DataScientest

London Fire Brigade

Analyzing Response and Mobilization Times: A Comprehensive Study

Jack Abajian, Ravindra Garde, Mergim Makolli

22.05.2023

1. Table of Contents

2. Introduction

Fire departments worldwide rely heavily on their response time to save lives and rescue individuals. Therefore, reducing response time is a crucial factor, and fire departments strive to examine potential influences on it. Although fire departments gather data on their operations, only a few uses data science to create a data-driven decision-making strategy.

To address this issue, this project will introduce a data science framework that predicts turnout time using London Fire Brigade (LFB) data, highlighting its effectiveness. The LFB is the fire and rescue service for the Greater London area. The LFB is run by the London Fire Commissioner (LFC) who is the fire and rescue authority for London. It is the busiest fire and rescue service in the UK and one of the largest in the world, with complex emergency responses that require specialized skills. However, unexpected disruptions such as absences due to sickness can make it challenging for teams to respond effectively to emergencies without substitutes for missing members, as not all firefighters possess the same skills.

Consequently, the LFB has to allocate a significant number of firefighters throughout London daily to cover for such absences, which is resource-intensive and inconvenient for the firefighters. To avoid potential problems and be better prepared, it is crucial for the LFB to anticipate and predict potential staff shortages. Therefore, the objective of this project is to analyze and estimate response and mobilization times for the London Fire Brigade using a data-driven approach to improve emergency response times. By creating a data-driven decision-making strategy, the LFB can better predict and prepare for potential staff shortages, ultimately improving their ability to save lives and rescue individuals.

The LFB releases various information showcasing how decisions were made, how public funds were utilized, and how resources were employed, in accordance with the Freedom of Information Act of 2000 (UK-Government, 2000). The project will use two of the publicly available datasets: the first dataset contains details of every incident handled since January 2009, providing information on the date and location of the incident as well as the type of incident handled. The second dataset contains details of every fire engine dispatched to the scene of an incident since January 2009, providing information on the unit mobilized, its location of deployment, and the times of arrival at the incident scene.

The main objective of this project is to forecast the incidents' response time for initial mobilizations of the LFB in the Greater London area using only information available at the time of the incident. The response time, also called attendance time, is the total of crew's turn-out time and fire engine's (pumps) travel time. Turn-out time refers to the time taken for the firefighters to leave the station after they have been alerted to a call. This time includes the time taken to put on protective clothing, gather equipment, and board the fire engine. The measurement of the travel time is taken from the time the fire engine is dispatched to the time it arrives at the scene of the incident. LFB's standard operating procedure for calculating turn-out time and travel time is limited to regular fire engines, excluding any special appliances like aerial appliances.

3. LFB Emergency Response Process

To begin our analysis of the emergency response process we will refer to the diagram depicted in Figure 1 (LFB, London Datastore, 2023). This diagram outlines the essential stages involved in responding to an emergency call, from the initial occurrence of the emergency to the arrival of LFB's vehicles and firefighters at the incident scene to address the situation.

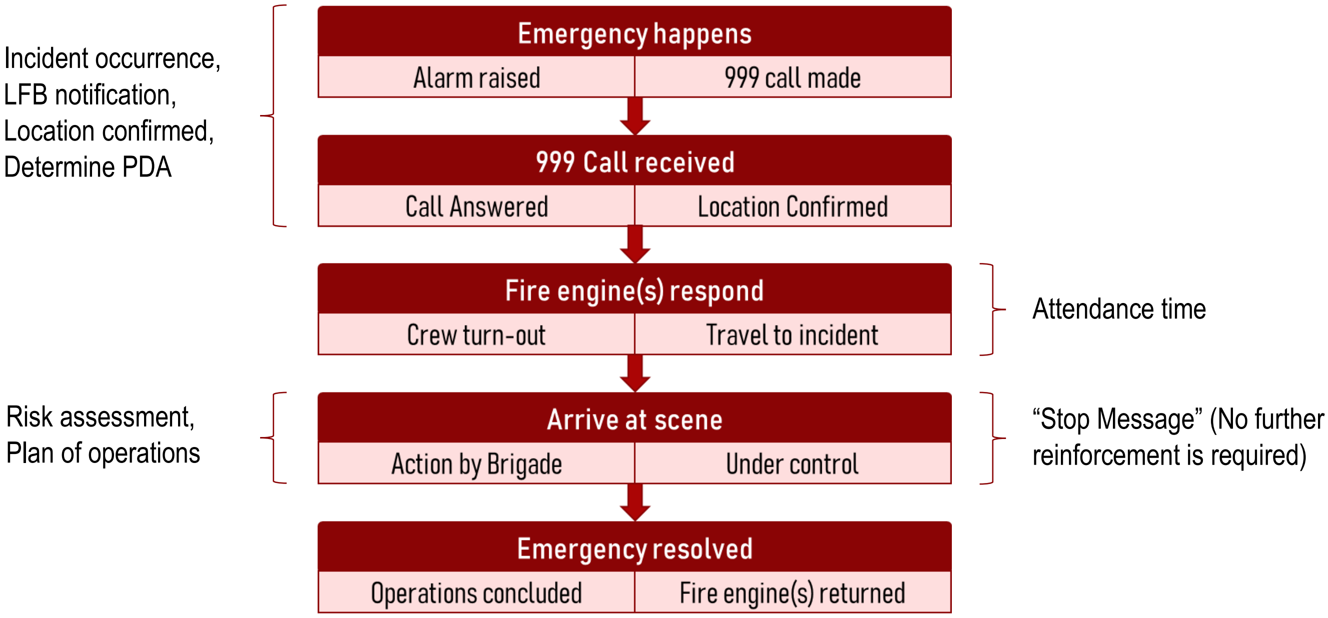


Figure 1 (LFB, London Datastore, 2023)

There is a crucial period of time between the incident occurring and the LFB being notified in any emergency scenario. This time frame is beyond the fire department's control and might significantly affect how the situation plays out.

Once a call has been received and the incident's location is confirmed, the LFB may dispatch a varying number of fire engines and appliances to the scene, depending on the nature of the incident and the location's characteristics (indoor, in street, etc.). The number of appliances initially sent to an emergency call is known as the "pre-determined attendance" (PDA), and it varies according to the type of incident. For instance, a PDA of two fire engines is the standard response for a dwelling fire. However, for complex buildings or situations that involve complicated firefighter logistics, the PDA may be higher.

To make sure the LFB team and the public are safe, they perform a dynamic risk assessment of the situation as soon as they arrive. A strategy for how to deal with the problem will also be included in this evaluation. When the situation is under control, the commander at the site sends a "stop message" to headquarters, informing them that no further strengthening resources are needed. Although the incident may persist for several minutes, hours, or even days, after the stop message is sent, it is considered contained at that point. The incident is formally closed when the last resource at the scene departs and returns to its home location.

The LFB's mobilizing system records important details related to emergency calls made to 999. These details are then transferred to the Incident Management System (IMS), where additional information is added by the responding crew and other personnel to create a comprehensive record of the incident.

To ensure the quality of data, various checks and quality assurance processes are in place. Despite these efforts, there are still some inaccuracies present in the system. For instance, some timings may be incorrect due to the failure to utilize the fire engine recording systems at the right moment. Similarly, there may be instances where arrival response times are not available.

4. Data Exploration and Analysis

The project relies on two datasets that are stored in two comma separated value files (csv). The incidents dataset is composed of 1,542,670 records and 39 feature columns, while the mobilizations dataset includes 2,152,581 records and 22 feature columns.

4.1. Incidents Data

Table1 summarizes the incidents data with a description, data type and percentage of missing values for each descriptor of the dataset. The rows highlighted in blue indicate features that were not considered in the modeling whereas the grey rows indicate features whose information was assimilated in the model after encoding.

Column	Description	Type	%
IncidentNumber	A unique identifier for each incident.	object	0
DateOfCall	The date on which the incident was reported.	object	0
CalYear	The calendar year in which the incident occurred.	int64	0
TimeOfCall	The time at which the incident was reported.	object	0
HourOfCall	The hour at which the incident was reported.	int64	0
IncidentGroup	Fire, False Alarm or Special Service	object	0
StopCodeDescription	Type of incident that occurred (False alarm - good intent, Flood, etc.)	object	0
SpecialServiceType	Type of special service required (Medical, Evacuation, etc.)	object	68.35
PropertyCategory	Category of property (Residential, Aircraft, etc.)	object	0
PropertyType	Specific type of property involved (House, Warehouse, etc.)	object	0
AddressQualifier	Location (in street, within building, etc.)	object	0
Postcode_full	Full postcode of the address where the incident occurred.	object	48.62
Postcode_district	District of the postcode (first part of the column Postcode_full).	object	0
UPRN	Unique Property Reference Number of the property.	float64	9.16
USRN	Unique Street Reference Number of the street.	float64	10.56
IncGeo_BoroughCode	A code identifying the borough where the incident occurred.	object	0.0
IncGeo_BoroughName	Name of the borough where the incident occurred.	object	0.0
ProperCase	Name of the area (based on IncGeo_BoroughName).	object	0.0
IncGeo_WardCode	A code identifying the ward where the incident occurred.	object	0.0
IncGeo_WardName	The name of the ward where the incident occurred.	object	0.0
IncGeo_WardNameNew	The new name of the ward where the incident occurred.	object	0.0
Easting_m	OSGB easting coordinate of the location	float64	48.62
Northing_m	OSGB northing coordinate of the location	float64	48.62
Easting_rounded	OSGB easting coordinate of the location rounded up to nearest 50	int64	0.0
Northing_rounded	OSGB northing coordinate of the location rounded up to nearest 50	int64	0.0
Latitude	The latitude of the location	float64	48.62
Longitude	The longitude of the location	float64	48.62
FRS	Name of the Fire and Rescue Service (London)	object	0
IncidentStationGround	The name of the LFB station that attended the incident	object	0
FirstPumpArriving_AttendanceTime	Time taken for the first fire engine to arrive at the incident	float64	8.05
FirstPumpArriving_DeployedFromStation	name of the station from which the first fire engine was deployed.	object	8.05
SecondPumpArriving_AttendanceTime	Time for the second fire engine (if needed) to arrive at the incident.	float64	64.24
SecondPumpArriving_DeployedFromStation	Name of the station from which the second fire engine was deployed.	object	64.24
NumStationsWithPumpsAttending	Number of stations that attended the incident.	float64	0.76
NumPumpsAttending	Number of fire engines that attended the incident.	float64	0.76
PumpCount	Number of fire engine crews that attended the incident.	float64	0.58
PumpHoursRoundUp	Total duration of attendance rounded up to the nearest hour	float64	0.6
Notional Cost (£)	Time spent multiplied by notional annual cost of a pump.	float64	0.6
NumCalls	Number of 999 calls received for the incident	float64	0.12

Table 1

4.2. Mobilization Data

Table 2 summarizes the mobilizations data with a description, data type and percentage of missing values for each descriptor of the dataset. The rows highlighted in blue indicate features that were not considered in the modeling whereas the grey rows indicate features whose information was assimilated in the model.

Column	Description	Type	%
IncidentNumber	A unique identifier for each incident	object	0
CalYear	The year in which the incident occurred	int64	0
HourOfCall	The hour at which the incident was reported	int64	0
ResourceMobilisationId	A unique identifier for each resource mobilization event	int64	0
Resource_Code	The code for the resource that was mobilized	object	0
PerformanceReporting	A code indicating the order in which a fire engine arrived at the location	object	0
DateAndTimeMobilised	Date and time at which the firefighters were notified	object	0
DateAndTimeMobile	Date and time at which the resource began travelling to the incident location	object	1.25
DateAndTimeArrived	Date and time at which the resource arrived at the incident location	object	0
TurnoutTimeSeconds	Prepare time for firefighters after they have been alerted to a call	float64	1.25
TravelTimeSeconds	Travel time from its station to the incident location	float64	1.26
AttendanceTimeSeconds	Turn out time + travel time	int64	0
DateAndTimeLeft	Date and time at which the resource left the incident location	object	0
DateAndTimeReturned	Date and time at which the resource returned to its station	object	0.02
DeployedFromStation_Code	The code for the station from which the resource was deployed	object	54.79
DeployedFromStation_Name	The name of the station from which the resource was deployed	object	0
DeployedFromLocation	The location from which the resource was deployed	object	0
PumpOrder	The order in which the resource (fire engine) was mobilized	int64	0.05
PlusCode_Code	The code for the mobilization (Initial, rca, add)	object	0
PlusCode_Description	A description of the Plus Code	object	0
DelayCodeId	reason for any delay in the resource mobilization	float64	0
DelayCode_Description	A description of the reason for any delay in the resource mobilization	object	75.26

Table 2

4.3. Data Analysis

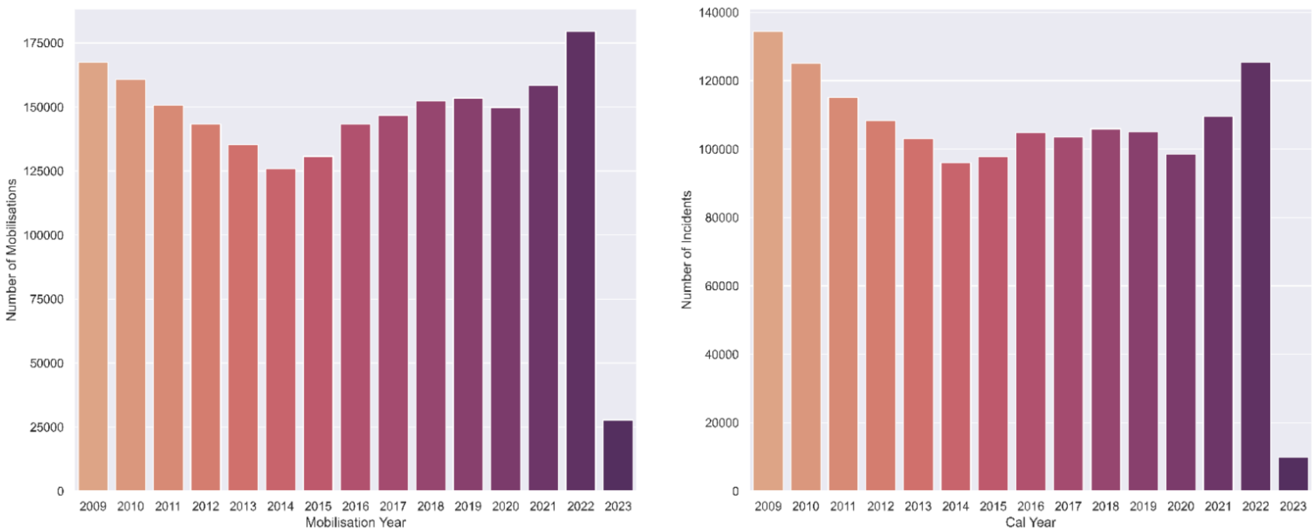


Figure 2

The Datasets cover the period from January 2009 to 2023 and during this time, the London Fire Brigade responded to just over 100,000 incidents annually in London. In Figure 2 we can see that the number of mobilizations and incidents increased over the years from 2014 to 2022, compared to a decrease from 2009 to 2013. The number of mobilizations peaked in 2022 at 179,458, which is significantly higher than any other year in the dataset, while the number of incidents peaked in 2009 with 134,379 recorded incidents.

Some possible factors for the increment in mobilizations and incidents could include an increase in population, urbanization, and/or the number of properties, as well as changes in fire safety regulations or public awareness campaigns. Additionally, external events such as extreme weather conditions or natural disasters also played a role in the number of incidents and mobilizations.

In 2022, the Brigade also responded to 703 incidents in neighboring county areas, at their request. However, for the purpose of this project, we are only focusing on the incidents that occurred within the boundaries of Greater London. While fire engines from neighboring counties also responded to some incidents in London, we are only considering the response provided by the LFB fire engines in this project, as we don't have data for the incidents attended by fire engines from other Brigades (LFB, London Datastore, 2023).

