**Analyzing Chicago attractions along with crime data**

Yasmeen Abu-Kheil

October 1, 2019

# **Introduction:**

Tourists find it difficult to select the optimal locations for their Hotels. They would like to select locations that are nearby attraction areas, safe to stay and walk in as well as near to famous restaurants and transportations. In this project, we will analyze and cluster Chicago city based on tourism attractions, crime rates, public transportations and famous restaurants. Then, we will provide recommendations on the top best hotel locations that optimizes tourists’ needs and requirements.

# **Data acquisition and cleaning**

## **2.1 Data sources**

In this section, the data used to solve the above problem are described as follow:

* Chicago Crime Data to determine the crime rate at different Chicago areas
* Top attractions and restaurants acquired from FourSqaure website
* Within the top attractions, find the restaurants, transportations and hotels and their geographical locations

### **2.1.1 Crime Data**

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days

(Source: <https://ibm.box.com/shared/static/svflyugsr9zbqy5bmowgswqemfpm1x7f.csv> )

### **2.1.2 FourSquare Data**

We will use the FourSquare website to request the top 50 sites in Chicago

### **2.1.3 Restaurants, transportations and hotels**

After getting the top 50 cites in Chicago, we will find category restaurants, transportations and hotels

## **Data content and cleaning**

The content of the crime data is described below:

['ID',

'CASE\_NUMBER',

'DATE',

'BLOCK',

'IUCR',

'PRIMARY\_TYPE',

'DESCRIPTION',

'LOCATION\_DESCRIPTION',

'ARREST',

'DOMESTIC',

'BEAT',

'DISTRICT',

'WARD',

'COMMUNITY\_AREA\_NUMBER',

'FBICODE',

'X\_COORDINATE',

'Y\_COORDINATE',

'YEAR',

'UPDATEDON',

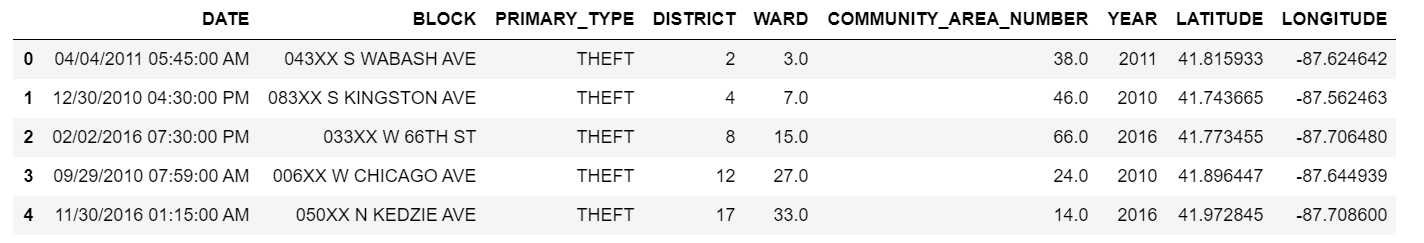
'LATITUDE',

'LONGITUDE',

'LOCATION']

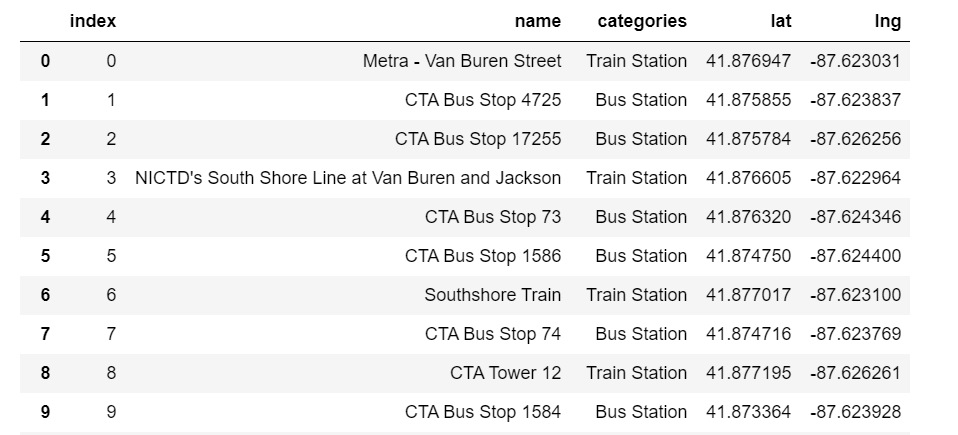
The following information were extracted from the crime data for the years from 2009:

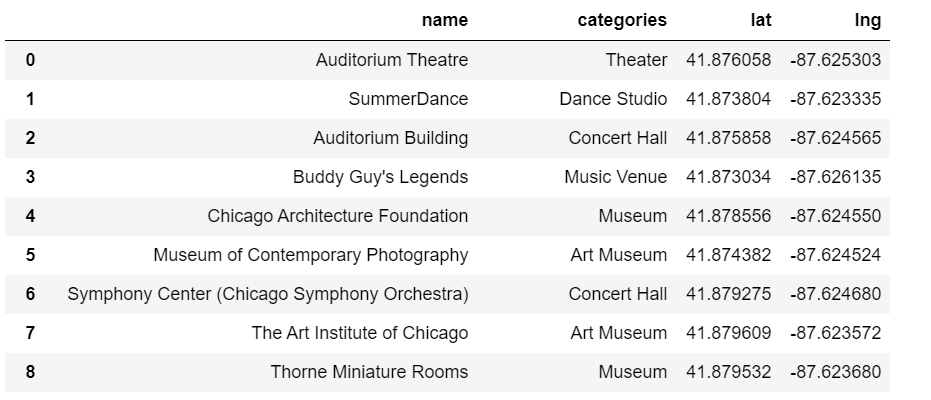
Crime\_Columns\_Interest = ['DATE', 'BLOCK','PRIMARY\_TYPE', 'DISTRICT','WARD', 'COMMUNITY\_AREA\_NUMBER', 'YEAR','LATITUDE', 'LONGITUDE'];

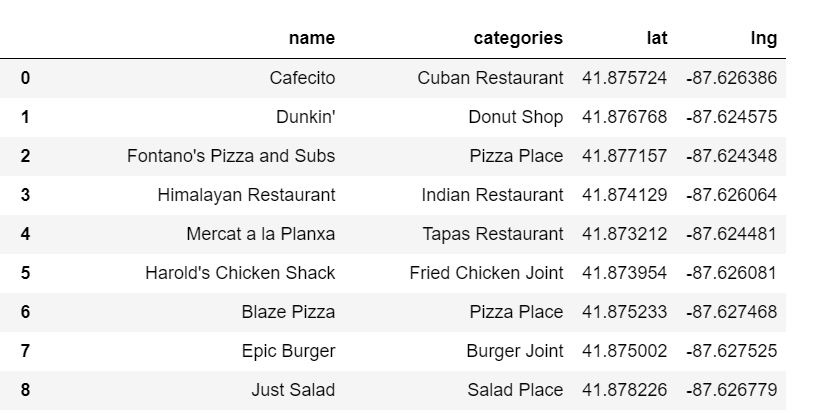


The foursquare website were used to obtain the top 50 attractions, restaurants, transportations and hotels within 500 radius of Chicago city. Samples of obtained data are shown below:







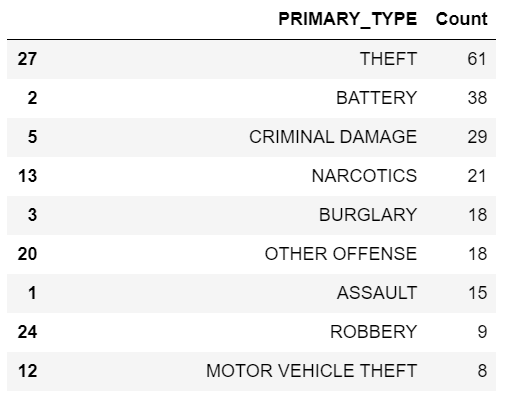


# **Methodology**

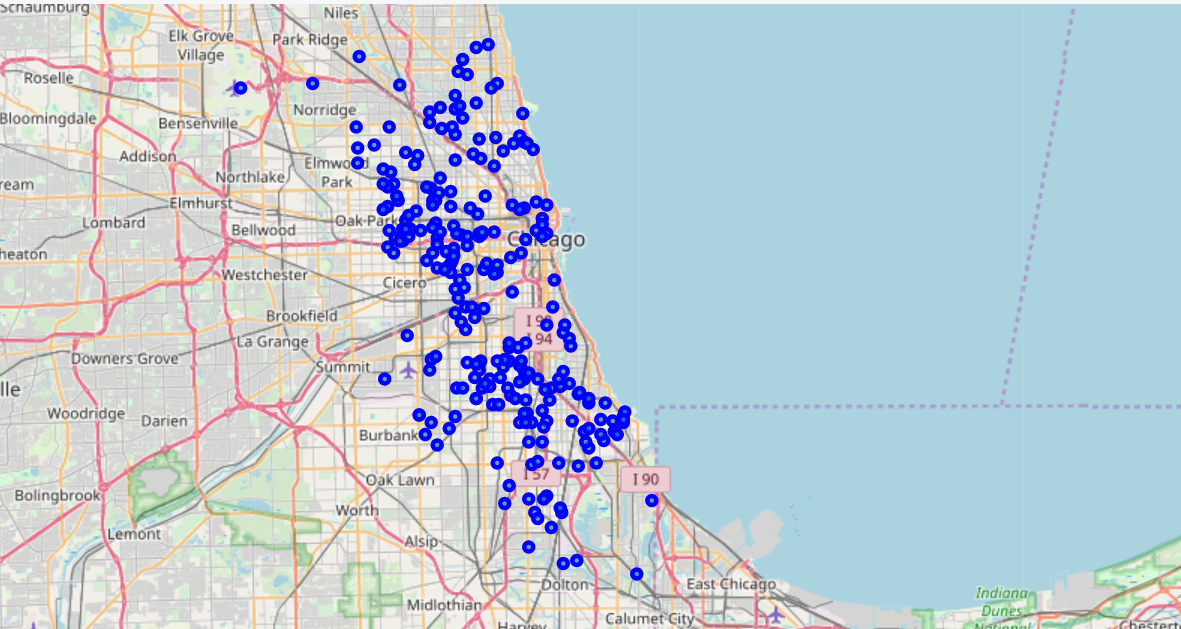
# **3.1 Data Visualizations**

# **- Crime Data**

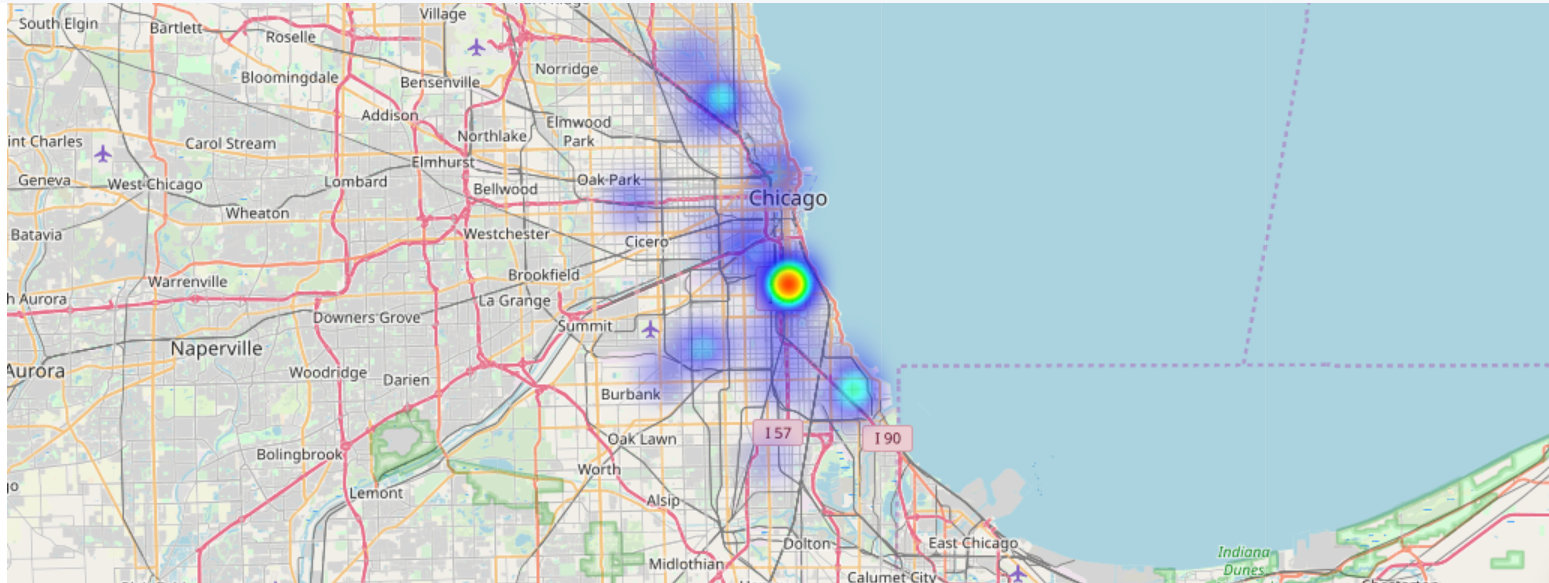
There are 29 unique crime category from 2009 and these category were ordered from the most common occurring category to the least occurring ones. Samples of the categories are shown below:



The location map that presents the crimes locations from 2009 till now in shown below:

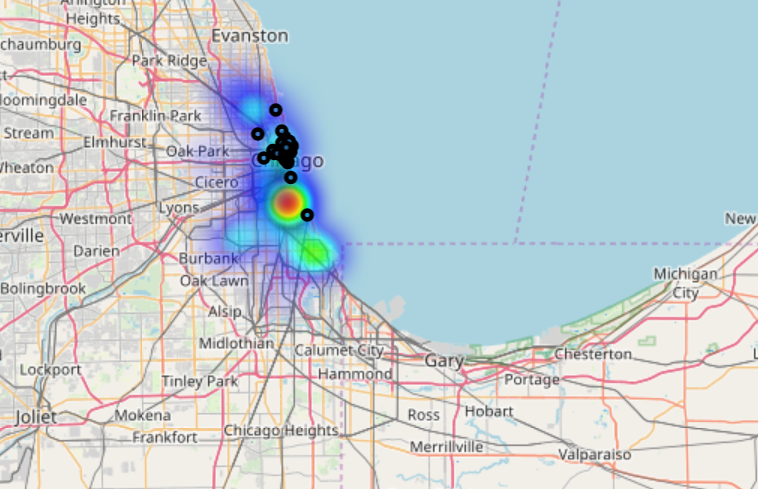


The heat map for the top three crime categories were generated:

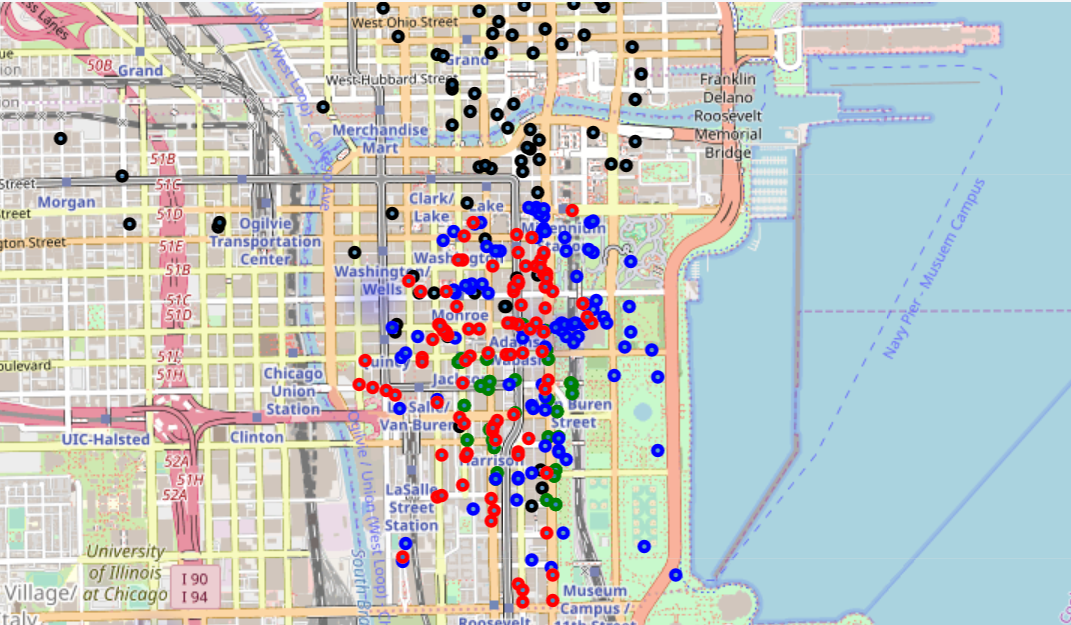


# **- Location Data**

On top of the above heat map, the extracted hotel locations were added and combined with the crime occurrence data:

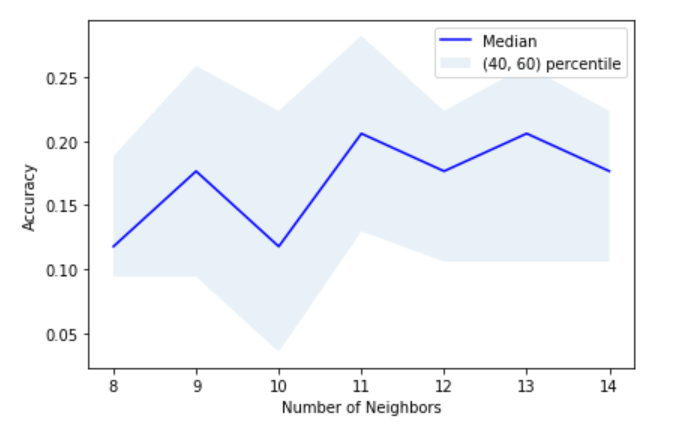


On top of the above heat map, the extracted attractions, food and transportation locations were added and combined with the crime occurrence data:

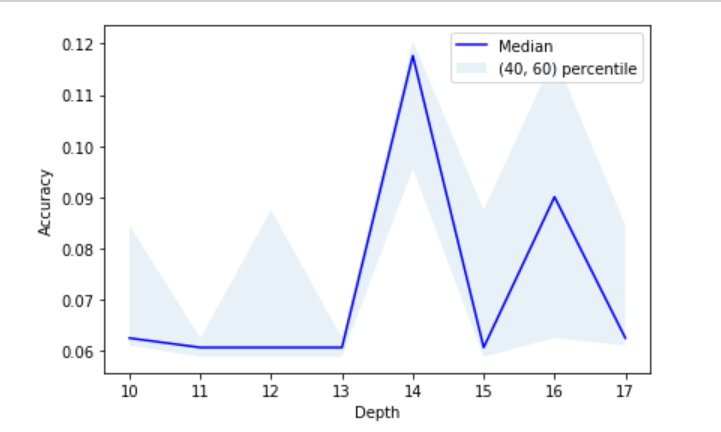


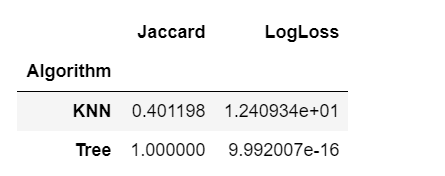
# **Data Analysis**

### **K Nearest Neighbor (KNN)**



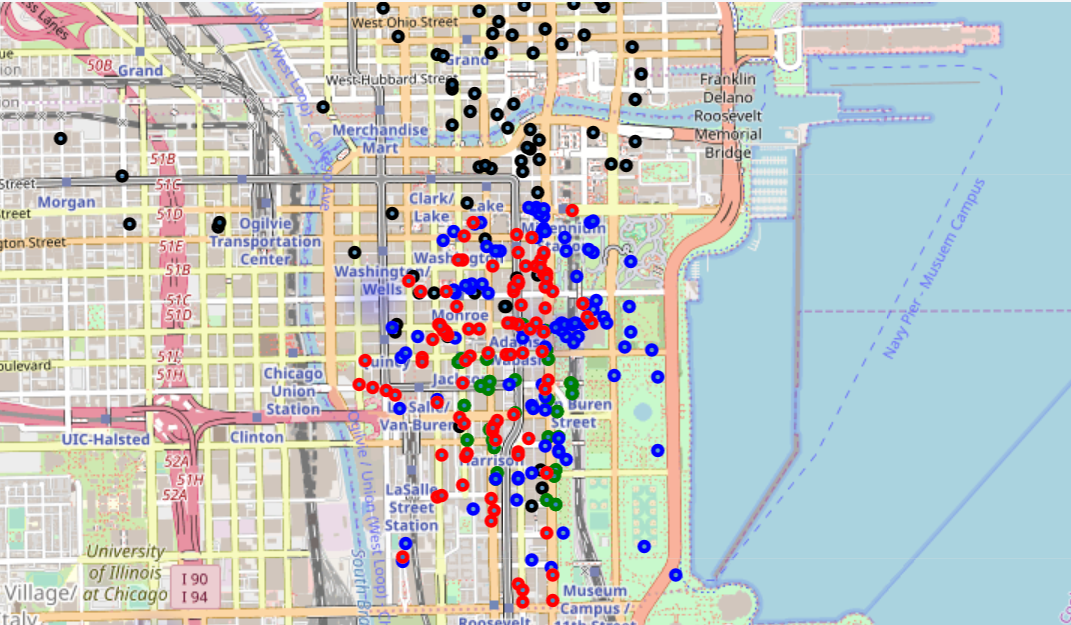
### **Decision Tree**

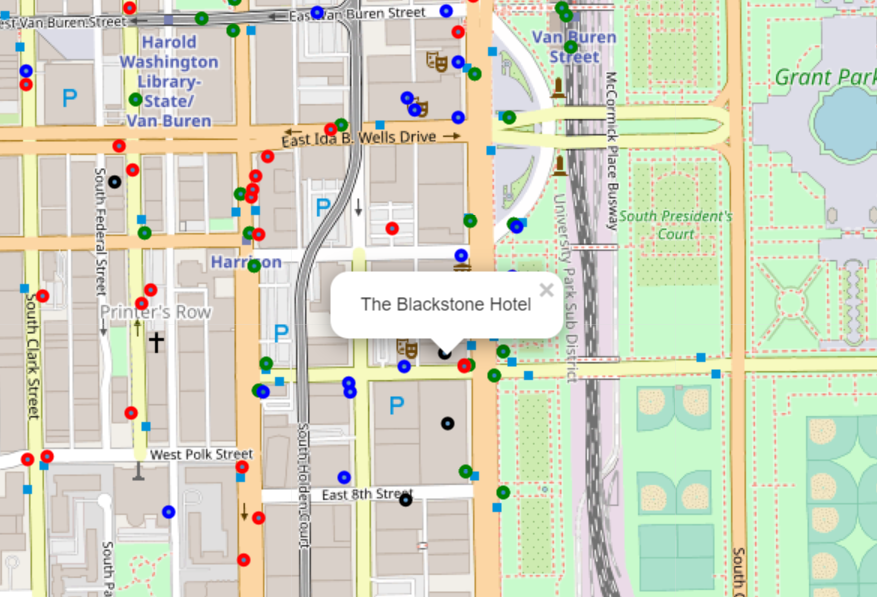




# **Results & Discussion**

Based on the below figure, tourism attractions, crime rates, public transportations and famous restaurants were visualized in the map. The best hotel that optimizes tourists’ needs and requirements is the black stone hotel.





# **Conclusions**

The use of data science can help tourists to find and select the optimal locations for their Hotels. They would like to select locations that are nearby attraction areas, safe to stay and walk in as well as near to famous restaurants and transportations. In this project, we analyzed and clustered Chicago city based on tourism attractions, crime rates, public transportations and famous restaurants. We also provided recommendations on the top best hotel locations that optimizes tourists’ needs and requirements.