DS8007 Project Documentation:

Fire Incidents in Toronto 2011-2018

Done By:

Faiza Noor

1 Problem Definition

Emergency fire services are often dispatched with emergency 911 calls. General fire services are also utilized for fire safety inspections, and fire hazard maintenance, removal and prevention. Fire related emergencies make up a large amount of emergency 911 calls. For example, in 2021, fire services responded to 133,533 emergencies in Toronto, one of Canada's largest cities. Although fire services are a necessary service to always have on high alert, high activity of these services can be a cause for concern. Exploring visualizations of the data on Toronto fire incidents can help us better gather multi-factor determinants of common fire issues, correlations between causes of fire and other factors, and impacts such as the kind of loss, damage and costs that arise from them. The insights gathered from these data points can reinforce existing strategies and help come up with more effective ones for fire services and residents to reduce the risk of fire related emergencies.

2 Research Questions

- What are the areas of origin where fires most occur?
- What are the top determinants (probable cause) of residential fire in Toronto over the 9 years?
- Which kinds of fires are most costly? What leads to more dollar loss?
- When do casualties most occur?
- Correlations between casualties and other factors

3 Publicly Available Dataset

This "Fire Incidents" dataset is from Kaggle. It's a dataset containing data points describing fire incidents in the City of Toronto over a 9 year period from 2011-2018. The data consists of 27 columns, and over 11,000 rows.

Kaggle Source URL: https://www.kaggle.com/datasets/reihanenamdari/fire-incidents/code

The datapoints within this dataset range in various types of data including mostly categorical data, numeric discrete, and timestamp data.

The columns and their data types is stated below:

Area_of_Origin - Categorical
Business_Impact - Categorical
Civilian_Casualties - Numeric (Continuous)
Count_of_Persons_Rescued - Numeric (Continuous)
Estimated_Dollar_Loss - - Numeric (Continuous)
Estimated_Number_Of_Persons_Displaced - Numeric (Continuous)
Ext_agent_app_or_defer_time - Numeric (Timestamp)
Extent_Of_Fire - Categorical
Fire_Alarm_System_Impact_on_Evacuation - Categorical
Fire_Alarm_System_Operation - Categorical
Fire_Alarm_System_Presence - Categorical
Fire_Under_Control_Time - Numeric (Timestamp)
Ignition_Source - Categorical
Incident_Station_Area - Numeric (Discrete)
Incident_Ward - Numeric (Discrete)
Last_TFS_Unit_Clear_Time - Numeric (Timestamp)
Latitude - Numeric (Location Coordinates)
Longitude - Numeric (Location Coordinate)

Material_First_Ignited - Categorical Method_Of_Fire_Control - Categorical Possible_Cause - Categorical Property_Use - Categorical Smoke_Alarm_at_Fire_Origin_Alarm_Failure - Categorical Smoke_Alarm_at_Fire_Origin_Alarm_Type - Categorical Status_of_Fire_On_Arrival - Categorical - Categorical TFS_Alarm_Time - Numeric (Timestamp) TFS_Arrival_Time - Numeric (Timestamp)

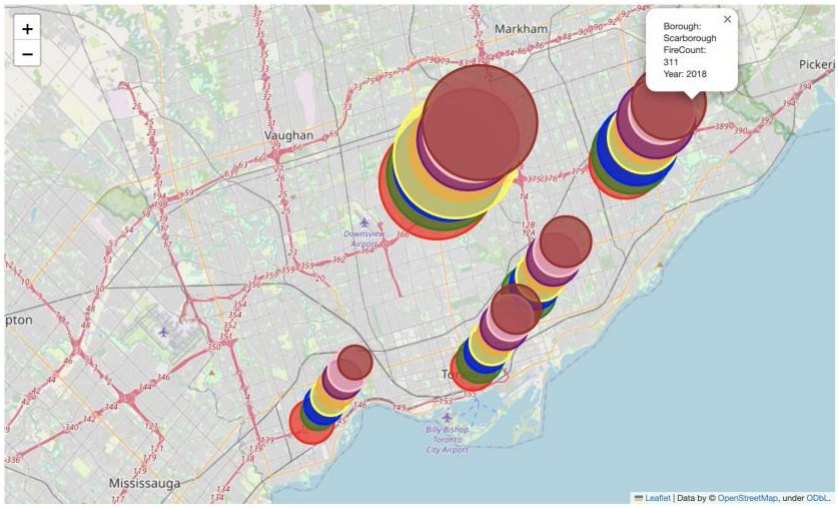
4 Data Pre-processing

The data pre-processing involved the following steps:

