Evaluation of Crime Incidences in Boston

Calvin Liang and Kevin Leung

CS591 – Data Mechanics, Fall 2016

Introduction

According to the Boston Globe, "Major crime in Boston dropped 9 percent between 2014 and 2015, bringing the figure to its lowest point in a decade." While the diminishing crime rate is a feat for Boston, it is still necessary to analyze the effect certain factors have on crime incidences. To evaluate crime incidence in Boston, it is helpful to look at aspects that may potentially significantly influence crime rate such as location, time, day, liquor stores.

It is also necessary to analyze the specifics of crimes that are particularly harmful to the city such as sexual assault. According to the US National Library of Medicine, "place-based" factors are particularly useful for evaluating crime in that "states and cities can build on such information to strengthen their alcohol control and policing policies."

Project

For our project, we examined Crime Incident Reports provided by the Boston Police Department, as released by the City of Boston, and linked them to various other datasets to see if we can determine if they are somehow correlated and how safe areas might be based on the businesses, public services, etc. that encompass these areas. We used Crime Incident Reports (August 2016). Liquor Licenses and Boston Public District Stations as our data sets (below).

Crime Incident Reports	Boston Public District Stations
- Contains Information for over 140,000 crime incidents including location, time, offense code group, reporting area, etc.	Liquor Licenses

Figure 1. Important data sets used for this project.

- **crime_time_correlation.py** uses Statistical Analysis to determine if there is a correlation between crimes and the hour they were committed.
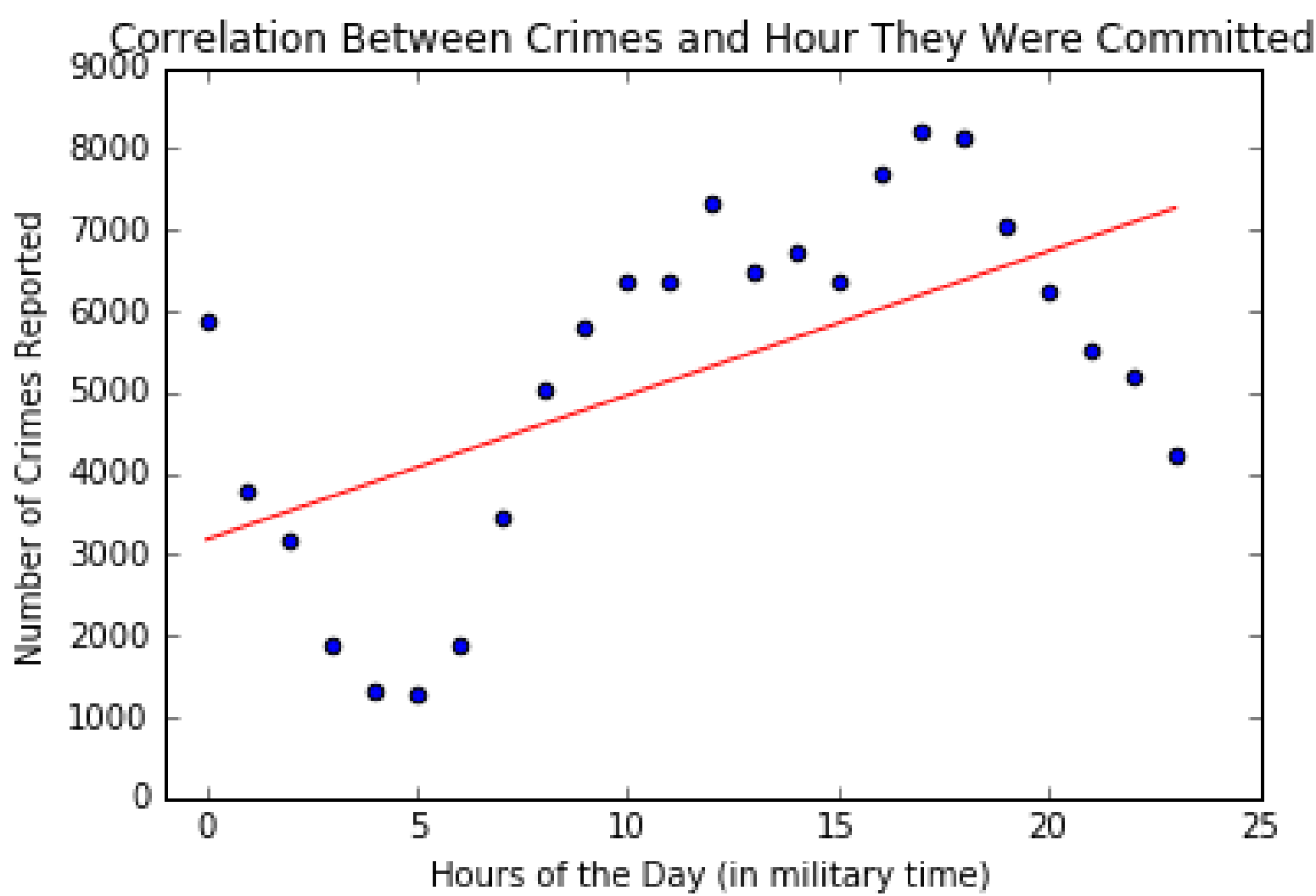

