**Analysis of 911 Call Reason in Toronto 2017**

**Kyle Lee**

**May 2019**

**Introduction**

Under provincial jurisdictions, the operation of the 911 services is practiced by the hands of local police, paramedics or firefighting authorities. The internationally accepted standard for answering 911 calls suggests 90% of all calls should be picked up within 10 seconds, or 95% calls should be answered within 20 seconds even during the busiest time of day. There have been many attempts to solve a common, yet the most important, problem of the 911 service in Toronto. When there is a large volume of calls are coming at once, there is a significant 911 wait times at the communication centre (maximum wait time up to 5 minutes), leading many delays in police, paramedic, and firefighter dispatch. It has been reported that nearly 1 in 5 calls to 911 in York Region so far in 2017 were unintentional.

This project will analyze information based on what is sent to the Ontario Fire Marshal relating to all incidents to which Toronto Fire responds. The type and frequency of all incidents will be categorized by the regions in and around the city of Toronto. The conclusion may help determining the number of hiring more call operators or providing ideas which regions the communication centre or the 911 control head quarter should put more focus in order to reduce the wait time on 911 calls.

**Data Description**

I retrieved all data set from Open Data Catalogue from the City of Toronto. The two data set were described as followed:

1. Fire Station Locations - City of Toronto fire station location (XY coordinates)

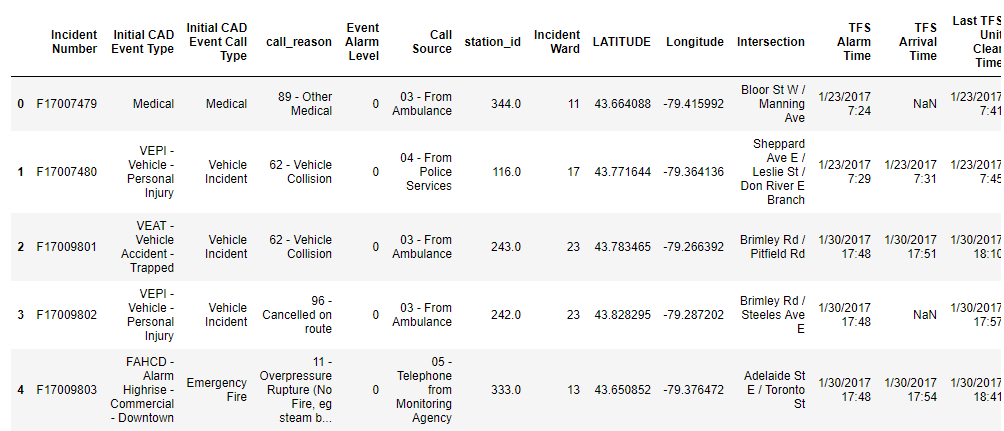
2. 2011-2017 Basic Incidents Details - All 911 call response from 2011 to 2017

3. Used Geopy to get coordinate of Toronto, ON

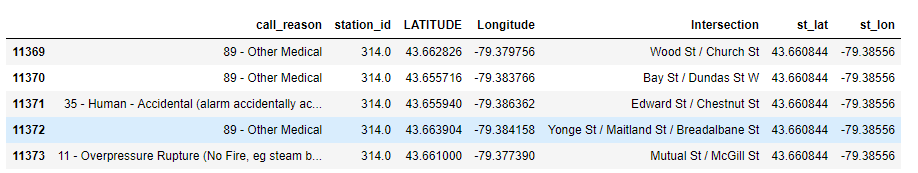
**Methodology**

For my project data base, I used GitHub repository for this project. First, I performed data cleaning of incident detail data frame by dropping all null values in some key features in the data frame: 1) call reason, 2) latitude and 3) longitude. More cleaning was done by renaming the columns of my interest.

The two data frames were then merged to see a list of top 5 stations responded the 911 calls in 2017.