Pre-processing Steps	**Example**
Replace nan values	df["Incident_Ward"].fillna(-1, inplace=True)
Convert dataframe data types to correct type	df = df.astype({"Area_of_Origin":"string”})
Refine categorical variables	df["Area_of_Origin"].nunique()
Feature Engineering	df['Year'] = df['TFS_Alarm_Time'].dt.year

5 EDA and Results

Image 1:



Number of fire incidents per Toronto’s 6 Boroughs:

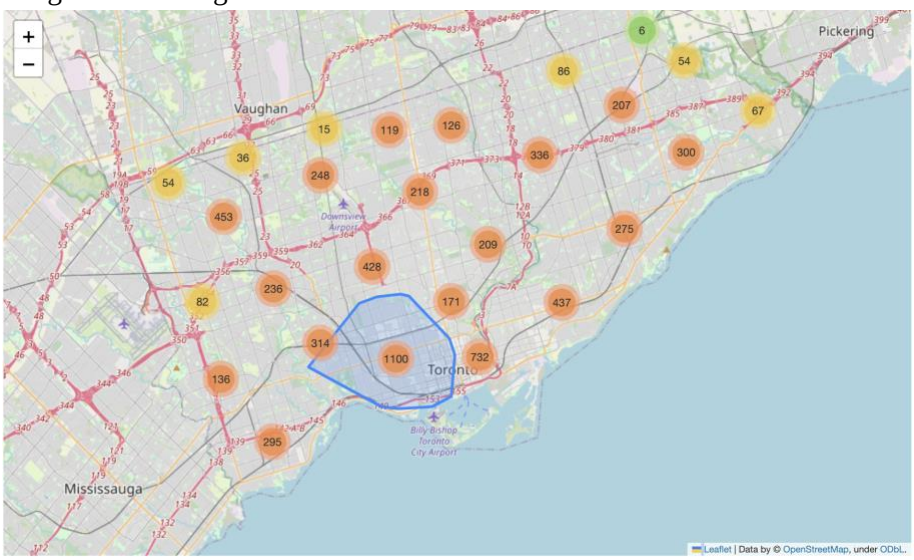
1. Scarborough
5. The Junction (combined with North York in the visualization)
2. North York
3. East York
4. Etobicoke
6. Old Toronto

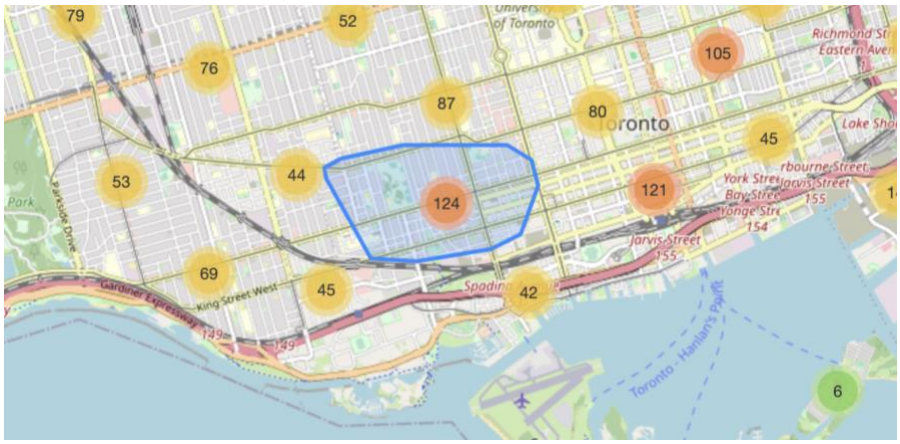
Above is a bubble map showing bubbles overlaying each borough of Toronto (with data on North York combined with data on The Junction). Each bubble represents the total number of fire incidents for each year from 2011-2018, in that order, the sizes of the bubbles represent the total counts of incidents. We

can see that there's been a slight decrease in incidents over the last 2 years (2017/2018) but all counts for each borough are consistently similar each year, give or take 50-100 incident increase or decrease.

