ANALYSIS & MODELING OF 911 CALLS ASSIGNED TO SAN FRANCISCO FIRE DEPARTMENT (SFFD)

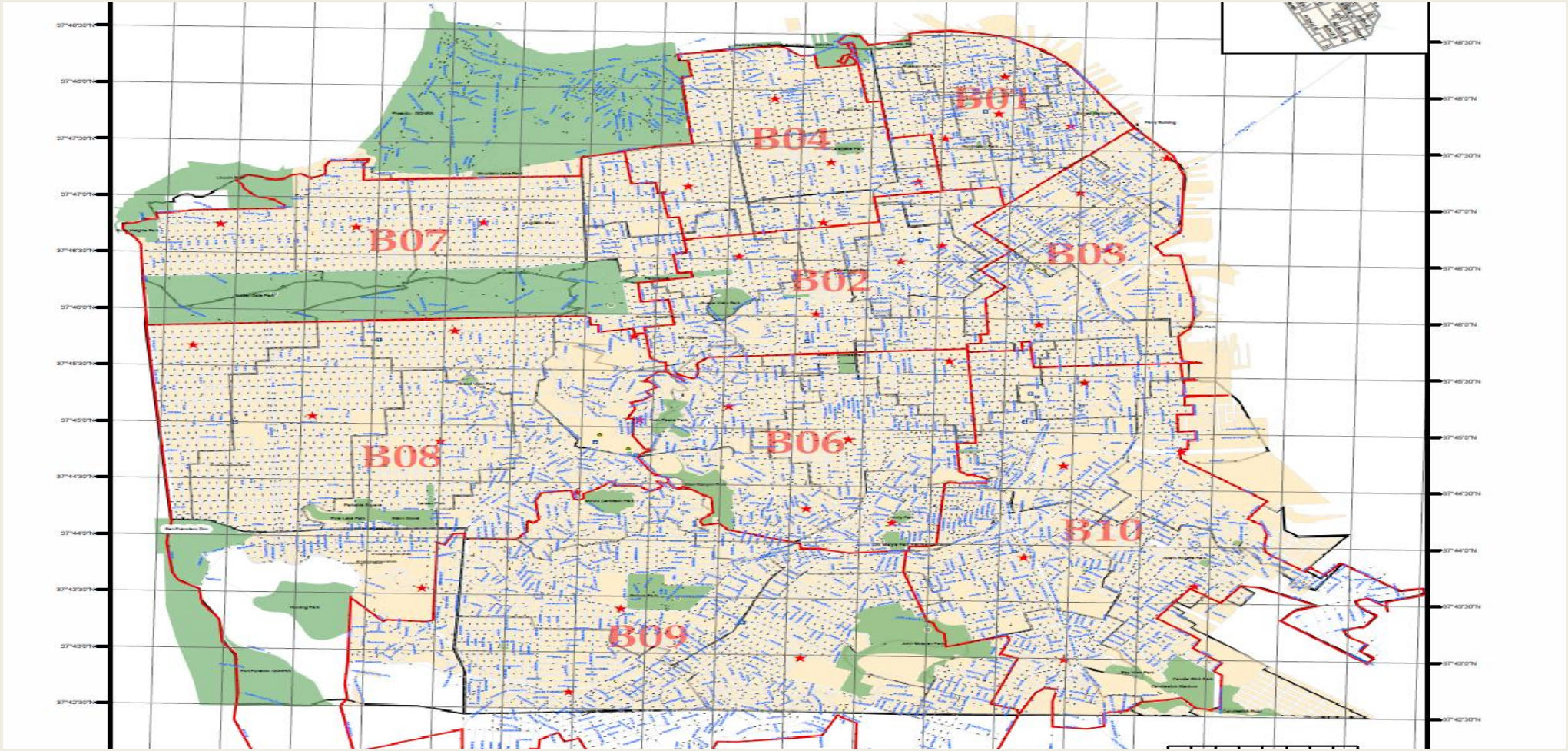
DS-SF-30 Final Project Presentation

Nikita Attiguppe Dasharath

