

# Exploring Austin for young professional and students

## 1. Introduction

### Background

The world has gotten a lot smaller and it is a lot easier for people to travel the world and move to a new place. Many employers also require you to have international experience. It can be hard though to get familiar with a big city and stay safe. The crime rate in the United States is most likely one of the highest in the developed world. According to Wikipedia, the total violent crime rate was 382.9 and total property crime rate was 2,362.2 per 100,000 population in 2017. Therefore, it is necessary to get information beforehand to see which areas are safe and which areas one should avoid. [[https://en.wikipedia.org/wiki/Crime\\_in\\_the\\_United\\_States](https://en.wikipedia.org/wiki/Crime_in_the_United_States)]

### Problem

How can one make sure to move in a safe area with interesting venues around and to find a good apartment to live with no experience or knowledge of the new city?

### Interest

This will be of high interest for every person who is moving to a new city or even for people who just plan a vacation.

## 2. Data

### Data Description

Crime data for Austin is available at <https://data.austintexas.gov/Public-Safety/Crime-Reports-2018/vmn9-3bvu>. This data showed crime incidents for the year 2018. It included offense description, crime occurrence date and time, reported date and time, longitude and latitude and several other variables that were of no interest for this report. Unfortunately, I could not find data about Austin neighborhoods; I therefore used Google to get longitude and latitude of neighborhoods in Austin and entered them manually into a pandas data frame.

The crime data consisted of 27 variables and has 102,631 observations. I also used Foursquare data for this project to find venues of interest.

After examining the data, I realized that there were crimes in the data that occurred before 2018. These crimes got reported in 2018, but had occurred long before that. I therefore subset the data to only crimes that occurred in 2018. After taking this subset, the new sample size was 101,090 crimes.

### 3. Methodology

#### Exploratory Data Analysis

First, I looked at counts and percentages of what crimes happened the most in Austin in 2018. Family disturbances and burglary of vehicles were the top crimes with each making up over 10 percent of all crimes in 2018 followed by theft (9.1%), criminal mischief (4.5%) and assault with injury (3.3%).

I was interested to see if there was a specific time of the year or if there was any pattern when these top five crimes occurred. To do so, I extracted the month of the occurrence date variable and plotted the count by month for each crime using a line plot.

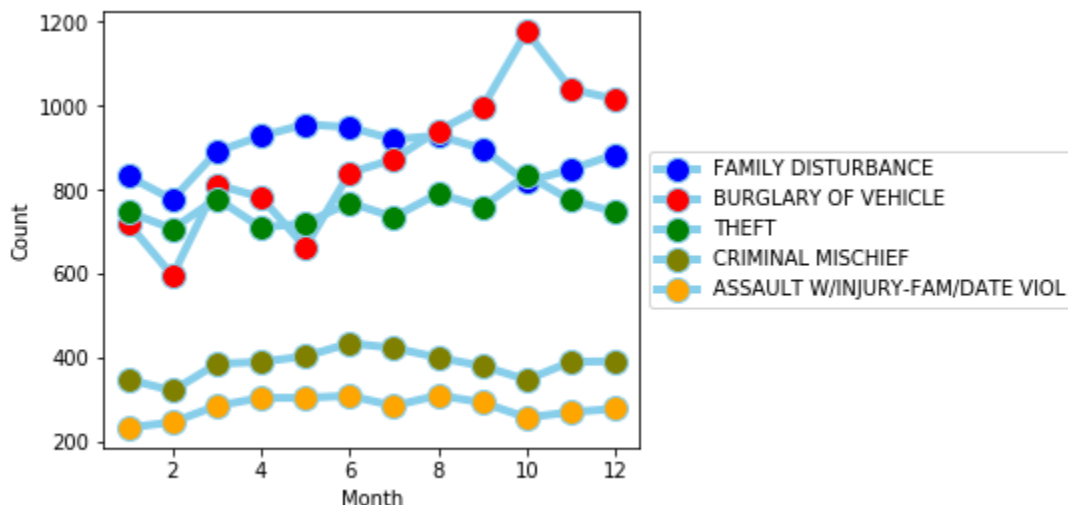


Figure 1. Line plot of crime count by month broken down by crime

One can see that burglary of vehicle peaked in October and seemed to be less in early 2018. Criminal mischief and assault with injury occurred around 50% less than family disturbance, theft and burglary of vehicle. I also looked at total crime during the day. For this analysis I used the variable "Occurred Date Time", extracted the

