Venue options in the Safest and Most Dangerous Neighborhoods of Atlanta

Ana Flávia Santos Souza

September 26, 2020

1. Introduction

In Atlanta, criminality has been a significant concern to the population for the past decades[1]. Atlanta is considered only safer than 2% of all cities and towns in the United States, and the chances of a person become a victim of "violent or property crime" in this area is one to 17[2]. This information opens to discuss the need to evaluate the crime rate more recursively. Fortunately, data have become accessible as well as methods for visualization and analysis. One use for a crime rate analysis would be to provide a reliable report to society and stakeholders. For example, locals and tourists visiting dangerous neighborhoods are more likely to get somehow assaulted. Sometimes, tourists only visit venues in certain places for not being aware of other options. Identifying venues in safer areas could give either locals and tourists more alternatives to consider and avoid unsafe places. Besides that, this could also benefit non-well-known businesses by making them get more publicized.

[1] https://en.wikipedia.org/wiki/Crime_in_Atlanta
[2] https://www.neighborhoodscout.com/ga/atlanta/crime#description

1.1. Problem

Considering how safety affects people's lives, it is essential to understand and to identify the crime rate tendencies and what areas are more critical about this topic. This project aims to describe the crime profile in the last ten years (2009-2019) and offer alternative venue options in less dangerous places in Atlanta, GA.

1.2. Target Public

The key point of this project is to provide alternative venues for both locals and tourists visiting Atlanta. Locals could better understand the safety of their city, and tourists could use these results to plan trips, be aware of risky areas, and maybe include places that they wouldn't know of without this kind of guide.

2. Data collection

To examine crime rates in Atlanta, I used a CSV file from the Atlanta Police Department website[3]. The original dataset had 19 columns and contained information about all the crime occurrences from 2009 to 2019, including date and time of crime occurrence, date and time of police report, the location the crimes occurred, type of crime, each neighborhood's coordinates, and so on. With this data, I expected to get a perspective about the safest and most dangerous areas, explore the crime rate tendency in the last ten years, and search for specific patterns in crime occurrences.

A GeoJSON file was downloaded from the Atlanta Department of City Planning[4] and used to visualize the relationship between crime rates and neighborhoods, using a choropleth map. The file had information about different administrative divisions, such as counties, cities and quadrants. Since I was only interested in the neighborhoods, I selected a file's part, referring to them.

Finally, I used the Foursquare API to get the venues in each of the safest and most dangerous neighborhoods.

[3] https://www.atlantapd.org/i-want-to/crime-data-downloads
[4] https://dpcd-coaplangis.opendata.arcgis.com/datasets/neighborhoods

3. Methods

For cleaning and analyzing the data, I used the *Pandas* library. The dataset had about 340 thousand rows and 19 columns. I decided to use only selected columns: *report_number*, *occur_date* (occurrence date), *location*, *ucr_literal* (type of crime), *neighborhood*, *latitude,* and *longitude*. After defining the attributes, the first thing I did was to check duplicated rows and delete them. One hundred and sixty-nine duplicates were dropped.

Next, I looked for missing data. In the column *neighborhood*, over 12,300 rows had missing data. Since this could impact my study results, I filled missing data with reverse geocoding, available in the *geocoder* library. However, the geocoder search revealed that these rows didn't refer to a specific community, so I also dropped them.

I created a new column called *year* to validate the date the crimes occurred. The column showed there were a few cases that had happened before 2009. Due to the selected timespan 2009 to 2019, these cases were removed as well. I also created a variable called *week_day* to analyze crimes per day of the week. All the columns had their types corrected and validated.

After that, I validated the number of neighborhoods there were in the data. According to Wikipedia[5], there are 243 neighborhoods in the city of Atlanta. The data also had that number of communities..

