Firefighters Fight What?

An Analysis of San Francisco Fire Department

Apple Li, Derek Fu, Kai Hung, Lucia Lu

04/25/2021

Contents

Introduction

Purpose and Questions

To a typical resident of the United States, the Fire Department is often viewed as a government faction primarily responsible for dealing with structural fire, wildfires, fire rescue, and other emergency situations with the word “fire” attached. However, what do the Fire Department really deal with on a daily basis? How rapidly do they respond to various types of emergencies? In addition, the Police were subjected to various allegations regarding discrimination against racial minorities in the past year. Does our nation’s very own Fire Department carry an underlying level of discrimination in their response rate for high minority populated regions?

Main Dataset: Fire Department Calls for Service

Our main dataset is the Fire Department Calls for Service from the official San Francisco government open source data website (link: https://data.sfgov.org/Public-Safety/Fire-Department-Calls-for-Service/nuek-vuh3). The dataset contains all fire unit response to calls in the San Francisco area from the year 2000 to 2018, amounting to over 5 million calls. For each call, the dataset records

- CallType - The type of call the incident falls under
- ReceivedDtTm - Date and time call received at the 911 Dispatch Center
- ResponseDtTm - Date and time this unit acknowledges the dispatch and records that the unit is en route to the location of the call.
- Priority - Classification of the Call (2 if non-emergency, 3 if emergency)
- CallTypeGroup - Classification into four main groups: Fire, Alarm, Potential Life Threatening and Non Life Threatening.
- UnitType - Classification for the unit responding to the incident
- City - City in which the incident is reported
- Zip Code of Incident - Zip code in which the incident is reported

Auxillary Dataset: Racial Distribution by Zip Code in San Francisco

We obtained this data using web scraping from the website ZipDataMaps (link: https://www.zipdatamaps.com/). The collected data is then used to build a data frame with 27 rows and 10 variables. Each row represents the racial distribution in a distinct zip code area. All racial distribution data are from 2014. The variables are zip code and the percentage of each ethnicity group within a region. We computed our own variable “minority” by aggregating all non-white ethnicity group percentages.

Research Direction

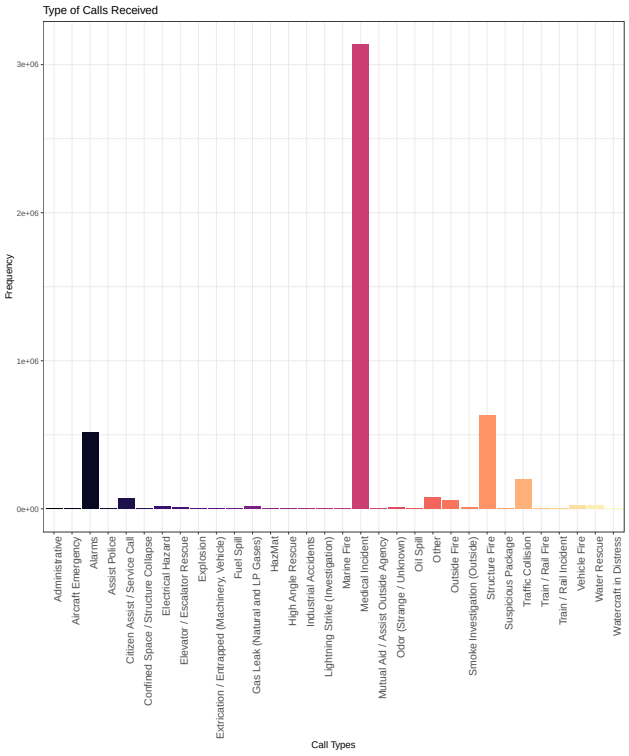
In our analysis, we hope to identify the incidents that Fire Departments actually deal with most frequently, explore whether the pattern of incidents differ by month or location, and explore potential latency in areas populated mostly by underrepresented minority groups.

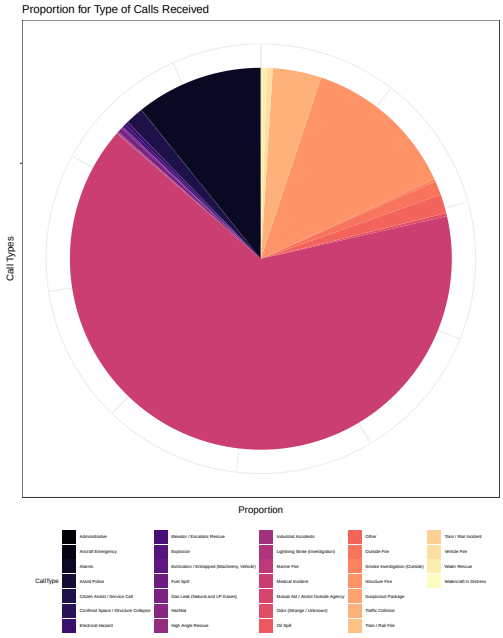
Data Investigation and Visualization

Call Type Analysis

Distribution of Call Types

A common myth for the fire department is that they primarily deal with fire incidents since the general public's impression with the fire department lies in their reputation as "firefighters". With this analysis, we hope to verify whether the myth is true and obtain a better picture for the types of calls the Fire Departments actually respond to.