time, and calculated the counts by hour of the day. Crimes seemed to increase starting at 7am until midnight, before starting to decrease. Crimes hit a minimum at around 5 am. Interestingly crime count spiked at 12 pm, which might be related to people going out for lunch.

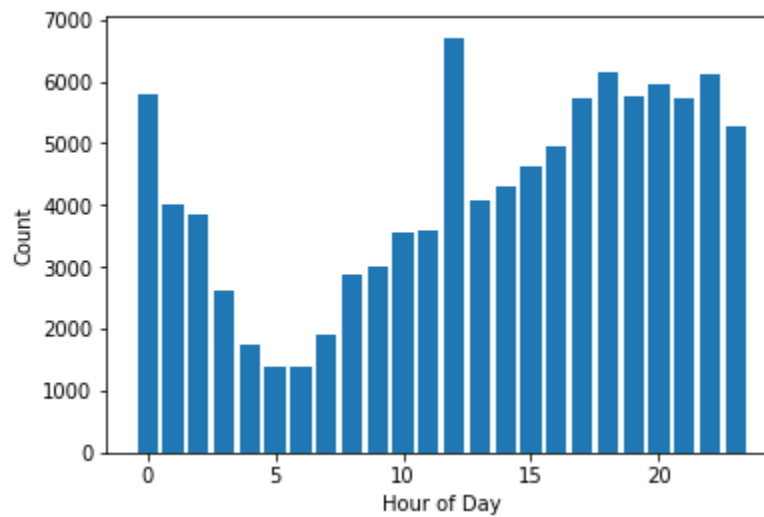


Figure 2. Crime Count by Hour of Day

Next, I wanted to see, if there was a specific area in Austin where crime is concentrated. This is important to figure out in which area you should move. I only looked at assault with injury and burglary of vehicle, since they seemed the most important to me. I had to restrict the data to 1000 crimes, because my computer was not able to load the map if it was more.

