Introduction

Crime has often been analyzed in the context of racial or economic demographics. In order to study a new connection, I opted to look at crime density in relation to both population count, and change in Market Value Analysis (MVA), as subjectively measured in numeric rankings. I had originally planned to include bus stop points and dangerous buildings (condemned to be torn down). For the buildings, I found that they provided nothing of value to the analysis, and the same can be said for the bus stops, as they were so numerous as to be almost uniform in study area.

As well, there is always the potential for biases, as well as inaccuracies, in any dataset. With the crime data in particular, there exists the potential for the data to be 'sanitized' to force or hide any truth in the data. This bias can begin as early as generating the report, and is nearly impossible to root out. Regarding the MVA data, bias can be presented in the observer's personal opinion of the neighborhoods, or financial incentives. As a third party to these datasets, I can only hope they were compiled as accurately and honestly as possible.

Kansas City is unique in that the municipality spans two states, and is actually two separate but conjoined cities. The data were provided only for Missouri, and I opted to focus on just this city. The Core-Based Statistical Area (CBSA) encompasses both cities, and an area significantly greater than the MVA neighborhoods or the relevant tracts. It is almost certain that the CBSA covers the surrounding suburbs.

Data and Methods

Data for Kansas City, Missouri was provided by Open Data KC, and the tracts were provided by the United States Census Bureau. The crime points span nationwide, but are specific to Kansas City, MO. This is likely due to the fact that the victim and perpetrator were located away from each other and maybe even the crime itself (i.e. the identify theft point in Miami). The MVA dataset is comprised of neighborhoods that are often similar to the tracts, but evidently distinct.

Rstudio was utilized to acquire the data from Application Programming Interfaces (APIs). From there I filtered out the important tracts, transformed data by renaming and adding fields, and performing basic analysis by calculating new data (i.e. CperPop = a count of crime points in a tract divided by that tract population). Kernel density was determined in Rstudio, but did not provide any insight. The Kolmogorov-Smirnoff Test was also applied, followed by Moran's I. Finally, I exported my data as shapefiles, and utilized QGIS for a visual presentation.

One of the greatest difficulties I encountered was in properly displaying the 'Crime per Population' (CperPop) field. I wanted to make it as a percent (100 * value), but the resultant numbers were too small to visualize properly. Multiplying it by 1000 means I can say 'X crimes per 1000 people', and any other number is too large to be valuable.

Analysis, Conclusions

The Kolmogorov-Smirnoff Test(K-S Test) is a 'goodness of fit' test that tells us if the data is normally distributed, aka if our data is randomly spaced or not. Using both the longitude and crime points, we see a very large gap between the lines representing our data and the expected observations. In other words, the gap is large, and the data are not random.

Moran's I is a global statistic, a single value for the whole spatial pattern. It doesn't tell us the location of clusters, but it does tell us that we have them. It also does not identify any cause for clustering. In my plot, red indicates that there is no clustering (aka it's random), and blue indicates that there is strong clustering.

What did surprise me is that the crime density hotspot was not centered on the tracts with the highest population count. One possible reason for this may be found by looking at time of incident, and at zoning laws for those tracts. I may find that a majority of crimes happen in commercial sectors, and at a certain time of day. For further analysis, I opted to look at the racial breakdown of suspects. I found that crimes were very evenly committed by Caucasian and African American suspects. Looking at tracts showed that only a few held a large number of crimes.

Overall, this project highlighted the value of analyzing your data in multiple ways. A map provides context that no other format can, but it can also tell an untrue story. Bar charts depicting crimes by race and tract showed another facet of the story. Statistical methods other than the K-S Test may also show another perspective. One analysis that I should have considered was distance between crime points and bus stops. Another is zoning laws. Still another is the neighboring city, and how that may affect Kansas City, MO.

More importantly, I did not find any connection between MVA neighborhood rating changes and crime numbers. I would conclude that crime is linked to something else, and not to any MVA statistic. Overall, the class was great for teaching me a lot about Rstudio, and some of QGIS. I personally would prefer if there was more time spent on QGIS, but for the amount of hands-on work I got with Rstudio, I won't complain. I definitely feel more confident with both, and would like to use them in a professional setting. The project went well, but I mostly stuck to what I learned from past labs. I know that my charts and final poster could look more aesthetically pleasing, but I'm very happy with the content.