I chose this bubble map to best illustrate counts per region to give the viewer an intuitive understanding of frequency of fire incidents across geographical areas.

Image 2a and Image 2b:





The above two images are other bubble maps indicating the total number of fire calls over smaller regions across Toronto. In Image 2a, we can see the lower part of downtown Toronto received the most legitimate fire calls, at 1100 calls. If we zoom in the 1100, in image 2b, we see counts over smaller regions of that area. As highlighted in the orange circle, the neighborhoods around Trinity Bellwood's park received the most calls for legitimate fires for that zoomed in downtown area.

I chose this bubble map to best illustrate counts per region to give the viewer an intuitive understanding of frequency of fire incidents across geographical areas.

Image 3:

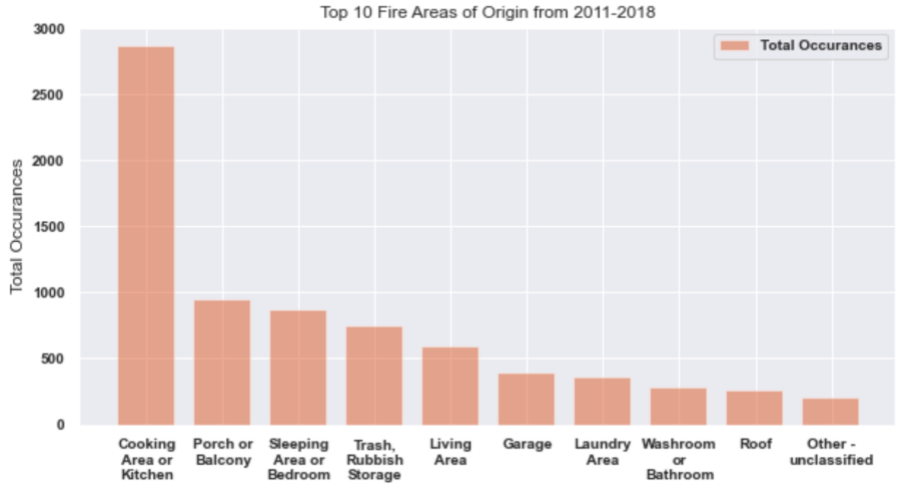
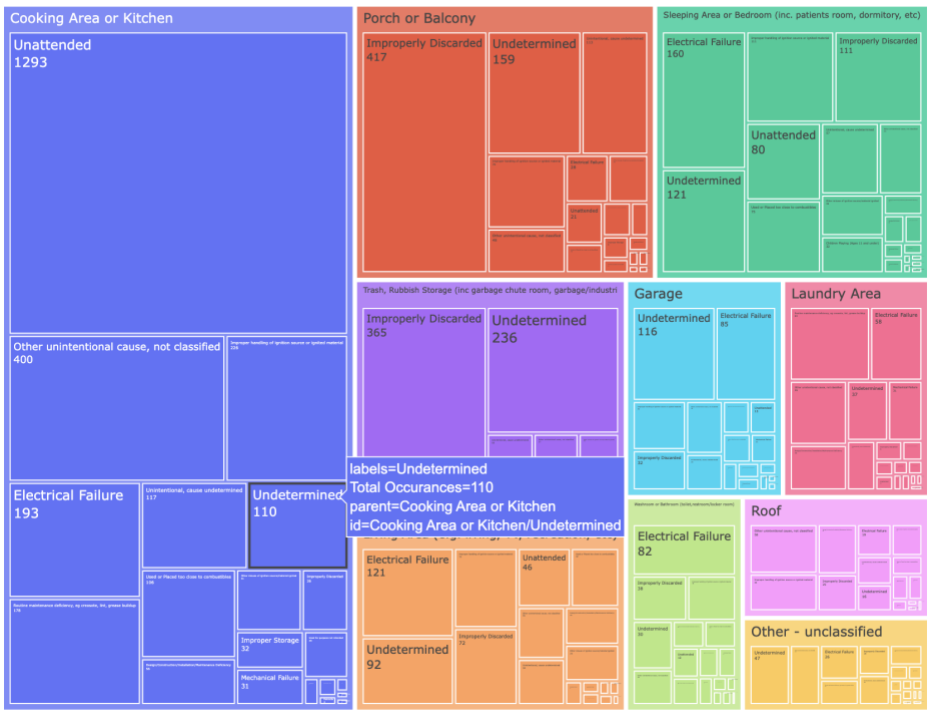


