + \beta_3^{\text{FireAlarmSystemOperation}} \\ +\beta_4^{\text{MethodofFireControl}} + \beta_5^{\text{SmokeAlarmatFireOrigin}} + \beta_6^{\text{SprinklerSystemPresence}} + \beta_7^{\text{Alarmtoarrivaltime}} \\ +\beta_8^{\text{Arrivaltofirecontroltime}} + \beta_9^{\text{Numberofrespondingapparatus}} + \beta_{10}^{\text{Numberofrespondingpersonnel}}\end{aligned}$$

where p is the probability of a fire incident being high risk, and β_0, β_1, ..., β_p are the regression coefficients estimated by the model.

2.2.2 Linear Regression Models:

I used two linear regression models, for estimated dollar loss and for number of casualties in a fire incident. Let X1, X2, ..., Xp be the independent variables used in the model, The linear regression model can be written as:

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) \\
Estimateddollarloss = \beta_0 + \beta_1 \times \text{BuildingStatus} + \beta_2 \times \text{BusinessImpact} \\
+\beta_3 \times \text{FireAlarmSystemOperation} + \beta_4 \times \text{MethodofFireControl} \\
+\beta_5 \times \text{SmokeAlarmatFireOrigin} + \beta_6 \times \text{SprinklerSystemPresence} \\
+\beta_7 \times \text{Alarmtoarrivaltime} + \beta_8 \times \text{Arrivaltofirecontroltime} \\
+\beta_9 \times \text{Numberofrespondingapparatus} + \beta_{10} \times \text{Numberofrespondingpersonnel}
\end{aligned}
$$

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) \\
Numberofcasualties = \beta_0 + \beta_1 \times \text{BuildingStatus} + \beta_2 \times \text{BusinessImpact} \\
+\beta_3 \times \text{FireAlarmSystemOperation} + \beta_4 \times \text{MethodofFireControl} \\
+\beta_5 \times \text{SmokeAlarmatFireOrigin} + \beta_6 \times \text{SprinklerSystemPresence} \\
+\beta_7 \times \text{Alarmtoarrivaltime} + \beta_8 \times \text{Arrivaltofirecontroltime} \\
+\beta_9 \times \text{Numberofrespondingapparatus} + \beta_{10} \times \text{Numberofrespondingpersonnel}
\end{aligned}
$$

where β_0, β_1, ..., β_p are the regression coefficients estimated by the model.

The selection of these models is deemed suitable for the given dataset and research inquiry, as they facilitate the determination of the impact of numerous independent variables on a binary or continuous dependent variable. The models are constructed utilizing the stepwise elimination technique, a method employed for the purpose of variable selection. The applicability of this method to the dataset is justified by the presence of a substantial quantity of independent variables. The final models were constructed by incorporating only those variables that demonstrated a statistically significant impact on the dependent variable. This approach was taken in order to streamline the models and enhance their comprehensibility.

3 Results

The Figure 4 shows that 75% of incidents that may not resume business are classified as moderate to high risk, while only 25% are classified as low or very low risk. On the other hand, 50% of incidents that may resume business are classified as moderate to high risk, while 25% are with no interruption in business classified as high or very high risk. The fire control category is also strongly associated with the risk type, the automated fire control was about 25% high risk type compared to others. Also, more than 50% of incidents with a no sprinkler system are classified as moderate risk, while 25% are classified as high or very high risk. Overall, the table shows that the risk type of a fire incident is strongly associated with several independent variables, including business impact, fire control, and system presence.

The chi-square tests Table 3 show that building status, business impact, fire alarm system operation, method of fire control, smoke alarm at fire origin, and sprinkler system presence all have p-values less than 0.05, indicating that there are significant differences in the distribution of these variables between high-risk and low-to-moderate risk fire incidents.

