

Data Science for Social Good : A Geo-Spatial Atlas of Crime Hotspots in Houston

Kevin Peng - Department of Bioengineering
Jack Duryea - Department of Computer Science
David Gan - Department of Computer Science
Emre Yurtbay* - Department of Statistics
Corresponding Author*

September 2018, HackRice8 - Rice University, Houston, TX

Contents

1 Abstract	2
2 Background	2
3 Methods	2
4 Exploratory Data Analysis	3
4.1 All Crime Comparison	3
5 Results	6
5.1 All Crime Comparison	6
5.2 Violent Crime Comparison	7
5.3 Crimes By Weekday	8

1 Abstract

Human behavior is notoriously hard to predict and forecasting crime is no exception. In this project, a team composed of computer scientists, engineers, and statisticians used geo-spatial analysis and data visualization to contextualize crime in the city of Houston, Texas. Our results show shifts in major crime hotspots in Houston over the last decade, which will help law enforcement agencies allocate resources more effectively to stop crime before it happens.

2 Background

As one of the largest cities in the United States, Houston is not without its fair share of crime. The Houston Police Department (HPD) possesses a vast repository of historical crime data, dating back to 2010. These data contain descriptions of crimes along with the geo-spatial and temporal measurements of these crimes. Gaining a better understanding of these data is crucial to the safety and well being of the city, and analyses could greatly aid crime prevention agencies. Additionally, comparing recent and historical datasets could allow us to track the city's progress in its fight against crime. We seek to provide a geo-spatial atlas that provides visualizations of crime hotspots in Houston in order to help law enforcement distribute their efforts in a more efficient way in order to help prevent crime.

3 Methods

We analyzed the geo-spatial and temporal aspects of crimes across Houston in both 2010 and 2018. Types of crimes include theft, burglary, robbery, auto theft, aggravated assault, rape, and murder. We broke up the data into violent and non-violent crimes, as well as the day of the week crimes were committed. Subsequently, we created visualizations superimposed on a map of Houston. All visualizations were built using the R programming language. Graphics were primarily built with the package ggplot2, and map visualizations were created using ggmap, a package built by David Kahle and Hadley Wickham. More information can be found here: <https://cran.r-project.org/web/packages/ggmap/ggmap.pdf>.

4 Exploratory Data Analysis

4.1 All Crime Comparison

In 2010, we see that crime is centralized to central Houston, just south of Market Square Park and Main St. In the last 8 years, the center of crime has moved gradually southwest. This includes both violent and non-violent crime.

