Larceny/Theft Are Most Common In the Afternoon & Evenings near San Francisco Hall of Justice/Country Jail/Traffic Court

The data used in the analysis was the San Francisco Summer 2014 Incident data set. Prior to visualization I augmented the data by adding 'morning', 'afternoon', and 'evening' category. Morning is defined between the hours of 6AM and 12PM. Afternoon is defined as 12PM to 7PM. Evening is defined at 7PM to 6AM.

The approach parsed and formatted the data to be ingested into Elasticsearch using python in Jupyter Notebook. The data was then rapidly visualized through Kibana pulling data from Elasticsearch. The data is 28,993 records and spans June – August of 2014. When the data was ingested into Elasticsearch both time and location were indexed so that aggregations could be done in time and space.

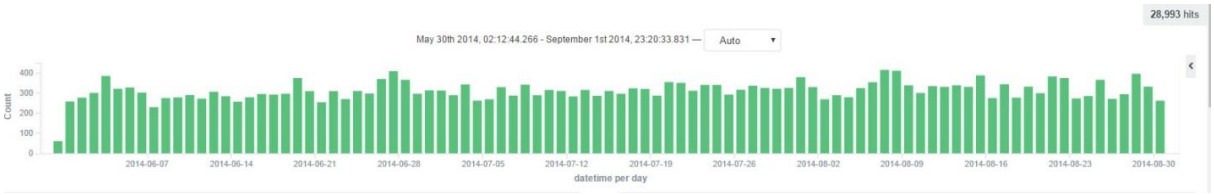

Figure 1 - Incident Volume By Day

When visualizing the overall counts by incident type, Larceny/Theft stands out as the largest. Figure 2 & 3 show only the top eight incidents. Figure 2, shows that 'Incident Category Count by Day of the Week' that there is not a specific day when Larceny/Theft occurs. The chart does show that Saturdays are bit higher than others but not by much.