4.3.1. Incidents Categories and Groups

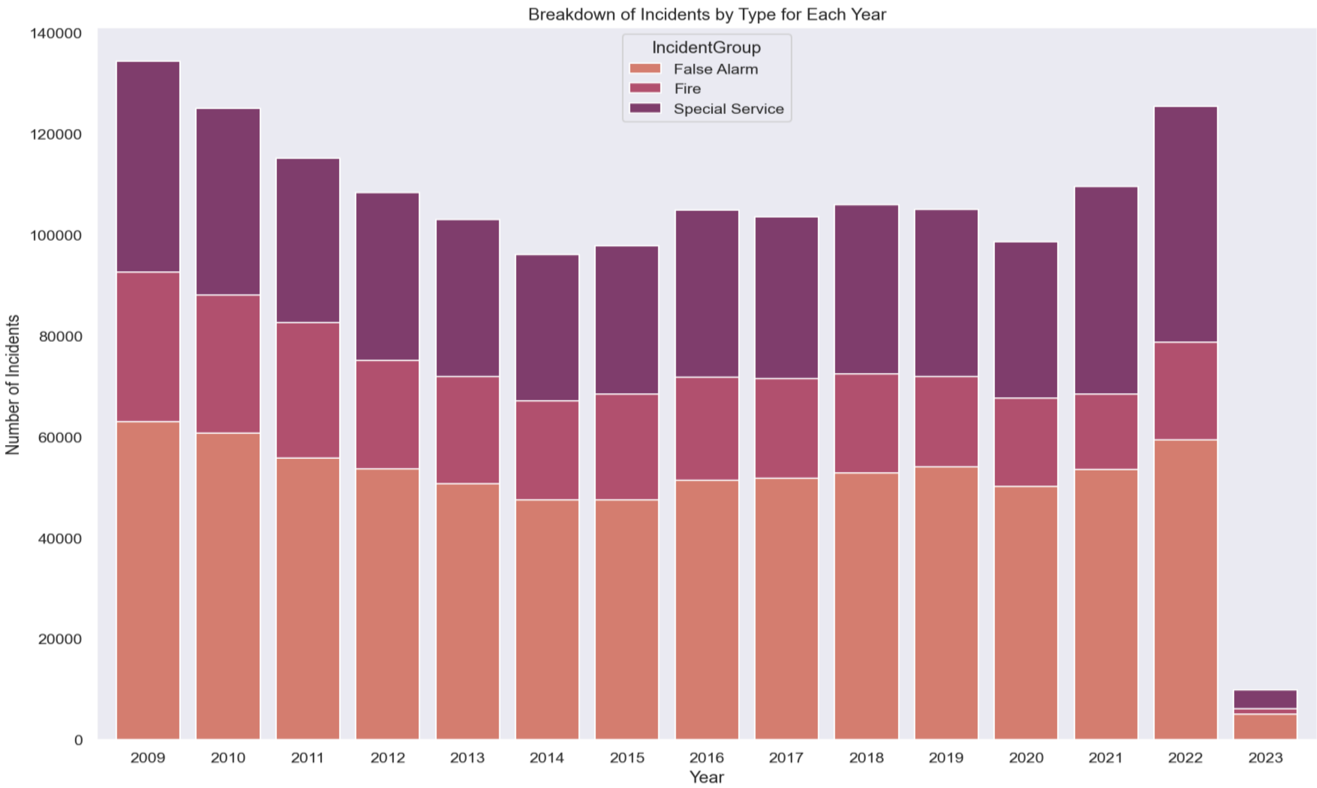


Figure 3

There are three categories of incidents: False Alarm, Fire, and Special Service. The datasets show that nearly 50% of the incidents recorded are actually false alarms, totaling 756,775 instances. The second most common incident is Special services, with 488,259 instances recorded, while Fire is the least common incident, with only 297,636 instances recorded. A breakdown of incidents by type for each year is depicted in Figure 3**Error! Reference source not found.**.

In Figure 4 we can see that around 75% of false alarm incidents were recorded as Automatic False Alarms (AFA). They are incidents that are triggered automatically by non-emergency activities like cooking or smoking. The remaining false alarms fall into two categories: those with good intention and those with malicious intent. Fire incidents are classified into two categories: primary fires, which include serious fires that cause harm to people or property, and secondary fires.

Primary fires occur in buildings, vehicles, or certain outdoor structures, involve fatalities, casualties or rescues, or are attended by five or more pumping appliances. They are classified into four sub-categories: dwelling fires which occur in residential properties, other building fires which occur in non-residential or institutional residential buildings (hostels/hotels), road vehicle fires which occur in vehicles used for transportation (cars/vans), and other outdoor fires which occur in primary outdoor locations that include aircraft, boats trains and outdoor structures such as post or telephone boxes, bridges, tunnels etc.

Secondary fires are typically minor outdoor fires that do not cause harm to people or property. This category includes fires in refuse (scraps or garbage), grasslands, and abandoned buildings or vehicles. However, if such fires involve casualties, rescues, or require the attendance of five or more pumping appliances, they are categorized as primary fires.

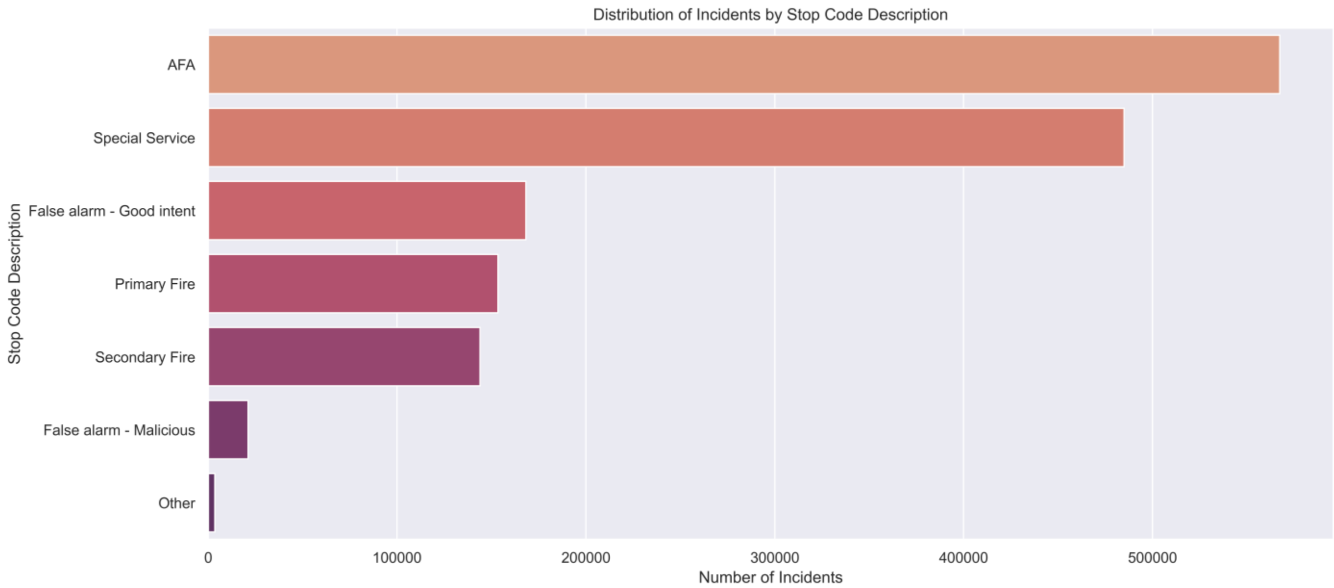


Figure 4

The incidents dataset contains nearly 500,000 cases of non-fire incidents, recorded as special service incidents, such as flooding, individuals stuck in lifts, or medical emergencies as shown in Figure 4.

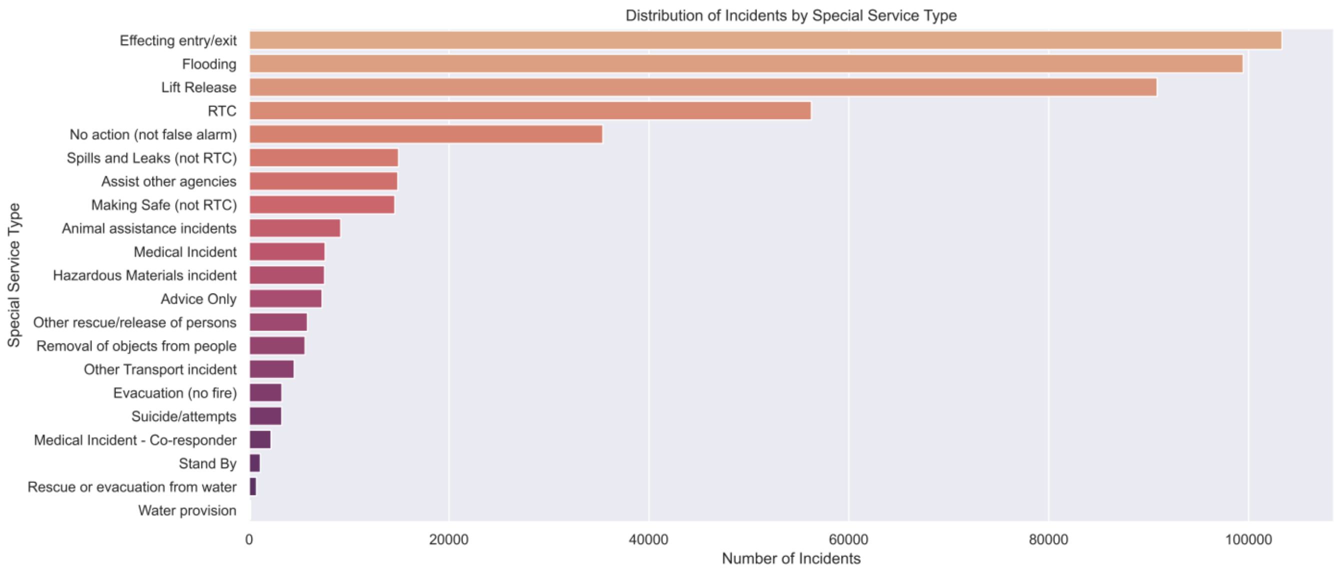


Figure 5

In Figure 5 we show the number of incidents related to different types of special services provided by the London Fire Brigade. The highest number of incidents is related to "Effecting entry/exit" with 103,350 incidents. The LFB employs Forced Entry Equipment (FEE) to enter buildings or premises for firefighting or emergency purposes and to assist other agencies, such as the London Ambulance Service (LFB, London Fire Brigade, 2023). The second most common special service incidents are “Flooding” with 99,473 incidents, followed by “Lift Release” with 90,845 incidents, which refers to the process of releasing people who are trapped in lift cars, typically in buildings such as apartment blocks, commercial buildings, and public buildings (LFB, London Datastore, 2023). “RTC” which stands for road traffic collision has 56,270 incidents recorded, while “No action (not false alarm)” has 35,405 incidents recorded.

The data indicates that the LFB responds to a wide variety of incidents besides fires, such as medical emergencies and animal assistance. The high number of incidents related to flooding and lift release suggests the importance of these services in the London area, while the relatively low number of incidents related to suicide and attempts indicates a lower occurrence of such incidents in the area.

4.3.2. Response Time

As described earlier, the turn-out time refers to the time taken for the firefighters to leave the station after they have been alerted to a call. This includes the time taken to put on protective clothing, gather equipment, and board the fire engine. The travel time is the time taken by the fire engine to reach the location of the incident from the station and the attendance time refers to the time taken by the fire engine to arrive at the location of the incident and contact the caller or person in charge.

Figure 6

In Figure 6 we can see that the average turnout time decreased from 2009 to 2012, with the biggest drop between 2009 and 2010. The time it takes for crews to turn out has significantly improved since 2008, owing to the emphasis placed on ensuring that they are able to respond as quickly as possible (LFB, London Datastore, 2023). From 2012 until now, the average turnout time remains relatively stable. Furthermore, the turnout times for fire engine crews across London are similar, although the layout of the fire stations are different.

The average travel time for a mobilization increased since 2009 and was highest in 2015, with 289.35 seconds (or approximately 4 minutes and 49 seconds). It decreased from 2015 to 2020, reaching a low of 266.51 seconds (or approximately 4 minutes and 26 seconds). The availability of fire engines to respond to incidents may have improved in 2020 due to Covid-19 related cancellations of community activities that are carried out by fire engine crews. With fewer people commuting and attending schools, fire engines may have been able to travel to incidents more quickly during this period. However, in 2021 and 2022, there has been an increase in average travel time due to more activity on roads and more community activities taken on by fire crews.

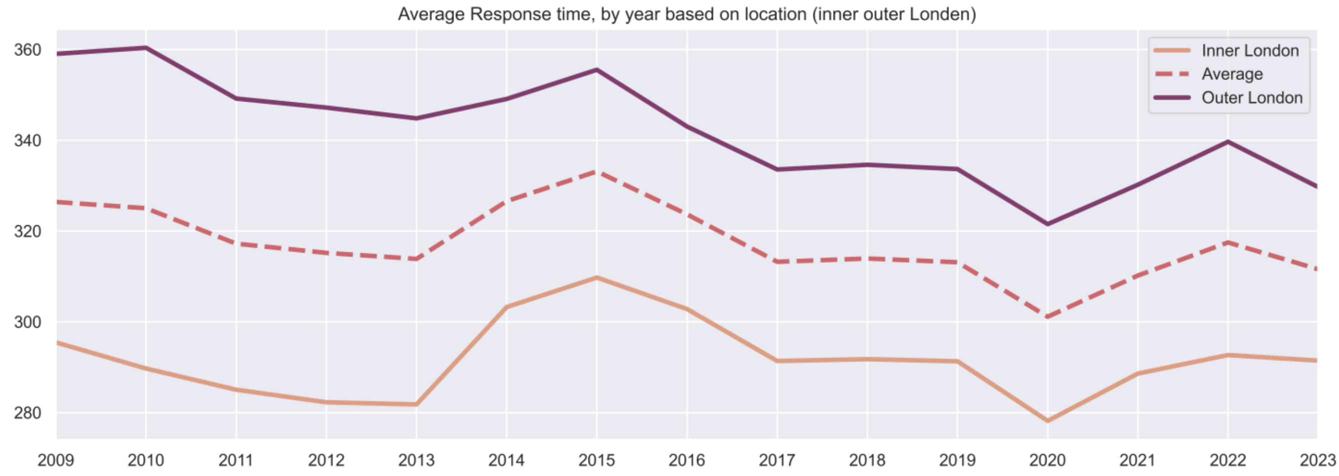


Figure 7

Overall, there seems to be some variation in the average travel time from year to year, but no clear trend. The primary factor influencing the LFB's response time to an incident is the geographic placement of their fire stations and fire engines as shown in Figure 7, Figure 8 and Figure 9.