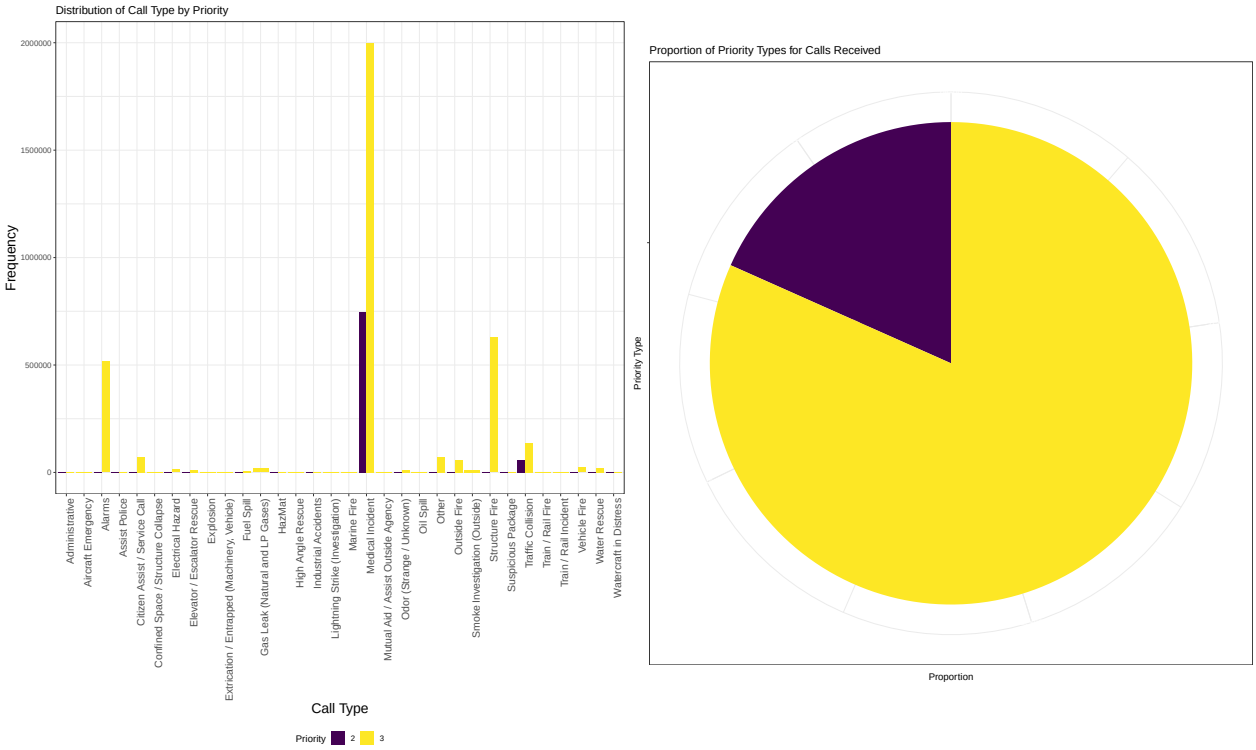




In order to examine the distribution of the various call types present in the dataset, we decided to plot the distribution of the number of calls by their respective call type category. From the bar plot on the left, we discovered that a majority of the calls received by the San Francisco Fire Department fall under the "medical incidents" category. This is an unexpected result which contradicts the usual perception that fire departments primarily deal with fire incidents. Furthermore, the bar plot shows that the distribution of the call types are relatively sparse. The most frequent call types are "medical incident", "structure fire", "alarms", and "traffic collision". The pie chart to the right provides a better representation for the proportion of each of these call types relative to all call types made to the San Francisco Fire Department.

