Alexander BeVier
Dr. Ajogbeje
QMSS 301 003
28 September 2023

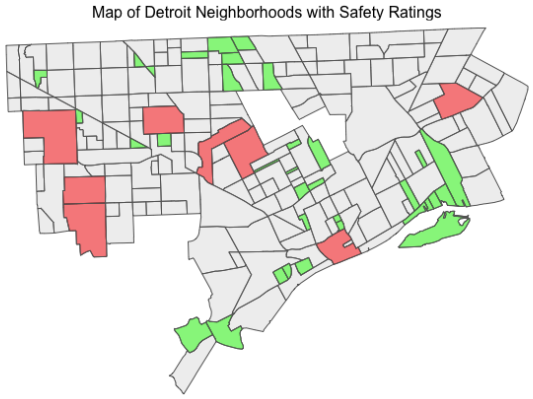
Project 1: Geospatial Analysis Lab Report

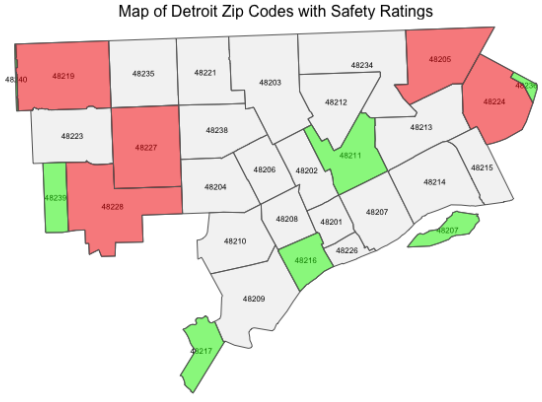
INTRODUCTION

Although the city of Detroit is going through a phase of growth and popularity, it is still considered to be one of the more dangerous cities in the United States. For many Detroit residents, navigating around the city safely to complete essential tasks can pose quite a challenge. The present project attempts to use data collected on grocery stores in Detroit as well as crime in Detroit to produce geospatial maps of the city. The goal of the project is to identify which areas of Detroit are safest, and which are the most dangerous. Additionally, this project will identify the grocery stores within these "safe" and "dangerous" areas, and all crimes within half a mile and mile of each grocery store (during the year 2021).

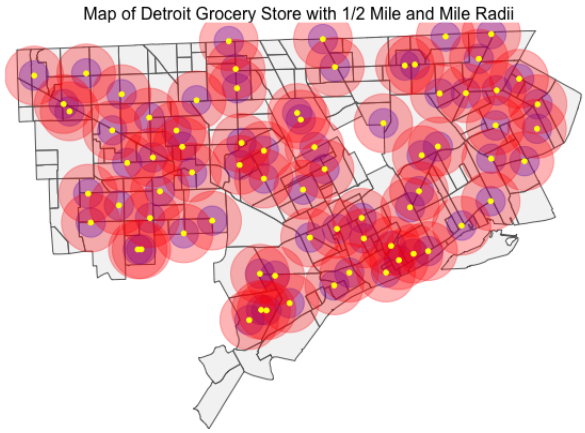
METHODS

Packages 'dplyr', 'sf', 'tmap', and 'writexl' were used for this project. There were 4 root data sets that were used for this project: a shape file that gives us information about Detroit neighborhoods, a shape file that gives us information about Detroit zip codes, a shape file that gives us information about the nature and location of crimes in the Detroit area, and a shape file that gives us information about locations of different grocery stores. After data was imported, cleaned, and projected, crime and neighborhood data were combined to create a map of crime by neighborhood (due to the size of the crime shape file, crime was filtered to only crimes in 2021). The neighborhoods with less than 100 crimes in 2021 were highlighted in green, and the neighborhoods with over 1,200 crimes were highlighted in red. A similar map was made, except organized by zip code instead of neighborhood (as zip codes are bigger than neighborhoods, zip codes with less than 1,000 crimes were highlighted green, and zip codes with over 5,000 crimes were highlighted red). This gives us a general idea of which areas in Detroit are "dangerous" and which areas are "safe".