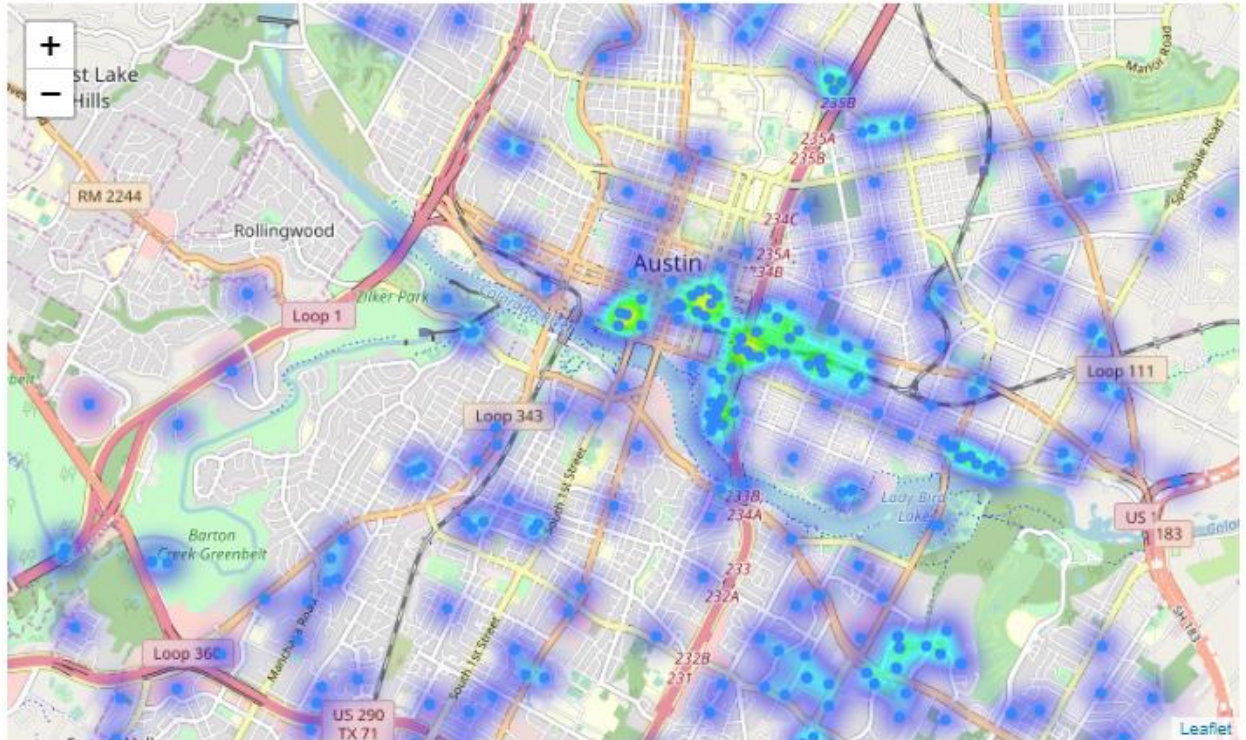


Figure 3. Crime heat map of Austin, TX for 2018

One can clearly see that crime was high around downtown Austin and just east of highway I 35. It seemed that not much crime occurred south of the Colorado River, between Zilker Park and South Lamar Blvd. For this reason, I wanted to explore the area around Barton Springs Rd, Austin TX to find apartments in a radius of 500 meters.

Using the code from the labs, I found six apartments of interest.

name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	postalCode
SoCo Apartments On The Lake	Residential Building (Apartment / Condo)	222 E Riverside Dr	US	Austin	United States	South Congress	578	[222 E Riverside Dr (South Congress), Austin, ...	[{"label": "display", "lat": 30.2560275, "lng": ...	30.256027	-97.742883	78704
Congress Square Apartments	Residential Building (Apartment / Condo)	500 Congress Ave	US	Austin	United States	NaN	152	[500 Congress Ave, Austin, TX 78701, United St...	[{"label": "display", "lat": 30.25740136890664, ...	30.257401	-97.747867	78701
Timbercreek Apartments	Residential Building (Apartment / Condo)	614 S 1st St	US	Austin	United States	Barton Springs	450	[614 S 1st St (Barton Springs), Austin, TX 787...	[{"label": "display", "lat": 30.25674011599587, ...	30.256740	-97.752039	78704
422 At The Lake	Residential Building (Apartment / Condo)	422 W Riverside Dr	US	Austin	United States	S 1st Street	124	[422 W Riverside Dr (S 1st Street), Austin, TX...	[{"label": "display", "lat": 30.259674, "lng": ...	30.259674	-97.748756	78704
Timber Creek Apartments Pool	None	614 S 1st St	US	Austin	United States	NaN	449	[614 S 1st St, Austin, TX 78704, United States]	[{"label": "display", "lat": 30.25683413117821, ...	30.256834	-97.752087	78704
Crescent Austin	Residential Building (Apartment / Condo)	127 E Riverside Dr	US	Austin	United States	Congress	414	[127 E Riverside Dr (Congress), Austin, TX 787...	[{"label": "display", "lat": 30.2555472, "lng": ...	30.255547	-97.745829	78704



Table 1. Apartments of Interest

Using Foursquare, I pulled the ratings of the apartments. Unfortunately, none of them had a rating. Next, I looked at tips and saw that users provided tips for Timbercreek Apartments. All the tip said was that the apartment complex needs an application from every person.

I conducted further analysis on venues around the apartment. Using Foursquare again, I received a good list of venues ranging from restaurants, bars, movie theaters and parks. All in a good distance from Timbercreek Apartments.