As the main point of interest, I observed the most dangerous and safest neighborhoods. These neighborhoods were selected according to the number of crime occurrences and then, compared. The most frequent crimes also were observed, and these two pieces of information were combined to check the incidence of most frequent crimes in dangerous neighborhoods. Crime frequency was also analyzed over time (crime rate per year, month, and day of the week from 2009 to 2019).

I plotted a crime rate choropleth map using the *folium* library. For this plot, it was necessary to use a GeoJSON file, which I cleaned with the *geopandas* library. *Geopandas* allows us to open a GeoJSON file in a *Pandas* data frame format, facilitating file manipulation. Englewood Manor, a neighborhood presented in the Atlanta Police Department data, wasn't present in the Department of City Planning data. Because of that, the choropleth map showed 242 neighborhoods.

[5] https://en.wikipedia.org/wiki/Neighborhoods_in_Atlanta

For the venues' search, I used the Foursquare API. I set a value of 1,500 meters for the radius of search and 100 for the maximum number of calls. And finally, I searched for the ten most common venues in each neighborhood.

4. Results

My analysis showed that in 10 years, Atlanta's crime rate has decreased, as shown in Figure 1. This result could represent that the police have been adopting better policies to fight criminality.

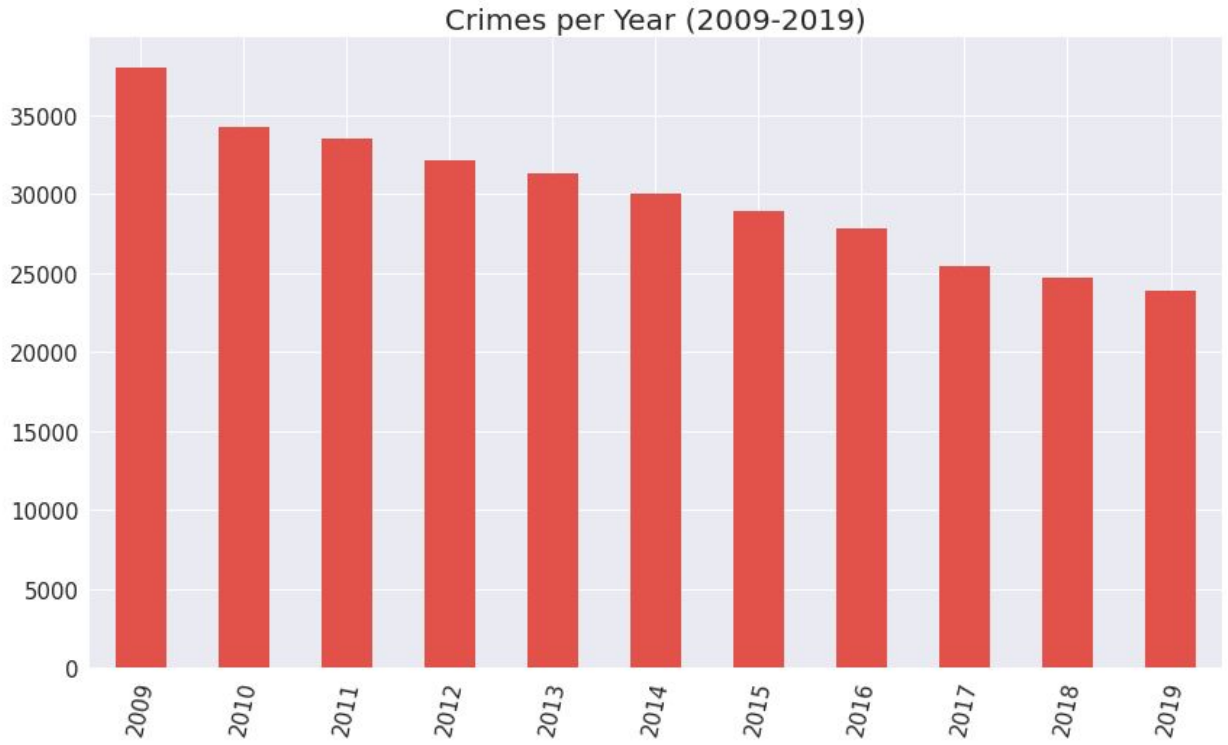

Figure 1: Crime rate per year, from 2009 to 2019.