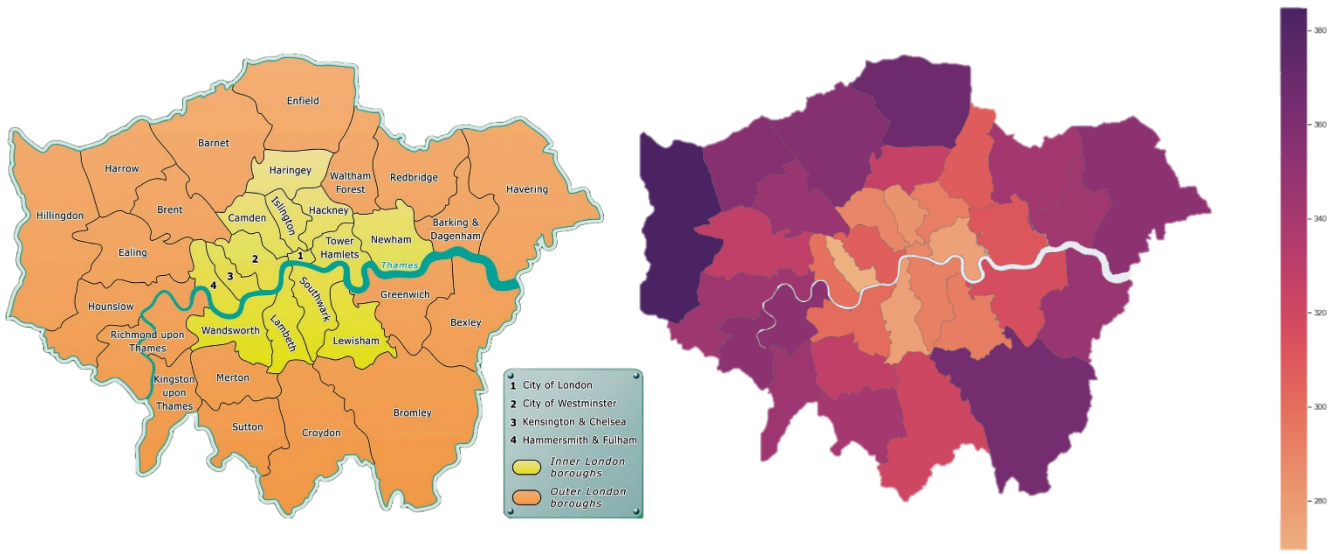


Figure 8 on the left: Inner and outer London; on the right: The average response time in outer London is higher than inner London

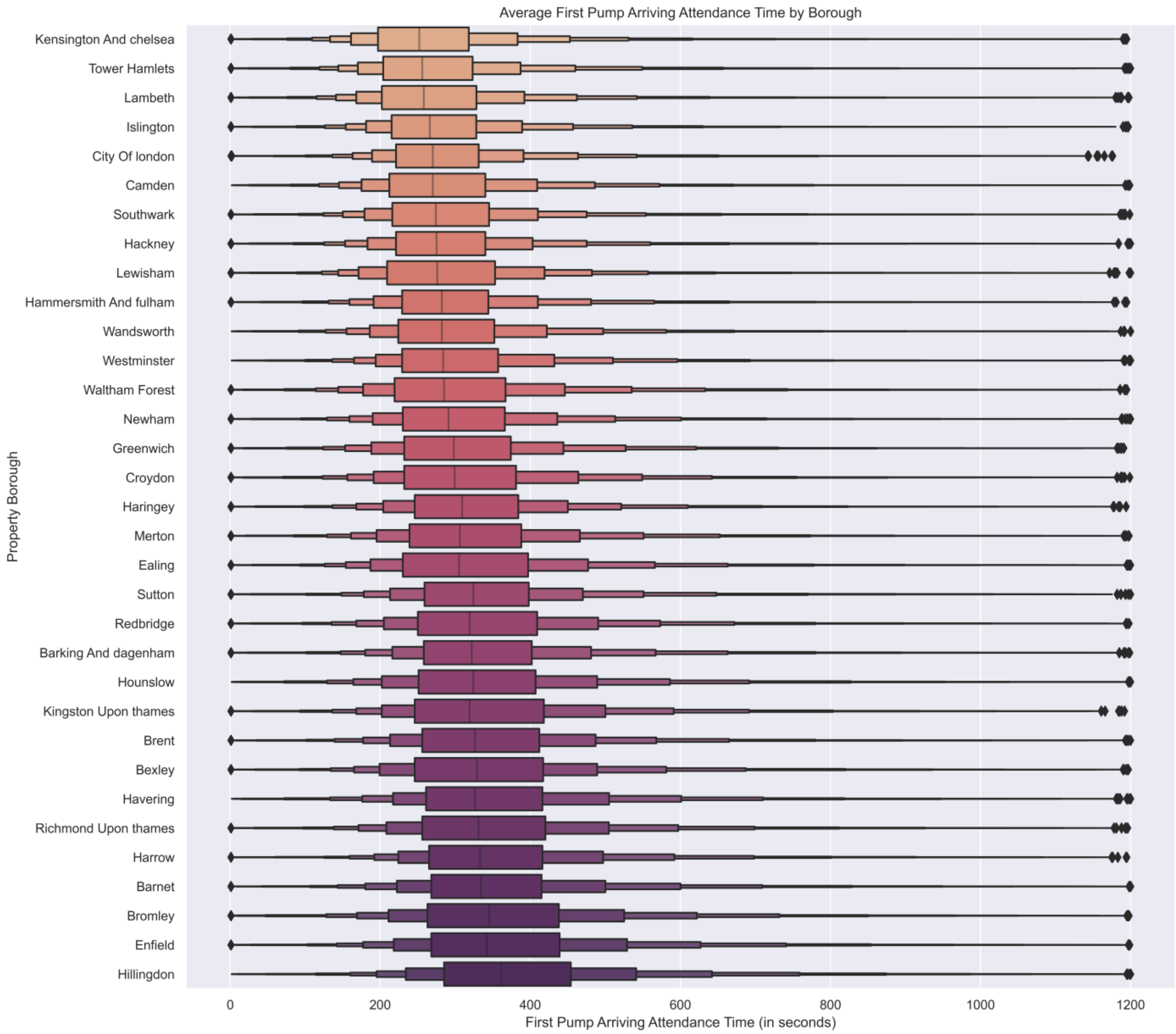
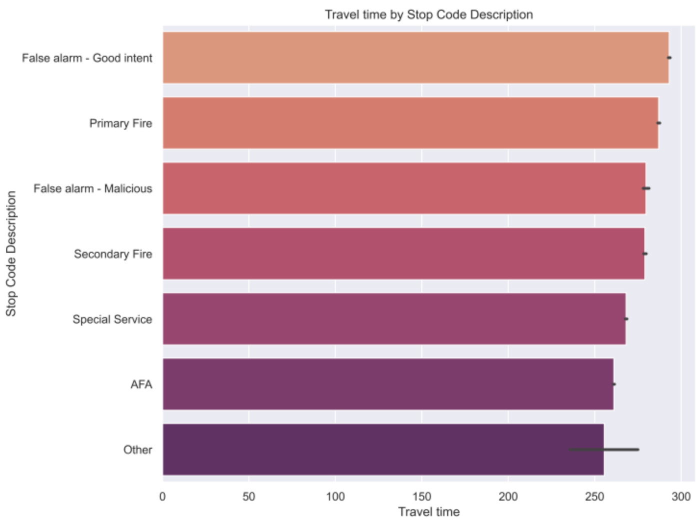


Figure 9

Our analysis revealed that the response time of the LFB is not highly affected by the type of incident.



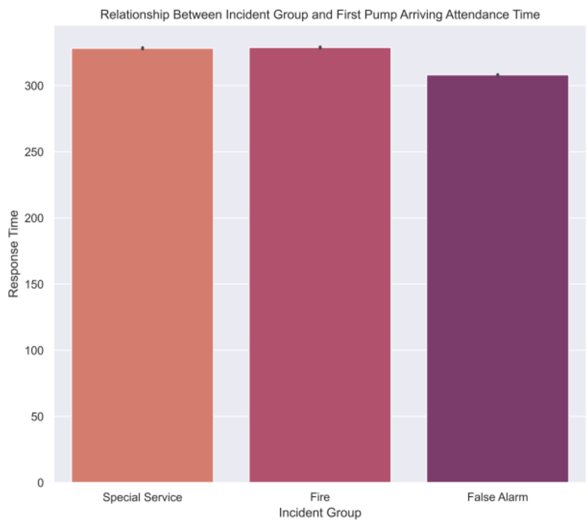


Figure 10

As depicted in Figure 10, the average response time for AFA incidents is lower compared to special service and fire incidents. This is because automatic alarms, such as smoke detectors, not only alert the LFB about an incident but also provide the exact location, reducing response time. Some incidents had longer response times due to geographical challenges in locating the incidents. This information is provided in the column `AddressQualifier` of the incidents-dataset.

Although the average response time of special service incidents is higher than that in the AFA, in Figure 11 we show that the average response time for medical emergencies (Medical or object removal incidents) is the lowest.

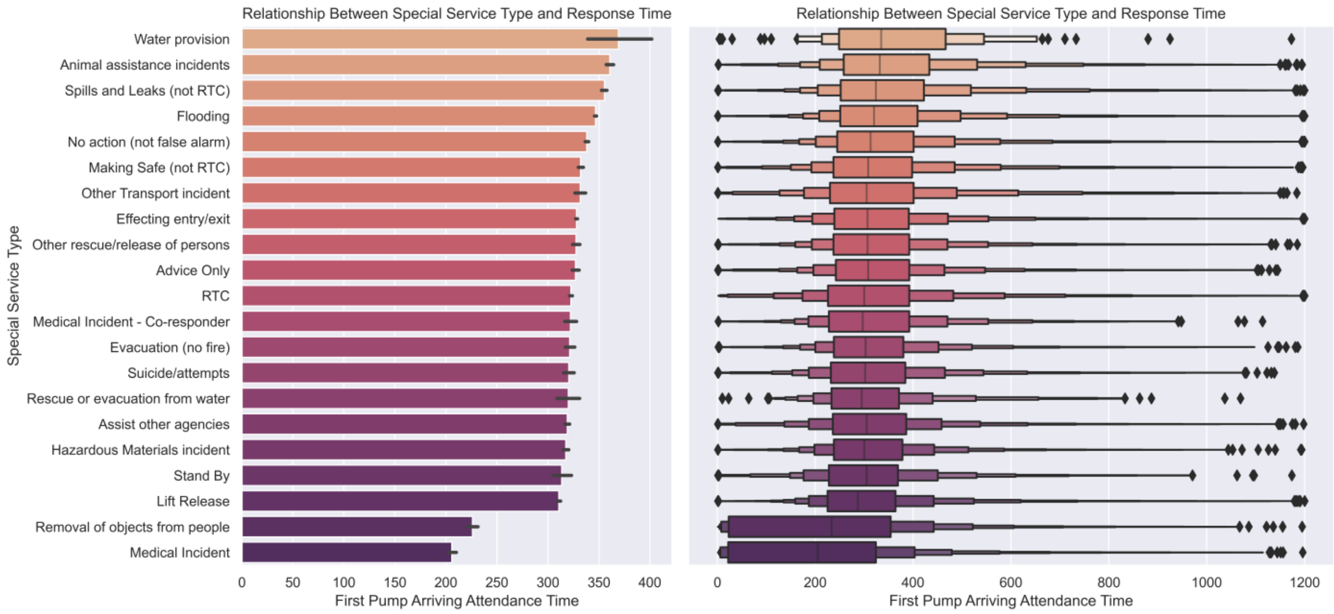


Figure 11

In Figure 12 and Figure 13 we explore some very important dependencies:

- One can see that there is a clear trend of increasing mobilizations and incidents from the early morning hours until the afternoon, peaking around 18:00, then decreasing in the evening hours. The number is highest between 16:00 and 19:00, which may correspond to peak rush hour traffic when accidents are more likely to occur. The lowest mobilization numbers are in the early morning hours between 0:00 and 05:00.
- Despite the expectation that response times would correspond with the number of incidents, there is no clear correlation. While it would be logical to assume that response times would be higher during rush hour and lower between 00:00 and 06:00, this is not always the case. Nonetheless, the shift system in London's fire stations ensures that they are fully staffed by full-time firefighters, with no part-time firefighters employed. This guarantees that there are always firefighters available to respond to incidents, as highlighted by the Fire Brigades Union (FBU, 2023). Despite this, the data reveals that the turnout time is consistently higher between 23:00 and 07:00, which is a contributing factor to the higher response times during the early morning hours between 00:00 and 06:00.
- It is evident that the travel time of the London Fire Brigade to an incident location is closely linked to the time of day and day of week. Figure 12 illustrates this correlation, indicating that the travel time of the fire brigade is linked to traffic congestion and delays during peak hours.
- The LFB response time is generally lower on Sundays and Saturdays between 08:00 and 09:00 compared to other times of the week. One possibility is that there may be less traffic on the roads during this time, allowing fire engines to reach incidents more quickly.

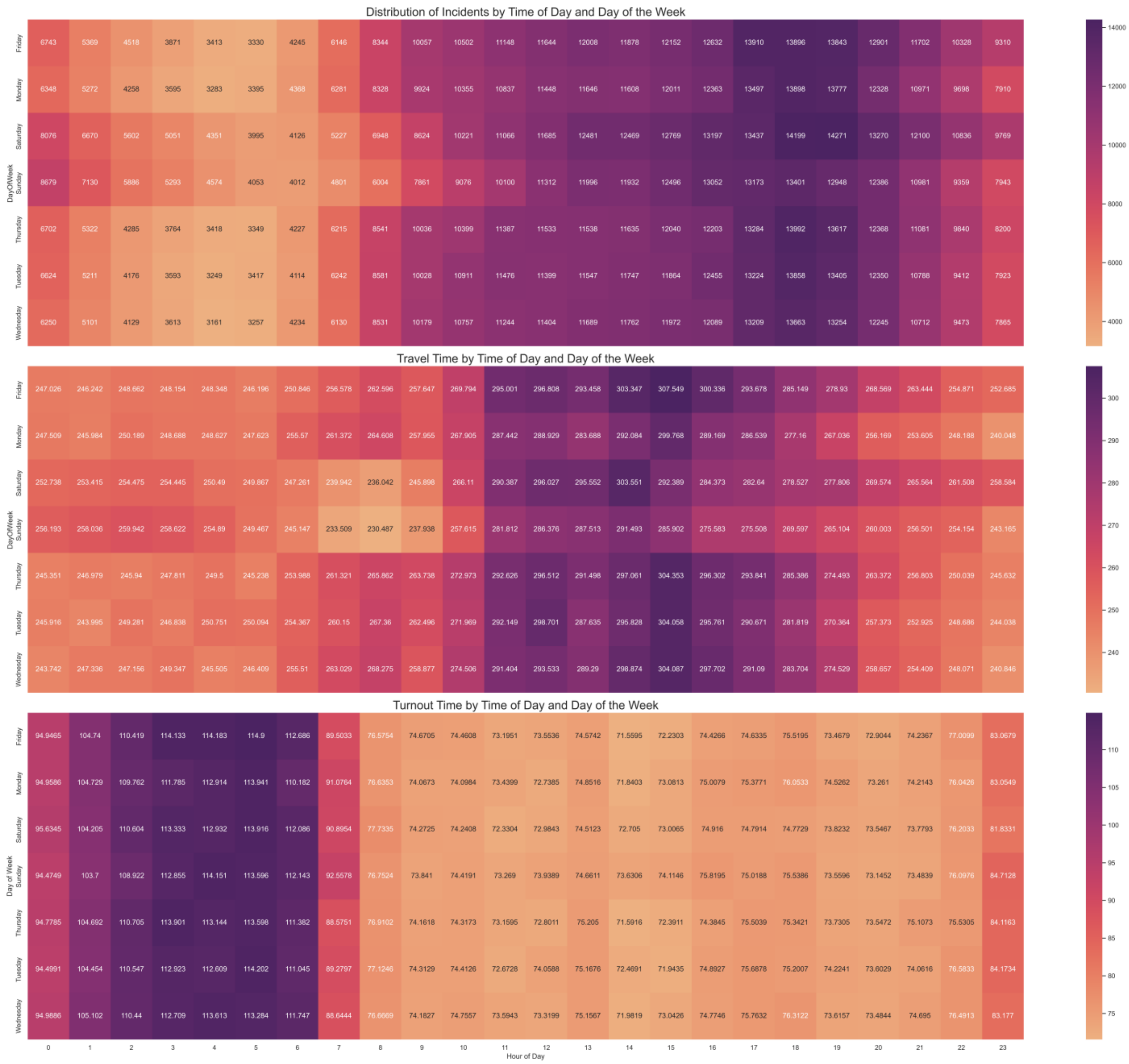

Distribution of Incidents by Time of Day and Day of the Week
Travel Time by Time of Day and Day of the Week
Turnout Time by Time of Day and Day of the Week
DayOfWeek
Day of Week
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday
Hour of Day
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
14000 12000 10000 8000 6000 4000
300 290 280 270 260 250 240
110 105 100 95 90 85 80 75