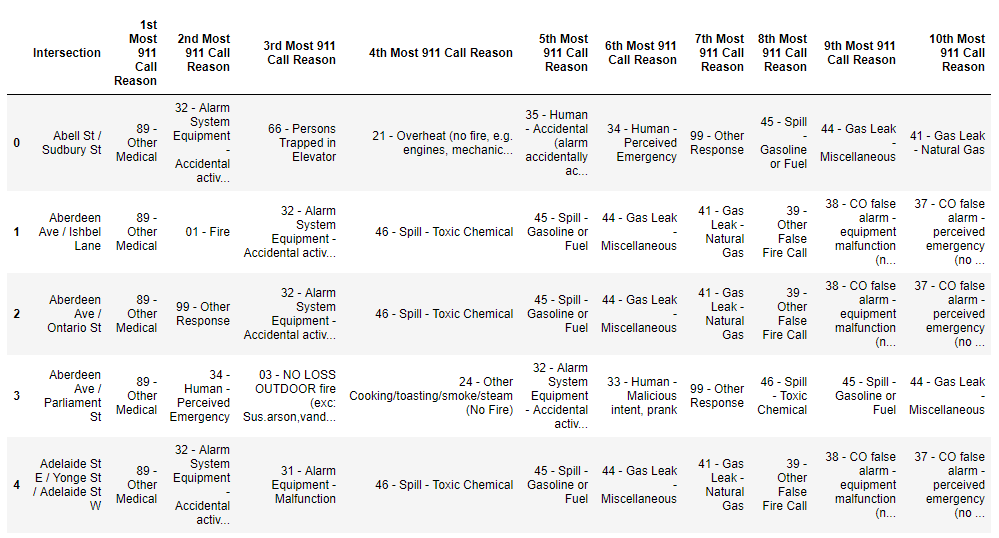


Matplotlib and seaborn were used to create histogram and heat map from the data frame for further EDA. The data frame was then filtered so that only It contains the columns of interests as below:



From this data frame, the top 5 call reasons were measured. Also, the frequencies of each call reason types were grouped by each intersection that the fire station covered.

A new data frame was created to see the top 10 most 911 call reason that were reported on the intersection:

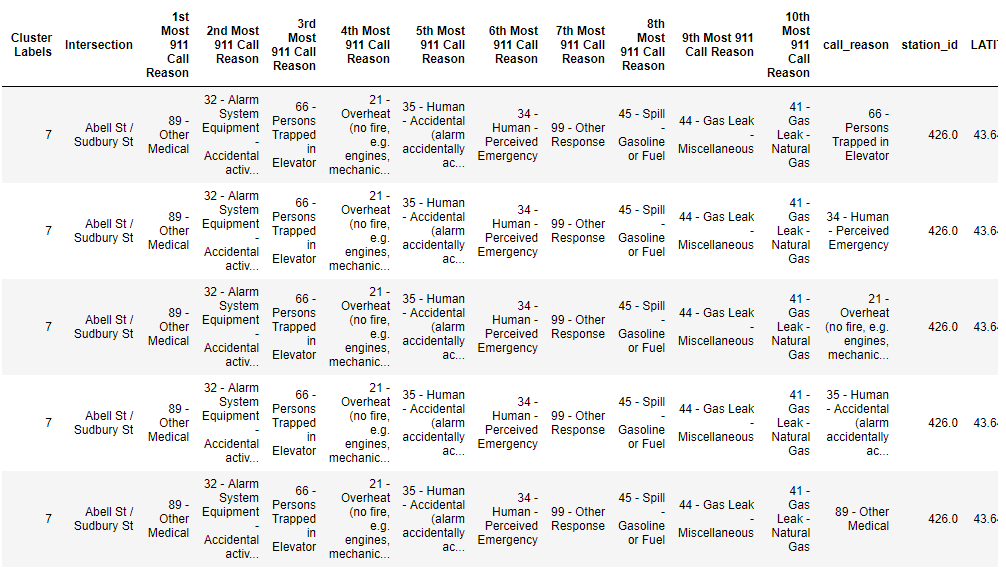


A K cluster is determined as 10 (due to the 10 different call reason types) and k-means clustering was performed to generate a cluster label for each row. The new data frame was then merged with the main data frame so that each intersection with coordinate has top 10 call reasons.