Distribution of Priority Type

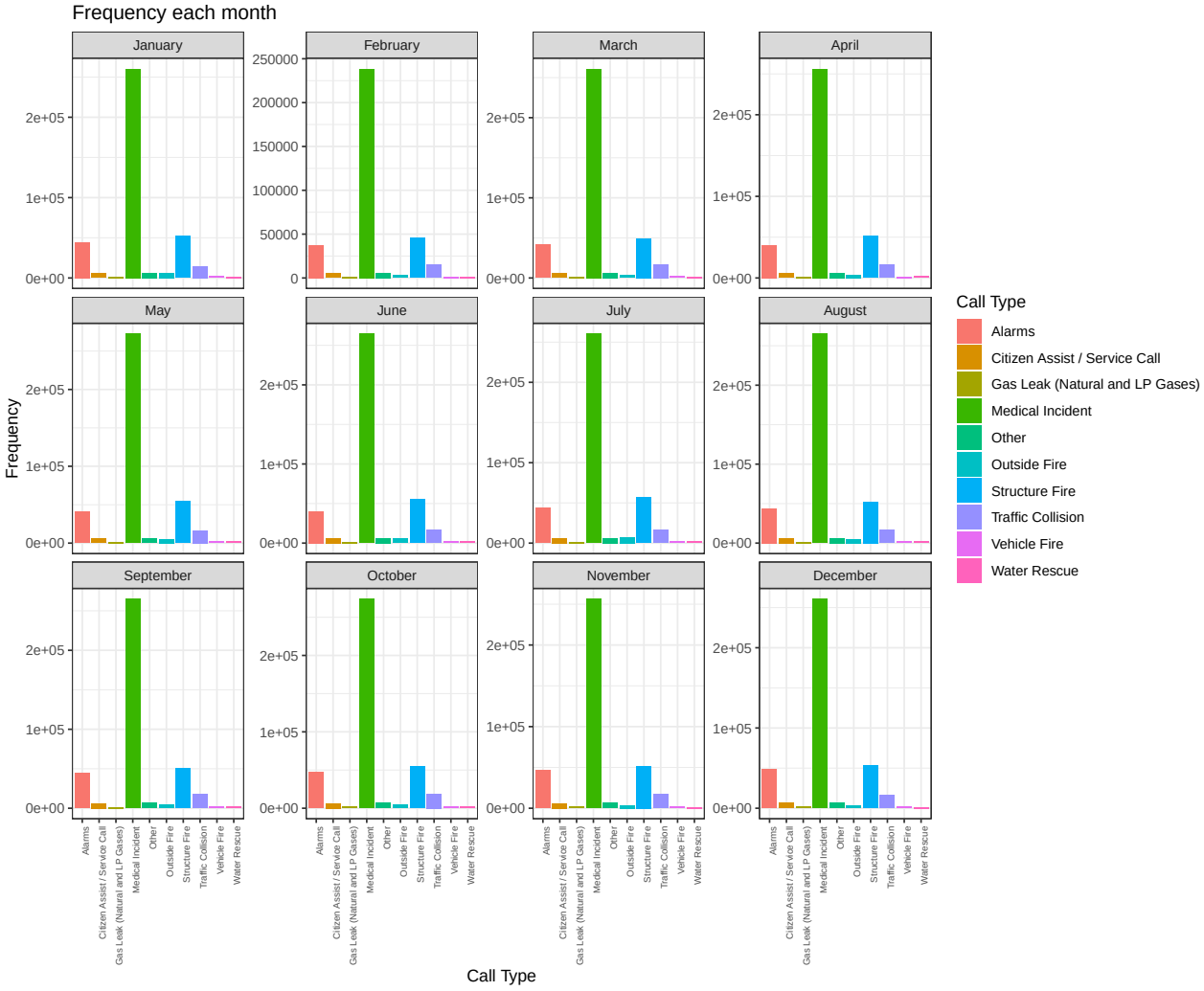
Dialing the emergency number is often seen as a last resort by common citizens since we are often taught at a young age to dial the number when you are in real emergencies and, often, there are repercussions for those who abuse the system. However, how do the fire departments decide what to prioritize? How often are calls classified as emergencies?



Based on the data dictionary provided, the San Francisco Fire Department classifies each call with a specific priority type: 2 for non-emergency and 3 for emergency. To further understand the proportion of each priority type, we decided to create a pie chart for the priority type distributions. From the pie chart to the right, we see that a majority of the call types are labeled as emergencies. To further dissect the distribution of priority type for each call type group, we plotted a bar graph split by priority type and call type. From the bar plot to the left, we see that most calls are classified as emergency. For medical emergencies, which accounts for the largest number of calls, there is a sizable amount of non-emergency calls.