Figure 2. Correlation between crime & hour committed

Scripts Written

- **crimesNearBPDS.py** is a non-trivial transformation that counts of crimes in each districts and matches the counts with their corresponding police district station. Refer to Figure 3 (below) and Figure 4 (right).

District	Counts	Population	Crime to Popul. Ratio
A1	14041	55,971	0.2927
A15	2346		
A7	5454	40,508	0.1346
B2	20407	76,917	0.2653
B3	13515	36,480	0.3704
C6	8908	35,200	0.2530
C11	17044	91,982	0.1852
D4	16372	77,773	0.2105
D14	7753	74,997	0.1033
E5	5136	50,983	0.1007
E13	6819	37,468	0.1819
E18	6608	30,631	0.2157

Figure 3. Districts in which each Boston Police Station serves. We used this map, in addition to crime counts, to see if there may be a relation between size of a district to crimes. *(Source: Harvard)*

- **liquorAndCrime.py** takes the subset of crimes meeting the constraints: that the crimes be within 100 meters from a liquor store and belongs to one of the following selected offense categories - "Harassment", "Aggravated Assault", "Simple Assault", "Sex Offender Registration" or "Homicide".

Selected Crimes Within 100 m of Liquor Store Within a Subset of 10,000 Entires

Day Range	Total Crimes	Average Crimes
Mon-Thurs	313	78.25
Fri-Sun	369	123.0

Total Selected Crimes Within 100 m of Liquor Store
682

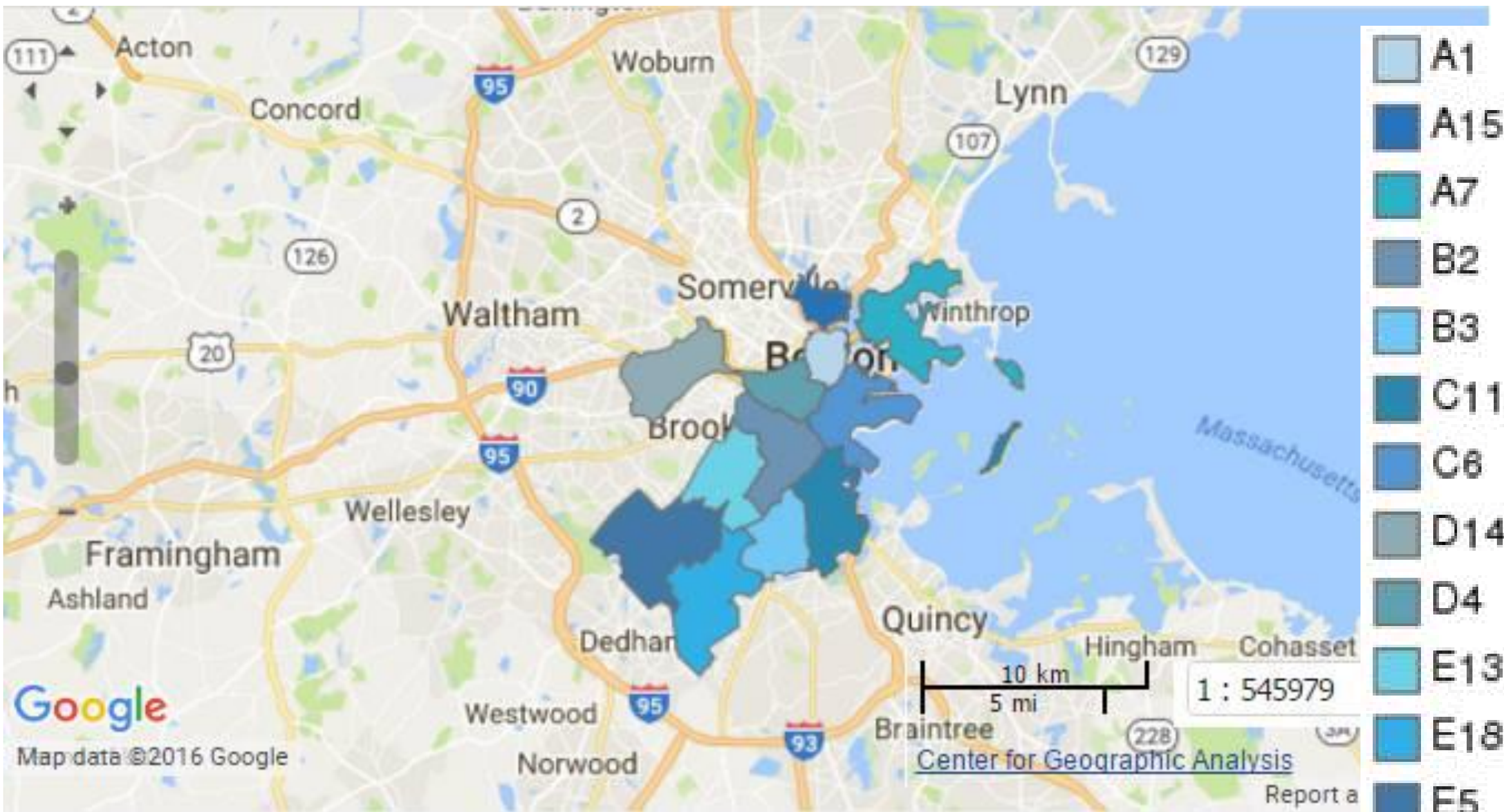Figure 5. Table comparing average crimes during weekdays & weekends.

Figure 4. Counts of crimes and population served for each BDPS *(Source: Wiki)*

Data Analysis and Visualization

Figure 2 shows the graph of the correlation between the crimes in the data set and the hour they were committed. After calculations, we received a correlation coefficient with a magnitude of 0.589, which tells us that the two variables have a positive, moderately strong correlation.