I also found that crimes have a small better chance to occur on Fridays and Saturdays (Figure 2), which are usually the busiest days of the week in venues like bars and restaurants. In terms of months, July has been the month with the most significant criminal record from 2009 to 2019, with 30,444 occurrences. February had the smallest crime record, with 21,381 events (Figure 3). It seems that crimes tend to occur more often in hot months.

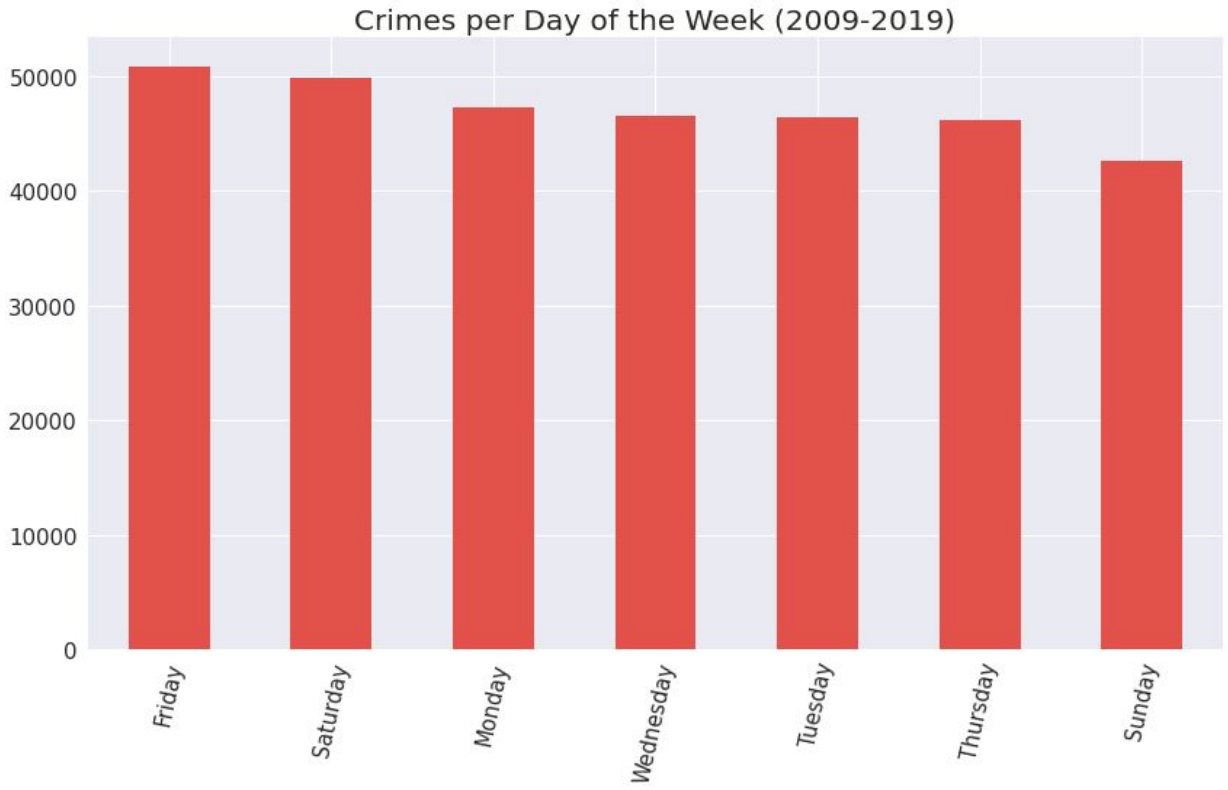

Figure 2: Crime rate per day of the week, from 2009 to 2019.