Table 3: Chi-square results for association between independent variables and risk type

	Variables	DF	Chi-square value	p-value
Building Status	Building_Status_Cat	12	115.00	0
Business Impact	Business_Impact_Cat	12	1,098.61	0
Fire alarm system operation	Fire_Alarm_System_Operation_Cat	12	457.43	0
method of fire control	Method_Of_Fire_Control_Cat	12	1,731.15	0
smoke alarm at fire origin	Smoke_Alarm_at_Fire_Origin_Cat	12	304.32	0
Sprinkler system presence	Sprinkler_System_Presence_Cat	12	494.97	0

The line graph in Figure 5 shows the association between the risk type and alarm to arrival time in minutes for fire incidents. The alarm to arrival time is grouped into intervals of 100 minutes. The chart shows that as the alarm to arrival time increases, the proportion of very low and low-risk incidents decreases, while the proportion of moderate and very high-risk incidents increases. Specifically, there are no very low-risk incidents with an alarm to arrival time greater than 15 minutes, and the moderate to very high-risk incidents increases at an alarm to arrival time more than 15 minutes. Overall, the chart suggests that a longer alarm to arrival time is associated with a higher risk of fire incidents outcomes.

The line graph in Figure 6 shows the association between the risk type and arrival to fire time in hours for fire incidents. The arrival to fire time is grouped into intervals of 10 hours. The chart shows that as the arrival to fire time increases, very low and low-risk incidents decreases, while the proportion of moderate and very high-risk incidents increases. Specifically, there are no very low-risk incidents with an alarm to arrival time greater than 10 hours, and the moderate to very high-risk incidents increases at an alarm to arrival time more than 10 hours. Overall, the chart suggests that a longer arrival to fire time is associated with a higher risk of fire incidents outcome.

For the continuous variables, the F-tests show in Table 4 that arrival to fire control time, alarm to arrival time, number of responding apparatus, and number of responding personnel all have p-values less than 0.05, indicating that there are significant differences in the distribution of these variables between high-risk and low-to-moderate risk fire incidents.

