Interactive Map with Optimal Location for Atlanta Police Station

Yufei Liu, Gyeongdeok Seo, Luis Lascano

Introduction

Atlanta is in the 29th percentile for safety in the U.S., making it one of the most dangerous cities among the nation ("The safest and most dangerous places in Atlanta Metro, GA", 2021). Recent events have driven crime rates even higher with an increase of 62% in the number of homicides from 2019 to 2020 ("Crime Data", 2020). This spike in crime addresses the importance of crime prevention and response time in order to ensure the city's safety. Studies have found that crime incidence increases as distance from police stations increases (Ajala & Owabumoye, 2018; Ackerman& Murray, 2004). Therefore, this project will provide recommendations on the optimal location for police stations in an effort to reduce crime rates and increase response efficiency.

Problem Definition

The goal of this project is to determine the optimal location and staff composition for police stations in Atlanta to minimize response time and ensure appropriate assistance based on the type of crime prevalent in the neighborhoods they serve. It is new to label the neighborhood in Atlanta by the demand for police force to propose a more precise shape of the police station. Previous studies mainly focus on location, but we also make an emphasis on the characteristics of each neighborhood. It makes it possible to more concrete estimate of location, and further can make it possible to propose staff composition.

Survey

Studies in this area use various techniques to map crime locations such as considering accessibility and unemployment rate in order to recommend the optimal location for police stations (Tengbeh, 2006; Turedi, 2012). However, these studies focus on locating police stations close to crime hotspots but do not give details such as the staff composition for each station. As different neighborhoods may have different types of crime, we aim to provide specific recommendations for each police station.

When deciding which variables to select, we can make use of prior studies. The prior research finds the optimal place by clarifying the objectives: making it close to the crime prone locations, keeping distance from orignal stations, and being on appropriate ground (Hasan Rahmani & Golmehr, 2015). We can use these concepts when choosing variables. Response time is also an important factor when it comes to effectiveness of criminal apprehension. Jordi Blanes and Tom Kirchmaier argue in their research that "minimising response time is a highly effective policy in terms of apprehending a larger percentage of criminals" (Blanes & Kirchmaier, 2015). According to research by the National Police Foundation, variables most likely to have an effect in response time are: "the distance an officer must travel," and "the officer's driving speed" (Pate et al,. 1976).

In addition, real world police officers often spend time doing work not directly related to crime. Research suggests that officers spend a lot of time in traffic-related work, so our variables will also include traffic information (Terrill *et al*., 2014). This study only ends with investigating how they spend time, but we will further apply this to determine the optimal location of police stations. We will also evaluate how much time it takes for citizens to get

help from the local police by neighborhood. The study first divides the time as three intervals for police officers from getting an emergency call to clearing the work of call: waiting time, travel time, and on-scene time (Zhu *et al*., 2021). Because each time interval has a different meaning, our research also divides time more precisely. This study is about redesigning the district for each police station in Atlanta, and we'll apply this concept of the study to select optimal locations for police stations.

Method

Data Preprocessing - We downloaded data for Atlanta 2009-2021 crime from the Atlanta Police Department, Atlanta 2009-2020 car crash from Georgia Department of Transportation, and Atlanta 2019 demographic data by neighborhood from DataNexus. We decided to only use crime and crash data from 2015-2019 for the most recent five years of data excluding COVID. For crime and car crash data that had null values for the neighborhood attribute, we used the Google geocoding API to find the neighborhood using the latitude and longitude. Below is a brief summary of the dataset.

	Size	Number of Columns	Number of Records
Crime Data	44.3MB	10	367795
Car Crash Data	120MB	22	289624
Demographics Data	13KB	20	103

Neighborhood Clustering & Zone Formation - We reduced the number of variables by choosing those with standard deviation > 0.01, then we used K means to cluster the neighborhoods using these factors. The elbow method was used to find the optimal range of K. For each k in the optimal range, we plotted the clustered neighborhoods and determined the final k value by choosing the k that provided the most clear separation so that we can group similar neighborhoods into zones easily.