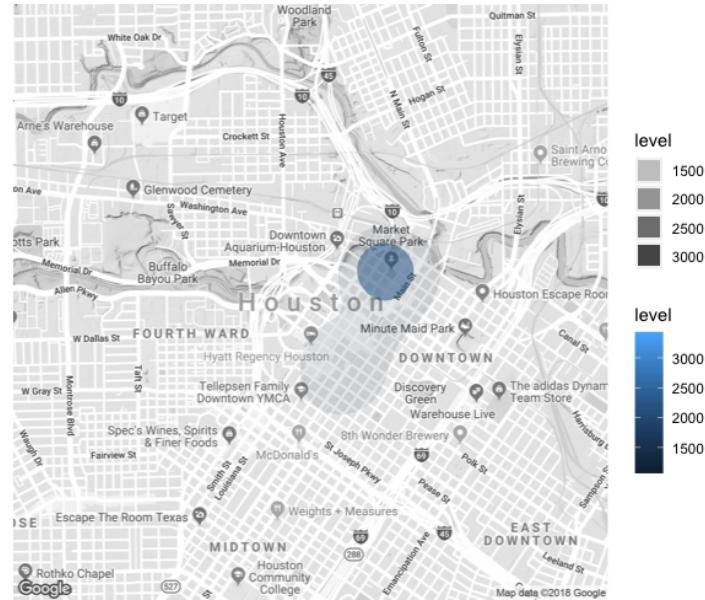


Figure 1: 2010 Non-Violent Crimes

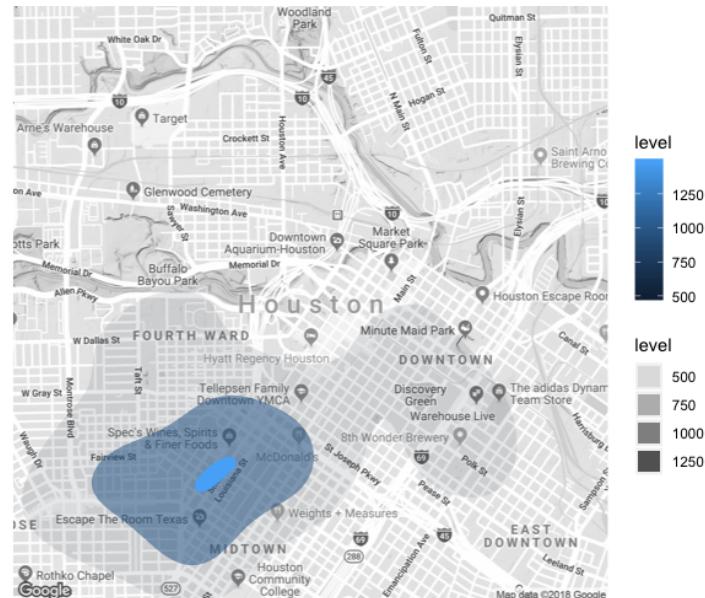


Figure 2: 2018 Non-Violent Crimes

In Figure 3, each dot represents a non-violent crime committed in the downtown Houston area in the time frame given by the dataset.



Figure 3: 2010 Violent Crimes Bubble

Further, when analyzing both 2010 and 2018 data on crime committed on specific premise types, it becomes evident that apartments, apartment parking lots, homes, and streets continue to be the primary areas where crime is targeted.

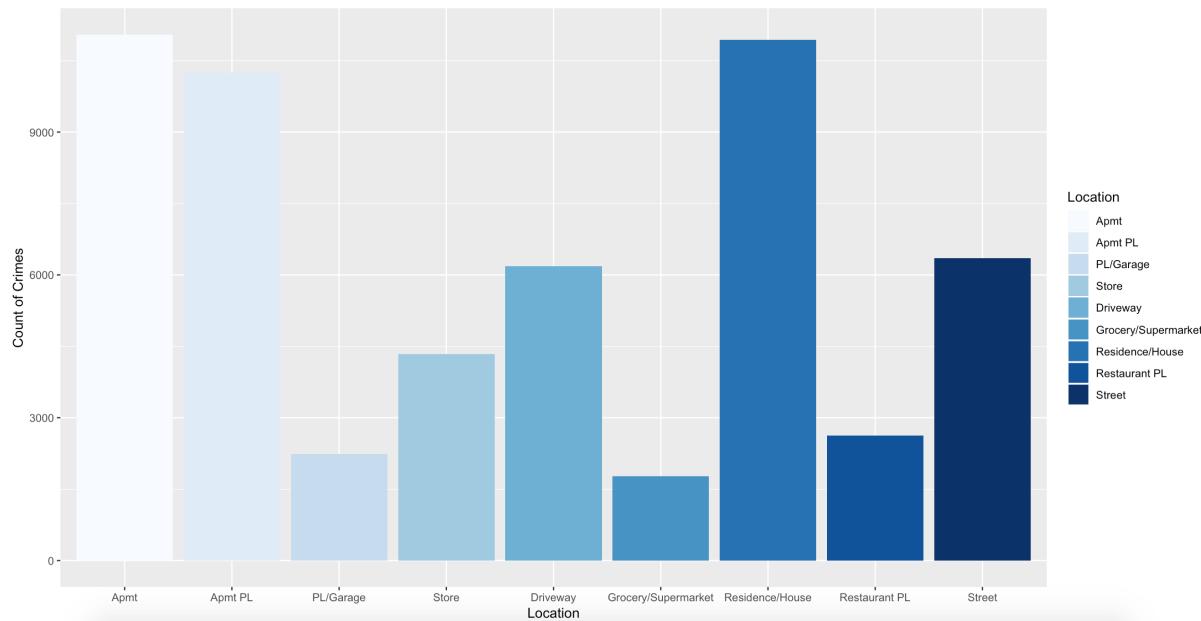


Figure 4: Counts of Crime by Premise Type (2010)