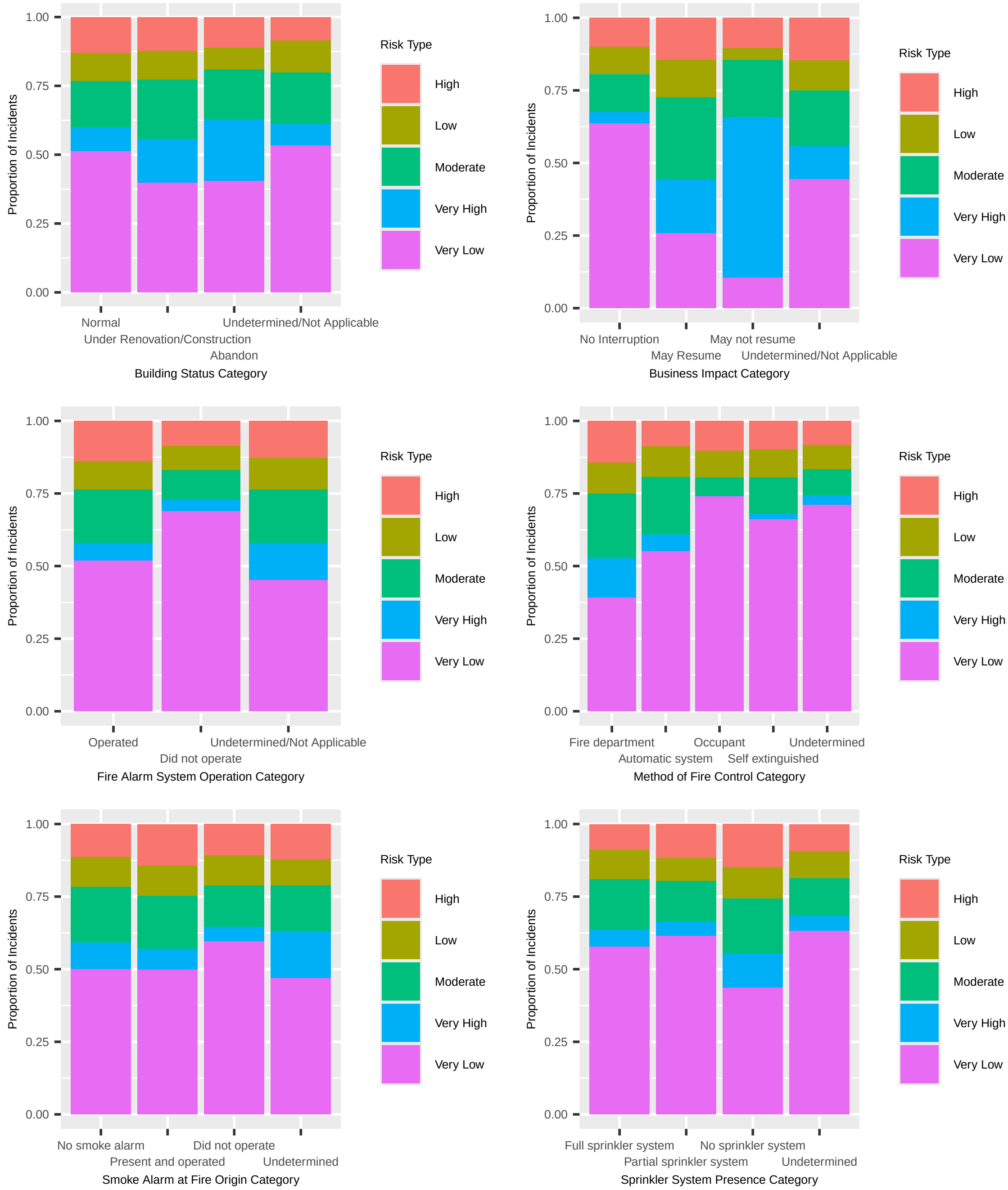

Figure 4: Bar graphs for assocition between risk type and independent variables

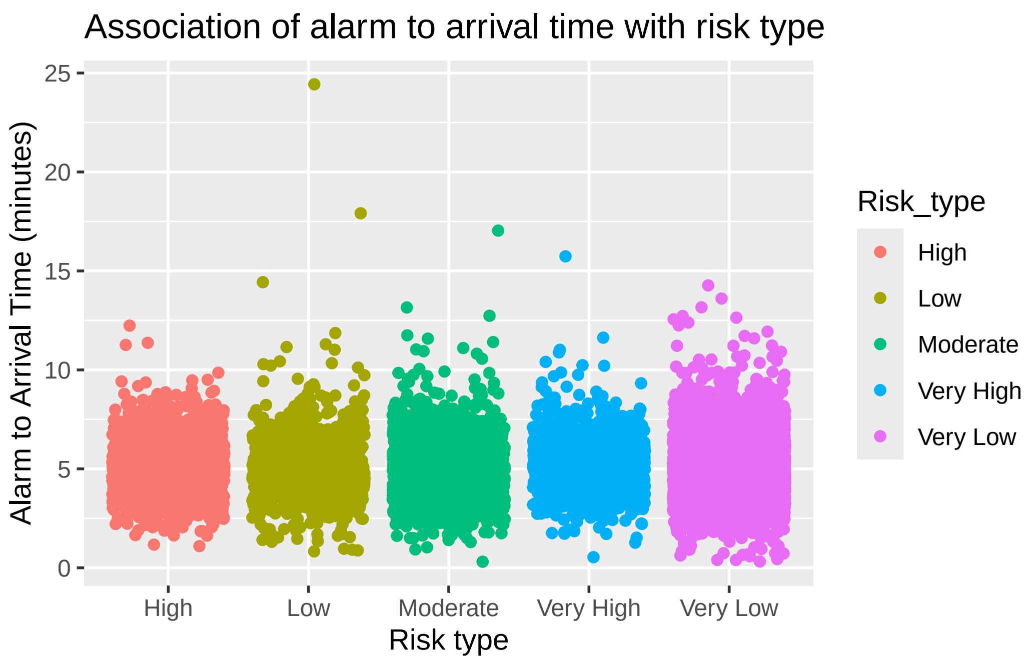

Figure 5: Line graph shows comparison of alarm to arrival time between various types of risk