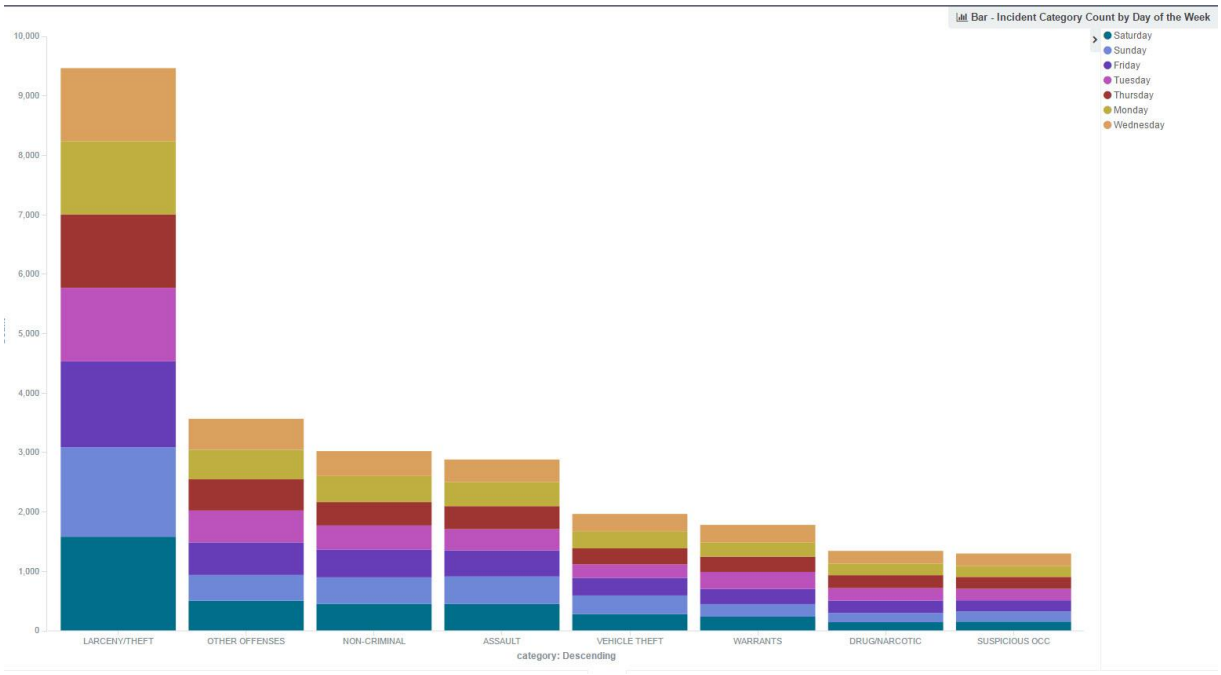

Figure 2 - Incidents by Day of the Week

With regards to time of day, Figure 3 shows that the afternoons have the highest counts of larceny/theft as well as evenings. Larceny/Theft is also happening in the mornings. These incidents do not occur only under the cover of night as one might expect.

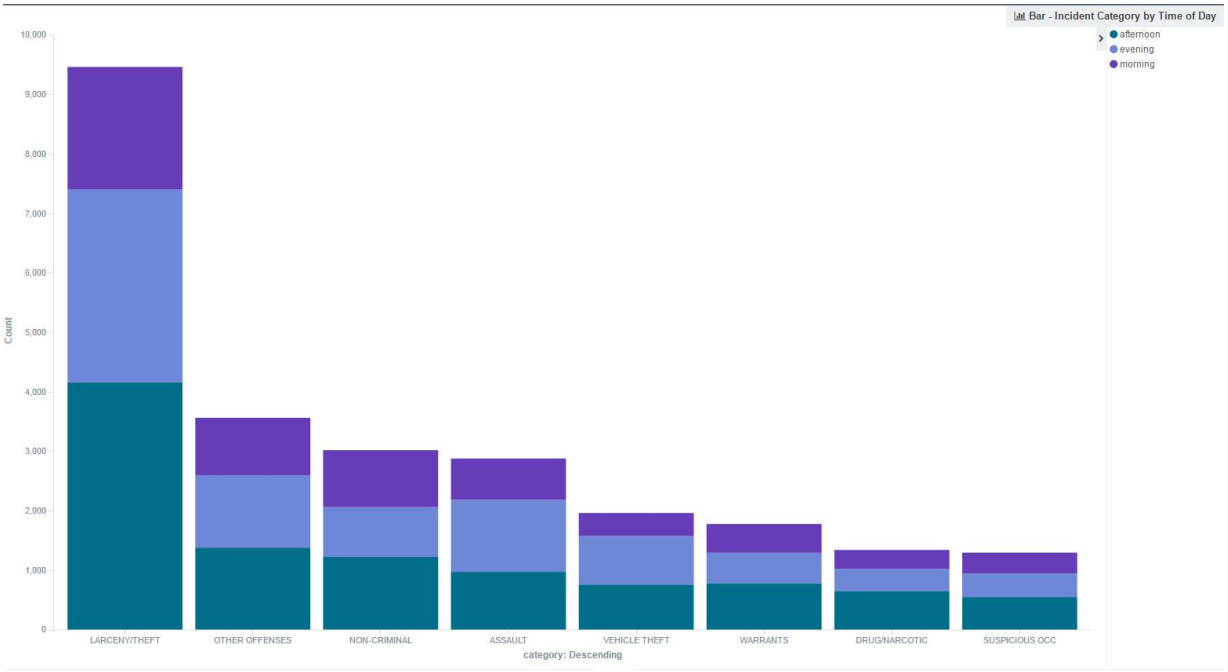

Figure 3 - Top 5 Incidents By Time of Day

When looking closer at Larceny/Theft geospatially we see that 64% of the events are happening at the geohash '9q8yysu9cby5'. When calculating the latitude and longitude for this geohash we see that the incidents are occurring near the Hall of Justice/County Jail/Traffic Court.

