To: Los Angeles Board of Police Commissioners, and James P. O'Neill, Police Commissioner of New York City
Subject: Unsolved Murders: What are the Factors?
Date: July 1, 2018

Introduction

This research was inspired by an article in the Washington Post titled 'Where Cases Go Unsolved', published on June 6, 2018[1]. This reporting sectioned out cities into high arrest and low arrest neighborhoods and found there are clear spatial variations in rates. Those from low arrest areas interviewed in various cities described feeling like the police are not trying hard enough, and the police say that cases are harder to solve in areas with low police trust and cooperation. This leads to a vicious cycle of people in neighborhoods where police trust is low, trusting the police even less as time passes as they see murderers committing crimes with impunity. The writers of the article also released their dataset, which is used in this report.
According to the FBI, in 2015 the clearance rate for cases across the nation was 64.1%, down from 90% 50 years previously[2]. For a case to be considered clear it has to end in arrest.
Our research looked at median household income of census tract the homicide occurred in, if it was within .1 miles of subsidized housing, and the race of the victim to try to determine if any or all of these factors had an impact on percentage of cases that fell under a certain case status. The three case statuses we divided the data into were 'open', 'closed by arrest' and closed without arrest.
We did not find a clear connection between either of the location income variables and case-status. However, these seems to be compelling evidence for the race of the victim influencing case-status, with Whites and Asians having the highest percentage of cases cleared and Blacks the lowest, in both cities.
We hope this research will help the leadership of the police force of both cities train their force to be better aware of underserved populations and work towards changing patterns of inequality in arrest rates.

Methodology

To start to answer the question of what populations were experiencing more unsolved homicide cases, we chose three variables- race of victim, proximity to subsidized housing, and median household income bracket of census tract.
We downloaded information the Washington Post had made available on 52,000 homicides across the country. This dataset included characteristics of the victim, status of the case, and exact location of the crime, among other factors. We narrowed my analysis to New York City and Los Angeles, with the possibility of expanding to other cities after getting an idea of what factor seemed to most influence case status. We created 3 binary variables for the homicide case status, corresponding to if a case was open, closed by arrest, or closed without an arrest.

---

[1] https://www.washingtonpost.com/graphics/2018/investigations/where-murders-go-unsolved/?utm_term=.69c68e428e0e
[2] https://www.npr.org/2015/03/30/395069137/open-cases-why-one-third-of-murders-in-america-go-unresolved

There were a different amount of years available for each city. NYC has data from 2016 and 2017 while LA has data all the way from 2010, also ending in 2017, giving us a lot more cases to analyze.
We downloaded information on Median Household Income by Census Tract for NYC and LA. Using the income information I divided the tracts of each city into below and above average income, and noted the count of cases of each status for my created income brackets (high and low).
For NYC we created a set of points for all the subsidized housing properties listed on the Furman Center website, then circled a .1 mile area of interest around them. As can be seen in Map 1, this area covers a large portion of the city due to the proliferation of subsidized housing. For LA we downloaded points for public housing and created the same .1 mile area of interest circles around these properties. To note, public housing is only one type of subsidized and we'd like to repeat the analysis with a more comprehensive subsidized housing dataset. I noted the counts of each case status near and not near subsidized housing.
Additionally, we used the homicide data table to gather case-status counts by victim race.
After gathering all these counts, we calculated percentage of total cases for each case status for each variable of interest and graphed the results (see figures).

Results

For reporting results we use 'percentage' to mean percent of total cases for a certain set of homicides corresponding to a case-status. The divisions of homicides are by income bracket of tract, proximity to public housing, and victim race.
LA has no cases listed as 'closed with no arrest', so the only type of unsolved cases in that city are open cases.
In LA there were no homicide cases within .1 mile of public housing so that analysis not doable (Fig 4). In NYC there was a slightly higher percentage closed in this area with no arrest, but a lower percentage of open cases (Fig 1).
Median Household Income of Census Tract had a more pronounced result for LA than NYC. In NYC there was no difference over 5% for above average income versus below average income census tracts. In LA the largest gap was in percent closed by arrest, with a 7% higher occurrence in above average income census tracts. In both cities below average income census tracts had higher percentages of open cases and lower percentages of cases closed by arrest. However, in NYC, there was a slightly lower percentage of cases closed without arrest in tracts in this bracket (Fig 2 and 5).
The most conclusive correlation with case-status is not by income of surrounding area, but by race of the victim (Fig 3 and Fig 6). In both NYC and LA, Black victims have the highest percentage of open cases, and lowest percentage of cases closed by arrest. In both cities Asian victims have the lowest percentage of open cases, and in LA they also have the highest percentage of cases closed by arrest. In NYC Whites have the highest percentage of cases closed by arrest. The difference in cases closed by arrest between the race with the highest percentage and that with the lowest percentage is 17.55% in LA, and 12.51% in NYC. In both cities Hispanic victims have a higher rate of open cases than Whites and Asians.