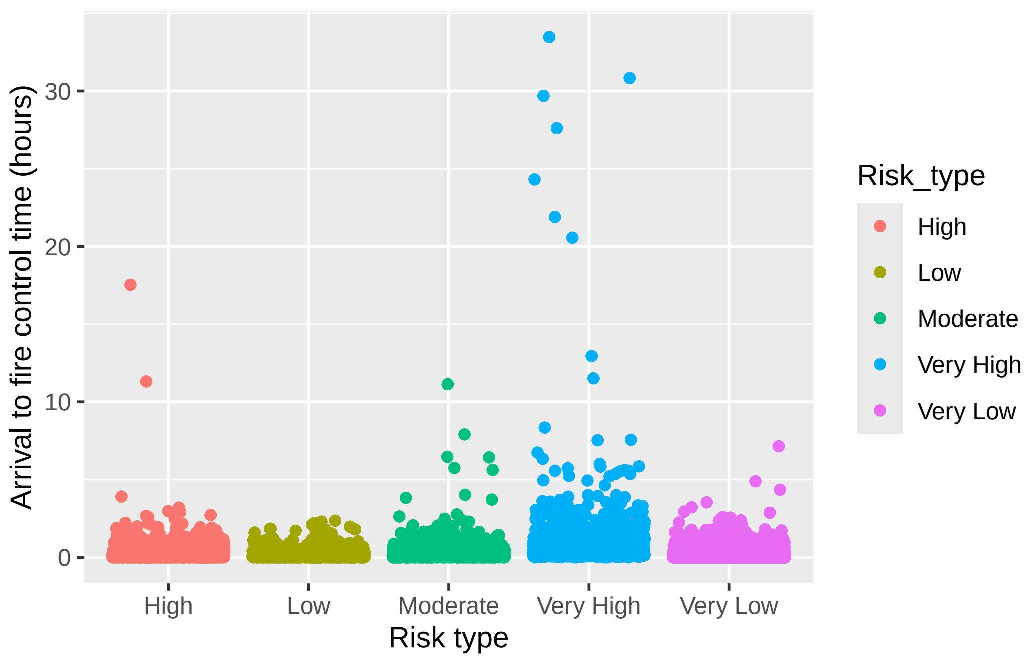

Figure 6: Line graph shows comparison of fire control time with type of risk

Table 4: ANOVA results from comparison between independent variables and risk type

	F-value	P-value
Arrival to fire control time	132.023	0
Alarm to arrival time	18.785	0
Number of responding appartaus	814.270	0
Number of responding personnel	795.786	0

These findings indicate that these variables could play a significant role in forecasting the potential for a fire incident. Analyzing the connections between these factors and the type of risk associated with a fire incident can help us create plans to minimize the risk and consequences of fire incidents.

Table 5 utilized a logistic regression model to establish the correlation between risk type (high versus low to moderate) and various independent variables such as building status, business impact, functioning of fire alarm system, method of fire control, presence of smoke alarm at fire origin, existence of sprinkler system, time taken to arrive at fire control, and time taken from alarm to arrival. The findings indicate that the factors considered in the model are strong predictors of the type of risk associated with a fire incident (p-value < 0.05). The final predictors after stepwise elimination can be seen in Equation 1. The building status coefficient states that the normal residential building has exp(1.058)= 2.88 times likelihood for high fire risk impact compared to abandoned building. Similarly, the building which was being renovated or under construction has odds of exp(0.8410) = 2.31 times high have fire risk impact compared to abandon building. The positive coefficients for arrival to fire control time and alarm to arrival time show that longer times for fire control and alarm response are linked to increased likelihood of a high-risk fire incident. More precisely, for each additional minute it takes for fire control to arrive, the likelihood of a high-risk fire incident increases by exp(0.462) = 1.59. Similarly, for every extra minute it takes for the alarm to arrival, the likelihood of a high-risk fire incident increases by exp(0.034)= 1.03, 3%. Apart from these, the coefficient for business impact shows that, the business which may resume after the fire incident has exp(-0.012)= 0.988, 2% less likely to have high fore risk impact compared to thr business which may not resume after the incident. On the other hand, the buildings where no sprinkler installed has exp(0.591)= 1.806, 81% more likely to have high fire risk imapct compared to buildings with full sprinkler system installed.

Table 5: Stepwise logistic regression model to predict high risk

term	estimate	std.error	statistic	p.value
(Intercept)	-4.931	0.489	-10.083	0.000
Building_Status_CatNormal	1.074	0.323	3.329	0.001
Building_Status_CatUnder Renovation/Construction	0.839	0.335	2.504	0.012