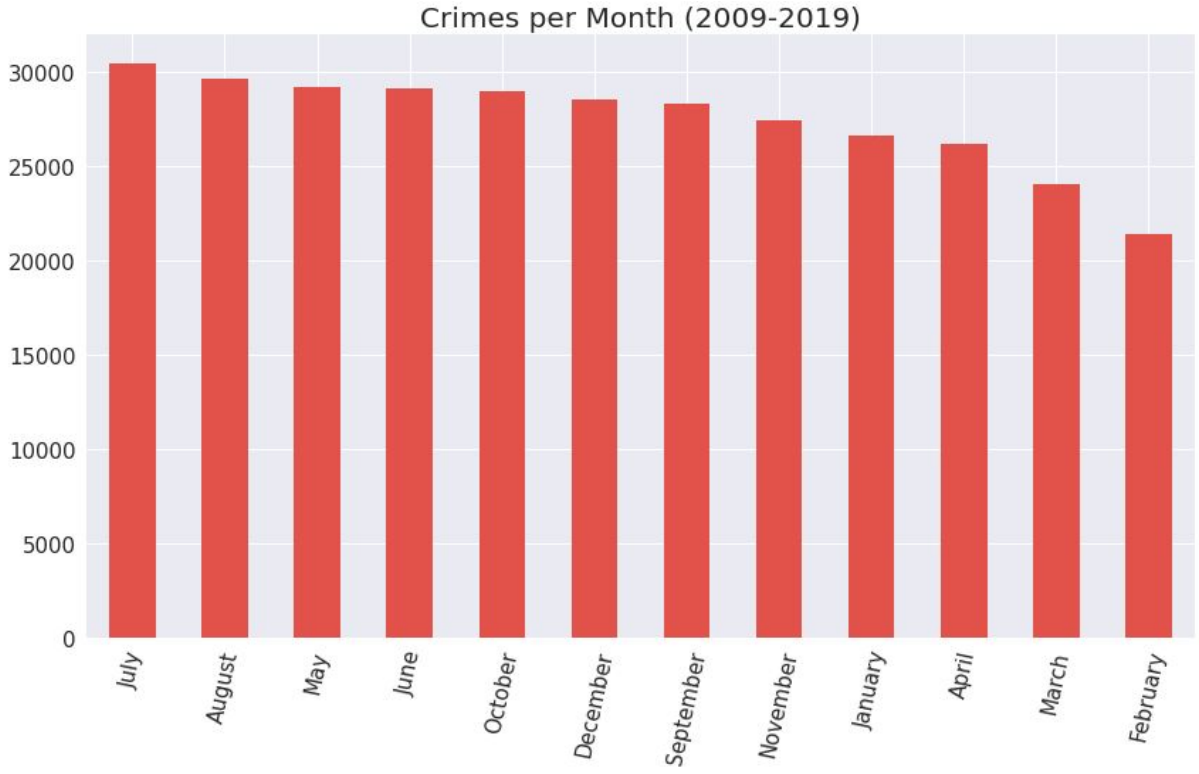

Figure 3: Crime rate per month, from 2009 to 2019.

Downtown, Midtown, Old Fourth Ward, and West End, were the most dangerous neighborhoods with an occurrence number of 25,350, 18,213, 9,924, and 8,492, respectively. The safest ones were Englewood Manor, State Facility, Edmund Park, Old Fairburn Village, Mt. Paran Parkway, and Chattahoochee, with less than 15 crime occurrences.