Frequency Check by Month

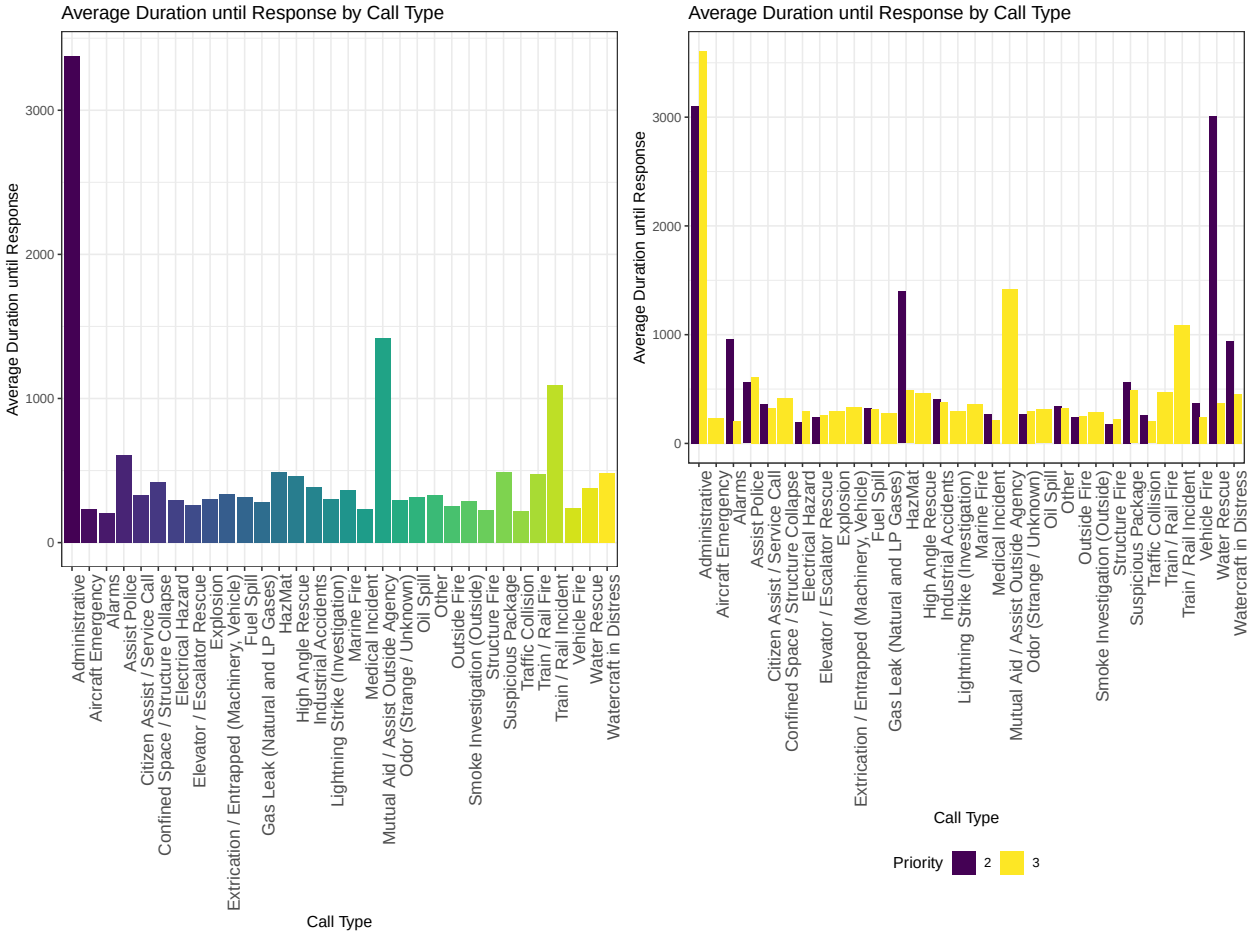
To examine the number of calls the SF Fire Department received each month on average, we created this series of graphs based on the frequency of calls each month. Here, we selected ten most frequent types of calls because they account for 98.468% of the total calls (4,750,031 out of 4,823,917 calls). This series of graphs can help us visualize the frequency of different types of calls, and through these graphs we can easily determine if there is a general trend or not. Thus, we are able to determine if in some of the months the SF Fire Department would receive more calls than other months.



As shown above, the graphs all appear to follow similar patterns. The most frequent type of calls is medical incidents, and there are slightly more calls for structure fire than alarms for all 12 months. These three types of calls are the most common ones as we can see from the graphs, and the frequency of receiving these calls each month is approximately the same - except for February, where the number of medical incident calls does not exceed 250,000. Still, the ratio between each call type is very similar across different months. Thus, we conclude that there is not a specific time period that receives more calls than other time periods do, which means that the SF Fire Department deals with roughly the same amount of accidents every month, and this should be intuitive.

Response Time

How fast does the Fire Department respond to certain types of events? What about events that are often classified as non-emergencies?