Image 3 is a bar chart of the top 10 fire incident "Areas of Origin". Areas of origin defines where the fire originally started for each incident. In this bar chart we are looking at frequency of how many times the fire is occurring in different areas of origin. We see that most fires originated in the cooking area or kitchen.

A bar chart was most appropriate to show counts of a categorical feature. It allows us to clearly compare the total amount of fire incidents when the fire happened in the different areas of origin above.

Image 4:



This tree map contains parent node: "Fire Areas of Origin" and child nodes "Probable Cause". It allows us to see the probable causes from each origin type. Here we can compare the common fire origins and common causes for those incidents. The cooking area, being the most common area where fires first start, we can see that the common probable cause for cooking area fires, is leaving items unattended. We can also see the leading cause for porch or balcony fires is improperly discarded items. Its important

to highlight that for many incidents, the cause is undetermined, which can pose the question: how do we better identify probable causes of fire incidents?

A tree map was a good way to visualize this type of data, as its to shows the hierarchal structure between the counts of each areas of origin and the counts of each probable cause types for each areas of origin. We can see the top areas of origin and top probable causes within those areas to better narrow down common determinants of fire incidents for different types of fires (i.e. Kitchen Fire, Bedroom Fire, etc.)

Image 5:

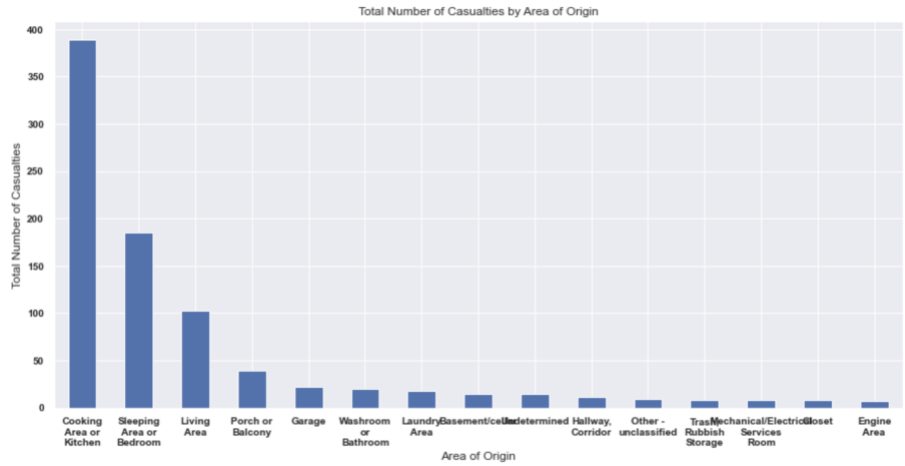
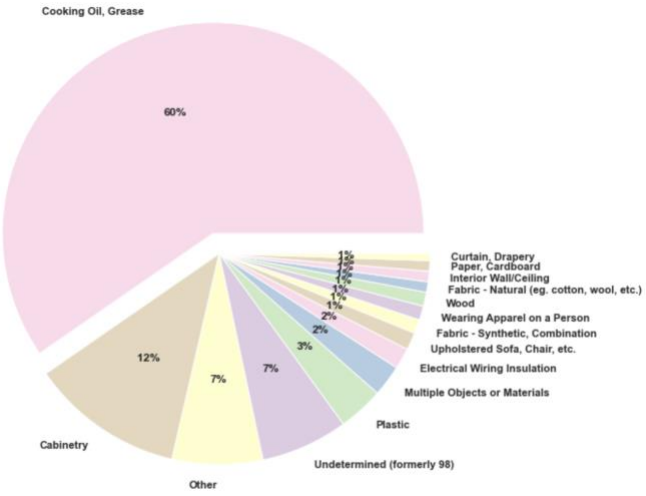


Image 5 is a bar graph of the counts of how many casualties occurred for each type of fire (areas of origin). Most casualties happened when the fire originates in the cooking area.