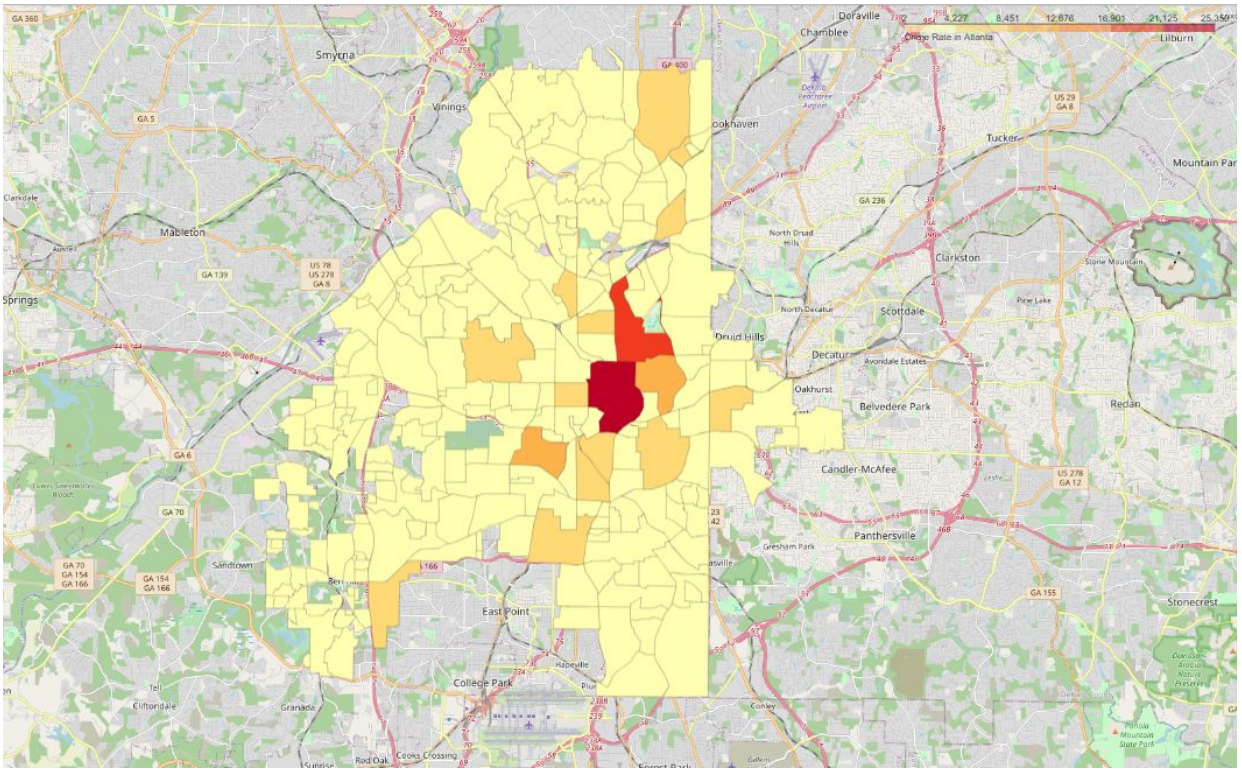

Figure 4: Choropleth Map of crime rate per neighborhood in Atlanta. The darker the area, the more dangerous it is. Those in the darkest shades of red are Downtown and Midtown.

The most common crime was larceny, both from vehicle and non-vehicle. In Downtown and Midtown, they represented over 74% of the crimes. Right after, we had burglary in residences and auto theft. The less common crime was Manslaughter, having only three occurrences in ten years. These results are shown in Figures 5 and 6.