Now, I tied in what we know about area safety ratings with the data on grocery stores. First, I created a range, or "buffer" of a half mile and a mile of each grocery store, then plotted a map with said buffers. Grocery stores were represented by bright yellow dots, radii of a half mile were represented by a blue-shaded circle, and radii of a mile were represented by a red-shaded circle. Filtering the 2021 crime data for just robberies, I then created two data matrices, telling us whether or not each robbery was within a half mile and/or within a whole mile of each grocery store. Those matrices were then applied (tallied) and bound together along with the grocery store data. This data frame now gives us the name, timestamp, and location information of each grocery store as well as how many robberies occurred within a half mile of that grocery store and how many robberies occur within a mile of that grocery store.



FINDINGS

Task 1. The first map represents the city of Detroit, split up and bordered by neighborhood. Each neighborhood was shaded in grey. All of the neighborhoods with less than 100 crimes were shaded in green, and all neighborhoods with more than 1,200 crimes were shaded in red. Generally, all of the "safe" neighborhoods were located on the North and Southeastern sides of Detroit, while the "dangerous" neighborhoods were located on the East and West sides of Detroit.

The second map also represents Detroit, although this time it is sectioned by zip code. All zip code areas were shaded grey, with zip codes with less than 1,000 crimes shaded in green and zip codes with more than 5,000 crimes shaded in red. The geolocational pattern of the first map loosely applies to this map; "safe" neighborhoods are located primarily on the Southeast side and "dangerous" neighborhoods are primarily on the East and West sides. The numbers that are in the middle of each area are the actual zip codes of each area.

Task 2. The map made in this task served a slightly different purpose than the first two. The city is still split by neighborhood and shaded grey, except different key points are shown. All of the yellow dots represent each recorded grocery store in Detroit. Each red circle represents a

one-mile radius within each grocery store, and each blue circle represents a half-mile radius around each grocery store. As described in the above 'Methods' section, these radii were applied and bound together with the grocery store and robbery data to produce two new variables: how many robberies within a half mile of each grocery store (robs_half), and how many robberies were within a mile of each grocery store (robs_whole). Descriptive statistics (min, max, median, mean, first and third quartile) were taken on each of the two variables. The results of the descriptive statistics are as follows:

For robs_half:	**For robs_whole:**
Minimum = 0.0	Minimum = 6.00
1st Quartile = 5.25	1st Quartile = 21.25
Median = 9.00	Median = 30.00
Mean = 10.11	Mean = 31.73
3rd Quartile = 14.00	3rd Quartile = 39.75
Maximum = 44.00	Maximum = 75.00

CONCLUSIONS

Task 1. As this was considered more of a preliminary study, not many conclusions can be drawn from the data specific to task 1. However, it is important to note that generally, the geospatial distribution of "safe" and "dangerous" areas between neighborhood and zip code data generally align (i.e., from the neighborhood map to the zip code map, "safe" and "dangerous" areas are generally in the same place). It should also be noted that "safe" and "dangerous" areas are more segregated when split up by zip code as opposed to split up by neighborhood. This could be due to zip codes covering more geospatial areas, and thus a "dangerous" neighborhood could be dissolved in the zip code map if it is surrounded by "safe" neighborhoods. For this reason, it may be beneficial to use the neighborhood map to give us the most detailed information.

Task 2. The descriptive statistics, at large, show that there are more robberies within a mile radius than within a half-mile radius. These results would necessarily stand as evidence against expectations or previous hypotheses but are still interesting to consider nonetheless. It should be noted that there are, on average, about three times as many crimes within a mile of each grocery store, and the IQR for a one-mile radius is two times that of a half-mile radius. This means that as you increase the radius surrounding the grocery store, there will be more reported robberies within the radius of the store. From the combination of data presented, Detroit grocery shoppers are able to make decisions on where to shop based on the relative safety of the area and how many robberies (in 2021) there were in the surrounding area around the grocery store.