Figure 5 shows a table with the selected crimes within 100 meters a subset of 10,000 entries out of over 140,000 crimes. We selected crimes that we deemed to be particularly threatening. The total crimes during the weekend (Fri-Sun) were greater than the total crimes on the weekday (Mon-Thurs). The average crimes on the weekend are greater than the average crimes on the weekday, as well.

Conclusion

From our results, while we cannot state any facts definitively, we may, however, confirm that there is certainly a strong correlation between crimes and the usual factors they're associated with (ie. time of day in which crimes transpire most and frequency (on average) throughout the week near locations they are more prone to occur). With regards to our results in Figure 3, we cannot confidently make any direct claims without knowing more about each district (ie. how many resources are allocated to each station, how many police officers each station has, etc.), but it is interesting to observe how districts with jurisdiction over smaller populations don't necessarily have fewer crimes (and the same goes with larger populations).

References

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3673267/
- https://www.bostonglobe.com/metro/2016/01/15/major-crime-drops-boston/TrO5ZAhmOD3bFDqdBX8vwN/story.html
- https://data.cityofboston.gov/Public-Safety/Crime-Incident-Reports-August-2015-To-Date-Source-/fqn4-4qap
- https://data.cityofboston.gov/dataset/Liquor-Licenses/hda6-fnsh
- https://data.cityofboston.gov/Public-Safety/Boston-Police-District-Stations/23yb-cufe
- http://worldmap.harvard.edu/data/geonode:boston_police_districts_f55
- https://en.wikipedia.org/wiki/Boston_Police_Department