Police Station Location Optimization - We created a mixed integer linear program to minimize the distance between the potential police station locations to crime locations. The candidate locations were generated by taking the intersection of vertical and horizontal lines plotted one mile apart from each other over each zone. The crime locations were divided into violent and non-violent crime types, and a weight of 0.7 and 0.3 were given to each type, respectively, so that more importance is given to minimizing distance to violent crime type locations.

The number of police stations n needed for each zone is calculated as follows. The ratios we used for guidelines were determined from Atlanta Police Department's number of total officers of 2,046 (Haney, 2020; Atlanta Police Department*)*.

$$n = population\ of\ neighborhood \times \frac{1\ police\ officer}{243.75\ people} \times \frac{1\ police\ station}{186\ police\ officers}$$

Below shows the objective function and constraints of the model.

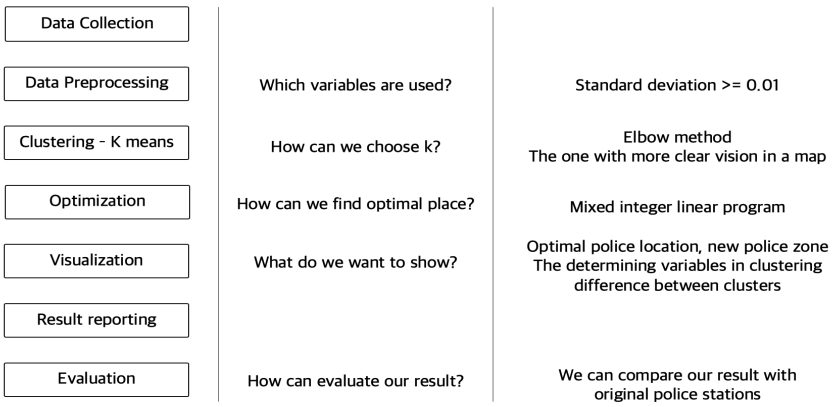
$$minimize \sum_{i \epsilon I} x_i \left(\sum_{j \epsilon J} 0.7 \times DH_{ij} + \sum_{k \epsilon K} 0.3 \times DL_{ik} \right), \; s.t. \sum_{i \epsilon I} x_i = 1$$

For each zone, I is the set of all potential locations, J is the set of all violent crime locations, and K is the set of all non-violent crime locations. The binary variable x_i is 1 if potential location i is chosen and 0 otherwise. The parameter DH_{ij} represents the distance in miles from location i to j, and DL_{jk} is the distance is miles from location i to k. The parameter n is the number of police stations needed for the zone. For zones with n>1, we divided them further into n sub-zones of roughly equal population for the purpose of optimization so that each sub-zone has one police station, and the number of police stations in sub-zones added together equals n. For each determined optimal location, we suggested the staff composition based on the characteristics of each cluster.

Visualization - We created a choropleth map using d3 that shows the clustered neighborhoods, the current police zones, new police zones, current police station locations, and optimal police station locations. Tooltips for each neighborhood provide additional information for crime in the area such as crime rate and breakdown of crime types. Tooltips for police stations provide information on the staff breakdown. Box plots of the most important features used in clustering are also shown.

There was no research which mixes clustering and optimization. This mixing helps to determine the location and staff composition of the police station. The prior research focuses on finding the optimal location, but we also take into account the staff composition by using the demographic and car crash data.

