Final Assignment
Battle of Neighborhoods

Introduction/Business Problem

The Boston Police Department is planning to open a new police station. In order to select the best location I'll use FourSquare so that the most suitable place is visible. The station should be in the place where most of crimes are committed and there's no police station nearby.

Data description

The Boston Crime data from open source https://www.kaggle.com/AnalyzeBoston/crimes-in-boston is used combining with the Foursquare location data. Dataset contains such characteristics of crimes as type, location, year, offence type.

I cleaned the data dropping missing values so that it doesn't affect the results. Only first 500 crimes were used as the dataset is too large itself.

SHOOTING	OCCURRED_ON_DATE	YEAR	MONTH	DAY_OF_WEEK	HOUR	UCR_PART	STREET	Lat	Long	Location
NaN	2018-09-02 13:00:00	2018	9	Sunday	13	Part One	LINCOLN ST	42.357791	-71.139371	(42.35779134, -71.13937053)
NaN	2018-08-21 00:00:00	2018	8	Tuesday	0	Part Two	HECLA ST	42.306821	-71.060300	(42.30682138, -71.06030035)
NaN	2018-09-03 19:27:00	2018	9	Monday	19	Part Three	CAZENOVE ST	42.346589	-71.072429	(42.34658879, -71.07242943)
NaN	2018-09-03 21:16:00	2018	9	Monday	21	Part Three	NEWCOMB ST	42.334182	-71.078664	(42.33418175, -71.07866441)
NaN	2018-09-03 21:05:00	2018	9	Monday	21	Part Three	DELHI ST	42.275365	-71.090361	(42.27536542, -71.09036101)
NaN	2018-09-03 21:09:00	2018	9	Monday	21	Part Three	TALBOT AVE	42.290196	-71.071590	(42.29019621, -71.07159012)
NaN	2018-09-03 21:25:00	2018	9	Monday	21	Part One	NORMANDY ST	42.306072	-71.082733	(42.30607218, -71.08273260)
NaN	2018-09-03 20:39:37	2018	9	Monday	20	Part Three	LAWN ST	42.327016	-71.105551	(42.32701648, -71.10555088)
NaN	2018-09-03 20:48:00	2018	9	Monday	20	Part One	MASSACHUSETTS AVE	42.331521	-71.070853	(42.33152148, -71.07085307)
NaN	2018-09-03 20:38:00	2018	9	Monday	20	Part Three	LESLIE ST	42.295147	-71.058608	(42.29514664, -71.05860832)

	INCIDENT_NUMBER	OFFENSE_CODE	OFFENSE_CODE_GROUP	OFFENSE_DESCRIPTION	DISTRICT	REPORTING_AREA	SHOOTING
0	I182070945	619	Larceny	LARCENY ALL OTHERS	D14	808	NaN
1	I182070943	1402	Vandalism	VANDALISM	C11	347	NaN
2	I182070941	3410	Towed	TOWED MOTOR VEHICLE	D4	151	NaN
3	I182070940	3114	Investigate Property	INVESTIGATE PROPERTY	D4	272	NaN
4	I182070938	3114	Investigate Property	INVESTIGATE PROPERTY	B3	421	NaN
5	I182070936	3820	Motor Vehicle Accident Response	M/V ACCIDENT INVOLVING PEDESTRIAN - INJURY	C11	398	NaN
6	I182070933	724	Auto Theft	AUTO THEFT	B2	330	NaN
7	I182070932	3301	Verbal Disputes	VERBAL DISPUTE	B2	584	NaN

Dropping NaN values

```
df2=df2.dropna(subset=['Long'])

df2=df2.dropna(subset=['Lat'])
```

Methodology.