TABLE OF CONTENT

OUTLINE

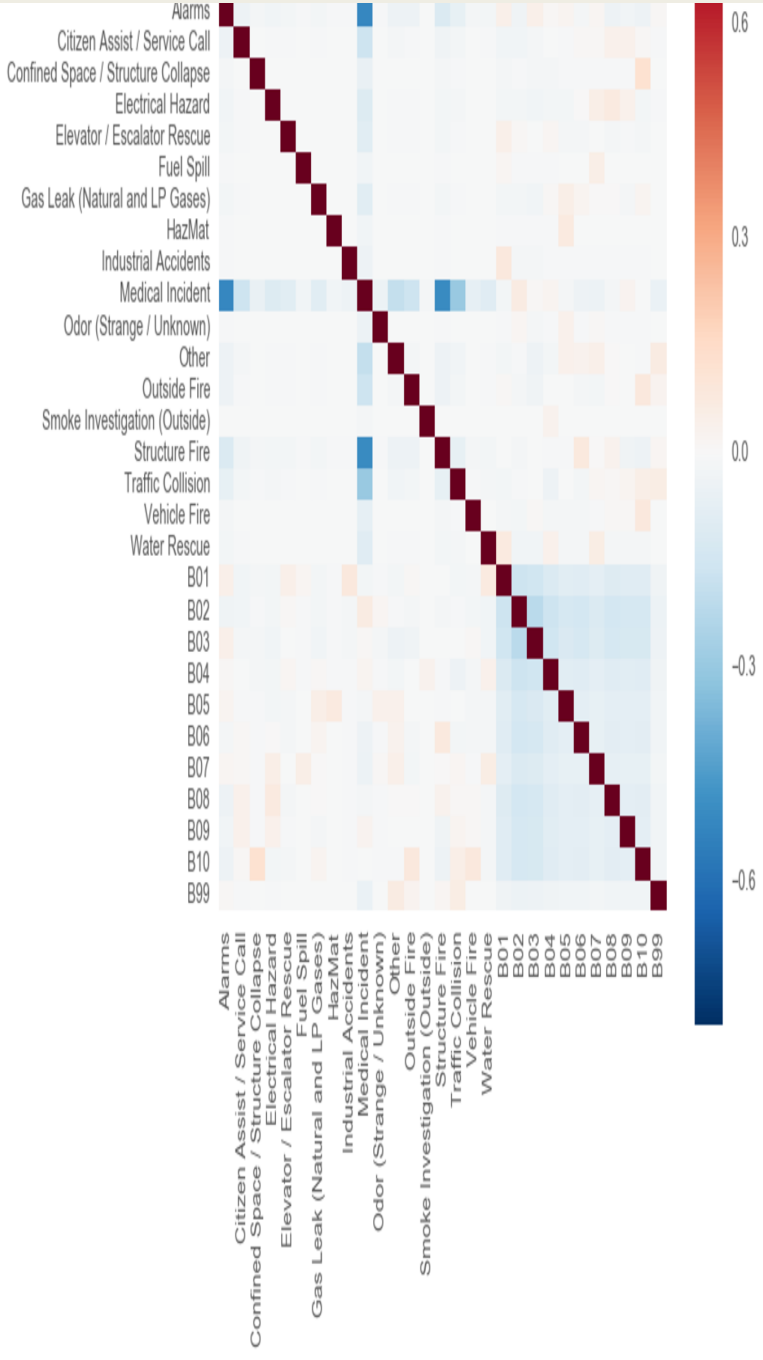
- Figure below shows the location of Fire engines and Battalion boundaries in San Francisco
- [Image source: https://data.sfgov.org]