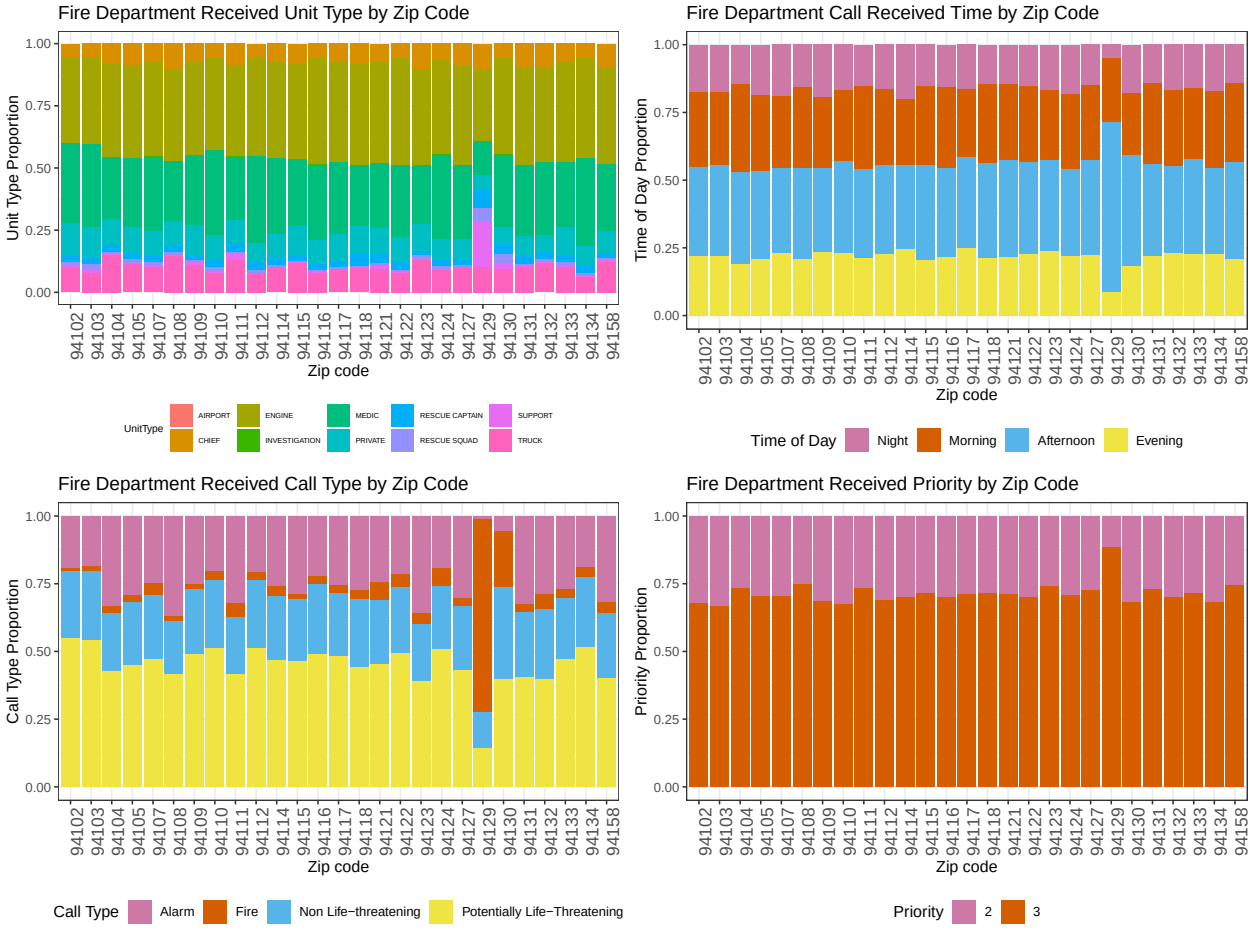
To better comprehend SF Fire Department's response rate, we computed a new metric, ResponseDuration, based on the difference in timestamp for when SF Fire Department received a call and responded to a call. By plotting the call types against the average response duration for each call type on the left barplot, there is clear indication that the SF Fire Department responds to most calls relatively quickly with the exception of administrative, mutual aid, and rail incident call types. Furthermore, we also plotted a bar plot of average response time versus each call type group further subsetted by priority to examine whether certain call types experience more delay. From the plot, we conclude that it is consistent with our intuition that calls classified as non-emergency typically experience longer response durations than their emergency classified counterparts. Notably, water rescue, hazmat, and watercraft in distress benefits from significant increase in response times when classified as emergency whereas assisting police and structure fire (!) actually takes slightly longer to respond to when classified as emergencies. The reasoning behind this, however, may attribute to the preparation and equipment needed for huge structural fires.

Deeper Analysis

Now, we hope to conduct a deeper analysis on the dataset and explore the average response duration of the SF Fire Department for regions with varying minority population. An obvious method would be to tag the zip code by their respective minority percentages and perform an analysis that way. However, we must first examine whether the distribution of incident types, call time, priority type, or responding unit type may vary from ZIP code to ZIP code.

Analysis by Zip Code

Zone Improvement Plan codes, or ZIP codes as they are commonly known, originated as a way of classifying street arrangements. Given that ZIP codes can be associated with most places of human habitation in the United States, they represent an alternative means of collecting, visualizing, and analyzing spatial information. Here, we are interested in whether the calls received by the SF Fire Department would vary by zip code. Thus we group the dataset by zip code and summarize data to produce the following plots.