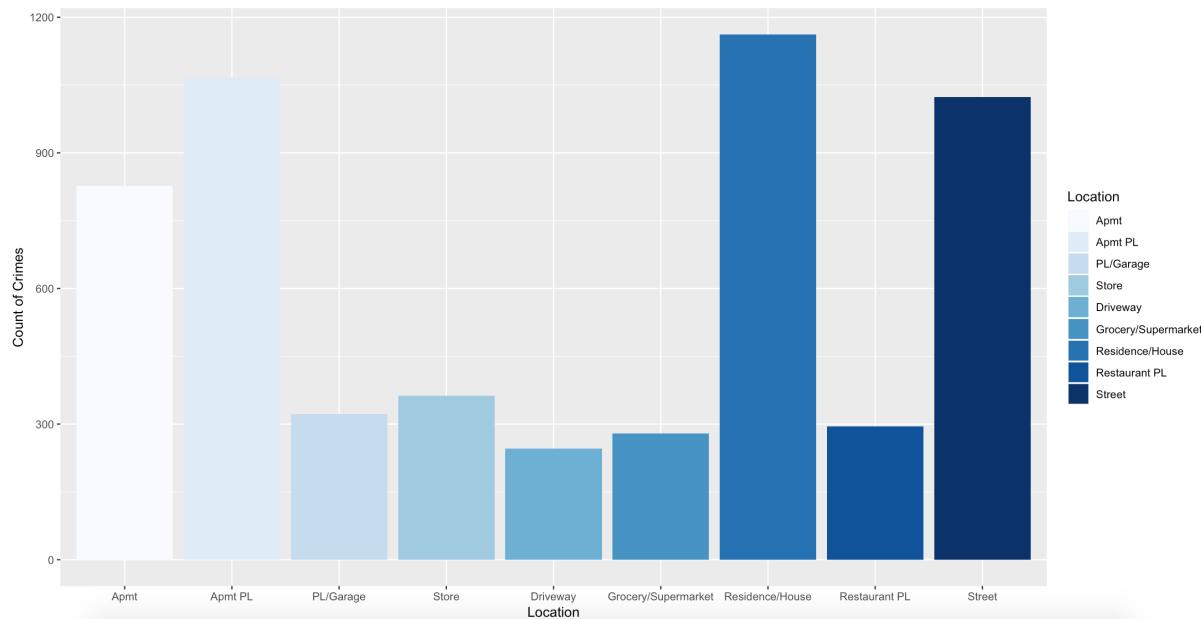


Figure 5: Counts of Crime by Premise Type (2018)