SUMMARY

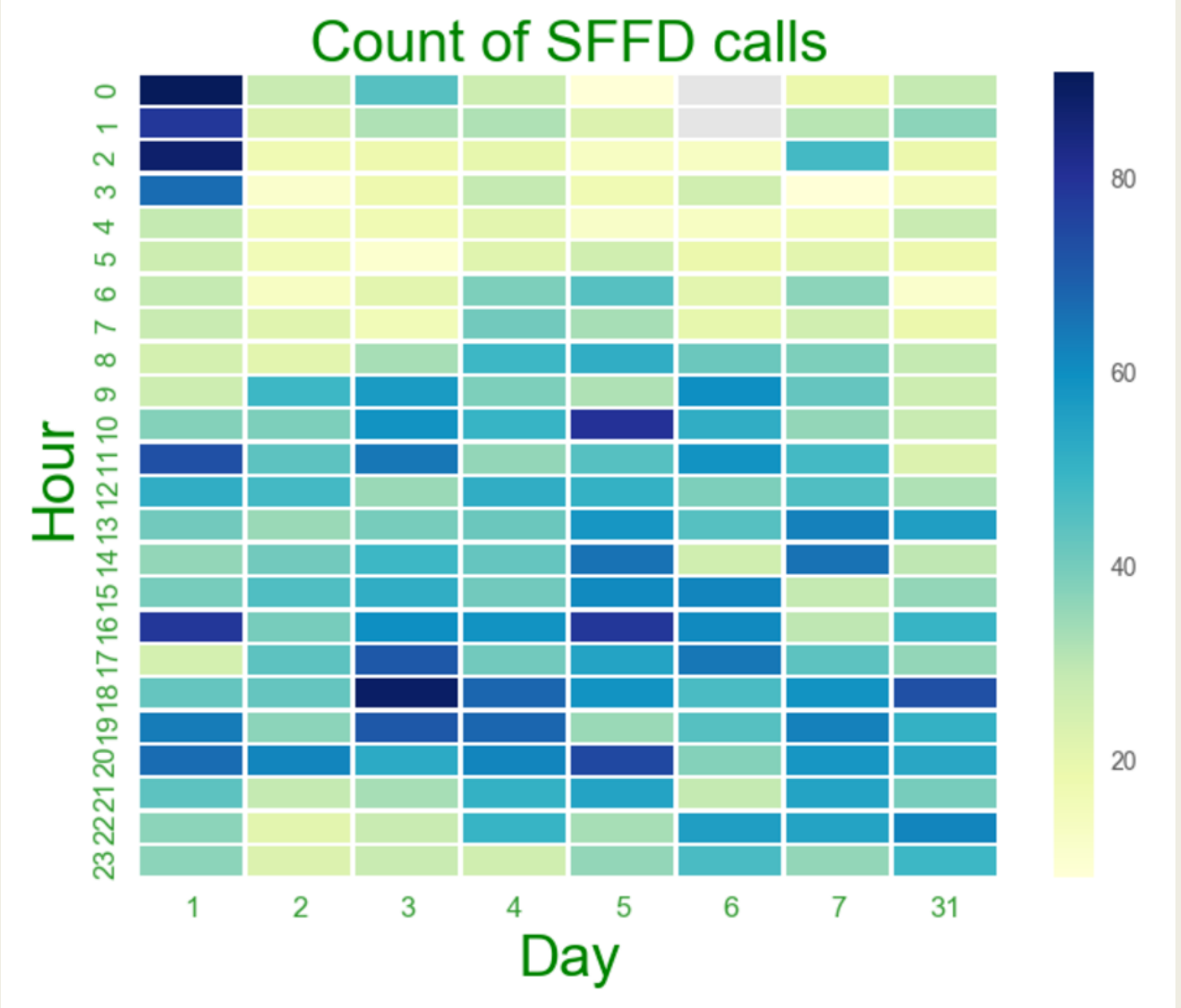
- Goal of the project is to determine the type, time and frequency of calls assigned to SFFD
- [Data source: https://data.sfgov.org]
- Time period: Dec 31, 2016 to Jan 7, 2017
- First figure on the right shows the features in the data set
- Second figure on the right shows the correlation between various types of calls and batalions

```
Call Number
Unit ID
Incident Number
Call Type
Call Date
Watch Date
Received DtTm
Entry DtTm
Dispatch DtTm
Response DtTm
On Scene DtTm
Transport DtTm
Hospital DtTm
Call Final Disposition
Available DtTm
Address
City
Zipcode of Incident
Battalion
Station Area
Box
Original Priority
Priority
Final Priority
ALS Unit
Call Type Group
Number of Alarms
Unit Type
Unit sequence in call dispatch
Fire Prevention District
Supervisor District
Neighborhood  District
Location
RowID
```