Assumptions

Our main assumption was to use location characteristics as a proxy for income. This was done in absence of information about the income of the victim themselves. In every geographical area there is a mix of people and we would like to delve deeper into the correlates between poverty and unsolved cases.
In ignoring the date of the cases we failed to filter out very recent (near when the dataset was published) cases, and therefore some or many of the cases still listed as open could be for lack of time.
We also did not do any statistical work to check if the percent differences we were seeing were within the range attributable to random variation. Here we could have factored in standard deviation and noted how many standard deviations away from the mean our results were. Further research can determine statistical significance.

Conclusion

This research is very preliminary but hopefully it gives the police commissioners food for thought on where to focus training. In the Washington Post article one mother in LA says she feels the police have ignored the murder of her son because they thought he was part of a gang. He was killed 7 years ago and the case is still open[3].
It is these kinds of stories that lead us to believe there is much work to be done in bringing up arrest rates for homicides across all populations in our cities. This research finds a compelling argument for racial disparity in arrest rates, but still more factors have to be examined. We hope to get funding and access to private police force data for further research into this topic, to look more at income and other demographic attributes of either the victims or the scene of the crime.
Thank you for your time.

[3] https://www.washingtonpost.com/graphics/2018/investigations/where-murders-go-unsolved/?utm_term=.6d070251c91a

Maps And Figures

Fig 1

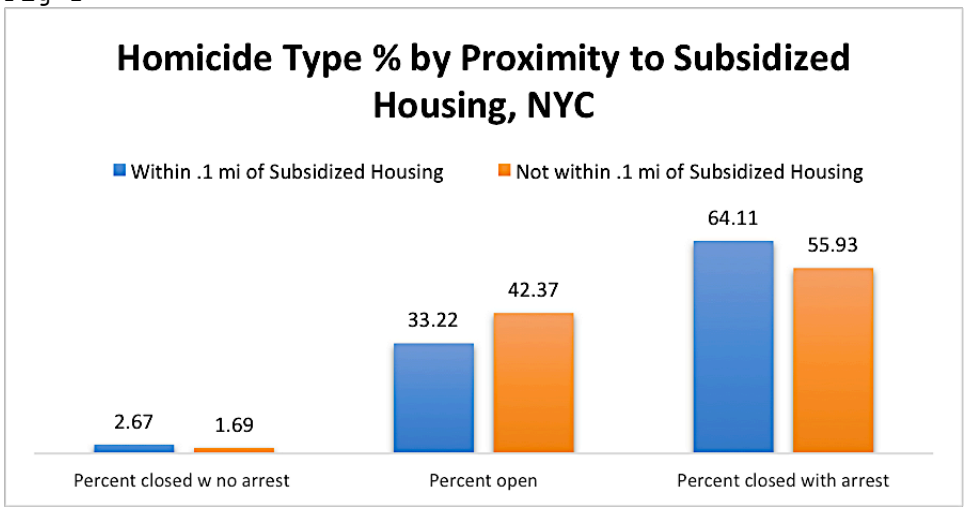


Fig 2

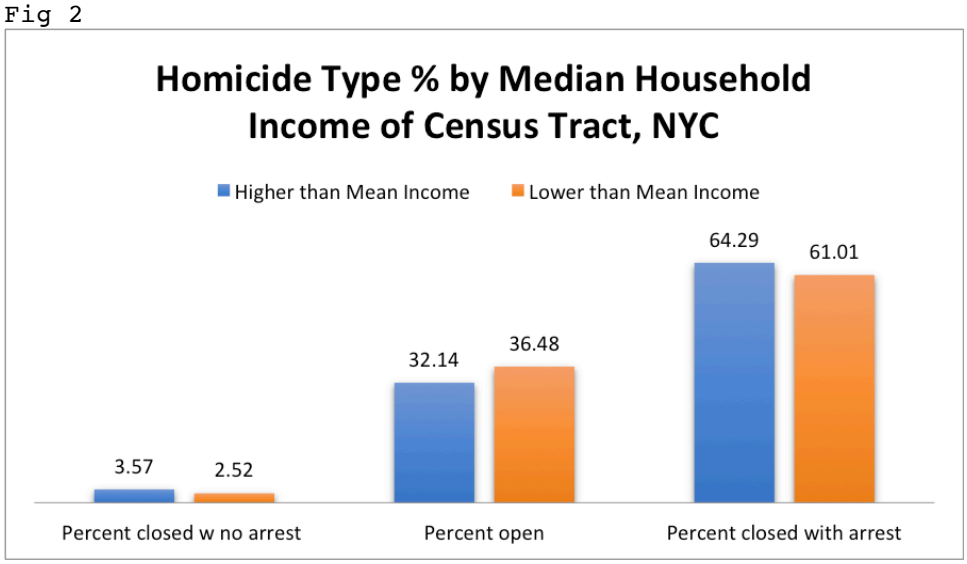
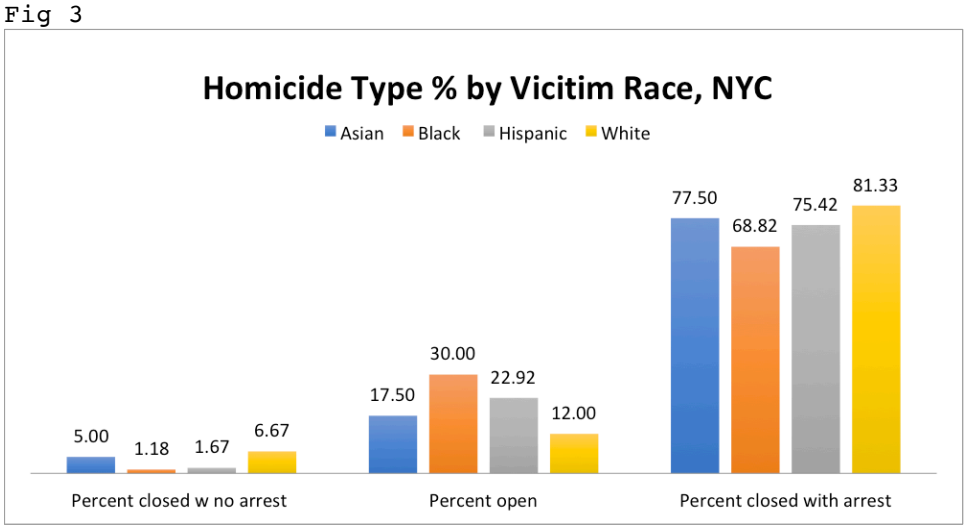