5 Results

5.1 All Crime Comparison

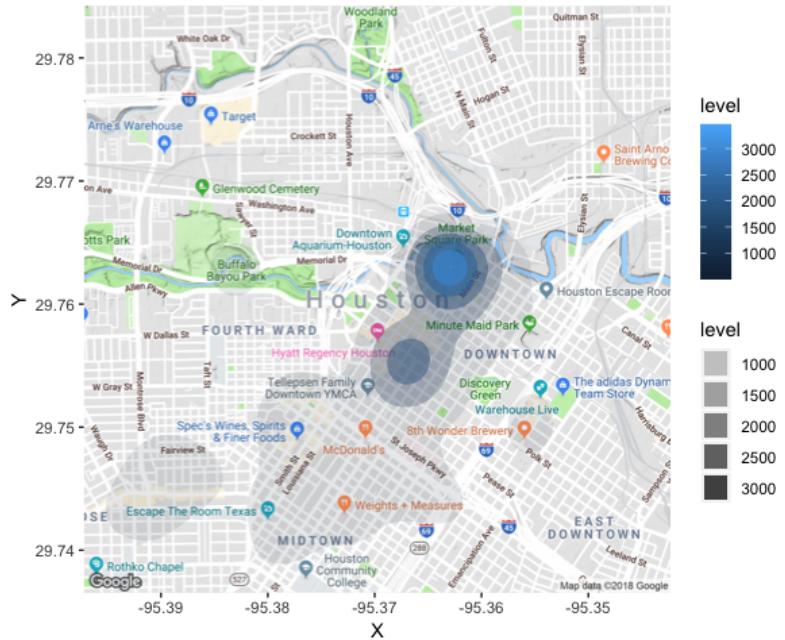


Figure 6: 2010 All Crimes

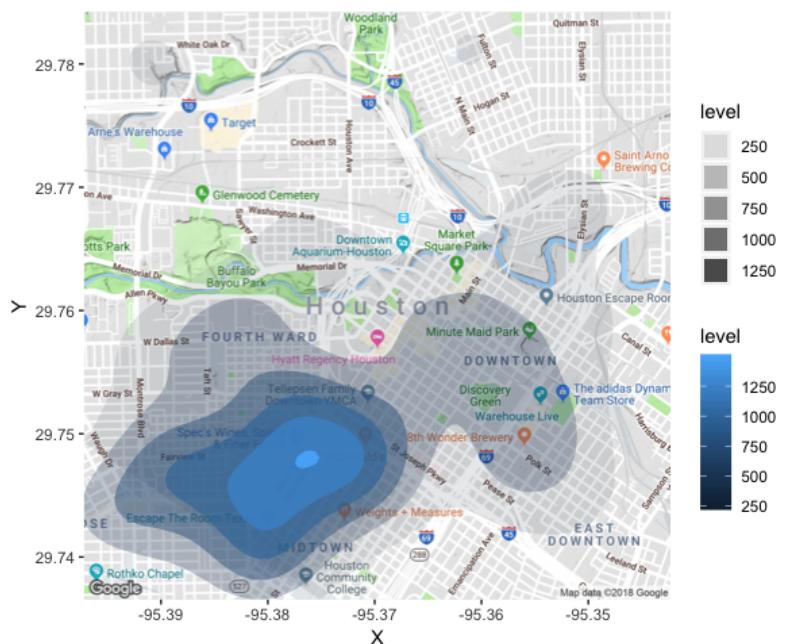


Figure 7: 2018 All Crimes

5.2 Violent Crime Comparison

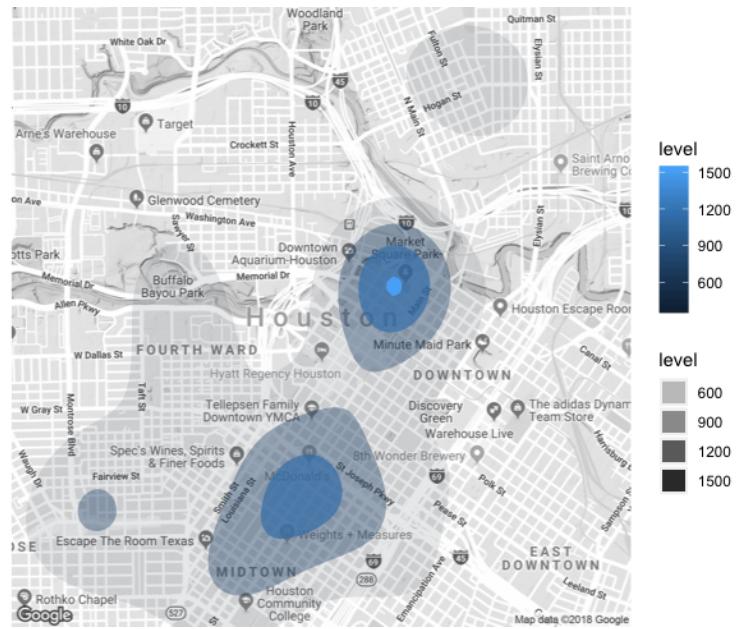


Figure 8: 2010 Violent Crimes