Figure 12

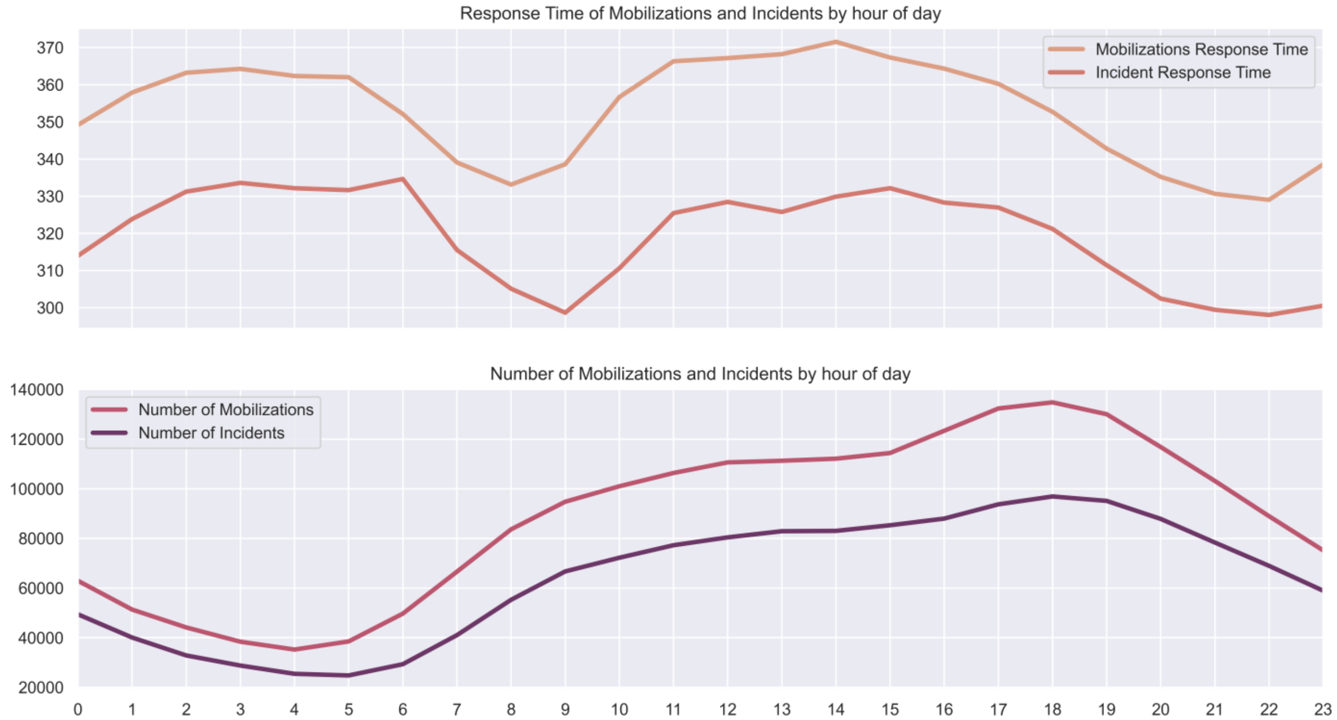


Figure 13

There is a correlation between the type of property where an incident occurred and the travel time of fire engines. The data in Figure 14 shows that incidents on boats and aircrafts tend to have higher median response times compared to other incidents. This suggests that there may be specific challenges or obstacles that firefighters face when responding to incidents on boats and aircrafts, which may increase the travel time.

According to our analysis, the station from where a fire engine is dispatched also has an impact on the turnout and travel times. While the turnout times for fire engine crews are generally consistent throughout London, there are some fire stations that experience slightly longer turnout times (about 10 to 15 seconds longer) than others. Moreover, the travel time is significantly influenced by the location of the fire station as depicted in Figure 15.

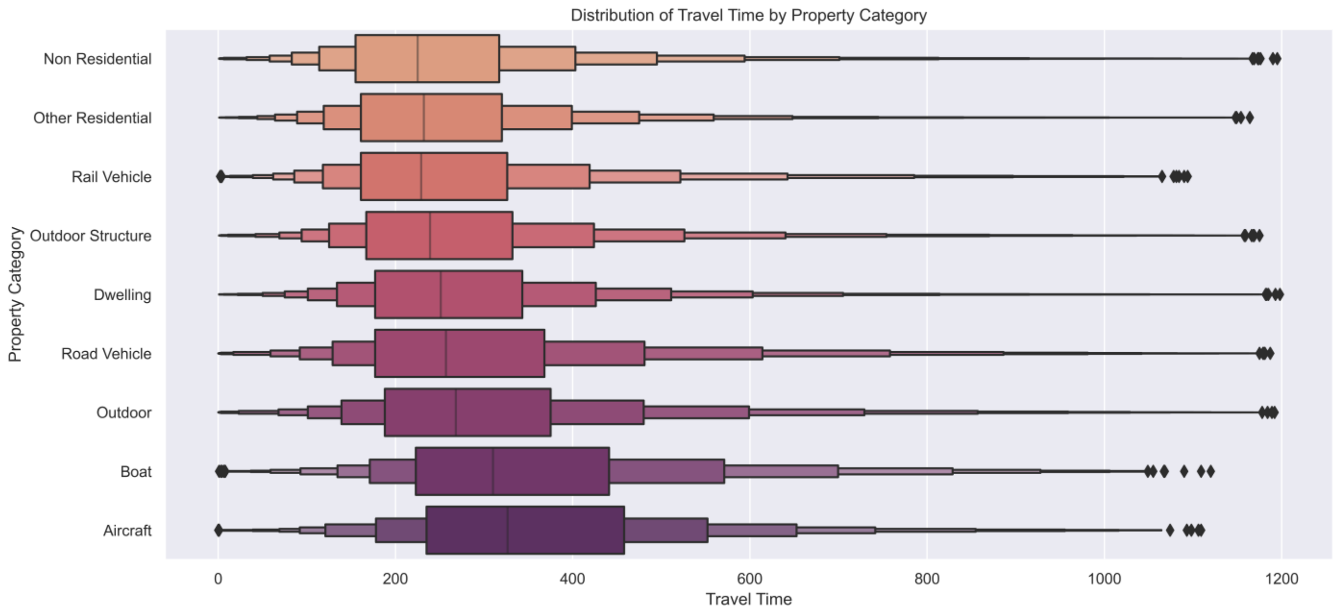


Figure 14

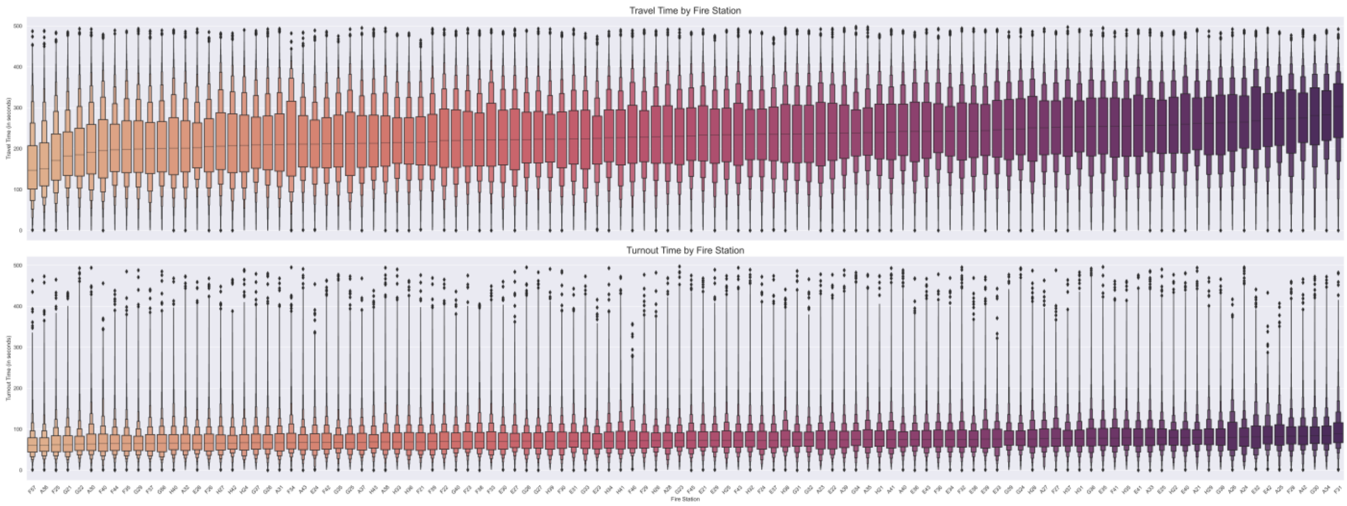


Figure 15 After accounting for outliers

5. Data Preprocessing

As previously stated, we will utilize data from the LFB website to forecast the response time of the LFB. The primary sources of our data are the incidents and mobilizations datasets, which are outlined in Table 1 and Table 2, respectively. Additionally, we have obtained a dataset that provides details about the locations of fire stations in London, which we will also incorporate in our analysis.

5.1. Incidents dataset

The incidents dataset contains 38 columns, many of which are not useful to explain the response time of the London Fire Brigade.

5.1.1. Remove features

The incident data set described in Table 1 contains some columns that are used to describe the values of other columns. Specifically, the columns `IncGeo_BoroughName`, `ProperCase`, `IncGeo_WardName` and `IncGeo_WardNameNew` are used to indicate the borough and ward names where the incident occurred. The columns `FRS` contains the name of the fire service that took care of the incident and contains one value (London). As a result, we have opted to exclude these columns from the dataset.

Unfortunately, not all information regarding the incident's category is accessible during the initial call. After receiving a call, firefighters are informed about the incident's category, which can be a fire or a fire alarm triggered by an automated system like a smoke detector, or a special service incident. The LFB is unaware if the call or fire alarm is a false alarm (such as AFA, malicious or good intent alarms) at the time of notification. Furthermore, the dataset does not include information about which incidents were generated by automatic alarms and which were actual calls. All incidents that were categorized as AFA were triggered by automated systems. The dataset contains fire incidents that were triggered by automated systems and turned out to be genuine fires. Additionally, some incidents that were originally classified as real alarms but were later determined by firefighters, upon arrival at the scene, to be false alarms, either with good or malicious intent. Hence, we decided to remove the columns `SpecialServiceType`, `StopCodeDescription` and `IncidentGroup` from the incidents-dataset.

The dataset includes an `AddressQualifier` column that provides information about the location of the incident, such as whether it occurred indoors, outdoors, or at the correct location. The `PropertyType` column provides specific details about the kind of property where an incident occurred, including but not limited to cars, lakes, houses, and so on. However, these columns are only added to the dataset after the location of the incident is inspected and the situation is concluded. Therefore, the information is not immediately available at the time of the incident. As a result, we did not include this column in the dataset.

The same implies to the columns `NumStationsWithPumpsAttending`, `NumPumpsAttending`, `PumpCount`, `PumpHoursRoundUp`, `Notional Cost (£)`, `FirstPumpArriving_DeployedFromStation` `SecondPumpArriving_DeployedFromStation`, and `NumCalls`. These columns are only available after concluding the incident and therefore, we decided not to include them in our final dataset.

The column `UPRN` contains the Unique Property Reference Number of properties in London where incidents occurred and has 9.16% missing values. However, the number of missing values is actually higher than initially estimated. The reason behind this is that the UPRN is only available for incidents that occur within "Dwelling" buildings. For other types of properties, the UPRN values are either stored as 0 or are not recorded in the dataset, which has resulted in a greater number of missing values. After addressing this issue, it was found that the missing values in the UPRN column were actually 57.78%, which is significantly higher than the previous estimation.

The `USRN` column represents the Unique Street Reference Number for the incidents. However, we found that approximately 10.56% of the incidents had missing values in the USRN column, all of which took place between 2009 and 2015. This raised questions about whether the USRN values were manually entered for incidents that occurred before 2015, which increases the risk of human error, or if there was a change in the navigation system's procedure for entering the number. Given the uncertainty surrounding the USRN data and the availability of a better location-related feature, we decided to remove the column from our dataset.

The United Kingdom is divided into sub-national electoral districts known as wards, each represented by one or more councilors. Greater London currently encompasses 32 boroughs, each containing a varying number of wards, totaling 624 in the region. Additionally, the City of London, a separate entity, comprises a single ward. The incident dataset includes three columns relevant to the ward level: `IncGeo_WardCode`, `IncGeo_WardName`, and `IncGeo_WardNameNew`. The `IncGeo_WardCode` column features 900 unique values, while the `IncGeo_WardName` and `IncGeo_WardNameNew` columns contain 770 and 774 unique values, respectively. Several factors contribute to this discrepancy. Ward names, codes and boundaries have changed over time, and the new names are reflected in the `IncGeo_WardNameNew` column. Additionally, human or system errors have caused some information about wards to be incorrectly stored in the data. Moreover, the London Fire Brigade dispatches fire engines to wards beyond the Greater London area. We attempted to update the ward names and codes in our dataset to align them with most recent ones. However, we encountered difficulties due to insufficient data and were unable to complete the task successfully. As a result, we decided to remove these columns from our dataset altogether.

We found that the `Postcode_district` column, which provides information about the postcode of the incident's address, is subject to changes in postcodes and their boundaries over time, much like the ward information. As a potential alternative, the column `Postcode_full` offers more comprehensive information about the postcode location of the incident, but is restricted to properties of the dwelling category and has almost 50% of its values missing. This column is also impacted by changes in postcodes and their boundaries over the years. Given this information and to ensure the accuracy and reliability of our model, we made the decision to remove this column from the dataset, because outdated or inaccurate postcode data could lead to incorrect or misleading findings.

Since we lack ward and postcode information in our dataset, we made the decision to use alternative location-based features to model response time accurately. Although information about the ward and postcode of the incident could be a valuable feature in some modeling scenarios, we recognized that other location-related data could be more critical for accurate predictions.

The Ordnance Survey Great Britain (OSGB) is one of the most popular coordinate reference system used for mapping Great Britain and providing precise geographic information. This system employs a two-dimensional Cartesian coordinate system with easting and northing coordinates to describe the positions of features on British maps (Ordnance-Survey, 2023).

The columns `Easting_m`, `Northing_m`, `Easting_Rounded`, and `Northing_Rounded` store the OSGB coordinates. Only incidents that occur in "Dwelling" buildings have the `Easting_m` and `Northing_m` columns available. On the other hand, `Easting_Rounded` and `Northing_Rounded` are available for all incidents and are calculated by rounding up the last two digits of the easting and northing coordinates to 50. The `Latitude` and `Longitude` columns contain GPS coordinates, and similar to the `Easting_m` and `Northing_m` columns, are only available for incidents that occurred in the "Dwelling" category. Unfortunately, the `Easting_m`, `Northing_m`, `Latitude`, and `Longitude` columns have a high number of nearly 50% missing values. Furthermore, the Latitude and Longitude columns have 623 entries with a value of 0, which we assume are errors in the system. We addressed these issues by computing the longitude and latitude GPS coordinates from the rounded Easting and Northing OSGB coordinates to replace the missing values in the dataset. For the 0 valued Longitude and Latitude entries, we calculated the correct values using the Easting and Northing coordinates, then replaced the incorrect entries with the accurate ones.

Although calculating the GPS coordinates based on the rounded easting and northing coordinates may not provide the exact incident location, it provides a good estimation of the location within a close proximity to the actual incident location. To test this hypothesis, we calculated the longitude and latitude using the rounded easting and northing coordinates for the non-missing non-zero values in the dataset. We then calculated the Euclidean distances (after accounting for the earth curvature) between the actual longitude and latitude and the approximate values. The average of the Euclidean distances, which was 37.28 meters, was used to determine the approximation error, with a standard deviation of 14.96 meters. The maximum distance was 70.81 meters, and it occurred in only 0.001% of the data.