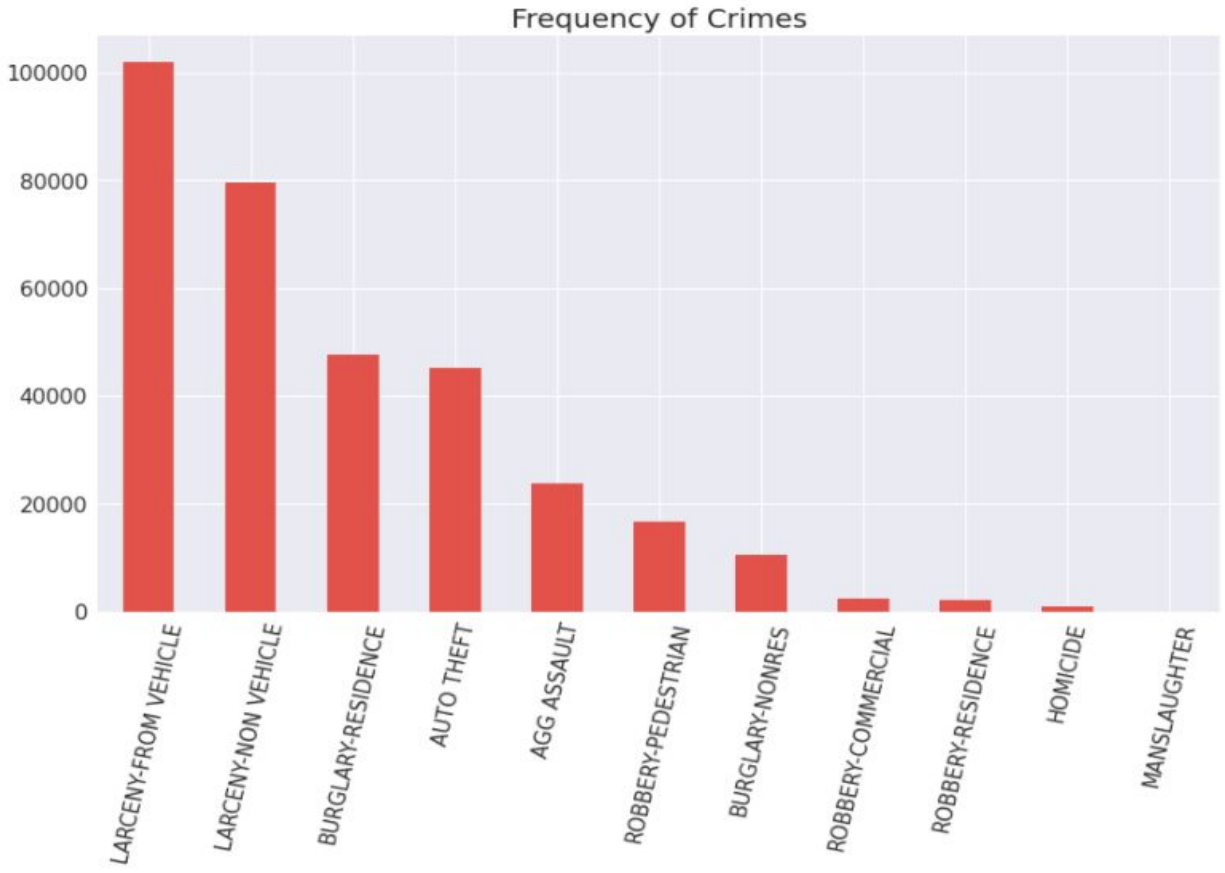

Figure 5: Frequency of each type of crime.