A bar chart was most appropriate to show counts of a categorical feature. It allows us to clearly compare the total amount of casualties for when the fire happens in different areas.

Image 6:



The pie chart of materials first ignited for fires that happen in the kitchen tells us that cooking oil is the most common material first ignited in kitchen fires as it makes up 60% of the incidents that happen in the kitchen.

A pie chart was a good visualization to show the percentage of different items first ignited when fires happen in the kitchen. It allows us to evaluate what's more common in terms of materials that are first ignited in the kitchen and how frequent that is compared to other kitchen materials.

Image 7:

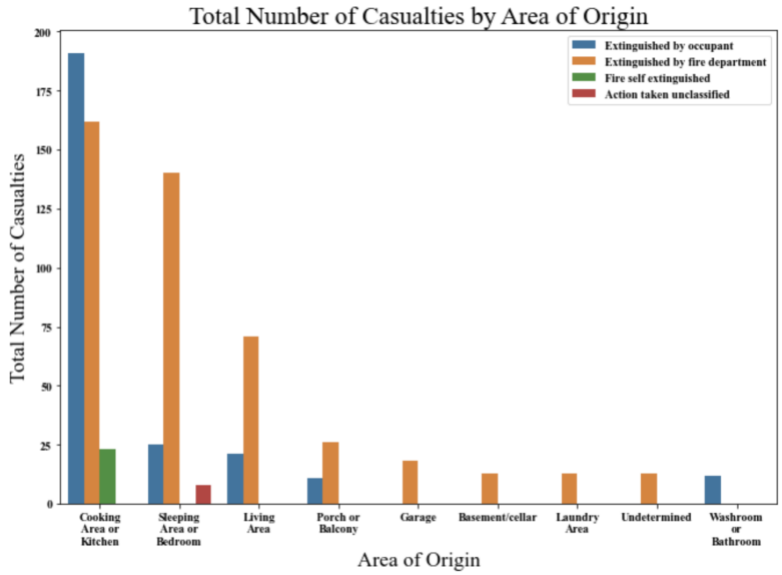


Image 7 is a bar graph that shows the total number of casualties by fire areas of origin and breaks it down further to show how the fire was extinguished for each area of origin. There are more casualties when the fire is attempted to be extinguished by the occupant for when fires occur in cooking areas.

A bar chart was most appropriate to show counts of a categorical feature. It allows us to clearly compare the total amount of casualties for when the fire happened in the different areas of origin above and allow another understanding of how that fire was extinguished for different kinds of fire.

Image 8:

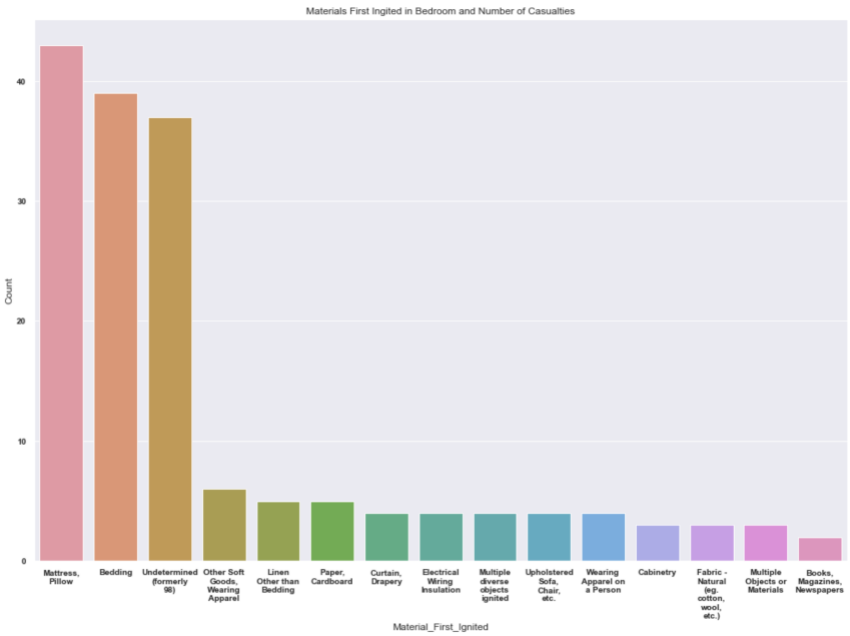


Image 8 is a count bar chart. It tells us that "Mattress/Pillow" is the most common material first ignited specifically in bedroom fires, which causes more casualties than other items in the bedroom.