Figure 16 illustrates the distribution of the location approximation error.

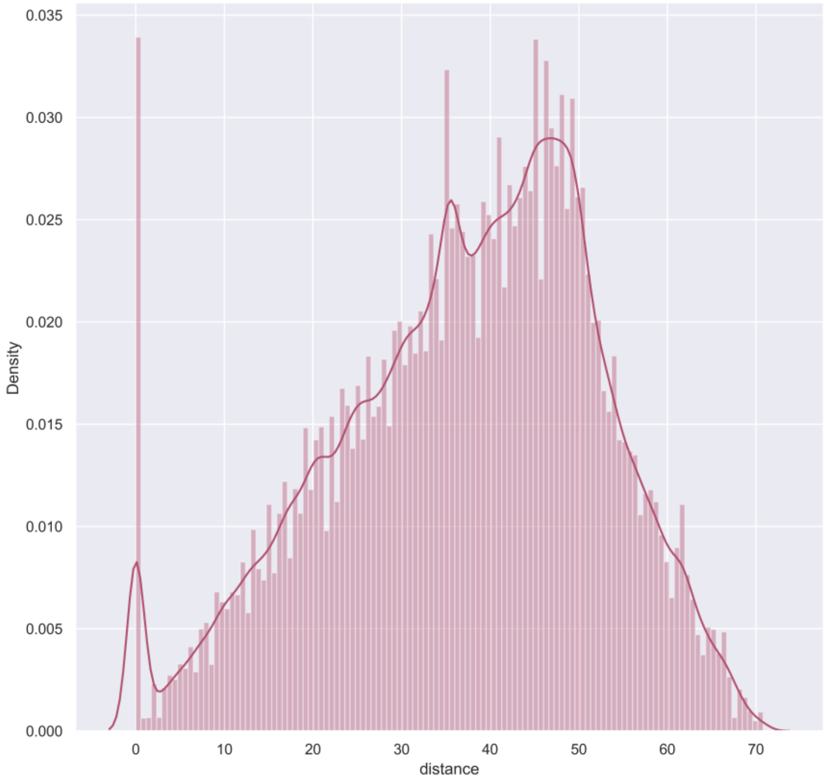


Figure 16

Further analysis (using simple physics calculations) showed that this approximation error would add a travel time of 3 to 6 seconds, based on the 0 to 70km/h speed, acceleration, and weight of fire engines used by the LFB (calculations performed to arrive at these results are beyond the scope of this project).

Once we confirmed our hypothesis, we proceeded to update the `Latitude` and `Longitude` columns with the newly calculated Latitude and Longitude values. with the GPS coordinates obtained from our calculations. Afterwards, we made the decision to eliminate the `Easting_m`, `Northing_m`, `Easting_Rounded`, and `Northing_Rounded` columns from the dataset.

The column `SecondPumpArriving_AttendanceTime` is only available if a second fire engine was dispatched to the incident's location and we are only interested in knowing the attendance time of the LFB, not the attendance time of every dispatched fire engine. To calculate the response time of the London Fire Brigade (LFB), we will not use the column `FirstPumpArriving_AttendanceTime` in the incidents-dataset. Instead, we will use the columns `TurnoutTimeSeconds` and `TravelTimeSeconds` from the mobilizations-dataset.

5.1.2 Features Encoding

After analyzing the data, it was found that the hour of the day, day of the week, year, and month all have an impact on turnout and travel times of LFB. To extract these temporal features, one-hot encoding was performed on the columns `DateOfCall`, `TimeOfCall`, and `CalYear`. These columns were then removed from the dataset.

The response time to an incident was influenced by its location and the category of the property, where the incident occurred. Certain boroughs, for example in outer London, have (on average) higher response times than others as previously seen in Figure 8 on the left: Inner and outer London; on the right: The average response time in outer London is higher than inner London. The travel times to some properties were higher than others due to some difficulties. To address these issues, we utilized One-Hot-Encoding to transform the categorical data in the `IncGeo_BoroughCode` and `PropertyCategory` columns into distinct columns that represent each specific borough and property category respectively where the incident occurred and then removed the borough and property category columns from the dataset.

The `IncidentStationGround` column contains the name of the fire station responsible for incidents in its area. We used Labe Encoding to convert the values of the column into numerical values. We did not use One-Hot-Encoding because of the features encoding that we performed on the mobilizations dataset, which is explained in the next section.

5.2. Mobilizations dataset

The mobilizations dataset contains information about each mobilization record of the LBF and consists of 22 columns.

5.2.1. Remove Features and Filtering

During our analysis of the mobilizations-dataset, we found that some feature columns were already present in the incidents-dataset, such as `CalYear` and `HourOfCall`. To avoid redundancy, we removed these columns from the mobilizations-dataset.

Another column we decided to remove was `ResourceMobilisationId`, which contained the mobilization ID. Similarly, we did not analyze the relationship between the firetrucks used by the LFB and the incident's response time, so we removed the column `Resource_Code`.

The `PerformanceReporting` column in the dataset provides details on the sequence in which the fire engines arrived at the incident location. Additionally, the `DateAndTimeLeft` and `DateAndTimeReturned` columns provide information about the time the fire engine left the incident location and the time it returned to the station, respectively. However, since this information is not pertinent to our project, we decided to remove these columns from our dataset. These columns were added to the dataset after the conclusion of the incident, making them irrelevant to our response times prediction.

The columns `DateAndTimeMobilised`, `DateAndTimeMobile`, and `DateAndTimeArrived` provided information about the times at which a crew at a fire station was notified of an incident, the time at which the fire engine left the station, and the time at which the fire engine arrived at the incident location, respectively. The turnout time and travel time are calculated using these columns and are stored in the columns `TurnoutTimeSeconds` and `TravelTimeSeconds`, respectively.

Initially, the `TimeOfCall` column was present in the incident dataset, but we removed it. Before we did, we wanted to calculate the time elapsed between the moment a control room officer receives a call and the moment firefighters get notified and start preparing to respond, since this information can be useful for our analysis. For this purpose, we tried using the `DateAndTimeMobilised` and `TimeOfCall` columns to calculate the difference. However, after merging the datasets and calculated the time difference, we observed that many entries in the newly generated column had negative values. This indicated that the dates and times in one or both of the columns were incorrect. As a result, we decided to remove the `DateAndTimeMobilised`, `DateAndTimeMobile`, and `DateAndTimeArrived` columns from the dataset as they did not provide us with any valuable information that we didn't already have.

We considered the column PumpOrder, which represents the order in which fire engines were dispatched, as it may provide valuable information. However, we ultimately decided to remove it from the dataset. Our decision was based on the following scenarios:

- Firstly, if two fire engines A and B were dispatched at the same time, and one arrived faster than the other, then the order in which they were dispatched would not matter.
- Secondly, if fire engine B was dispatched after fire engine A, and A arrived faster, then there would be no information lost if one were to predict the response time of fire engine A.
- However, if fire engine B was dispatched after fire engine A, and B arrived faster than A, there could be information loss. This could mean that fire engine A was either late, unavailable due to another incident, or unable to respond in time for other reasons.

Unfortunately, we cannot analyze the effect of dispatch order on response time because we are uncertain about the correctness of the information in the DateAndTimeMobilised, DateAndTimeMobile, and DateAndTimeArrived columns.

The column PlusCode_Code contains information about the type of dispatchment and has three unique values (Initial, rca, add) describing the dispatch code of fire engines (Initial Mobilization, Running Call, and Addition to first attendance). Since our purpose is to predict only the initial response time, we removed mobilizations that were not an initial dispatch, and we removed the columns PlusCode_Code and PlusCode_Description from the dataset.

The term "delay codes" refers to the reasons for delays in the arrival of fire engines, which can include factors such as adverse weather conditions, heavy traffic, or incomplete address information. The responding crew is allowed to enter a delay code for any vehicle that experienced an impeded response time, but they are specifically reminded to consider whether there was an avoidable delay if the first fire engine takes more than six minutes, or the second fire engine takes more than eight minutes to arrive at the incident location (LFB, London Datastore, 2023). This is considered a delay due to the 6 minute and 8 minute attendance standards in London from 2008.

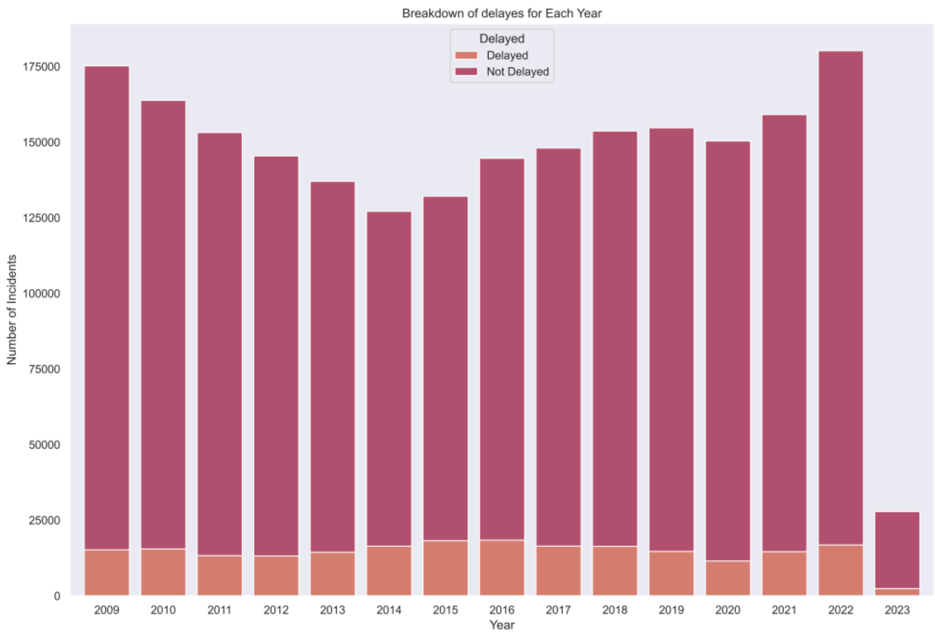


Figure 17

We decided to treat mobilizations with over 8 minutes and 20 seconds delays as outliers and removed them from the dataset, which amount to 8.5% of the dataset. In addition, we also eliminated the DelayCodeId and DelayCode_Description columns. The reason for this was twofold. Firstly, we were unable to find any datasets containing information such as weather conditions, traffic data or road network characteristics that could assist the model in predicting delays. Secondly, the columns that contained the reasons for delay were unreliable as they did not provide clear information about which mobilizations were delayed or not. After concluding an incident, firefighters were required to enter a delay reason into the system; if there was no delay, they could enter "Not held up". However, we found over 220,000 mobilizations with over an 8-minute delay that either had no reason entered or entered "Not held up" as a reason. Moreover, over 100,000 mobilizations with no delay reason entered a reason, even if they were not held up.

We first removed unnecessary columns and then utilized the fire stations dataset to filter out mobilizations dispatched to locations outside of London. We further filtered the fire stations dataset to only include active fire stations in London. We also limited our analysis to mobilizations present in the incidents dataset, as there were some dispatched outside of London. Since we will be using `TurnoutTimeSeconds` and `TravelTimeSeconds` to predict response time, we removed the `AttendanceTimeSeconds` column, which is the sum of both columns. We also eliminated any missing values from the dataset.

Since we are predicting the initial response time of the LFB, we filtered the dataset to retain only the fastest mobilization turnout and travel times for each incident.

5.2.2. Additional Features

It was determined that there is a correlation between the turnout and travel times and the fire stations. The incidents dataset already contain information about the fire station grounds. However, sometimes additional stations are called upon for serious incidents. To account for this, new features were generated in the dataset, with fire station as columns and a binary value (1/0) being assigned to indicate whether a fire engine was dispatched from that station to a specific incident.

As explained in the section regarding the LFB Emergency Response Process, the number of fire engines dispatched to an incident is determined by the LFB's pre-determined attendance (PDA), which considers the severity and type of the incident as well as the category of the property where the incident occurred. Each fire station is responsible for responding to incidents within its vicinity. In 73.6% of incidents, only the closest fire station responded. Due to this, applying one-hot-encoding to the fire station ground column would result in numerous duplicated values in the dataset. Instead, we used label encoding to encode the fire station grounds.

5.3. Final Dataset

Before merging the datasets, we calculated the aerial distances in meter between each fire station and each incident's location using the longitude and latitude of the fire station contained in the fire stations dataset and the longitude and latitude of the incidents using the haversine great circle distance.

To calculate the distance d in kilometers between two points A and B that have the latitudes ϕ_{A} , ϕ_B and longitudes λ_A, λ_B respectively, one can use the formula:

$$a = sin^2\left(\frac{\phi_B - \phi_A}{2}\right) + cos(\phi_A) \times cos(\phi_B) \times sin^2\left(\frac{\lambda_B - \lambda_A}{2}\right)$$

$$c = 2.\,arcsin\left(\sqrt{a}\right)$$

$$d = R \times c$$

Where R is the earth's mean Radius of 6,371 km. To convert d into meters, we simply multiplied by 1,000.

The earth is not a perfect sphere but rather an oblate spheroid and these distances are not accurate. However, we are calculating the distances only in a relatively small area (all in London) compared to the surface of the earth, which means that the error is negligible.

The distances calculated between fire stations and incident locations do not accurately reflect the actual travel distance, unless aerial appliances are being used by the LFB. Although the LFB does have access to aerial appliances, the datasets used in this project only include records of fire engines that were dispatched on land.

To obtain a more accurate measure of the travel distance between fire stations and incident locations, we utilized the OSMNX python package, which provides access to geospatial data from OpenStreetMap for analyzing real-world street networks. Our focus was on the Greater London area, and we obtained the network graph for this region, limiting our analysis to the drive street network. We then added speeds and travel times to the edges of the network, allowing us to estimate travel time between nodes. To evaluate the network's performance, we calculated the length of the shortest path between 5 incident locations and all fire stations in our dataset using the NetworkX package.

On average, it took us 84 seconds to calculate the actual travel distances between a single incident location and all fire stations. We then converted the graph to an igraph which was expected to be faster than using the OSMNX graph. However, even with this optimization, it still took an average of 82 seconds to perform the same calculations. To further improve our speed, we can use the cuDF and cuGraph modules of the RAPIDS library which provide GPU-accelerated (graphics processing unit) graph analytics using Nvidia's CUDA, a general-purpose computing on GPUs. This solution is reportedly 1000 times faster than using NetworkX. However, using this approach would still take a long time because of the amount of data that we have in our dataset. Therefore, we decided to use the great circle distances instead of the actual travel distances due to the time constraint.

Following the distance calculations, we combined the incidents and mobilizations datasets using the incident number. Then, we merged this result with the new features that indicated which fire station dispatched a fire engine to an incident using the incident number columns and we removed the incident number column from the dataset. We also included the great circle distances that we had computed to the dataset.

The final dataset now contains 1,279,831 rows representing individual incidents that the LFB responded to and occurred in the Greater London area and had a response time less than 8 minutes and 20 seconds. The dataset contains 327 features column:

- 1 Column represents the incident Station Ground.
- 15 columns representing the year (2009 to 2023).
- 24 columns representing the hour of day.
- 7 columns representing the day of week.
- 12 columns representing the month.
- 33 columns representing the boroughs.
- 9 columns representing the category of the property.
- 112 columns representing the distances between fire stations and incidents' locations.
- 112 columns representing which fire station responded to the incident.
- 2 column containing turnout time and travel time respectively.

Now we can now use this dataset to train a model to predict the response time of the LFB.

6. Modelling

The main goal of our project is to accurately predict the response time of the LFB. We have a preprocessed dataset containing turnout time and travel time, and we want to predict their sum.

6.1. Model Selection and Evaluation

Since the turnout time is independent of the distance between fire stations and incident locations, we have two options: either train separate models to predict the turnout and travel times and then sum the predicted values or train a single model to directly predict the response time.