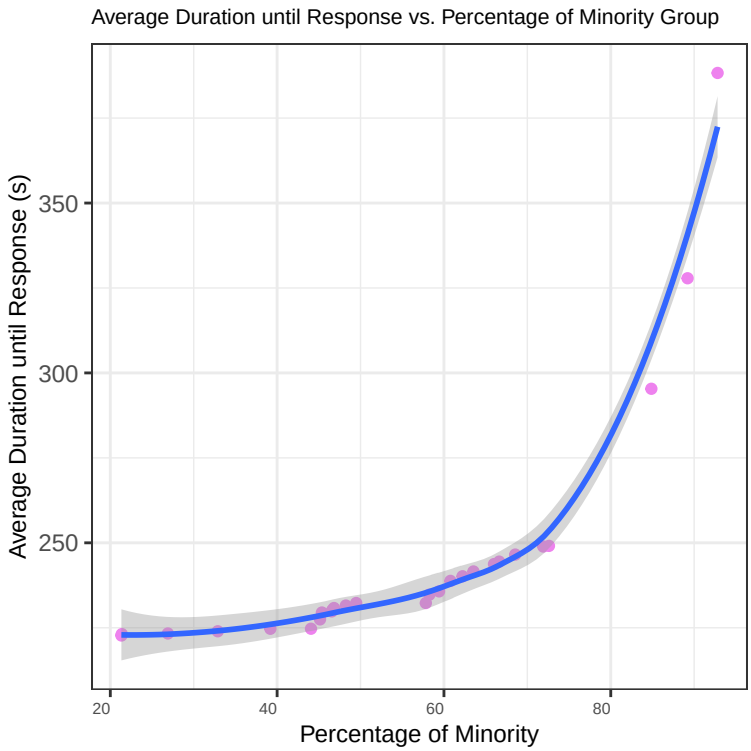


As shown in the graphs, almost all the zip code areas share a similar pattern for the four categories (UnitType, Received Time, Call type, Priority type). The top two most frequent unit types that the SF fire department has received are Engine and Medic. However, we also notice that zip code 94129 is an outlier for all four plots. After research, we realize that the Presidio of San Francisco, a 1500-acre outdoor park and part of the Golden Gate National Recreation Area, is located at 94129, which explains why it has especially high fire type calls. For the rest of the zip codes, the most frequent call type is “Potentially Life Threatening” and “Alarm”. There is no clear pattern of when the SF fire department would receive more calls throughout the day. However the afternoon (12:00 PM - 18:00 PM) and evening (18:00 PM - 23:59 PM) categories seem to have slightly higher counts. The SF fire department would receive relative the same amount of calls from different zip codes throughout the day except zip code 94129 which has a high number of calls in the afternoon. Last but not least, priority group 3 is the most frequent priority group for calls in all areas. Thus we can conclude that there is no specific correlation between zip codes and the unit type, call time, call type and priority group that the SF fire department has received.

Analysis by Race and Ethnicity

We are curious as to whether race plays a role in the way fire stations respond to incidents. We decide to quantify this problem by looking at the relationship between average response time for incidents and the percentage of racial minorities in regions with different zip codes. In this report, we define racial minorities as races that are not white. We do acknowledge that this might not be the most thorough way of classification.

We fail to find any free, downloadable dataset that contains information regarding racial distribution in different zip code areas in San Francisco. Therefore, we use web scraping to obtain data from ZipDataMaps (https://www.zipdatamaps.com/). We gather data about racial distribution in San Francisco by zip code from 2014, which we then use to calculate the percentages of minorities by zip code. To make our analysis more accurate, we only use the data from 2014 in our main dataset in this analysis. We focus on the average response time by zip code, a data that we obtained from calculating the time difference between columns “ReceivedDtTm” and “ResponseDtTm”. In the plot below, we used percentages of minorities by zip code for the x-axis and average response time by zip code for the y-axis.



This graph shows that there is a non-linear positive correlation between percentages of minorities in a zip code area and the average response time of incidents in that area. At a first glance, it seems evident that the San Francisco fire department discriminates towards racial minorities. A linear regression model in the form of