Table 5: Stepwise logistic regression model to predict high risk

term	estimate	std.error	statistic	p.value
Building_Status_CatUndetermined/Not Applicable	0.943	0.338	2.793	0.005
Business_Impact_CatMay Resume	-0.008	0.345	-0.024	0.981
Business_Impact_CatNo Interruption	-0.792	0.338	-2.345	0.019
Business_Impact_CatUndetermined/Not Applicable	-0.371	0.336	-1.104	0.270
Fire_Alarm_System_Operation_CatOperated	0.021	0.098	0.217	0.828
Fire_Alarm_System_Operation_CatUndetermined/Not Applicable	0.274	0.093	2.934	0.003
Smoke_Alarm_at_Fire_Origin_CatNo smoke alarm	0.075	0.089	0.847	0.397
Smoke_Alarm_at_Fire_Origin_CatPresent and operated	0.307	0.082	3.731	0.000
Smoke_Alarm_at_Fire_Origin_CatUndetermined	0.350	0.090	3.907	0.000
Sprinkler_System_Presence_CatNo sprinkler system	0.592	0.082	7.195	0.000
Sprinkler_System_Presence_CatPartial sprinkler system	0.067	0.107	0.623	0.534
Sprinkler_System_Presence_CatUndetermined	0.029	0.101	0.284	0.776
arrival_to_fire_control_time	0.460	0.071	6.495	0.000
alarm_to_arrival_time	0.037	0.018	2.098	0.036
Number_of_responding_apparatus	0.197	0.005	42.154	0.000

$$
\begin{aligned}
\operatorname{logit}(Risktype) = \beta_0 + \beta_1^{\text{BuildingStatus}} + \beta_2^{\text{BusinessImpact}} \\
+\beta_3^{\text{FireAlarmSystemOperation}} + \beta_4^{\text{SmokeAlarmatFireOrigin}} \\
+\beta_5^{\text{SprinklerSystemPresence}} + \beta_6^{\text{Alarmtoarrivaltime}} \\
+\beta_7^{\text{Arrivaltofirecontroltime}} + \beta_8^{\text{Numberofrespondingapparatus}}
\end{aligned}
\tag{1}
$$

Table 6 utilized a linear regression model to establish the connection between the projected financial loss in dollars and the variables acting independently. As the estimated loss of dollar was non normally distributed, we will used logarithm value of estimated dollar loss. The findings indicate that the factors considered in the model are strong predictors of the projected financial loss ($p < 0.05$). The final predictors after stepwise elimination can be seen in Equation 2. The coefficient for building impact states that projected financial loss increased by $\exp(0.424) = 1.528$, 53% for a building that is under construction or renovation than compared to abandon building. Similarly, a building business may resume the estimated loss increased by $\exp(0.806) = 2.23$ times compared to a building with no business operations. The negative coefficients for fire control methods such as a occupancy, self-extinguished, and undetermined indicate that these factors are connected to a lower estimated financial loss. More precisely, a building with occupant put off the fire and self-extinguished system the estimated dollar loss decreased by exp(-1.320) = 0.323, 32% and exp(-0.959)= 0.383, 38% respectively. The positive coefficient for arrival time to fire control indicates that longer response times are linked to higher estimated financial losses. Put simply, for each additional minute it takes for the fire control team to arrive, the predicted financial loss goes up by exp(0.081) = 1.08, 8%.

The positive coefficients for the number of responding apparatuses suggest that these factors are linked to a increase in estimated dollar loss. More precisely, an increase of one unit in the number of responding apparatus is associated with a increase of exp(0.073) = 1.076, 8% estimated dollar loss, while an increase of one unit in the number of responding personnel is associated with an increase of exp(0.014) = 1.014, 1% estimated dollar loss.