To explore both options, we first split the dataset using the `sklearn.model_selection.train_test_split` function into four subsets: explanatory train and test sets, and target train and test sets for each of the target variables (turnout, travel, and attendance). To use the turnout time as a target variable, we removed the columns that contained distances between fire stations and incident locations before splitting the dataset. This gave us a total of 12 datasets. We then scaled the datasets using a standard scaler to ensure that the models' results are not affected by variations in the data.

To predict the response time, we selected four linear models (Linear, Lasso, Ridge, and Elastic Net regression) and three non-linear models (Multilayer Perceptron, XGBoost Regressor, and Histogram-based Gradient Boosting Regression Tree) and trained them to predict the turnout time and travel time individually. We then calculated the models' predictions for both the training and test sets, summed the results, and calculated the root mean square error (RMSE) between the predictions and actual attendance time. After comparing the RMSE scores of the models on the test and train sets, we chose the model with the lowest RMSE score. Since we are only interested in predicting the response time, we will only present the results for predicting the sum of the turnout and travel times. The scores are presented in Table 3.

Model	**Test RMSE**	**Train RMSE**
Linear Regression	85.07626	84.9788
Lasso	86.6771	86.6001
Ridge	85.0759	84.9788
Elastic Net	86.7422	86.6644
Hist Gradient Boosting Regressor	74.9804	74.7885
Multi-Layer Perceptron	68.1043	66.7788
XGBoost Regressor	65.2970	62.2773

Table 3 Train and test results using the sum of turnout and travel times

Next, we trained the same models on the entire dataset, including all the columns, to predict the target variable of attendance time and assessed whether there was any enhancement in the model's performance. The results are presented in Table 4.

Model	Test RMSE	Train RMSE
Linear Regression	85.0734	84.9735
Lasso	86.6663	86.5910
Ridge	85.0732	84.9735
Elastic Net	86.7386	86.6626
Hist Gradient Boosting Regressor	74.9311	74.7204
Multi-Layer Perceptron	67.8509	66.6084
FCDNN	73.8610	69.7786
XGBoost Regressor	**64.9995**	**61.9085**

Table 4 Train and test results using attendance time

After training the models using all the columns of the dataset to predict the attendance time, we observed an improvement in their performance. Hence, we have decided to use this approach to estimate the response time.

Among all the tested regressors, we have identified that XGBoost outperforms other methods for predicting the response time since it exhibits the lowest RMSE.

Once we selected the model, we conducted several rounds of hyperparameter optimization and attained a **train RMSE of 58.6768** and a **test RMSE of 64.4578**. To illustrate the performance of the model, we randomly selected 70 instances from both the train and test sets and plotted them to visualize the degree of closeness between the actual and predicted values. We also used a scatter plot to compare the distances between the actual and predicted values. The visualization results are shown in Figure 18.

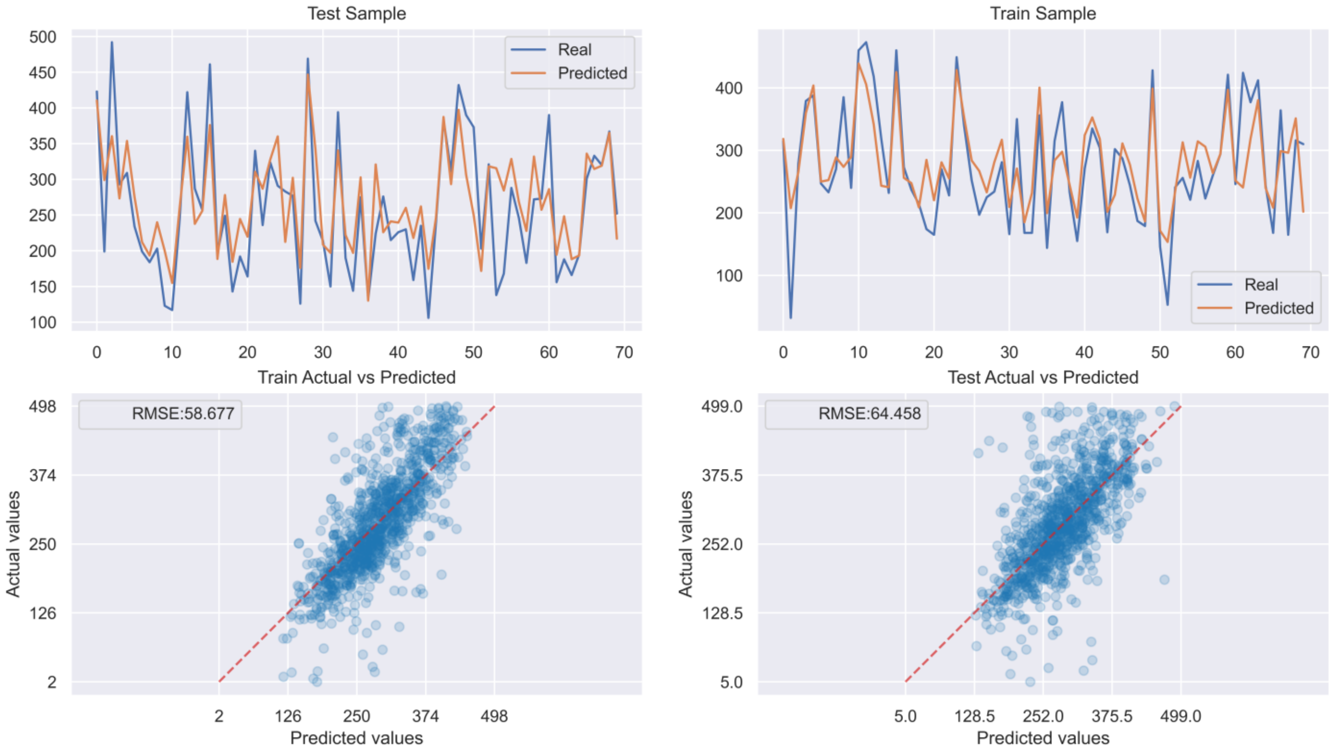


Figure 18

In an attempt to enhance the model's performance, we utilized cyclical encoding to encode the time-related features which were initially encoded using one-hot encoding. Since time features are naturally cyclical, employing cyclical encoding is essential to convey to the machine learning model that the feature exhibits cyclical patterns and to enable it to comprehend trends. We established two columns for each of the time of day, day of the week, and month features, which were calculated using the following formulas:

$$s = sin(\frac{x \times 2\pi}{p}), c = cos(\frac{x \times 2\pi}{p})$$

Where x is the value to be encoded and p is the period, which is 7 days for days of the week, 24 hours for hours, and 12 months for months. E.g., in Figure 19 we visualized the values of the hours columns after using cyclical encoding. The plot shows that the values of sin and cos oscillate smoothly between -1 and 1, forming a plot that resembles a clock face. This indicates that the cyclical encoding successfully captured the cyclical nature of the hours feature, which can now be used as input for the XGB model.

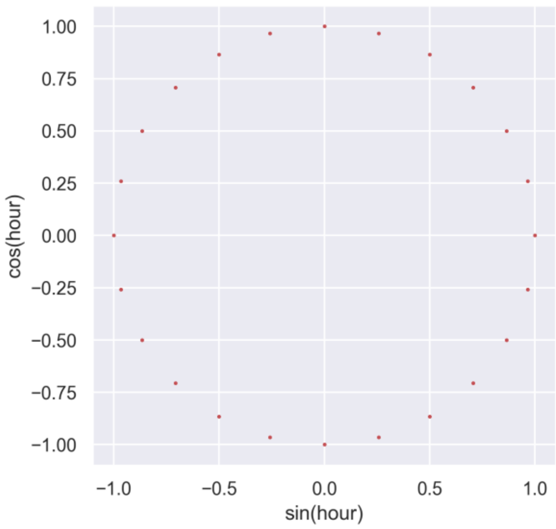


Figure 19

We trained the same XGBoost model after using cyclical encoding and obtained better results: **RMSE train: 58.5114**, **RMSE test: 64.352**. The training and test results are shown in Figure 20.

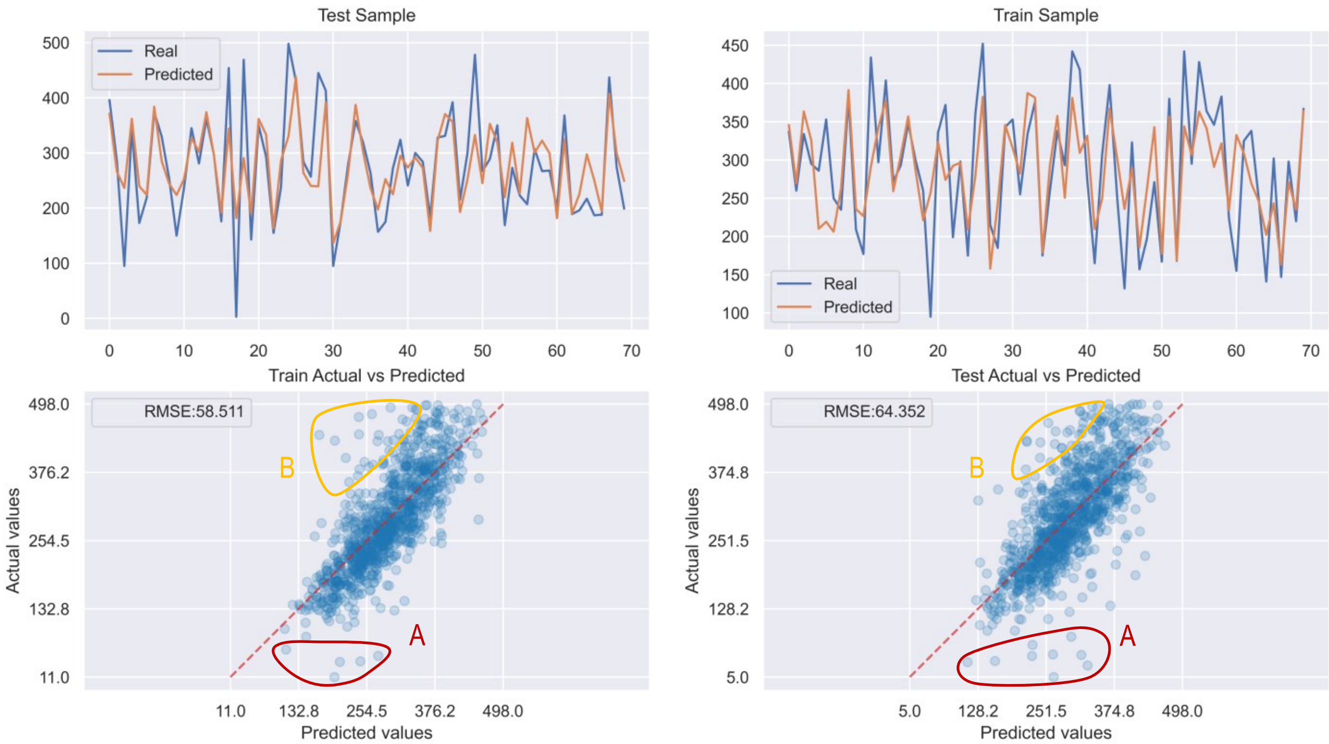


Figure 20

Overall, the model's performance is satisfactory. However, upon analyzing the results shown in Figure 20, we observed the formation of two distinct clusters (labeled as A and B) where the model seems to struggle in predicting the response times accurately. In cluster A, there were incidents with an attendance time of less than 20 seconds, which is implausible since attendance time is the sum of travel time and turnout time. On the other hand, cluster B consists of response times that exceed the LFB's standard of 6 minutes for the first fire engine, and the model performed poorly in predicting them. This outcome may be attributed to various factors:

- The incident location may be very close to the fire station resulting in low travel time.

- The fire engine is on the move and already in the vicinity of the incident resulting in low travel time and no turnout time.
- The LFB has received advance notice of the incident and prepared in advance, which we assume to be unlikely.
- There are issues regarding the quality of the data.

To investigate these anomalies, we started by analyzing the distributions of the turnout time, travel time and their sum, which is shown in Figure 21.

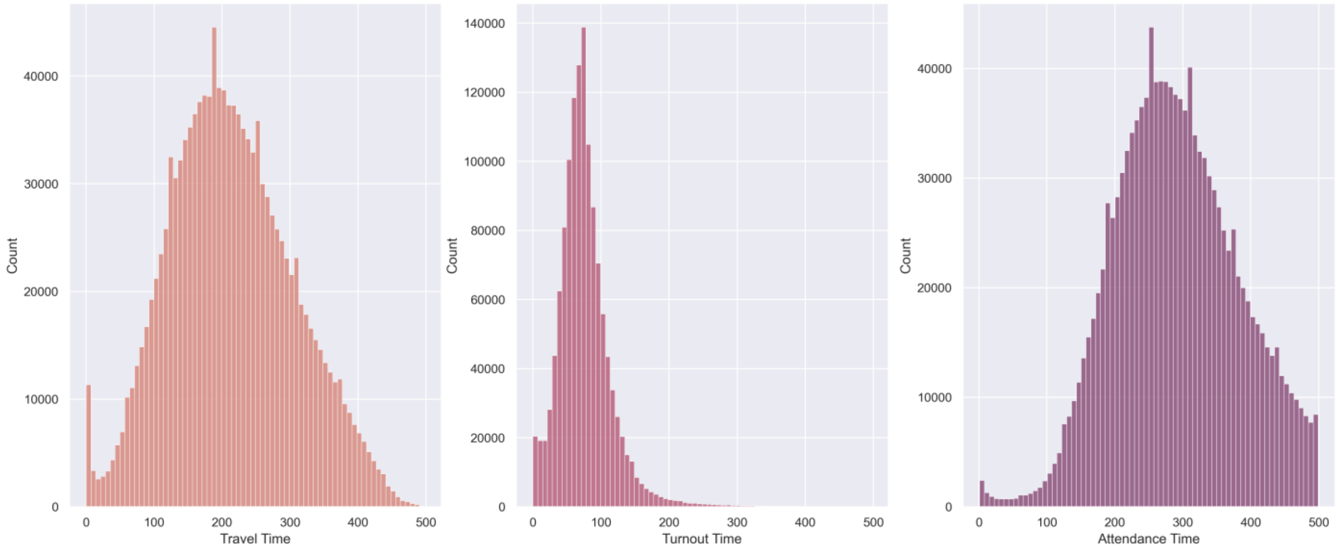


Figure 21

The distribution of the target variable poses a challenge for the model to identify a meaningful pattern. Some incidents have a turnout time of nearly 500 seconds and others have a travel time of nearly 500 seconds, making it difficult to determine which data points are erroneous. As we lack information about the data collection process for the turnout and travel times, it is difficult to ascertain the accuracy of each data point.

Turnout Time	Travel Time
1.0	498.0
1.0	497.0
1.0	497.0
4.0	495.0
5.0	494.0

Turnout Time	Travel Time
497.0	2.0
495.0	2.0
495.0	3.0
495.0	1.0
494.0	1.0

Turnout Time	Travel Time
0.0	2.0
0.0	3.0
0.0	22.0
0.0	32.0
0.0	36.0

To better understand the anomalies in the dataset, we sorted and printed out the turnout and travel times. The left table displays the travel time in descending order and reveals that some incidents had a turnout time as low as 1 second, while the travel times were almost 500 seconds. In the middle table, we observe the exact opposite - incidents with very short travel times but with very long turnout times. Finally, the right table shows the sorted turnout and travel times, and we can see that multiple incidents had no turnout time at all and very low travel times. As our dataset lacks information about the geographical location of fire engines during an incident, some of the data points that show low turnout time or no turnout time could be because the fire engines were already on the road responding to other incidents or a fire crew had just arrived at the fire station and are ready to be dispatched. Alternatively, these anomalous data points could be attributed to either human or system error while entering the data into the system. In order to eliminate the risk of having potentially anomalous data in our dataset and to reduce the model's RMSE, we can simply remove the data with both low and high turnout and travel times from the dataset.