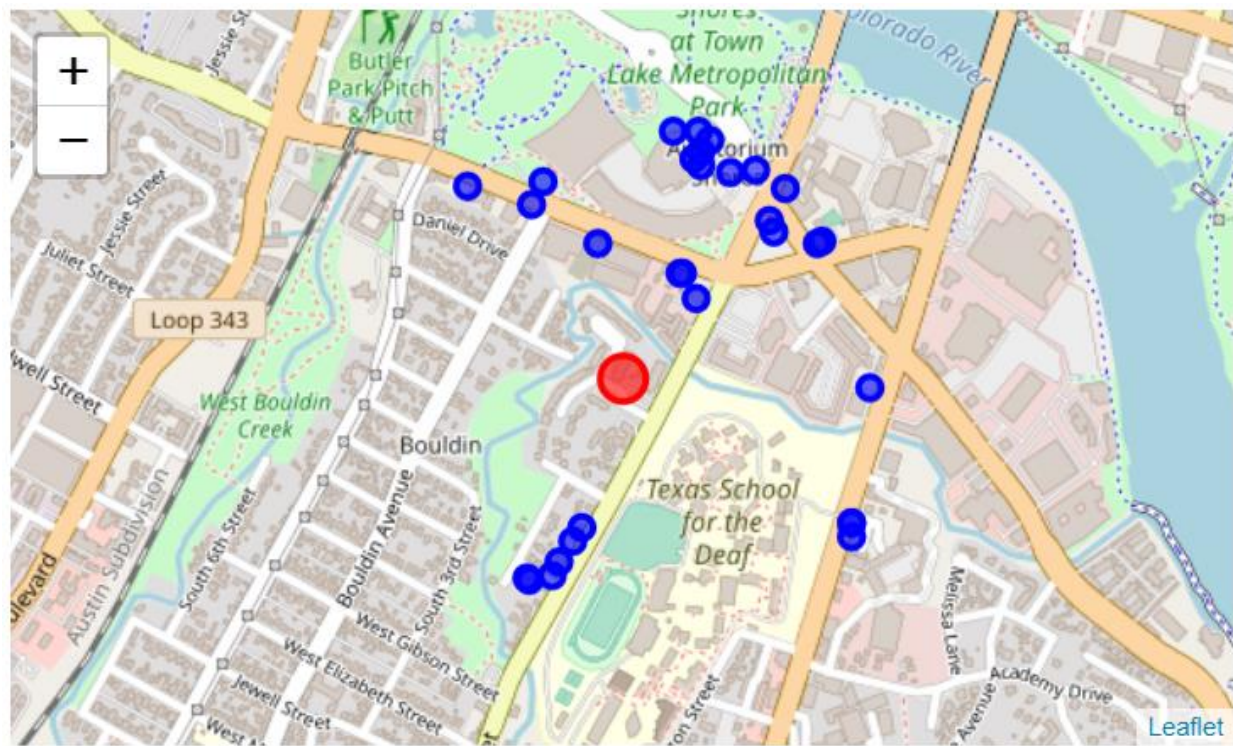


Figure 4. Red dot shows Timbercreek Apartments with blue dots showing venues such as park, restaurants and bars.

Next, I analyzed other neighborhoods in Austin to see if I can break Austin into clusters. That would make it easier for someone to orient himself or herself faster in Austin. Using Foursquare data, I pulled venues and venue categories and summarized the results in the table below which shows the neighborhood of Austin and the 1<sup>st</sup>, 2<sup>nd</sup>, etc. common venues. One can see that Downtown Austin's most common venue is hotels, East Austin's Boutique's, North, and South Austin's food places.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Downtown Austin	Hotel	Steakhouse	Capitol Building	Coffee Shop	Movie Theater	Wine Shop	Mexican Restaurant	Art Museum	Arts & Crafts Store	Bagel Shop
1	East Austin	Boutique	Café	Wine Shop	Comedy Club	Food Truck	Field	Fast Food Restaurant	Event Service	Donut Shop	Design Studio
2	North Austin	Fast Food Restaurant	Middle Eastern Restaurant	Donut Shop	Sushi Restaurant	Design Studio	Deli / Bodega	IT Services	Sandwich Place	Bus Station	Business Service
3	South Austin	Pizza Place	Mexican Restaurant	Bank	Furniture / Home Store	Convenience Store	Coffee Shop	Chinese Restaurant	Event Service	Grocery Store	Pharmacy

Table 2. x most common venue by neighborhood.