A bar chart was most appropriate to show counts of a categorical feature. It allows us to clearly compare the frequency of items first ignited in bedroom fires.

Image 9:

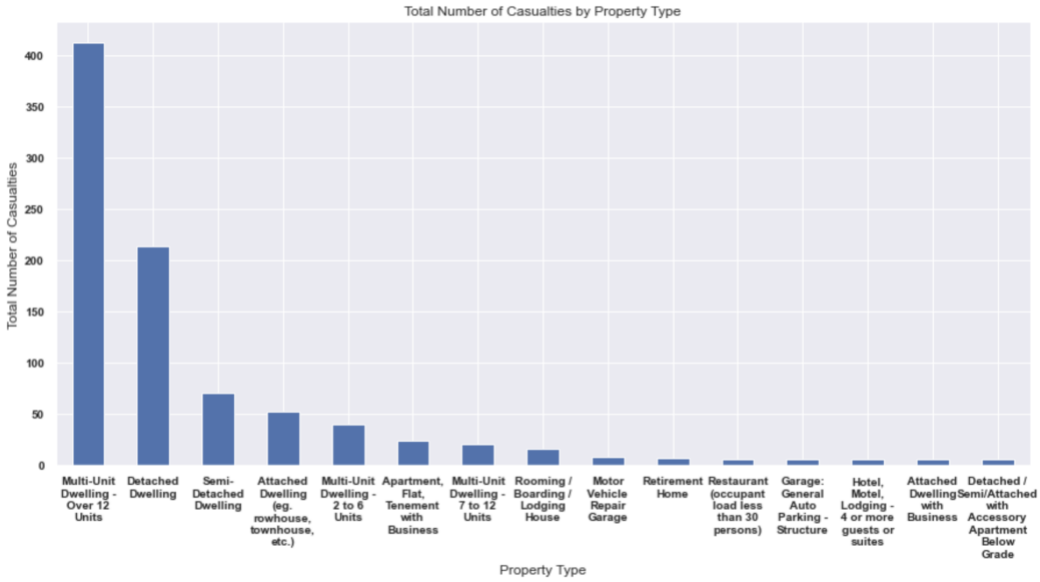


Image 9 shows the total number of casualties by property type. Its reporting that multi-unit dwellings seem to be the most dangerous property type over others.

A bar chart was most appropriate to show counts of a categorical feature. It allows us to clearly compare the total amount of casualties for when the fire happens in different property types.

Image 10:

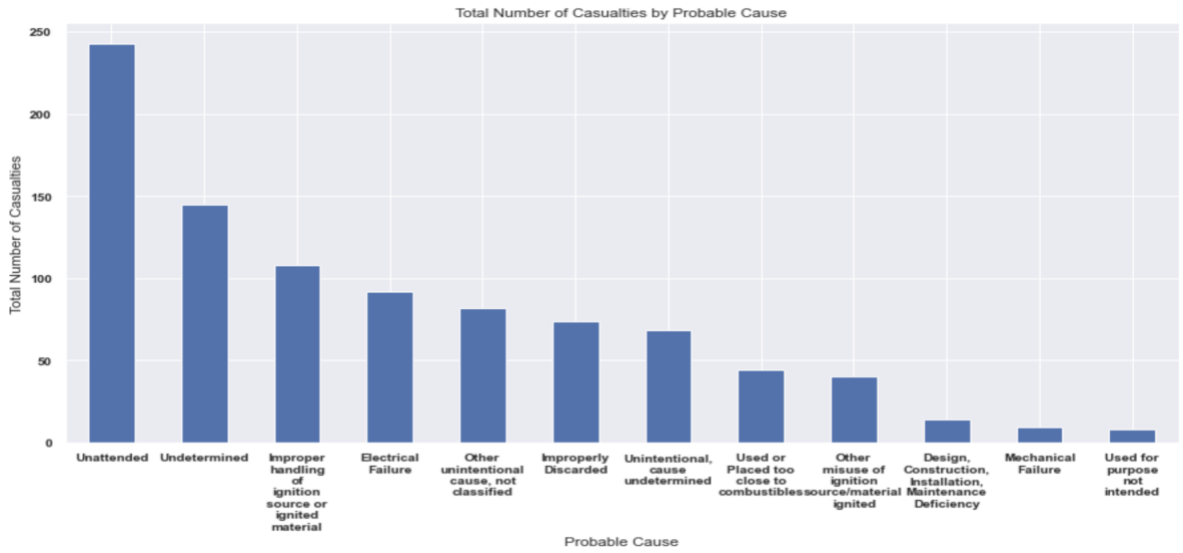


Image 10 illustrates the total number of casualties by probable cause. Here we can see the most common causes in fire casualties is inattentiveness and improper handling of ignition sources.

A bar chart was most appropriate to show counts of a categorical feature. It allows us to clearly compare the total amount of casualties for when the fire happens for different property types.

Image 11:

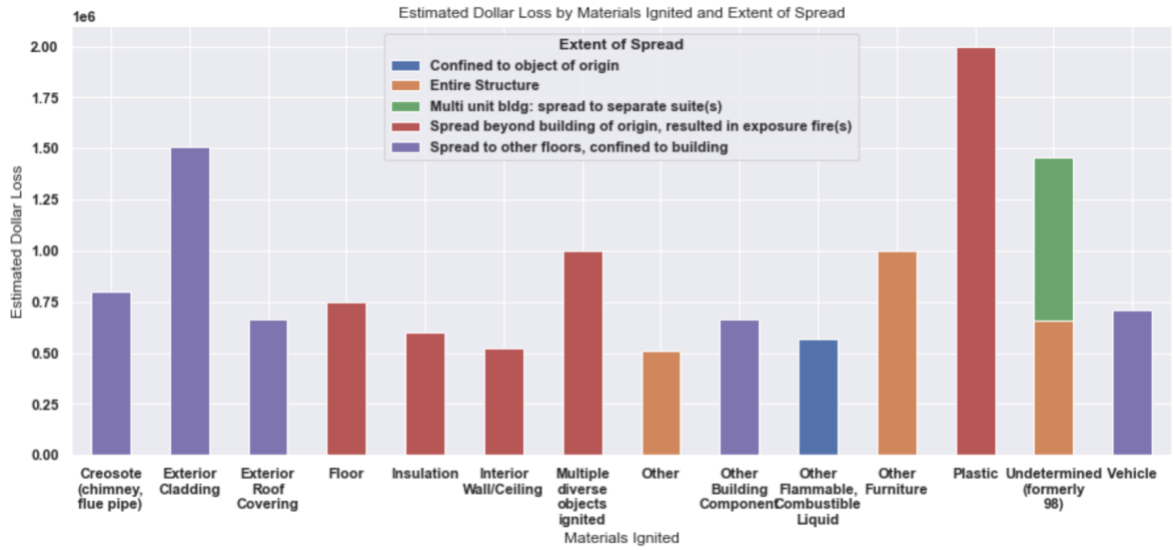


Image 11 shows the highest estimated dollar loss by materials ignited. Its further broken down to show the different kinds of spread occurred for each material type and how much loss occurred for it. The spread extent falls into the 5 categories seen in the legend. Out of the top 15 most costly materials ignited, plastic seems to both spread the most and is most costly.

A stacked bar chart was most appropriate to visualize total dollar loss for all materials first ignited. Here the stacks allow us to see ,for each material, how much loss in dollar value occurred for each kind of fire spread extent.