To determine the optimal range for the turnout and travel times, we employed the Bayesian Optimization algorithm. This approach offers advantages over manual search methods like grid or random search, as it is more accurate and converges faster (Wu et al., 2019).

Bayesian optimization is an algorithmic approach used for finding the optimal solution of an expensive noisy black-box function with a limited number of evaluations. In the case of hyperparameters optimization, the algorithm maintains a probabilistic model, typically a Gaussian Process (GP), of the function mapping from hyperparameter values to an objective $f: \mathrm{R}^{j} \mapsto \mathrm{R}$, where j is the number of hyperparameters, to capture the underlying unknown function f.

The hyperparameters being optimized in our case are the minimum and maximum turnout times, as well as the minimum and maximum travel times. The optimizer selects values for minimum turnout, minimum travel, maximum turnout, and maximum travel, and filters the dataset accordingly. This involves selecting data points with `TurnoutTimeSeconds` between the minimum turnout and maximum turnout times, as well as `TravelTimeSeconds` between minimum travel and maximum travel times. Subsequently, the dataset is split into training, validation, and testing sets with the target variable as the sum of turnout and travel times. An XGBoost regression model is trained on the training data. The objective function f, which is to be maximized, is defined as the negative root mean squared error (RMSE) between the predicted values of the trained model and the actual values on the test data.

Throughout the optimization process, the optimizer gathers observations to gain insight into the function f and determine the potential position of its maximum. It aims to strike a balance between exploration (to estimate a prior distribution), involving hyperparameters with uncertain outcomes, and exploitation, leveraging knowledge about hyperparameters with more definitive outcomes. An algorithmic representation of our Bayesian Optimization process:

Algorithm: Bayesian Optimization

Input: *Objective function* $f(min_{turnout}, min_{travel}, max_{trunout}, max_{travel})$
Acquisition function $A(min_{turnout}, min_{travel}, max_{trunout}, max_{travel})$**,**
Exploration iterations E = 30**,** *Maximum iterations* T = 80

Output: *The best observed solution:* $min_{turnout}, min_{travel}, max_{trunout}, max_{travel}$

Initialize: ***Set*** t = 0 *(iteration counter)*

Search Range: $min_{turnout} \in [20,50]$, $max_{trunout} \in [150,250]$,
$min_{travel} \in [20,100]$, $max_{travel} \in [350,450]$

GP prior: *Randomly select* $min_{turnout}, min_{travel}, max_{trunout}, max_{travel}$
$y_{best} = f(min_{turnout}, min_{travel}, max_{trunout}, max_{travel})$
$X = \{min_{turnout}, min_{travel}, max_{trunout}, max_{travel}\}$
$Y = \{y_{best}\}$

While (t < T) ***do***
  if (t < E) ***then*** *//Exploration*
    Randomly select $min_{turnout}, min_{travel}, max_{trunout}, max_{travel}$ *values from the search range*
  else *//Exploitation using the current GP model*
    $min_{turnout}, min_{travel}, max_{trunout}, max_{travel} = argmax\ A(min_{turnout}, min_{travel}, max_{trunout}, max_{travel})$
  $y_{next} = f(min_{turnout}, min_{travel}, max_{trunout}, max_{travel})$ *//Evaluate the objective function*
  $y_{best} = max(y_{best}, y_{next})$ *//Update the best observed value*
  $X = X \cup \{min_{turnout}, min_{travel}, max_{trunout}, max_{travel}\}$ *//Update the GP model*
  $Y = Y \cup \{y_{best}\}$
  Update the GP posterior using X and Y
  t = t + 1 *//Increment the iteration counter*
end

The optimizer found the optimal value for the objective function at max_travel: 350.0, max_turnout: 150.0, min_travel: 100.0, min_turnout: 41.69. The distribution of the new dataset is shown in Figure 22.

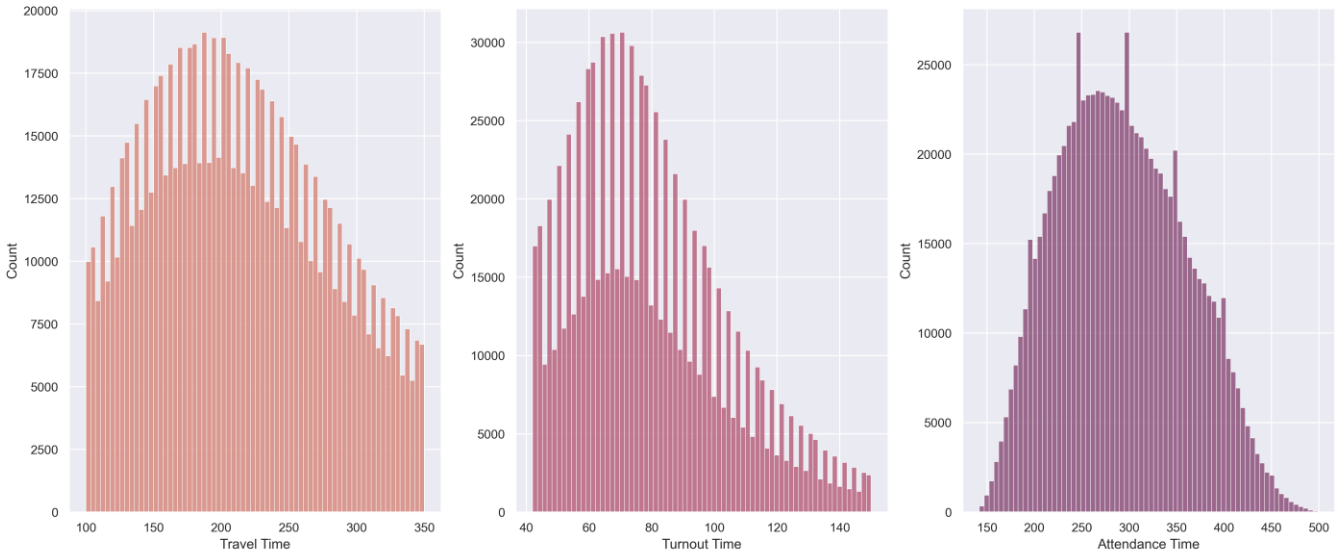


Figure 22

The distribution of the target variable looks normally distributed. Using the new filtered data, the XGBoost regressor performs a lot better than previous models: **RMSE train: 41.784**, **RMSE test: 48.035**. The training and test results are shown in Figure 23.

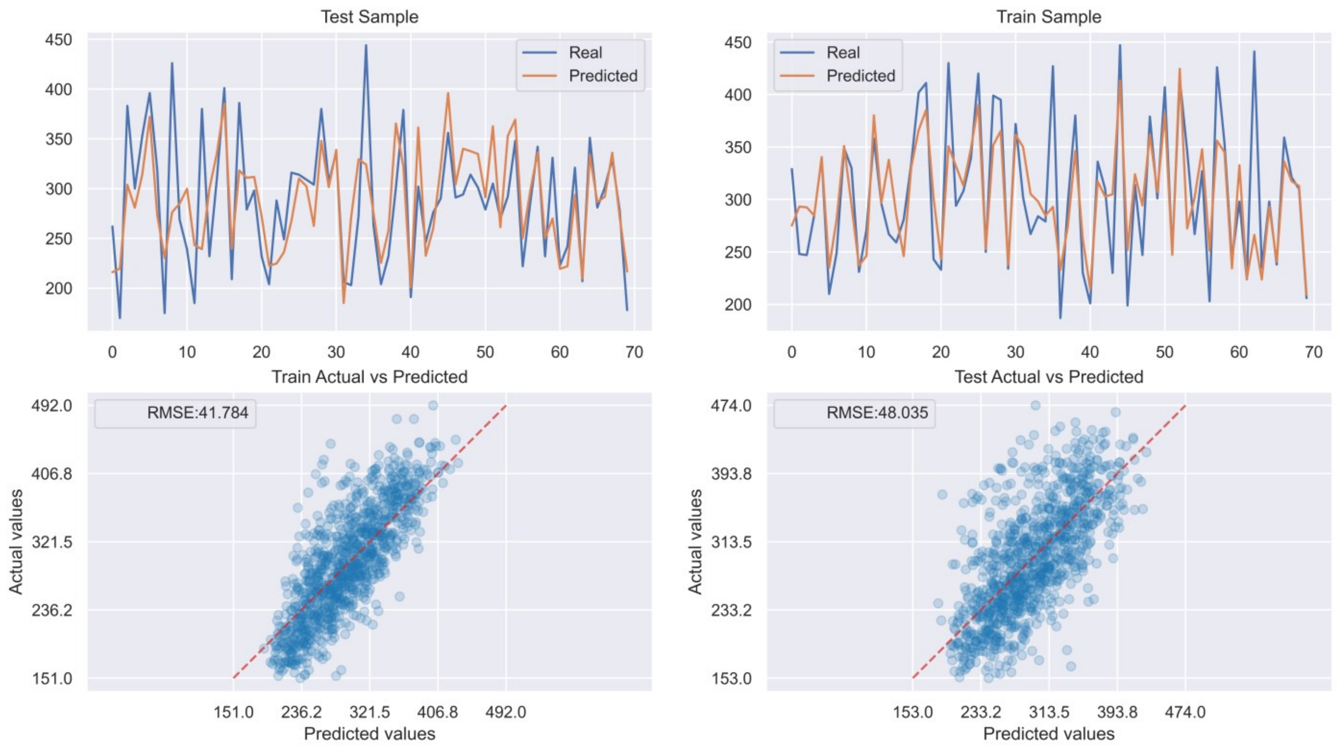


Figure 23

Despite our efforts to improve the model's performance, it still falls short in accurately predicting low or high response times. One of the key factors contributing to this challenge is the imbalance in the distribution of the target variable, which creates difficulties for the model to generalize effectively.

In our filtered dataset, the occurrences of extreme response times, both low and high, are relatively rare compared to the majority of moderate response times. This imbalance in the distribution poses a significant challenge for the model, as it tends to be biased towards predicting response times that are more prevalent in the data.

In an attempt to mitigate the impact of imbalanced distribution on the model's performance, we explored the use of a resampling technique called Synthetic Minority Over-Sampling Technique for Regression with Gaussian Noise (SMOGN). This technique, similar to the SMOTER method used in classification problems, aims to create a more balanced training dataset for regression tasks (Branco, 2017).

However, we encountered practical limitations in applying the SMOGN method to our dataset. Due to the extensive computational resources required for the resampling process, it was not feasible to implement this technique within the constraints of our available computational power. As a result, we decided not to pursue the resampling approach in this project.

6.2. Model Interpretability

In XGBoost (and tree-based models in general), the `.get_score()` method is used to retrieve the feature importance scores after training a model. The `importance_type` attribute specifies the type of importance score to be computed. Here's an explanation of each importance type and the importance scores of our final model after combining features that have similar characteristics (temporal features, borough features and distances):

Weight: The "weight" importance type corresponds to the number of times a feature is used to split the data across all trees in the ensemble. It is a measure of the relative number of times a feature appears in the trees. In Figure 24 we can see that the temporal features appear the most in the trees and are considered the most important features.

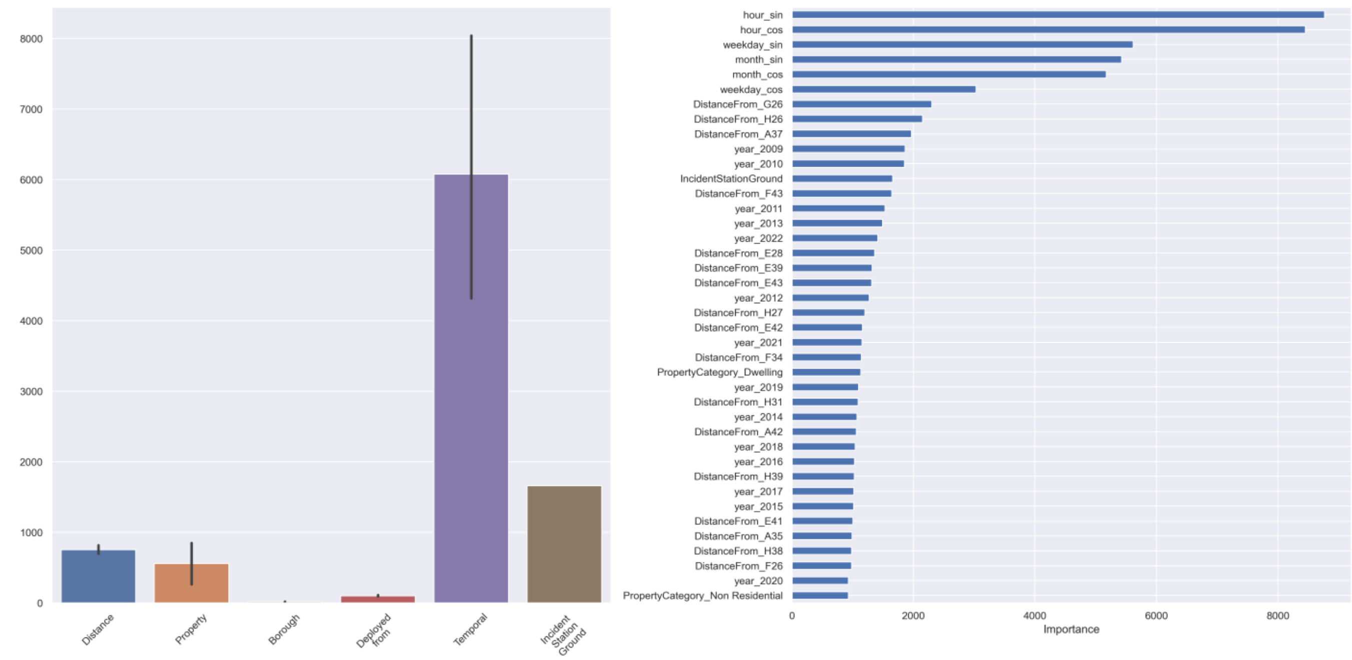


Figure 24

Gain: The "gain" importance type measures the improvement in the model's performance (the reduction in the loss function) contributed by a particular feature when it is used for splitting. The gain is calculated by summing up the improvement over all splits where the feature is involved. Our model shows that the borough features and the distances between fire stations and incidents locations contribute the most to the model's predictive power as seen in Figure 25. The black line in the middle of the bars corresponds to the 95% confidence interval, i.e., the borough features gain has a high standard deviation. The reason is that the response times are not consistent over all boroughs in London. Some boroughs have higher importance gain than others. For example, the Southwark borough (E09000028) is a borough in inner London, where the number of incidents is high and the response times are low.