Table 6: Stepwise linear regression model to predict estimated dollar loss

term	estimate	std.error	statistic	p.value
(Intercept)	6.059	0.326	18.568	0.000
Building_Status_CatNormal	0.255	0.206	1.239	0.215
Building_Status_CatUnder Renovation/Construction	0.424	0.215	1.978	0.048
Building_Status_CatUndetermined/Not Applicable	0.246	0.213	1.155	0.248
Business_Impact_CatMay Resume	0.815	0.233	3.491	0.000
Business_Impact_CatNo Interruption	-0.345	0.228	-1.512	0.131
Business_Impact_CatUndetermined/Not Applicable	0.090	0.227	0.395	0.693
Fire_Alarm_System_Operation_CatOperated	0.439	0.059	7.385	0.000
Fire_Alarm_System_Operation_CatUndetermined/Not Applicable	0.553	0.057	9.730	0.000
Method_Of_Fire_Control_CatFire department	0.127	0.087	1.455	0.146
Method_Of_Fire_Control_CatOccupant	-1.131	0.091	-12.456	0.000
Method_Of_Fire_Control_CatSelf extinguished	-0.764	0.109	-7.008	0.000
Method_Of_Fire_Control_CatUndetermined	-0.983	0.135	-7.268	0.000
Smoke_Alarm_at_Fire_Origin_CatNo smoke alarm	0.240	0.057	4.230	0.000
Smoke_Alarm_at_Fire_Origin_CatPresent and operated	0.289	0.054	5.381	0.000
Smoke_Alarm_at_Fire_Origin_CatUndetermined	0.373	0.059	6.357	0.000
Sprinkler_System_Presence_CatNo sprinkler system	0.241	0.054	4.447	0.000
Sprinkler_System_Presence_CatPartial sprinkler system	-0.238	0.066	-3.616	0.000

Table 6: Stepwise linear regression model to predict estimated dollar loss

term	estimate	std.error	statistic	p.value
Sprinkler_System_Presence_CatUndetermined	-0.283	0.062	-4.582	0.000
arrival_to_fire_control_time	-0.141	0.026	-5.467	0.000
alarm_to_arrival_time	0.080	0.012	6.907	0.000
Number_of_responding_apparatus	0.071	0.020	3.473	0.001
Number_of_responding_personnel	0.014	0.007	2.172	0.030

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) \\
Estimateddollarloss = \beta_0 + \beta_1 \times \text{BuildingStatus} + \beta_2 \times \text{BusinessImpact} & \\
+\beta_3 \times \text{FireAlarmSystemOperation} + \beta_4 \times \text{MethodofFireControl} & \\
+\beta_5 \times \text{SmokeAlarmatFireOrigin} + \beta_6 \times \text{SprinklerSystemPresence} & \\
+\beta_7 \times \text{Alarmtoarrivaltime} + \beta_8 \times \text{Arrivaltofirecontroltime} & \\
+\beta_9 \times \text{Numberofrespondingapparatus} + \beta_{10} \times \text{Numberofrespondingpersonnel} & \\
\beta_0 &\sim \text{Normal}(0, 38.91) \\
\beta_1 &\sim \text{Normal}(0, 25.59) \\
\beta_2 &\sim \text{Normal}(0, 27.85) \\
\beta_3 &\sim \text{Normal}(0, 7.14) \\
\beta_4 &\sim \text{Normal}(0, 16.18) \\
\beta_5 &\sim \text{Normal}(0, 7.02) \\
\beta_6 &\sim \text{Normal}(0, 7.85) \\
\beta_7 &\sim \text{Normal}(0, 3.09) \\
\beta_8 &\sim \text{Normal}(0, 1.43) \\
\beta_9 &\sim \text{Normal}(0, 2.50) \\
\beta_{10} &\sim \text{Normal}(0, 0.83) \\
\sigma &\sim \text{Exponential}(1.913)
\end{aligned}
\tag{2}
$$

The linear regression model in Table 7 was utilized to assess how the number of casualties relates to the independent variables. The findings indicate that the factors considered in the model are a strong predictor of the number of casualties ($p < 0.05$). The final predictors after step wise elimination can be seen in Equation 3. The casualty rate is 0.06 higher in a under construction or renovated building compared to abandon building. In the same vein, the casualty rate is 0.071 greater for a building where occupants themselves controlled the fire. Fewer casualties occur in areas where the smoke system is used and operated with a coefficient of 0.002. The positive coefficients for undetermined fire control method, undetermined smoke alarm at fire origin, and undetermined sprinkler system presence suggest that these factors are linked to an increased number of casualties. The negative coefficient for arrival time to fire control suggests that a decrease in casualties is linked to longer fire control times. Additionally, the data show that for each additional responding apparatus, there is a corresponding increase of 0.045 in casualties, while each additional responding personnel is associated with a decrease of 0.012.