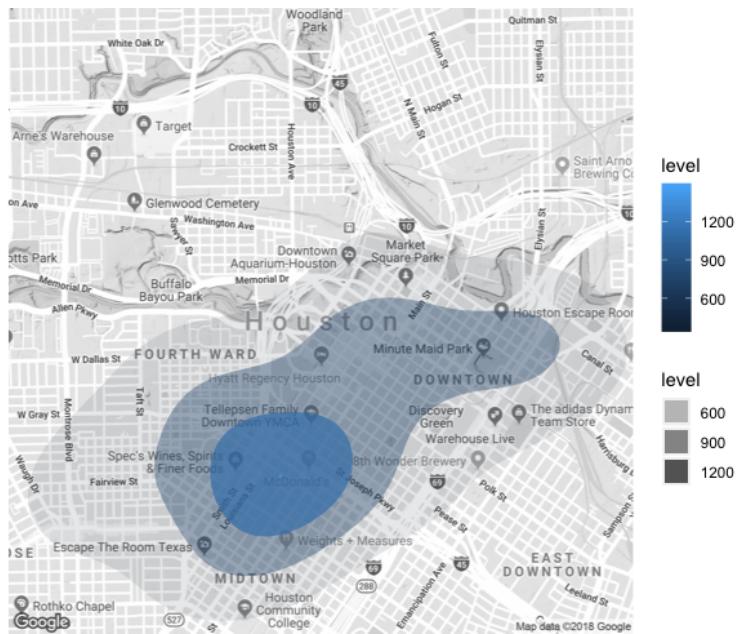


Figure 9: 2018 Violent Crimes

5.3 Crimes By Weekday

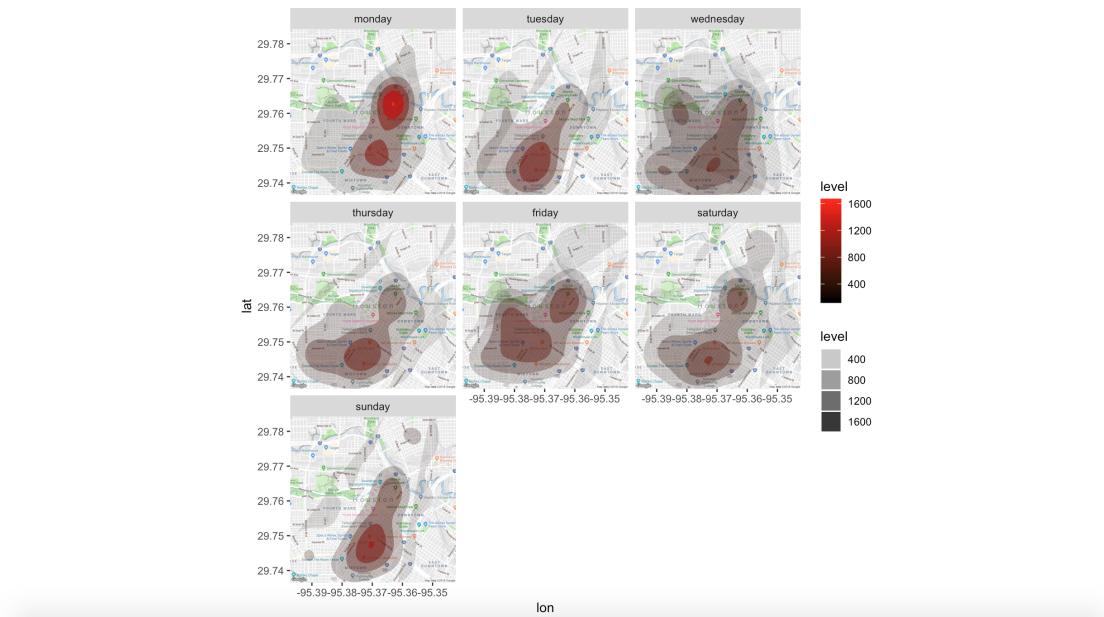


Figure 10: 2010 Violent Crimes by Weekday

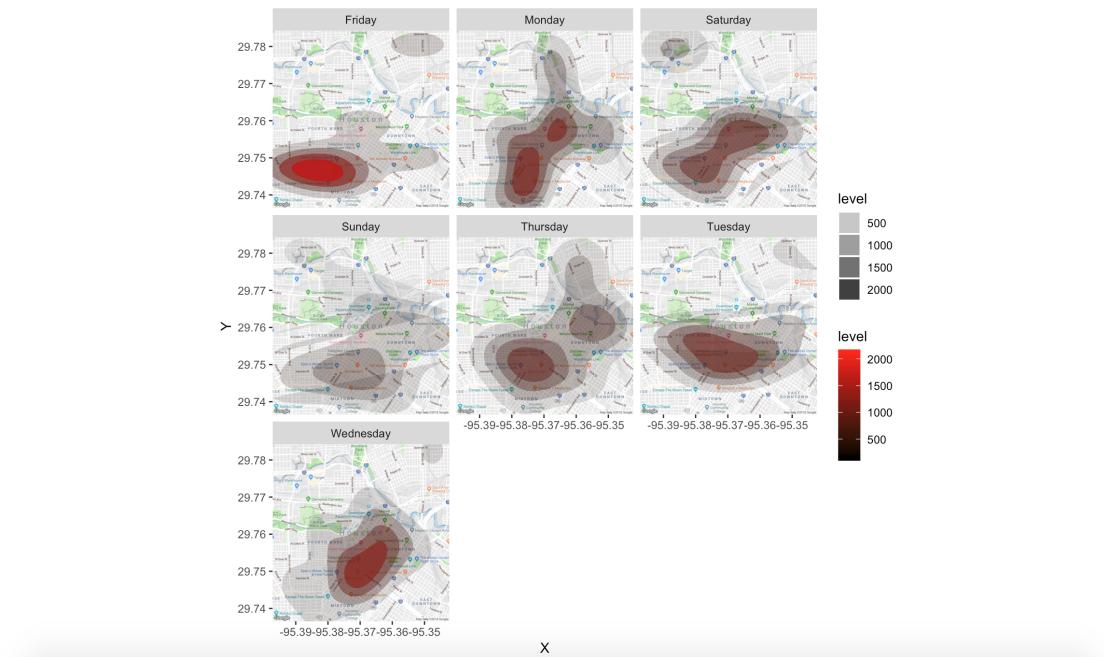


Figure 11: 2018 Violent Crimes by Weekday