$$y_i = 245.25 + 142.87x_i + 102.52x_i^2$$

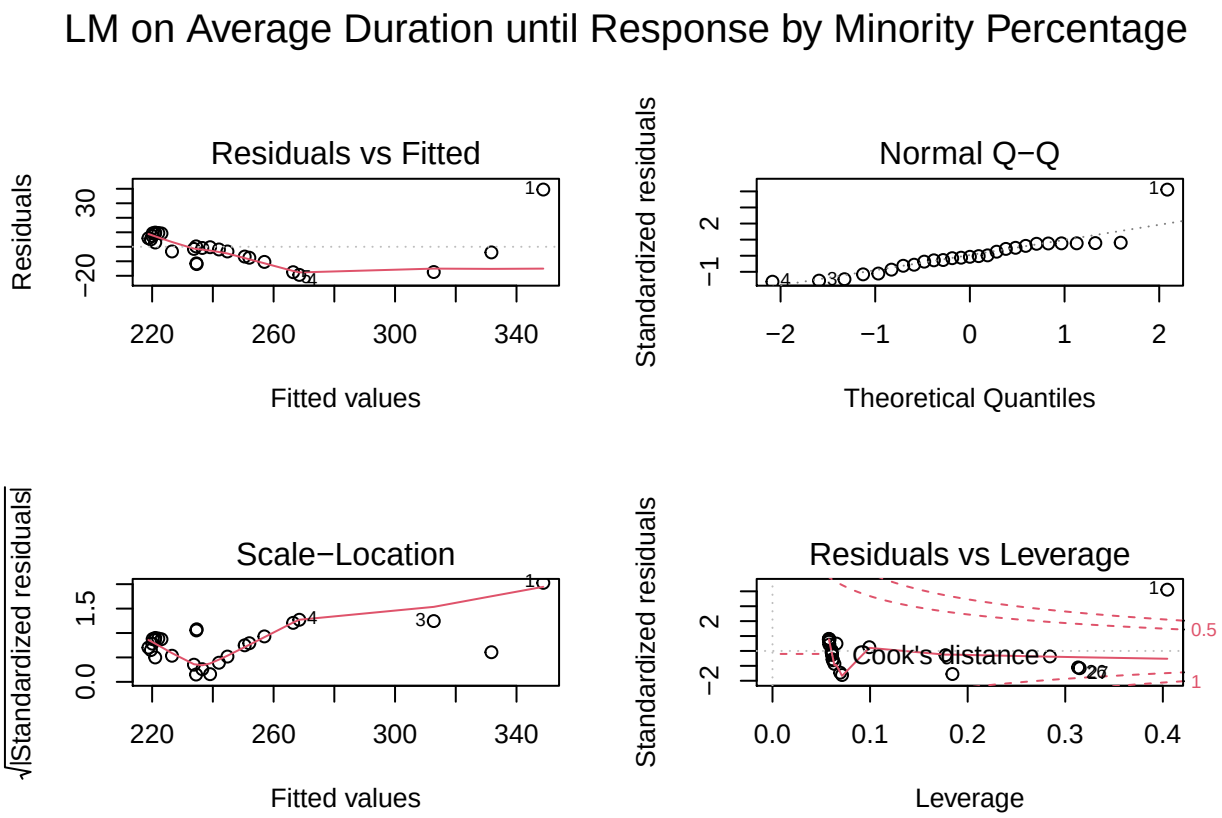
where x_i represents the minority percentage and y_i represents the average response time of incidents in the area displays the summary results below, indicating that minority percentage is a statistically significant indicator for the duration of response time.

Observations	27
Dependent variable	AverageResponseDuration
Type	OLS linear regression