Table 7: Stepwise linear regression model to predict number of casualties

term	estimate	std.error	statistic	p.value
(Intercept)	-0.320	0.074	-4.330	0.000
Building_Status_CatNormal	0.125	0.047	2.652	0.008
Building_Status_CatUnder Renovation/Construction	0.058	0.049	1.179	0.239
Building_Status_CatUndetermined/Not Applicable	0.087	0.049	1.770	0.077
Business_Impact_CatMay Resume	0.122	0.054	2.268	0.023
Business_Impact_CatNo Interruption	0.143	0.052	2.730	0.006
Business_Impact_CatUndetermined/Not Applicable	0.167	0.052	3.215	0.001
Fire_Alarm_System_Operation_CatOperated	0.040	0.014	2.943	0.003
Fire_Alarm_System_Operation_CatUndetermined/Not Applicable	-0.011	0.013	-0.833	0.405
Method_Of_Fire_Control_CatFire department	0.035	0.020	1.730	0.084
Method_Of_Fire_Control_CatOccupant	0.078	0.021	3.744	0.000
Method_Of_Fire_Control_CatSelf extinguished	0.046	0.025	1.845	0.065
Method_Of_Fire_Control_CatUndetermined	0.074	0.031	2.396	0.017
Smoke_Alarm_at_Fire_Origin_CatNo smoke alarm	-0.004	0.013	-0.320	0.749
Smoke_Alarm_at_Fire_Origin_CatPresent and operated	0.004	0.012	0.334	0.739
Smoke_Alarm_at_Fire_Origin_CatUndetermined	0.024	0.013	1.782	0.075
Sprinkler_System_Presence_CatNo sprinkler system	0.068	0.012	5.502	0.000
Sprinkler_System_Presence_CatPartial sprinkler system	0.029	0.015	1.916	0.055
Sprinkler_System_Presence_CatUndetermined	0.025	0.014	1.786	0.074
arrival_to_fire_control_time	-0.035	0.006	-5.949	0.000
Number_of_responding_apparatus	0.044	0.005	9.376	0.000
Number_of_responding_personnel	-0.011	0.002	-7.479	0.000

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) \\
Numberofcasualties = \beta_0 + \beta_1 \times \text{BuildingStatus} + \beta_2 \times \text{BusinessImpact} & \\
+\beta_3 \times \text{FireAlarmSystemOperation} + \beta_4 \times \text{MethodofFireControl} & \\
+\beta_5 \times \text{SmokeAlarmatFireOrigin} + \beta_6 \times \text{SprinklerSystemPresence} & \\
+\beta_7 \times \text{Arrivaltofirecontroltime} + \beta_8 \times \text{Numberofrespondingapparatus} & \\
+\beta_9 \times \text{Numberofrespondingpersonnel} & \\
\beta_0 &\sim \text{Normal}(0, 8.69) \\
\beta_1 &\sim \text{Normal}(0, 5.71) \\
\beta_2 &\sim \text{Normal}(0, 6.31) \\
\beta_3 &\sim \text{Normal}(0, 1.55) \\
\beta_4 &\sim \text{Normal}(0, 2.98) \\
\beta_5 &\sim \text{Normal}(0, 1.55) \\
\beta_6 &\sim \text{Normal}(0, 1.78) \\
\beta_7 &\sim \text{Normal}(0, 0.71) \\
\beta_8 &\sim \text{Normal}(0, 0.60) \\
\beta_9 &\sim \text{Normal}(0, 0.24) \\
\sigma &\sim \text{Exponential}(0.430)
\end{aligned} \tag{3}
$$

4 Discussion:

This paper evaluates the data on Fire Incidents Data for the City of Toronto to find the relationship between risk from fire incidents and various independent variables. The research find out that the risk of causalities and monetary loss is associated with the variables like building status, business impact, smoke alarm origin, method of fire control etc. The research utilized regression model like logistic regression and linear regression to predict the expected financial loss and human casualties associated with fire incidents. Moreover, our study utilized chi-square tests, ANOVA tests, and step wise elimination techniques to find the potential predictor for high risk. The results of this investigation have substantial ramifications for the protection of the public and the application of prevention strategies, and they significantly add to the body of knowledge already available on fire events.