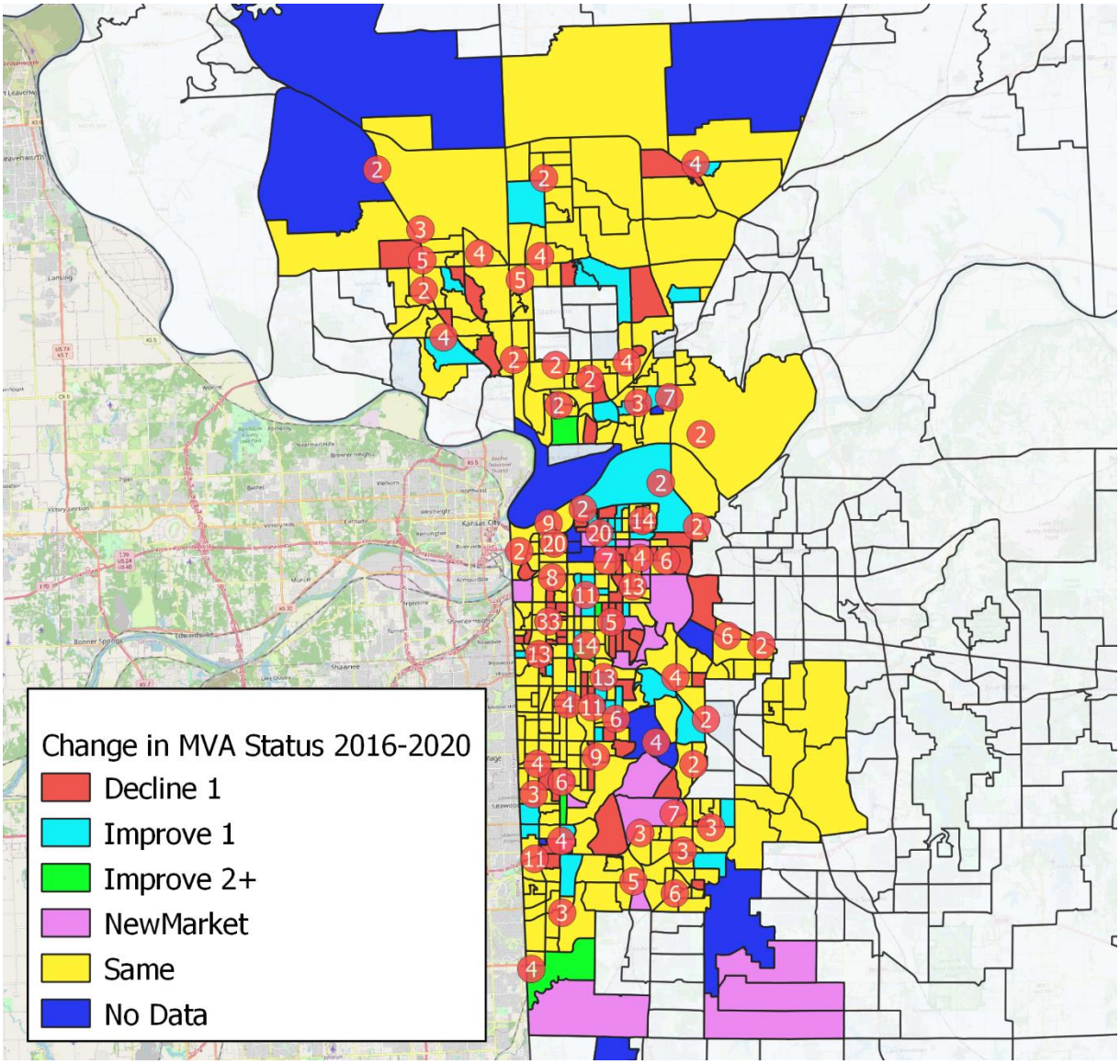

Figure 1. Market Value Analysis compared to the number of crimes. Contrary to expectations, the majority of crimes were not linked to neighborhoods of declining sales, but mostly those with no change in rank.