Folium library is used to demonstrate everything on map. Crimes' latitude and longitude were taken.

```python
# create map of Boston using latitude and longitude values
map_boston = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, label in zip( df2['Lat'], df2['Long'], df2['OFFENSE_CODE_GROUP']):
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_boston)

map_boston
```
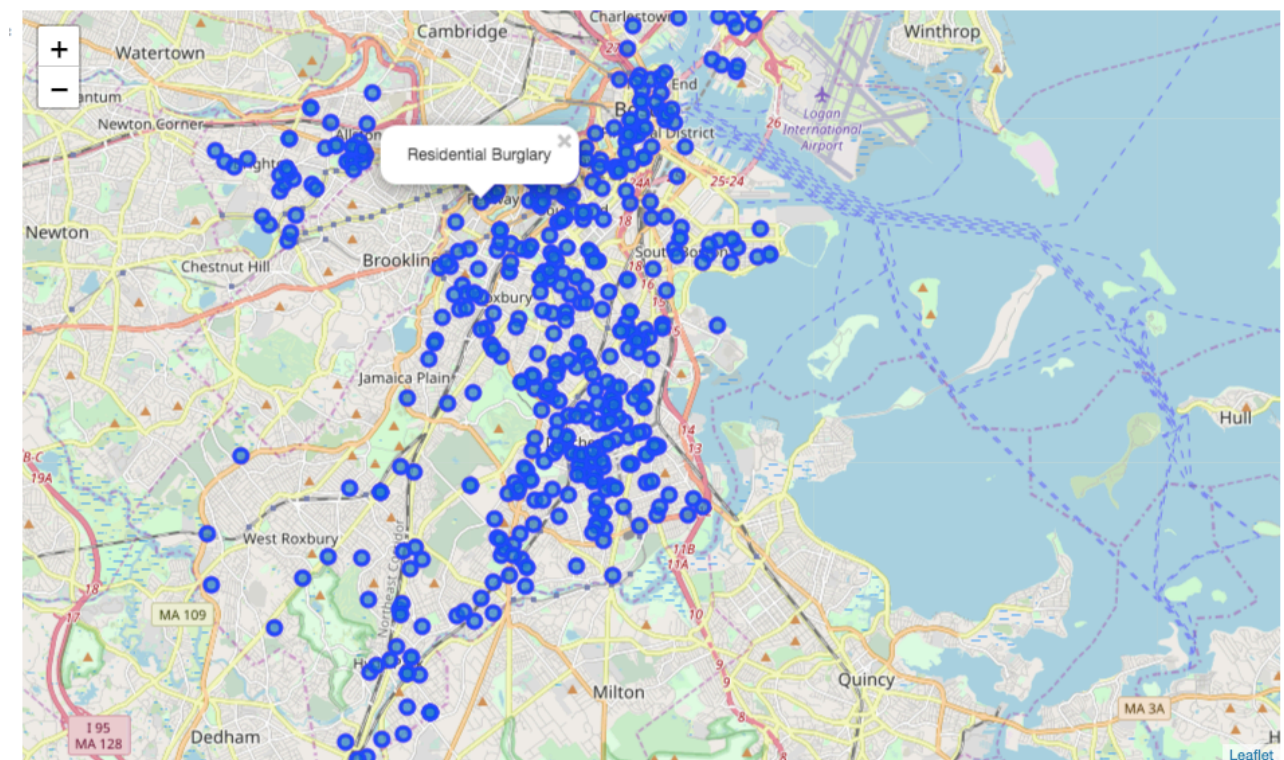
Boston map with crimes on it
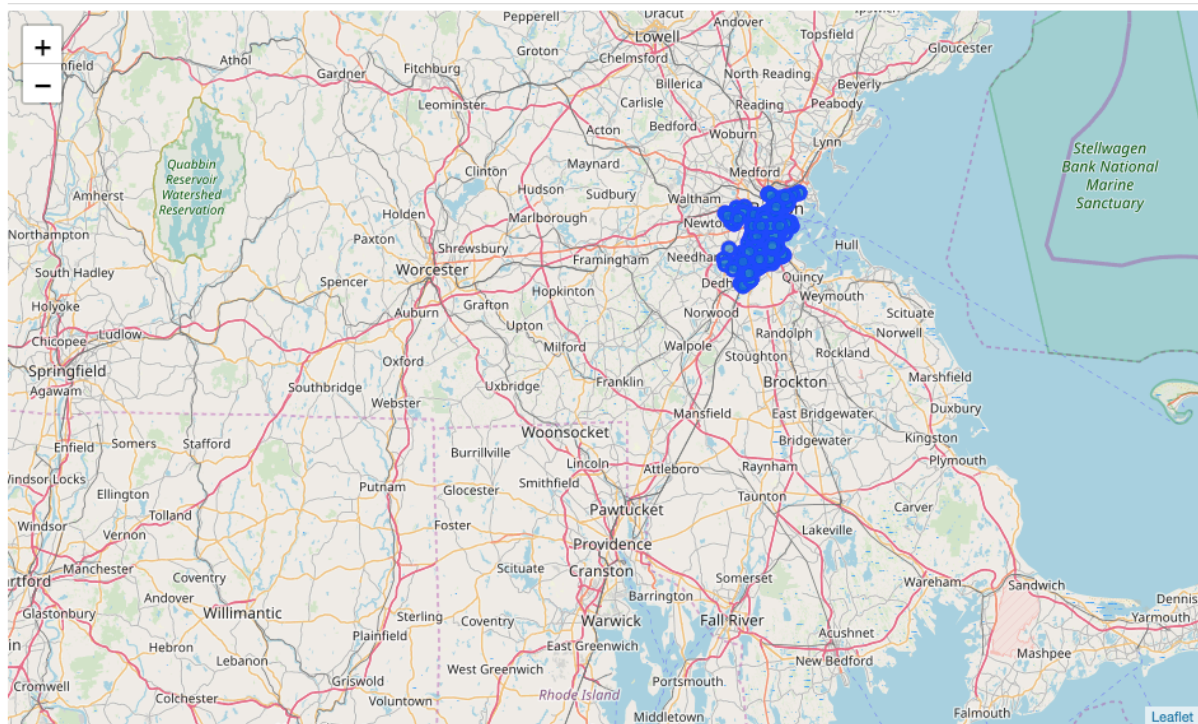
Zooming and we can see crime name.

Results

Foursquare API was used as well. With it we could see nearby venues which are in 500 meters from crime happened. We can use this data for the sake of security of these venues.

	name	categories	lat	lng
0	Jump On In	Athletics & Sports	42.359800	-71.138445
1	Kohi Coffee	Café	42.356692	-71.142516
2	CrossFit ONE Nation Boston	Athletics & Sports	42.359642	-71.138673
3	HomeGoods	Furniture / Home Store	42.356479	-71.138858
4	Warrior Ice Arena	Hockey Rink	42.357094	-71.143708