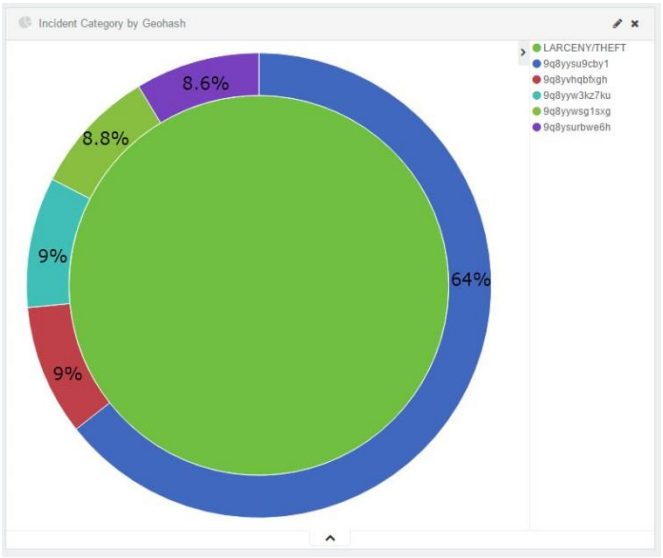

Figure 4 - Pie Chart for Larceny/Theft by Geohash

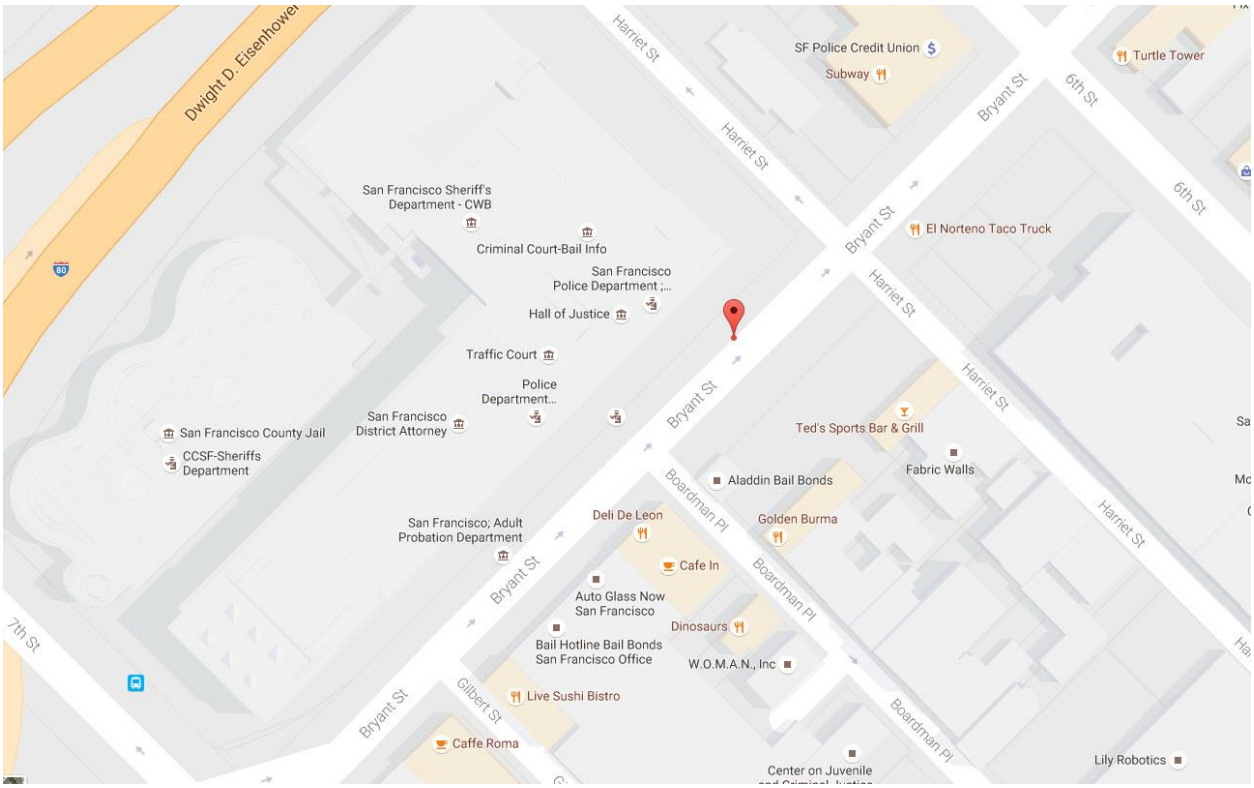

Figure 5 - Larceny/Theft Geohash Lat/Long Plotted