Image 12:

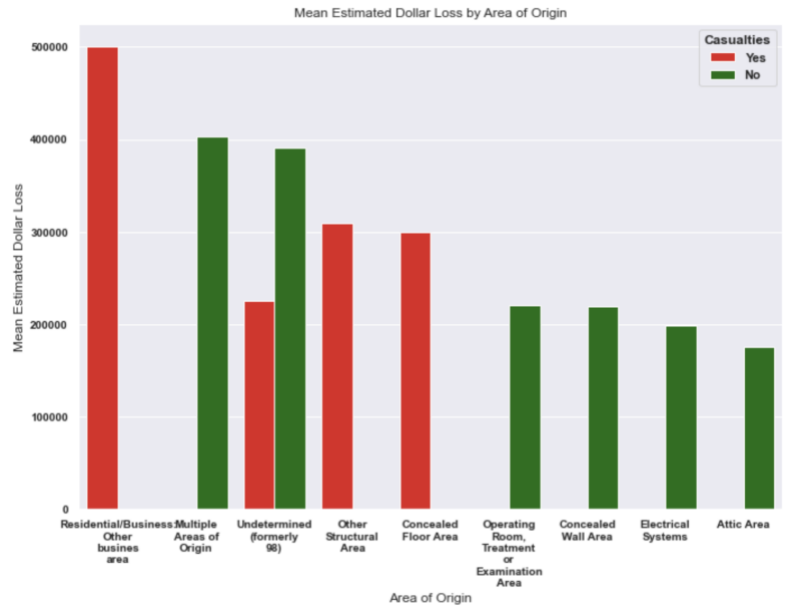


Image 12 shows the mean estimated dollar loss vs fire area of origin and breaks down the mean loss for when there were casualties and when there were none. We can say that fire in residential and business areas have seen more casualties and highest dollar loss.

A bar chart was most appropriate to show counts of a categorical feature. It allows us to clearly compare each kind of fire, what the total cost attributed to each kind of fore and furthermore broken down costs for when there were casualties and no casualties.

Image 13:

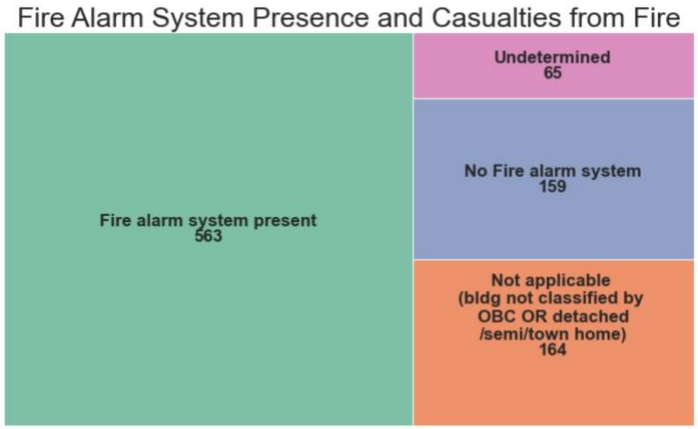


Image 13 is a tree map that illustrates the total counts of casualties and whether those incidents had fire alarm systems present. Surprisingly, there were more casualties when fire alarm systems were present. There isn't an indication to show otherwise and therefore we cannot conclude that the presence of having no fire alarm systems can directly impact on whether the fire incident is lethal or not. An interesting area to further explore.

A tree map was used to illustrate the total counts of casualties per alarm system type. It allows us to compare frequency of casualties with each type of alarm system type in a hierarchical way.

Image 14:

Image 14 gives a breakdown on the frequency of different kinds of fire spread and here I am only looking at incidents where the fire spread beyond the material/room of original. The child nodes in this tree map give shows us the counts of incidents by the material of origin for each spread type. This can

help us understand which common materials cause more spread and to which extent. When the spread is to the entire room, the common materials ignited for this type of spread is cooking oil, mattress and pillow. But for the type of spread that goes beyond the room of origin to the rest of the floor, the most materials first ignited is interior wall/ceiling.

A tree map here best illustrates the hierarchical information of types of fire spread extent and the materials first ignited corresponding to each fire type.

6 GitHub Repository

The GitHub repository for the python code for all the images in this document is the following:
https://github.com/tmudatavizstudent/DS8007-Final_Project/invitations
https://github.com/tmudatavizstudent/DS8007-Final_Project/invitations
https://github.com/tmudatavizstudent/DS8007-Final_Project