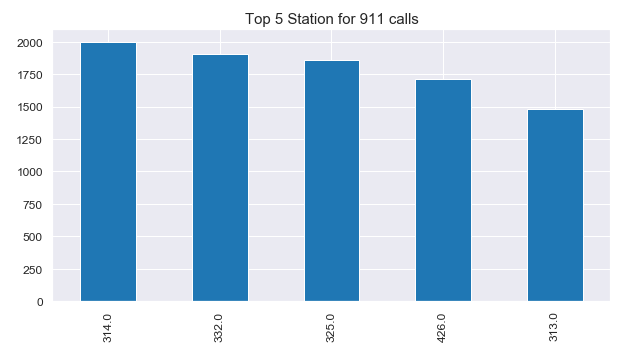
Folium library was used to visualize geographic details of the city of Toronto and its intersection was marked with popup label stating the name of the intersection, cluster label, and call reason.

**Results**

The main **master data frame**.



Top 5 Station ID



Top 5 Intersections Where 911 Calls Reported

A screenshot of a cell phone

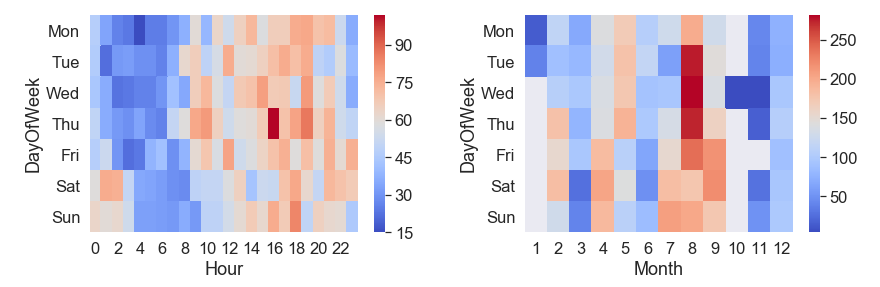
Description automatically generated

Top 5 Call Reasons

A screenshot of a social media post

Description automatically generated

911 Call Frequency on Weekdays by Hour and Month

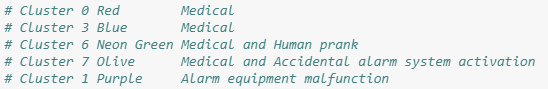


Approximately, 90 calls in total was made on Thursdays around 16:00. Significantly a larger volume of the 911 calls were reported during the summer season from Tuesday to Thursday.

Toronto Map with 911 Responses by Call Reasons

A close up of a map

Description automatically generated

Legends:

**Discussion**

The city of Toronto is considered a capital of Ontario state. It is a big city with a high population density in a narrow area. As there is such a complexity, very different approaches can be tried in clustering studies. The major 911 call reasons can be divided into 3 groups: 1) medical, 2) accidental or malfunction of the alarm system, and 3) Human prank or accidental press.

Based on the cluster map, the East Toronto area, especially along the Sherbourne St., the major 911 calls were reported. It would be more detailed analysis if we can imbed this marker on choropleth map based on the post code so that which area corresponding each fire station should focus more on 911 quality calls over the call volume.

The data set was also biased due to large amount of data. Therefore, I had to deleted some data in order to import and upload on GitHub repository. During the deletion, certain features of call types could have been deleted in a larger portion than other features.

**Conclusion**

The main problem of 911 call response in Toronto was to accurately and time-effectively manage in-coming 911 calls around the downtown Toronto area so that appropriate authorities are able to dispatch their agents to solve the problem. The major 911 call reasons can be divided into 3 groups: 1) medical, 2) accidental or malfunction of the alarm system, and 3) Human prank or accidental press. These calls were more likely reported in East Toronto area. The analysis would have been better if more optimization of Kmeans is performed based on the postcode for better clustering of the 911 call reason.