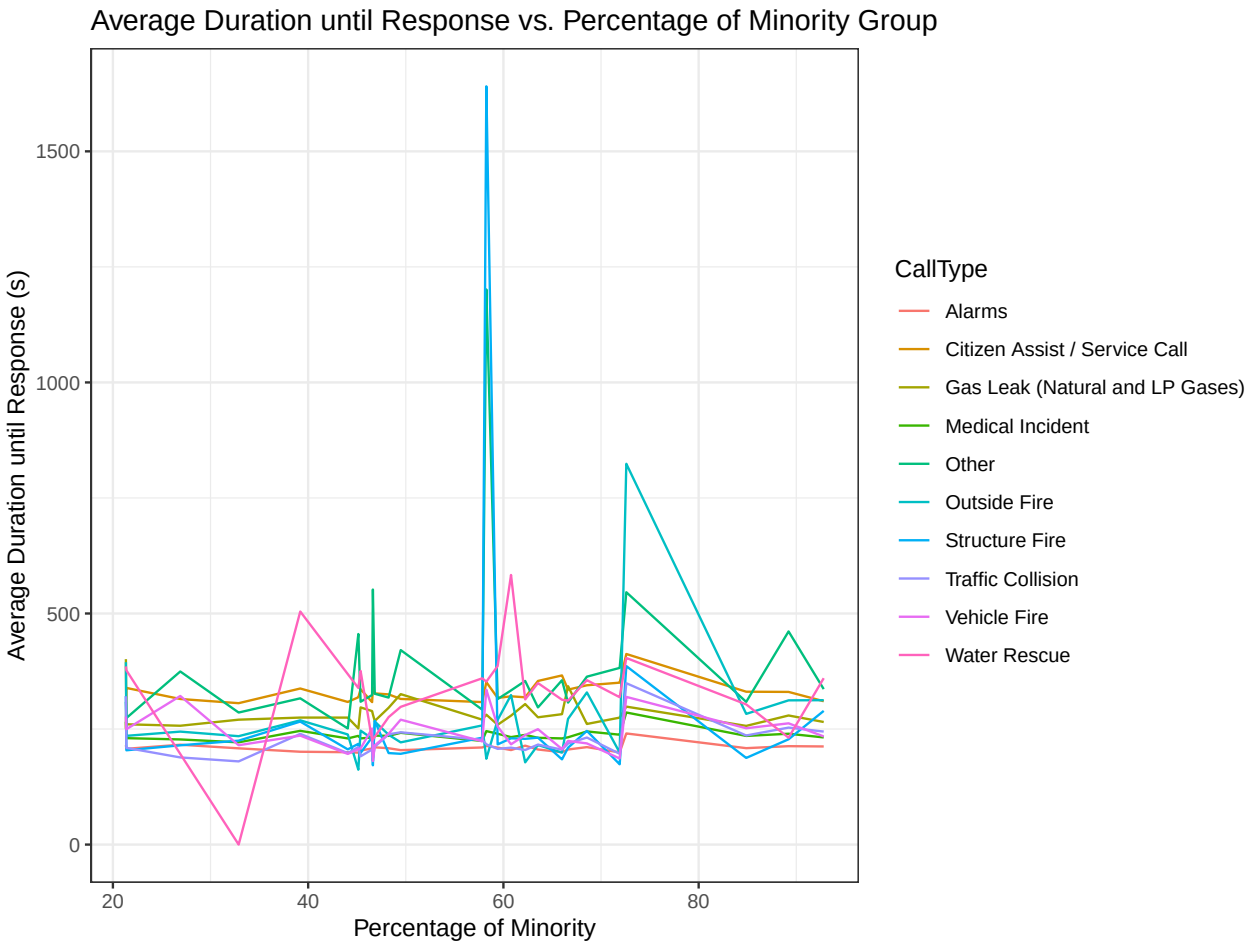
F(2,24)	99.77
R^2	0.89
Adj. R^2	0.88

	Est.	S.E.	t val.	p
(Intercept)	245.25	2.40	102.36	0.00
poly(MinorityPercentage, 2)1	142.87	12.45	11.48	0.00
poly(MinorityPercentage, 2)2	102.52	12.45	8.24	0.00

Standard errors: OLS



However, we also recognize that there might be many other factors contributing to this. Our first question is whether the same positive correlation holds across all call types. The graph below builds on top of the previous graph by identifying the average response time by zip code for each call type.



As we break down the average response time by call type, there is no clear positive correlation between percentages of racial minorities and the average response time. This can also be backed with a ANOVA test performed below.

Table 1: ANOVA Table for Average Response Time by Minority Percentage and Call Type

	SS	df	MS	F	P_val
Minority Percentage	18352.97	1	18352.97	1.31	0.253
Call Type	740164.26	9	82240.47	5.88	0.000
Error	3565843.57	255	13983.70		
Total	4324360.80	265	114577.14		

In the ANOVA test results above, it is apparent that when call type is included as an additional variable, minority percentage ceases to be statistically significant (p-value > 0.05). Still, we cannot rule out the correlation between minority percentage and response duration since the lack of strong relationship may attribute to the fact that our data only comes from the year 2014, as a result supplying a limiting sample size. A further analysis with data across various years may be able to advance our discovery further.

Conclusion

In conclusion, our data analysis debunks the urban myth that fire fighters primarily engage in fire rescue. While it is likely true that the firefighters remain the first department one should dial when dealing with bursting flames, the data clearly demonstrates that firefighters most often respond to medical incidents rather than structure fires. This is not a surprising result considering the fact that medical incidents probably occur more often than buildings on fire, however, the question remains as to why these medical incidents are directed to the fire department rather than the EMS.

Furthermore, our analysis on frequency of calls across months shows that the pattern of calls received by firefighters remain relatively stable across the months, with no notable spike in the summer time despite the data being collected from California, a location notorious for their frequent incidents of wild fires. The general response time analysis also provides us with better understanding for which incidents typically require longer time to properly respond to on average, providing us an understanding of incidents that are perhaps less urgent or harder to combat.

Lastly, the most interesting discovery may be related to the change in average response duration between call receival and call response with respect to an increase in minority percentage of a zip code. To better understand whether there are latent factors to our most interested question, we plotted various proportions of call types, priority groups, incident report times against the zip code and discovered that these factors remain relatively consistent for all zip codes. With that in mind, our analysis led us to a shocking discovery that there appears to be a strong positive correlation between the percentage of minority population at a region and the time it takes the fire department to properly address the calls. This is a shocking result since it may indicate an innate favortism towards the racial majority within the fire department system, similarily to how incidents of racial profiling by the police department enraged our nation in the past year. However, our further analysis with call type group included suggests that there remain plenty of variability for the positive correlation we observe, and that further analysis is necessary to truly understand whether there exists racial bias within the San Francisco Fire Department.