Experiments & Evaluation

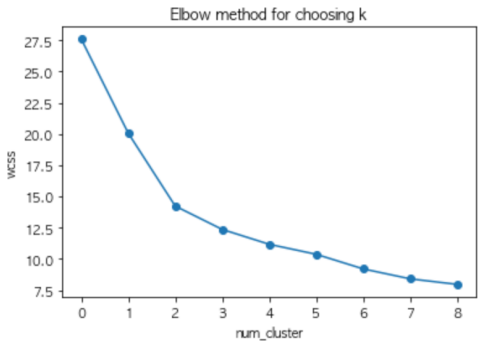


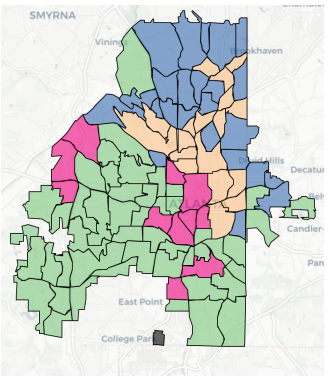
This is our process for visualizing the map, and the questions our experiments are designed to answer. The table below shows the variable used in clustering. The underlined values are deleted because the standard deviation is <0.01.

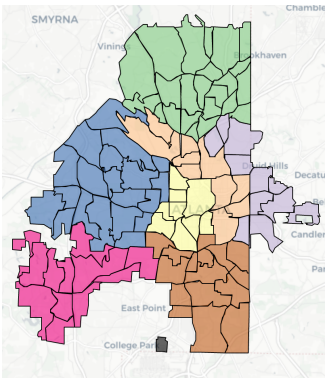
Type	Variables
Crime Type	Agg assault, Auto theft, Burglary, Burglary-nonres, Burglary-residence, Homicide, Larceny-from vehicle, Larceny-non vehicle, Manslaughter, Robbery-commercial, Robbery-pedestrian, Robbery-residence

Time	fall, spring, summer, winter, 1-7am, 12-17pm, 18pm-0am, 8-11am
Age	% Age 5-9 years 2019, 10-14, 15-19, 20-24, 25-34, 35-44, 45-54, 55-59, 60-64, 65-74, 75-84
Population	pop (The rate of population), count_pp (Number of crime per person)
Economy	% In labor force 2019, % Population 15 years and over divorced 2019, % Public transportation (excluding taxicab) to work 2019, Median household income 2019, % Unemployment Rate 2019
Education	% Population 25 years and over 9th-12th grade no diploma 2019
Crash	crash_pp (Number of crash per person)

We clustered the neighborhoods using k = 1 to 8 and using the elbow method (shown left), we found that the optimal range of k is between 4 and 6. By plotting out the results of each k in the optimal range, we found that k = 4 gives us the best divisions between clusters on the map (shown middle). We then grouped neighborhoods in the same cluster into police zones (shown right).







The number of police stations needed for each new zone is shown below.

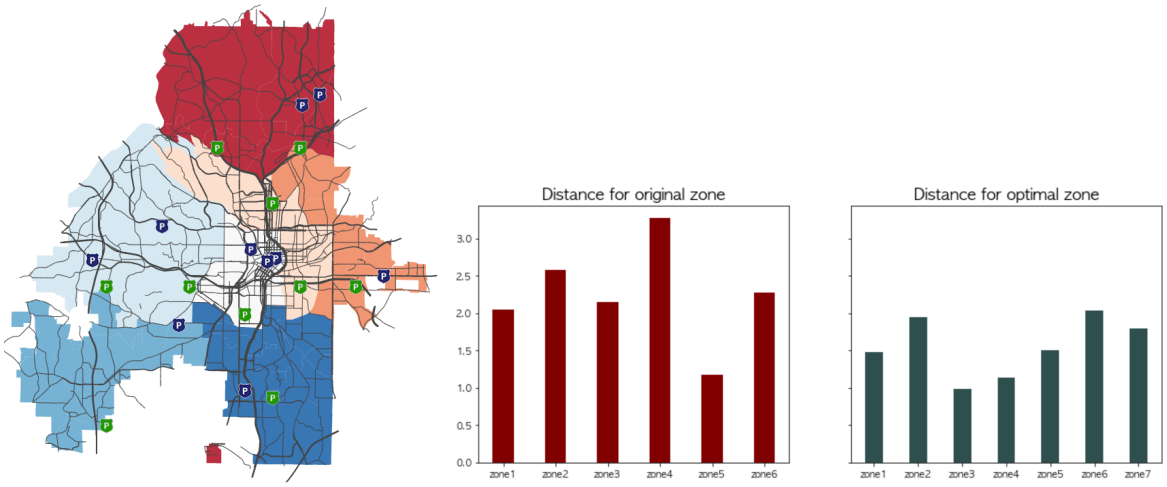
Zone	1	2	3	4	5	6	7
# Stations	2	1	2	1	2	1	1