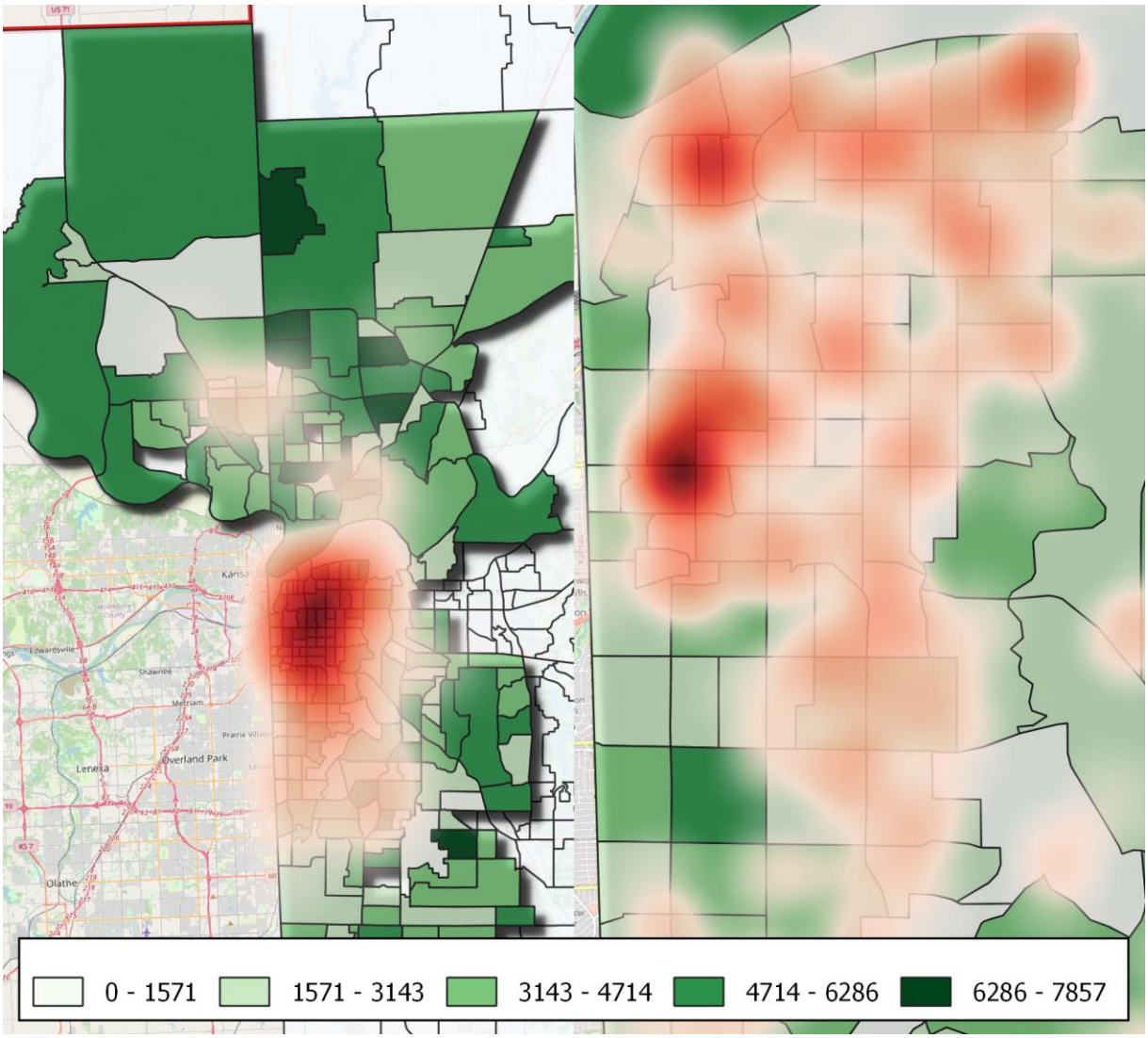

Figure 2. Crime density in relation to tract population counts. Again contrary to expectation, the hotspot is not located in tracts with particularly high populations.