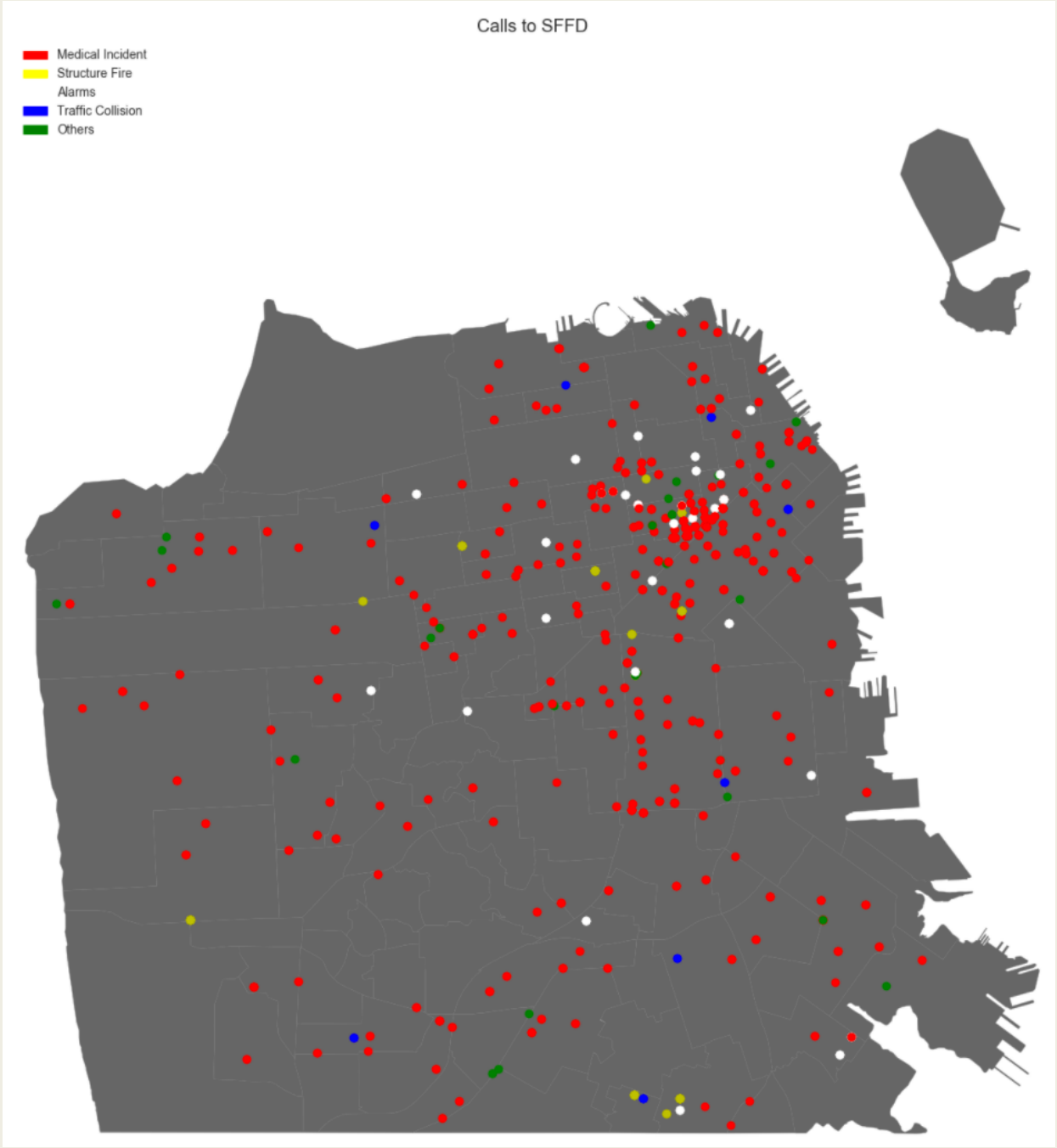
NUMBER OF CALLS PER HOUR

- X-axis shows the day
- Y-axis shows the hour
- January 1st had the maximum calls in the first week of January 2017
- Majority of the calls were between 12am – 4am and 5pm – 6pm on January 1st
- Dec 31, 2016 and January 2nd 2017 had the least number of calls in the first week of January.