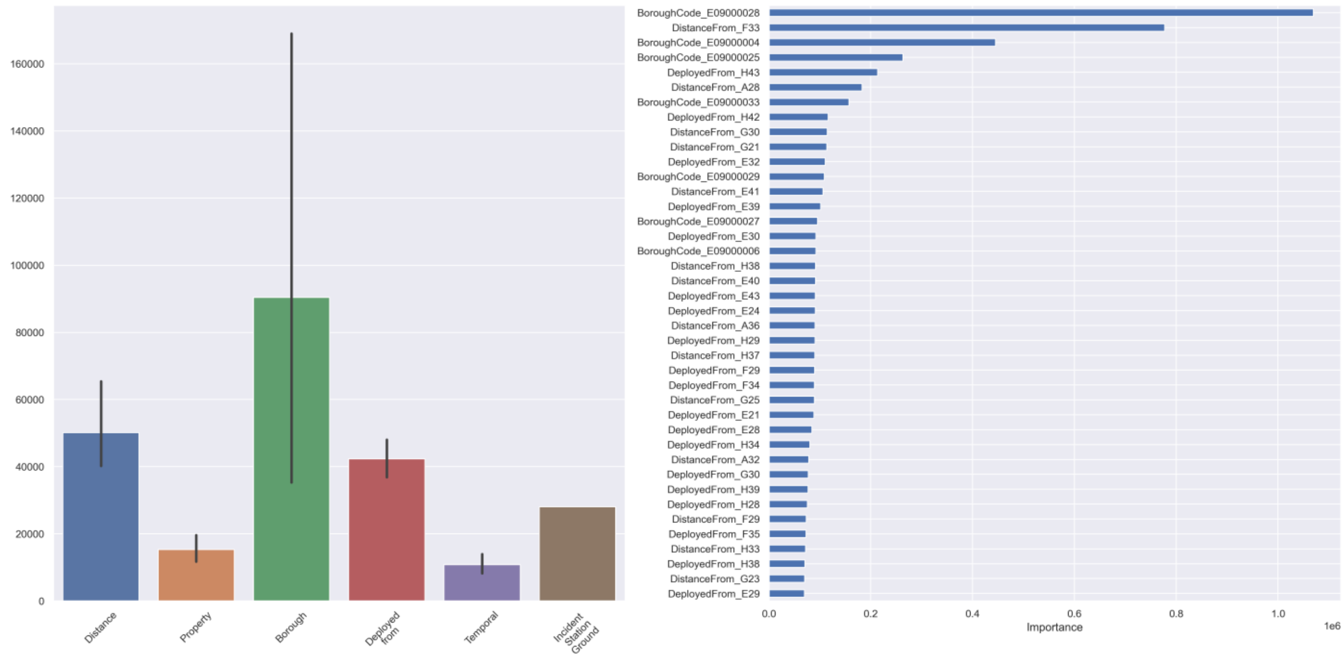


Figure 25

Total Gain: The "total_gain" importance type is similar to "gain" but represents the total gain of a feature across all the trees in the ensemble. It is calculated as the sum of the gain values over all splits where the feature is involved, without considering the individual gain contributions from different splits. In Figure 26 we see that the Temporal features have the highest total gain value compared to other features implying that they are more important for generating more accurate response time predictions. However, the high standard deviation of the importance of temporal features suggests that not all temporal features are equally important. Indeed, the hours features are more important than the month or weekday features. The distance between the Whitechapel Fire Station (F33) and incident's locations has the highest total gain importance.

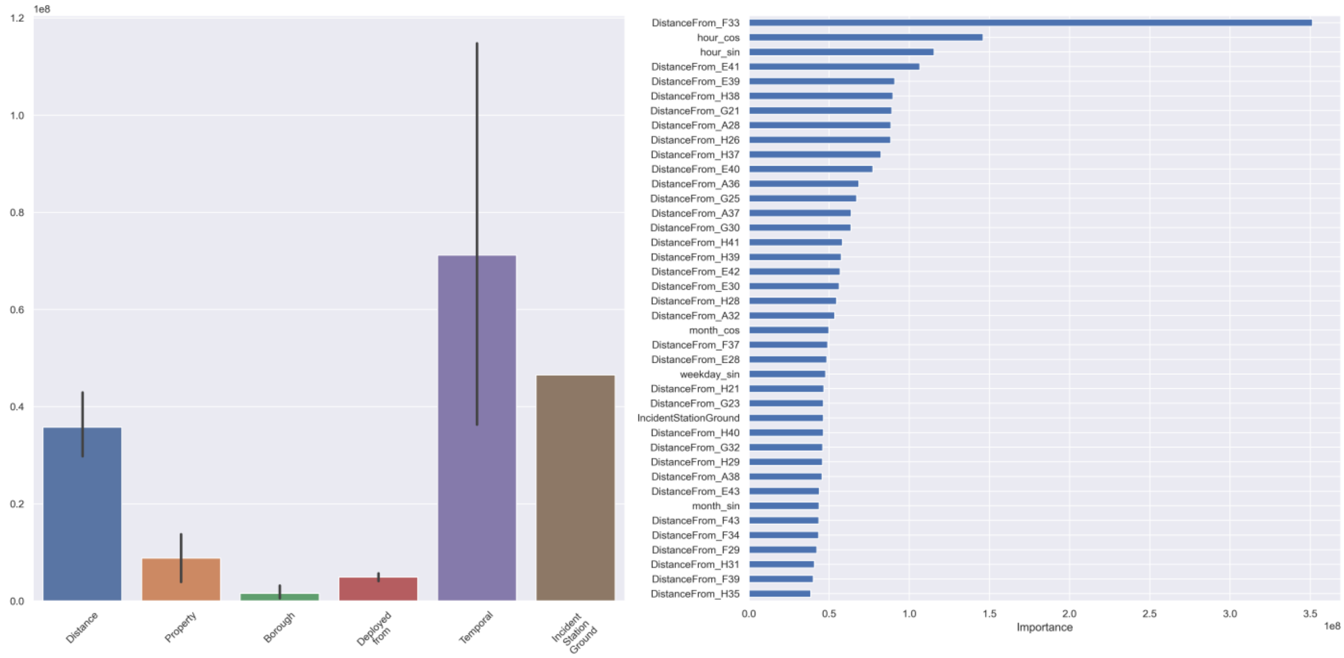


Figure 26

Cover: The "cover" importance type represents the average coverage of a feature throughout all the trees in the ensemble. It is calculated as the sum of the number of times a feature appears in the splits weighted by the number of training samples that go through those splits. Features with higher cover values are considered more important as they affect a larger portion of the training data.

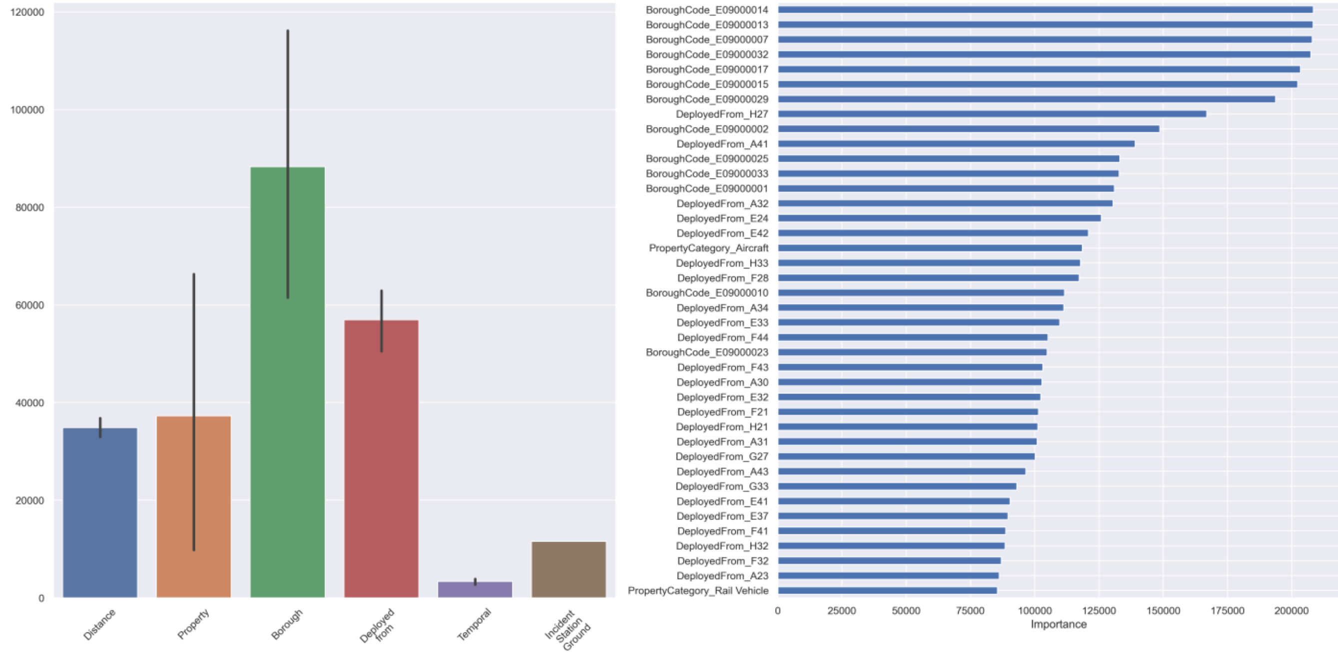


Figure 27

Total Cover: The "total_cover" importance type is similar to "cover," but it represents the total coverage of a feature across all the trees in the ensemble. It is calculated as the sum of the number of times a feature appears in the splits, without considering the weights of the splits or the number of training samples passing through them. In Figure 28 we can see that the distance features have the highest total cover importance followed by the temporal features.

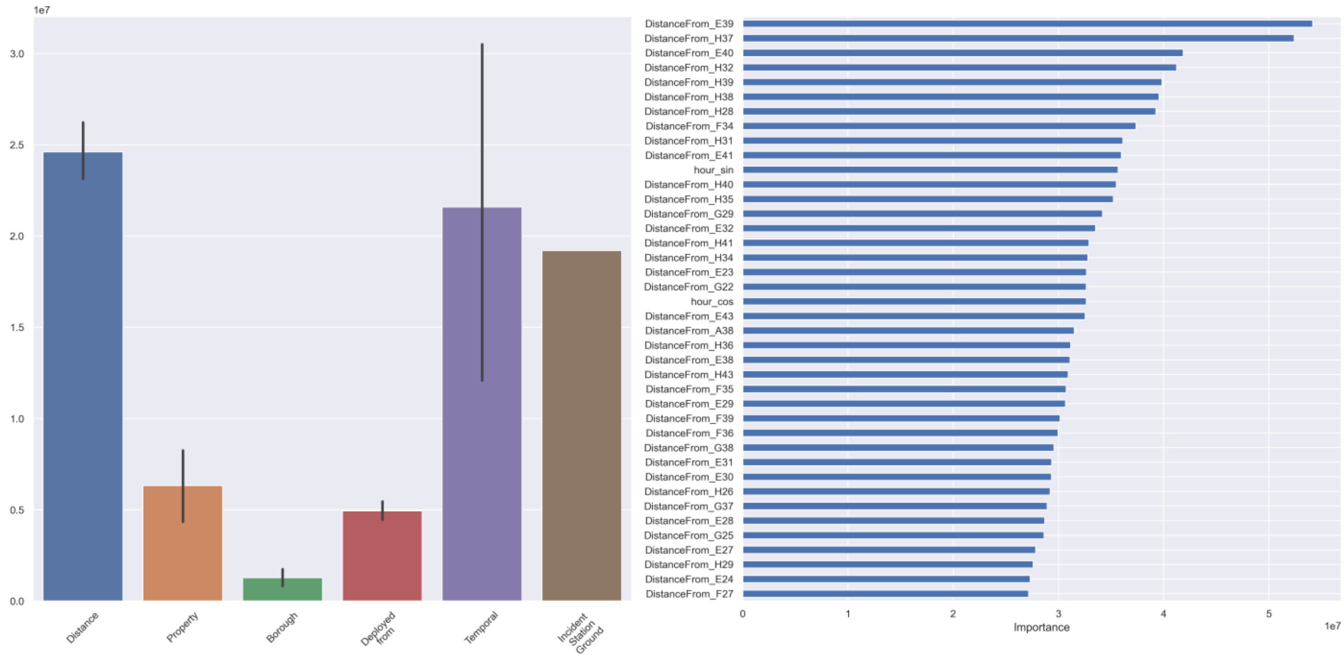


Figure 28

7. Conclusion and Final Remarks

Our project aimed to predict the response time of the London Fire Brigade by analyzing various factors that influence the turnout and travel times of the fire engines. We employed machine learning techniques, particularly XGBoost regression, to develop a predictive model.

During the course of this project, we devoted significant effort to data preprocessing and feature engineering. Our examination of the datasets provided by the LFB led us to the realization that many of the feature columns contained limited or unreliable information. For example, certain columns such as those pertaining to the type of the incidents lacked meaningful data, while others like delay codes were found to be inaccurate. Additionally, we identified errors in certain columns, such as mobilization times, further diminishing their utility for our analysis.

To address the missing Longitude and Latitude values, we employed the Ordnance Survey Great Britain (OSGB) coordinates and computed new feature columns that captured the travel distances between fire stations and incident locations. However, due to computational constraints, we were only able to calculate the Haversine distances, which provide an estimate rather than the actual travel distances.

We explored different approaches, including training separate models for turnout time and travel time, as well as training a single model to directly predict the response time. After thorough evaluation, we identified XGBoost as the best-performing model, exhibiting superior performance in terms of root mean squared error (RMSE) compared to other regression methods.

We observed two clusters of anomalies in the predicted response times. In the first cluster the model failed to predict unusually low attendance times, which could be attributed to fire engines already in motion or recent arrivals at fire stations; the other cluster displayed response times exceeding the standard 6-minute threshold set by the LFB for the first fire engine. The model struggled to accurately predict response times within these clusters, likely due to imbalances and anomalies present in the dataset.

To address these challenges, we employed Bayesian optimization to identify optimal ranges for the turnout and travel times. This optimization technique allowed us to refine our model by filtering the dataset based on specific ranges and train the model using these filtered data subsets.

However, despite our efforts, the model continued to exhibit suboptimal performance in predicting low or high response times. We attributed this difficulty to the inherent imbalance in the distribution of the target variable and the challenges associated with accurately estimating extreme response times.

In terms of feature importance, our analysis revealed the significance of temporal variables, borough of the incident and travel distances in the model's decision-making process. The "Weight" metric highlighted the influence of the temporal variables while the “Gain” metric showed that the borough where incidents occurred contributed significantly to the model’s predictive power.

While our model has shown promising performance, it is crucial to consider its predictions in conjunction with other evaluation metrics and beyond numerical measures alone. Interpreting the model's outputs and assessing its effectiveness should involve a holistic approach. that incorporating the OSMNX library into our model, which enables the calculation of precise travel distances and estimated travel times, would greatly enhance its performance. By considering additional factors such as weather conditions, traffic congestions, and collisions that can influence travel times and distances, we can provide more accurate predictions. This integration would enable our model to better account for real-world complexities and improve its ability to estimate response times effectively.

To improve the accuracy of our model, it is essential to conduct additional research and make necessary adjustments, focusing particularly on addressing the difficulties encountered when predicting response times in anomalous clusters. Furthermore, the LFB should strive to enhance its data collection procedures to minimize the presence of erroneous data. Including additional information in the datasets, such as distinguishing between call incidents and automated alarms or the actual geographical locations of fire engines at the time of the call, would also be beneficial. These improvements would contribute to a more robust and reliable prediction model for estimating response times accurately.

Despite these challenges, our project provides valuable insights into predicting LFB response times and highlights areas for further research and improvement. The accurate estimation of response times can have a significant impact on emergency services and help optimize resource allocation and emergency response planning.

8. Bibliography

Branco, P. T. (2017). SMOGN: A Pre-Processing Approach for Imbalanced Regression. *Proceedings of Machine Learning Research, 74*, 36-50.

FBU. (2023). *Fire Brigade Union*. Retrieved from FBU: https://www.fbu.org.uk/sites/default/files/publications/A%20Short%20Introduction%20to%20Firefighting%20in%20London.pdf

LFB. (2023, March 11). *London Datastore*. Retrieved from London Fire Brigade: https://data.london.gov.uk/dataset/incident-response-times-fire-facts

LFB. (2023, March 16). *London Datastore*. Retrieved from London Datastore: https://data.london.gov.uk/dataset/shut-in-lift-releases--lift-entrapments--attended-by-lfb

LFB. (2023, February 01). *London Fire Brigade*. Retrieved from Freedom of Information request reference number: 7125.1: https://www.london-fire.gov.uk/media/7349/71251_response.pdf

ons. (2023, February 27). *ONS*. Retrieved from Office for National Statistics: https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/changestolondonpostcodessince2000

Ordnance-Survey. (2023, 04 15). *Ordnance-Survey*. Retrieved from Ordnance Survey: https://www.ordnancesurvey.co.uk/documents/resources/guide-coordinate-systems-great-britain.pdf

UK-Government. (2000, November 30). *legislation.gov.uk*. Retrieved from legislation.gov.uk: https://www.legislation.gov.uk/ukpga/2000/36/contents