According to the table, I could have split Austin into three clusters. I used the K-means cluster algorithm to split the data into clusters. One problem of the K-means algorithm is to find the optimal k. I took advantage of the so-called “Elbow Method”. I simple ran the algorithm for a range of values of k and plot the k on the x-axis and sum of squared distances on the y-axis. The plot looks like an arm. The “elbow” shows the optimal k.

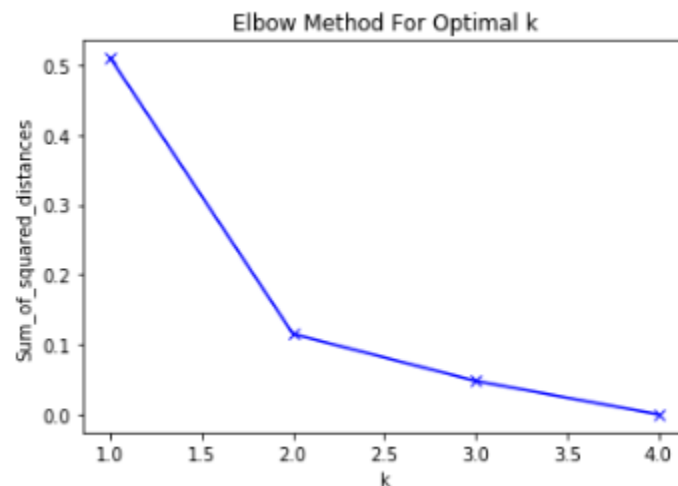


Figure 5. K-means elbow plot.

Using the elbow method I found the optimal  $k = 2$ .

Running K-means algorithm again with  $k = 2$ , it labelled Downtown Austin, South Austin and North Austin as one cluster. I named this cluster “hotel/restaurant area of Austin”. The second cluster consisted only of East Austin where you can find Boutique’s and cafes.

## 4. Results

Analyzing the crime data of Austin in 2018 I concluded that the area just south of the Colorado river was a safe area to live. I therefore explored the neighborhood around Barton Spring Rd, Austin, TX. Using Foursquare data and using a radius of 500 meters, I received a list of 6 apartments. Unfortunately, using another Foursquare data search for ratings about these apartments did not provide any results and therefore did not narrow down my search. I picked the apartment complex “Timbercreek Apartments”, which did not limit my further analysis, since all apartments were in the same area. Pulling venue data from Foursquare gave me an idea of what venues and venue categories are around the apartment complex. Bars, movie theaters, places to eat and a park were all close by, which fulfills all needs for someone who moves to a new city. Last, I conducted cluster analysis of several neighborhoods in Austin. I found that I could split Austin into two clusters. One cluster included Downtown Austin, South Austin and North Austin. Hotels and restaurants mostly defined this cluster. The other cluster included East Austin. This neighborhood embodied mostly boutiques and cafes.

## 5. Discussion

Crime analysis was limited by the fact that I could not plot more than 1000 data points on the map due to a computing power of my computer. Re-running the analysis with a more powerful computer might show a different image of the crime heat map. I could have also limited the crime analysis to a specific crime or broken it up by crime type. Furthermore, the crime data was only for 2018, which limits the decision-making. Crime hot spots can change over time and therefore could be different in the future, meaning that the area around Barton Springs Rd could become more dangerous. I also did not take into account affordability of the apartment and proximity to work place. These factors could be included in the analysis to make a clearer decision where to move.

## 6. Conclusion

With my analysis, I gave an overview over how to use crime data and data from Foursquare to find an apartment in a safe area in Austin. I tried to “simplify” the problem of orientation in a big such as Austin. I am sure it will benefit (young) employees or/and students to have good start in Austin.