To validate our results for station locations, we calculated the distance from police stations to crime locations. The distance for each zone is calculated as follows:

$$\sum_{i \epsilon I} \frac{x_i}{n} \left(0.7 \times \frac{\sum DH_{ij}}{n_h} + 0.3 \times \frac{\sum_{k \epsilon K} DL_{ik}}{n_l} \right)$$

The image of the map below shows the current police stations in blue and the optimized locations in green. The bar graph shows the calculated average distance from police station to crime locations for each zone. As shown, our optimized results have lower average distance and lower fluctuation. The police stations in the original system are located 2.25

miles from crime locations on average, and the police stations in the optimized system are located 1.55 miles from crime locations on average, which is a 31% reduction.



Each cluster shows a different pattern of how crime happens. Because of this, the staff composition in each police station must be different to deal with several types of crimes. Based on the cluster results, we suggest the staff composition for each cluster in the conclusion section.

Conclusion & Discussion

We obtain several observations from the cluster results.

Cluster	Characteristic
1	**High**: crime in summer, burglary residence, auto theft, unemployment rate, crime in 1-7 hour, public transportation usage, divorce population rate **Low**: labor force, crime rate, larceny from vehicle, crime in winter, car crash rate, median household income, age in 25-34, education level **Staff composition**: Theft is taking property belonging to another person by extortion or embezzlement, so it is the crime directly related with the safety of a person rather than larceny. Patrol officers should be increased and especially at dawn.
2	**High**: labor force, crime rate, crime in night, larceny from vehicle, crime in winter, age in 25-34, education level **Low**: crime in summer, burglary residence, auto theft, unemployment rate, crime in 1-7 hour, public transportation usage, divorce population rate **Staff Composition**: More nighttime patrol officers are needed. More security cameras nearby expensive properties are helpful for finding the properties. The officers who survey the whole city with this security camera are needed.
3	**High**: labor force, larceny from vehicle, crime in winter, median household income, education level **Low**: crime in summer, crime rate, car crash rate, unemployment rate, crime in 1-7 hour, public transportation usage, divorce population rate **Staff Composition**: By reducing the money spent on dealing with traffic, this region can spend more budget in preventing crime related with larceny. The officers who survey the whole city with security cameras are needed.
4	**High**: crime in summer, crime rate, car crash rate, auto theft, unemployment rate,

	crime in 1-7 hour, public transportation usage, divorce population rate **Low**: labor force, larceny from vehicle, crime in winter, median household income, education level **Staff Composition**: More officers dealing with traffic are needed. This region also needs more patrol officers at dawn, and in summer.

- There is a strong correlation between crime rate and the following variables.
    - Higher education & lower crime rate
    - Higher income & lower crime rate
    - Lower divorce rate & lower crime rate
    - Higher public transportation usage & higher crime rate
    - Higher unemployment & higher crime rate
- Cluster 1 & Cluster 4 have a similar trend with crime. However, cluster 4 has a higher crime rate and cluster 1 shows the highest burglary-residence rate. Cluster 4 also shows the highest car crash rate, but cluster 1 shows a very low rate in car crashes.
- Cluster 2 & Cluster 3 have a similar trend with crime. However, cluster 2 has a lower burglary residence rate and cluster 3 has a lower car crash rate.
- For areas with higher crime rate, the crimes usually occur in sumer; the crime types are more break ins, auto theft, pedestrian robbery and aggravated assault (more serious crimes)
- For areas with lower crime rate, the crimes usually doesn't occur at dawn (1am - 7am); the crime types are more larceny from vehicle (more passive crimes)