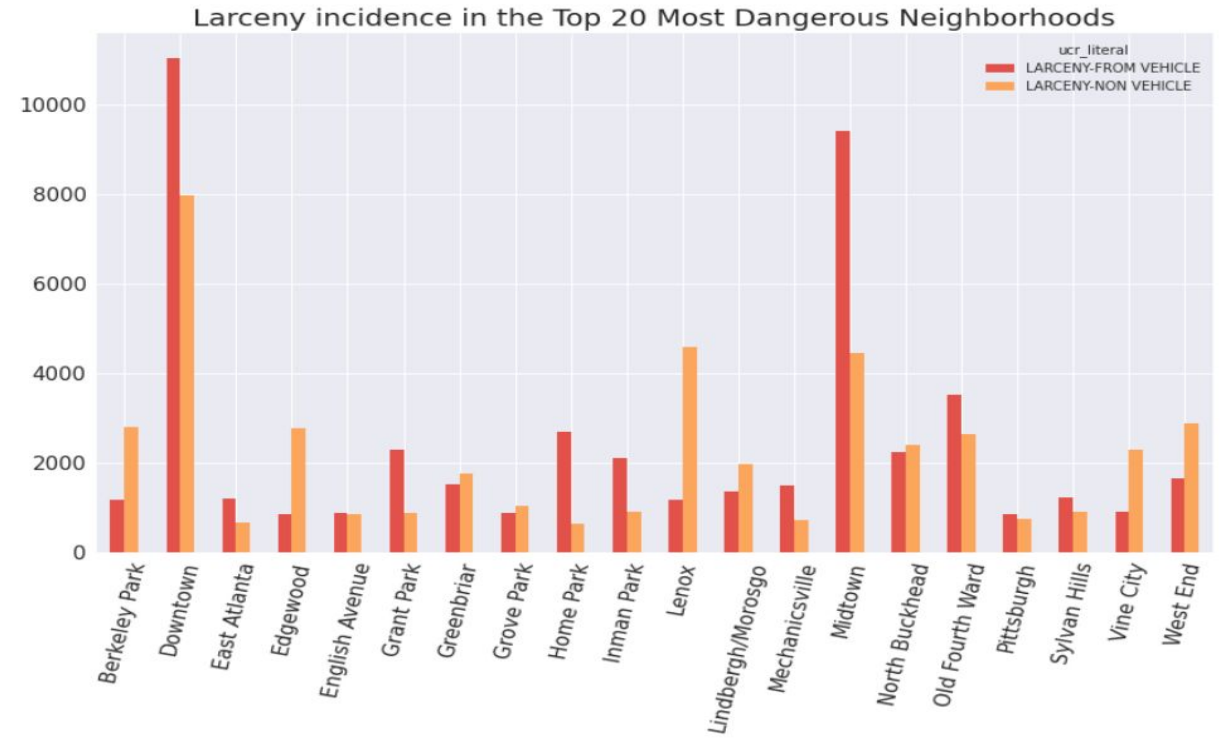

Figure 6: Larceny incidence in the top 20 most dangerous neighborhoods.

Finally, I found which venues are most common in both dangerous and safe neighborhoods[6]. As the city's central neighborhood, Downtown has most of its venues focused on bars, restaurants, and tourist places. Since the safest communities are scattered in less central areas, they are more diverse, like stores, nature-related places (parks, trails, zoos), restaurants, and so on. Overall, the types of venues between groups of neighborhoods are very similar, which indicates that either locals and tourists can enjoy their time in venues located in safer places.

5. Discussion and Conclusion

This project analyzed Atlanta's crime rate from 2009 to 2019. Using data provided by the Atlanta Police Department, it was possible to identify that the criminality has been dropping over the years. Fridays and Saturdays were the days with most crimes, on average.

Downtown is the most dangerous neighborhood in Atlanta. Because of its location and relevance, this result was already expected. Since larceny is the most common type of crime and Downtown centers some significant tourist and leisure venues, such as History museums and different kinds of restaurants, it would be recommendable that people, mostly tourists, stay alert when visiting the city center.

However, if people want to find similar types of venues in the less dangerous places, some recommendations would be Collier Hills North, which presents a wide variety of restaurants, and Fort McPherson, which also offers many types of restaurants and even an Art Gallery. For other types of venues, Mellwood has many kinds of stores, like Department, Music, and Shoes. Chattahoochee, Edmund Park, Englewood Manor, and Lake States are safe neighborhoods that present zoos, trails, and parks as venue options to be visited. They could be feasible alternatives to East Atlanta, Edgewood, and Vine City, which are more dangerous neighborhoods that also present these nature-related venues.

One suggestion for further studies is to include more variables in the analysis and to predict crime for the next years, using a regression model. Since this project's objective was to use crime data to compare neighborhoods' safety only, we didn't focus on that part, but this could be added to the present analysis.

---

[6]. It's possible to access the venues' table for each group of neighborhoods in https://github.com/anaflvss/Coursera-Capstone/blob/master/05_Capstone_Atlanta.ipynb.