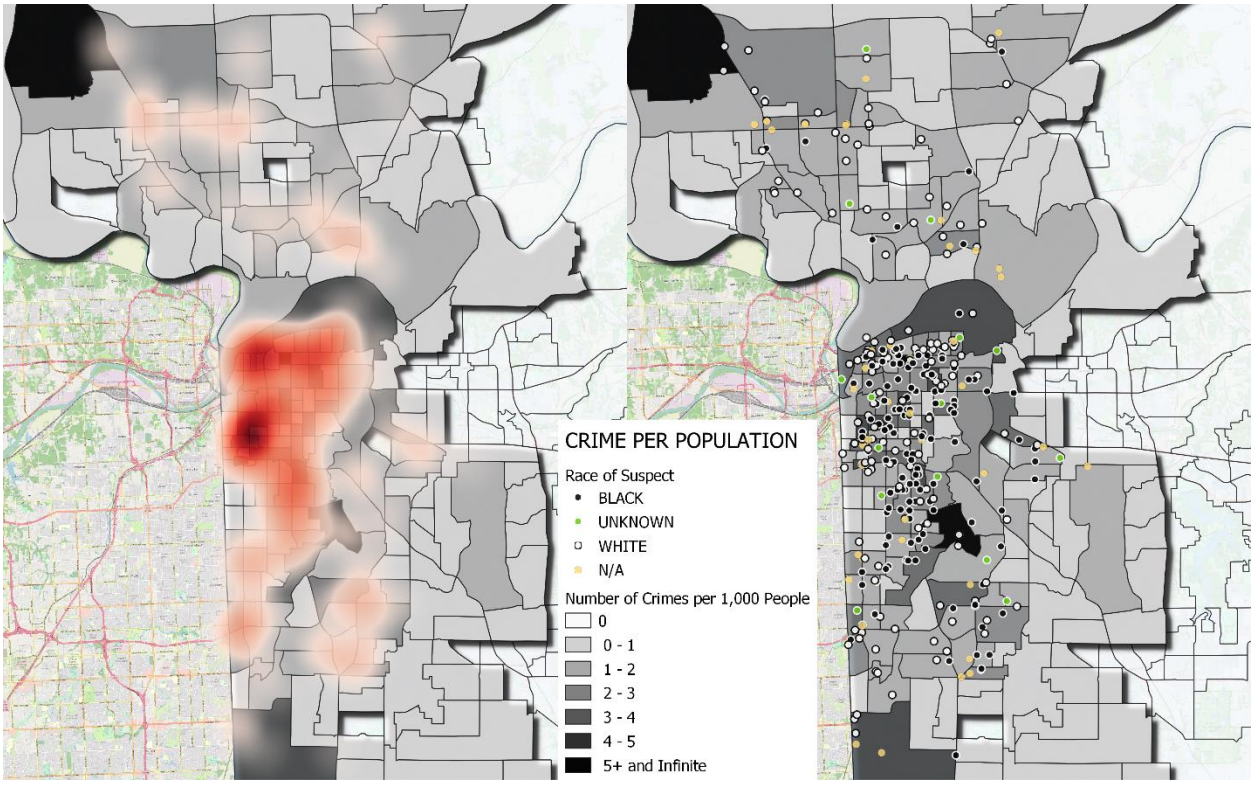

Figure 3. Crime per Population. Hoping that a ratio would shed some light on the data, I found no special connection.

Data Sources

Kansas City Data: Open Data KC | data.kcmo.org

Bus Stops: KCATA Bus Stops | Open Data KC | data.kcmo.org

API: https://data.kcmo.org/resource/bd2s-bfst.geojson

Dangerous Buildings: Dangerous Buildings List | Open Data KC | data.kcmo.org

API: https://data.kcmo.org/resource/ax3m-jhxx.geojson

Crime Data 2021: KCPD Crime Data 2021 | Open Data KC | data.kcmo.org

API: https://data.kcmo.org/resource/w795-ffu6.geojson

Market Value Analysis 2020: Market Value Analysis (MVA) 2020 | Open Data KC | data.kcmo.org

API: https://data.kcmo.org/resource/qhqc-4mi2.geojson

Table Description: Market Value Analysis (MVA) 2020 | Open Data KC | data.kcmo.org

Census Bureau: Census Bureau Data