Project Plan

		Assignment
Gather & clean data 10/18 - 10/29	Gather demographic data	Seo
	Gather crime reports	Liu
	Gather car crash data	Liu
	Clean & prepare data	Liu, Seo
Build model 11/1 - 11/5	Build clustering model	Seo
	Build optimization model	Liu
Write report	Write progress, final report	Liu, Seo
Analyze results 11/9 - 11/12	Determine staff composition	Seo
Visualize results 11/15 - 12/26	Graph police stations	Lascano
	Graph clustered results	Lascano
	Add interactive elements	Lascano

References

1. Ackerman, W. V., & Murray, A. T. (2004). Assessing spatial patterns of crime in Lima, Ohio. *Cities*, *21*(5), 423-437.
1. Ajala, O. A., & Owabumoye, B. R. (2018). Influence of police stations' location on crime incidence in developing countries like Nigeria. *International Journal of Society Systems Science*, *10*(2), 132-147.
2. Atlanta Police Department. (n.d.). *Crime Data*. Atlanta Police Department. Retrieved October 16, 2021, from https://www.atlantapd.org/i-want-to/crime-data-downloads.
3. *Atlanta Police Department*. About APD | Atlanta Police Department. (n.d.). Retrieved December 5, 2021, from https://www.atlantapd.org/about-apd.
4. Blanes, V., & Kirchmaier T. (2015). The Effect of Police Response Time on Crime Detection. *Centre for Economic Performance, CEP Discussion*(1376).
5. CrimeGrade.org. (n.d.). *The safest and most dangerous places in Atlanta Metro, GA ...* Crime Grade. Retrieved October 16, 2021, from https://crimegrade.org/safest-places-in-atlanta-ga-metro/.
6. Haney, A. (2020, December 9). *Atlanta police down 220 officers since start of January, Department says*. 11Alive.com. Retrieved December 5, 2021, from https://www.11alive.com/article/news/local/atlanta-police-numbers-2020/85-a54ace87-bed2-4d41-87b1-150a9b877672.
7. Hasan Rahmani, M., & Golmehr, E. (2015). The Ideal and Practical Pattern of Positioning Police Stations in Qazvin City and its Comparison with the Present Condition. European Online Journal of Natural and Social Sciences: Proceedings, 3(4 (s)), pp-320.
8. Pate T., Ferrara., Bowes R., Lorence J. (1976). Police Response Time Its Determinants And Effects. *National Police Foundation*. 21-23. Retrieved from: http://www.policefoundation.org/wp-content/uploads/2015/07/Pate-1976-Police-Response-Time.pdf
9. Peng Lu et al. (2019) IOP Conf. Ser.: Mater. Sci. Eng. 569 052106
10. Tengbeh, S. (2006). *Crime analysis and police station location in Swaziland: a case study in Manzini* (Doctoral dissertation, Stellenbosch: University of Stellenbosch).

11. Terrill, W., Rossler, M. T., & Paoline III, E. A. (2014). Police service delivery and responsiveness in a period of economic instability. *Police Practice and Research*, *15*(6), 490-504.
12. Turedi, S. (2012, July). Spatial analysis of Ohio police station locations using geographical information systems. In *Proceedings of the 3rd International Conference on Computing for Geospatial Research and Applications* (pp. 1-6).
13. Zhu, S., Wang, H., & Xie, Y. (2021). Data-Driven Optimization for Police Zone Design. arXiv preprint arXiv:2104.00535.

Appendix A

- Atlanta 2009-2021 crime data
    - https://www.atlantapd.org/i-want-to/crime-data-downloads
- Atlanta 2009-2020 car crash data
    - https://gdot.numetric.net/crash-data#/
- Atlanta 2019 demographics data
    - http://data.neighborhoodnexus.org/