GEOCODING

- Converted String Addresses to Latitude-Longitude using pygeocoder library
- Figure shows the map of San Francisco and its neighborhood boundaries
- Colored dots on the map are the majority types of calls assigned to SFFD
- Map shows calls for January 1st 2017
- 72% of calls are Medical Incidents
- 50% of calls are from only four neighborhoods : Tenderloin, SoMa, Mission and Financial District/SB



MODELING

- Linear Regression and Regularization models for Box and Battalion have around 75.5% accuracy
- Logistic Regression model for Medical Incidents and Battalion have around 69% accuracy

OLS Regression Results

Dep. Variable:	Box	**R-squared:**	0.755
Model:	OLS	**Adj. R-squared:**	0.755
Method:	Least Squares	**F-statistic:**	2369.
Date:	Tue, 21 Feb 2017	**Prob (F-statistic):**	0.00
Time:	12:51:12	**Log-Likelihood:**	-65233.
No. Observations:	7704	**AIC:**	1.305e+05
Df Residuals:	7693	**BIC:**	1.306e+05
Df Model:	10		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[95.0% Conf. Int.]
B01	1410.5198	40.527	34.804	0.000	1331.075 1489.965
B02	3224.8947	30.316	106.376	0.000	3165.467 3284.322
B03	1997.3649	31.696	63.017	0.000	1935.232 2059.497
B04	2892.2544	40.883	70.744	0.000	2812.112 2972.396
B05	4380.3296	49.713	88.113	0.000	4282.879 4477.780
B06	5380.2907	46.952	114.591	0.000	5288.252 5472.330
B07	6639.5866	57.314	115.845	0.000	6527.235 6751.938
B08	7682.5896	46.491	165.249	0.000	7591.454 7773.725
B09	7089.1927	49.346	143.662	0.000	6992.460 7185.925
B10	5529.7403	48.422	114.198	0.000	5434.819 5624.661
B99	4062.4638	138.685	29.293	0.000	3790.603 4334.324

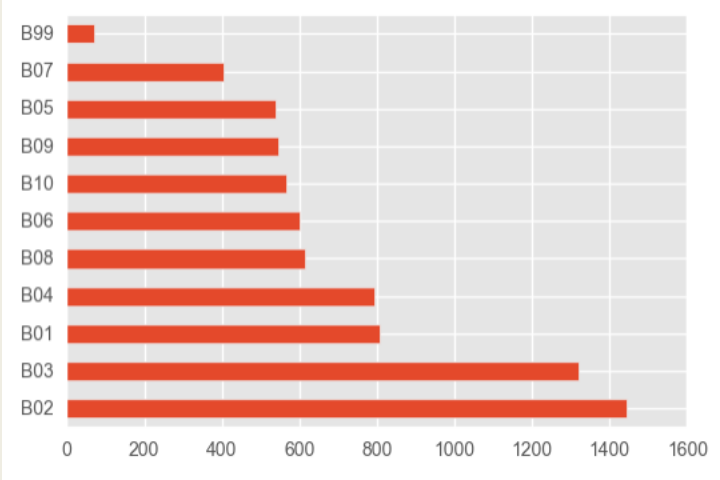
Confusion Matrix

True Class	**0.0**	**1.0**
Hypothesized Class		
0.0	40	2333
1.0	29	5302

Omnibus:	1710.125	**Durbin-Watson:**	0.935
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	6987.085
Skew:	-1.044	**Prob(JB):**	0.00
Kurtosis:	7.172	**Cond. No.**	4.57

RESULTS

- Time based analysis of the calls assigned to San Francisco Fire Department showed increase in calls during holidays such as New Years Eve. It also sowed less number of calls from 5am in the morning till noon
- Majority of the calls are assigned to four battalions (B02, B03, B01, B04) where there are more incidents as compared to other neighborhoods.



CONCLUSION

- Classification of call frequency based on neighborhood and time of the day describe the calls to San Francisco Fire Department.
- Although Medical Incidents are the most common type of calls assigned to SFFD, predicting such incidents using Battalion has an accuracy of only 70%. Analysis with few years of call data could improve the accuracy.
- Integrating call data with neighborhood data would help improvise the models.

REFERENCES

- General Assembly course material
- http://stackoverflow.com/
- https://www.kaggle.com
- **pandas**.pydata.org
- seaborn.pydata.org
- https://github.com/sfroid/sfcoord2nbh
- http://sensitivecities.com/so-youd-like-to-make-a-map-using-python-EN.html#.WKy_6BIrKRs

NEXT STEPS

- Incorporate more San Francisco Fire Department call Data
- Analysis of Response times of calls for call type and severity of the incidents