4.1 Summary of Key Findings

Furthermore, our study has demonstrated that the period of interval between the sounding of the alarm and the timing of arrival for fire officials as a major predictor of fire events, highlighting the crucial significance of an early reaction. The same has been find the time of fire fighter arrival and fire control time.The outcomes of our linear regression study show that these factors have a major impact on both the anticipated monetary damages and the overall number of casualties. In a similar manner a study by Challands (2010) also confirmed that with delay in fire fighter response time at the incident result in greater monetary loss. The report also found that with increase in number of fire fighter personnel at the fire incident location, the number of causality decreases. A study by Shoffner (2023) done in United states and Canada found the same results that with increase in fire fighter personnel the human causalities and property loss in dollars decreased significantly.

4.2 Insights into Fire Incident Risk Factors

The one of the significant factor is to reduce the response time to mitigate the impact of fire on human causalities and property loss. It is essential to make sure that the fire department has all necessary instruments and man power to respond quickly to these situations. The use of real time data analytics and predictive modelling can further help fire departments to respond timely to these situations. Apart from this the fire department should prioritize the necessary community education, fire drills and dispense resources to offices, school's and residential buildings.

4.3 Impact of Building Design and Safety Measures

The design of buildings and safety measures implemented by building management have a crucial impact on fire incidents. It is important for building designers and engineers to prioritize safety measures during the design and construction phases. Several measurable steps can be taken to prevent such incidents in the future, including the use of fire-resistant materials in construction and the proper installation of sprinkler systems and smoke alarms. Beyond installing necessary alarms, it is also important to regularly test these alarms to ensure they function properly. Hence, routine inspections, testing, and maintenance of fire safety systems, along with the use of fire-resistant materials and strategic smoke alarm placement, can significantly enhance building safety.

4.4 Limitations and Future Research

The present study is subject to certain limitations, as it did not take into account all relevant factors for the analysis of the risk impact of fire incidents. This study lacks key variables including building type and composition, accessibility to fire extinguishers, and the level of training of the fire fighters involved in reporting on fire incidents, thus hampering a comprehensive analysis. Future research endeavors may involve the analysis of extensive dataset and the inclusion of a broader range of geographical areas to yield more generalized results. It is important to thoroughly assess the effectiveness of various fire protection measures, such as fire evacuation education, construction regulations and norms, and inspections necessary fire safety using fire drills. In addition to evaluating the possible advantages of technical advancements in boosting the quickness and effectiveness of fire identification and response systems, future research may look at the impact of rising temperatures on the frequency and intensity of wildfires.

4.5 Conclusion

In conclusion, this study examined Toronto fire incident data and projected the anticipated financial loss and number of casualties using a number of independent variables. The findings indicate that a number of variables are significant in forecasting fire events, including building conditions, the influence on nearby companies, the efficacy of fire alarm systems, firefighting techniques, and the installation of smoke detectors at the scene of the disaster. The research project stressed the significance of prompt reaction times and demonstrated how the anticipated loss of revenue and fatalities were impacted by the amount of duration it took to get at the scene and contain the fire, as well as by the time it took to arrive after the alarm was raised. Comprehending these variables is essential for firefighters in order to effectively allocate personnel and formulate strategies for both prevention and response to fire incidents.

References:

Auguie, Baptiste. 2017. *gridExtra: Miscellaneous Functions for "Grid" Graphics*.

Challands, Neil. 2010. "The Relationships Between Fire Service Response Time and Fire Outcomes." *Fire Technology* 46 (3): 665–76.

Gelfand, Sharla. 2020. "Opendatatoronto: Access the City of Toronto Open Data Portal." *R Package Version 0.1* 4.

Pedersen, Thomas Lin. 2024. *Patchwork: The Composer of Plots*. https://patchwork.data-imaginist.com.

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

———. 2024. *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://github.com/apache/arrow/.

Robinson, David, Alex Hayes, and Simon Couch. 2023. *Broom: Convert Statistical Objects into Tidy Tibbles*. https://broom.tidymodels.org/.

Shoffner, Jeffrey Daniel. 2023. "Firefighter Staffing Model Implications on Fire Casualties and Fire Loss: Life Safety and Socio-Economic Impacts of the Fire Service."

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, Jim Hester, and Jennifer Bryan. 2024. *Readr: Read Rectangular Text Data*. https://readr.tidyverse.org.

William Revelle. 2024. *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois: Northwestern University. https://CRAN.R-project.org/package=psych.

Xie, Yihui. 2017. *Dynamic Documents with r and Knitr*. Chapman; Hall/CRC.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. http://haozhu233.github.io/kableExtra/.