Fig 3



(Sources: MapPluto, US Census Bureau, NYU Furman Center, American Community Survey, Washington Post)

Fig 4

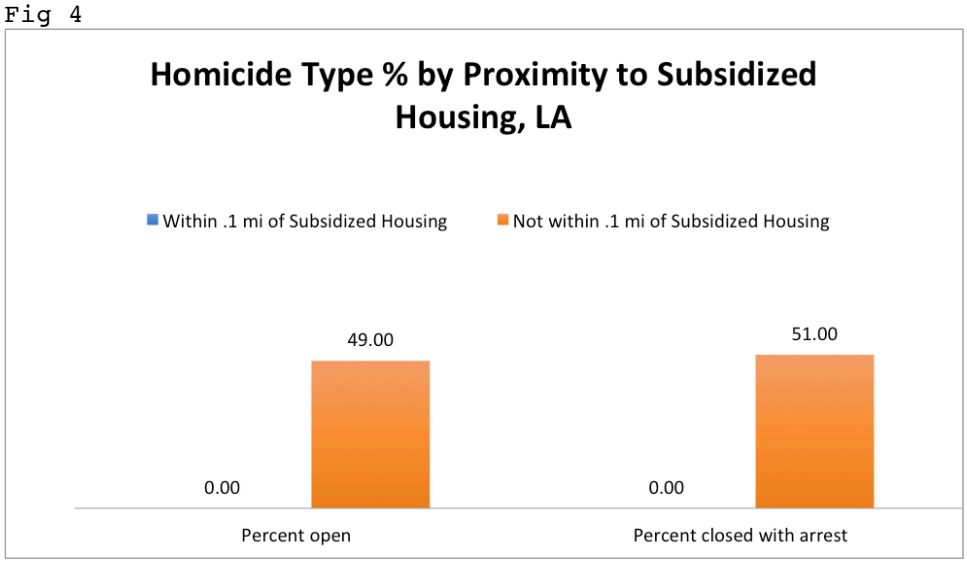


Fig 5

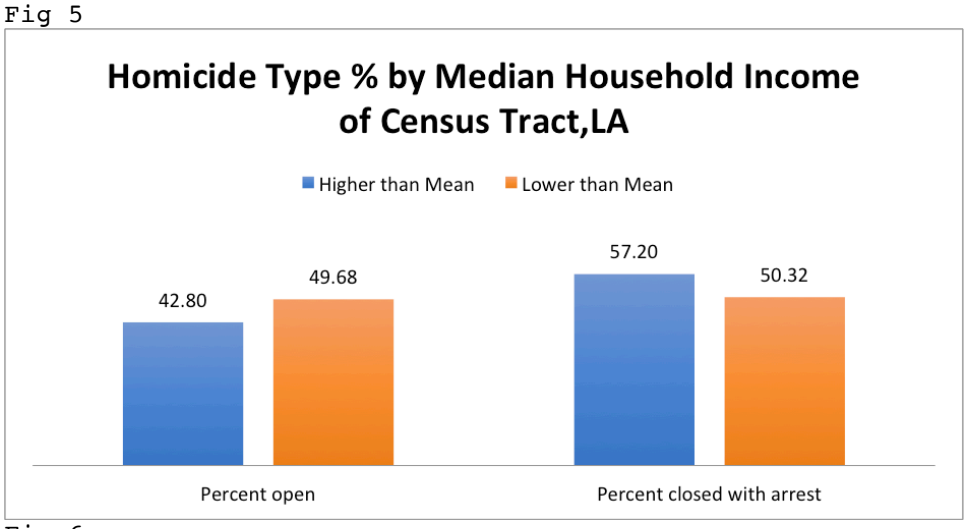
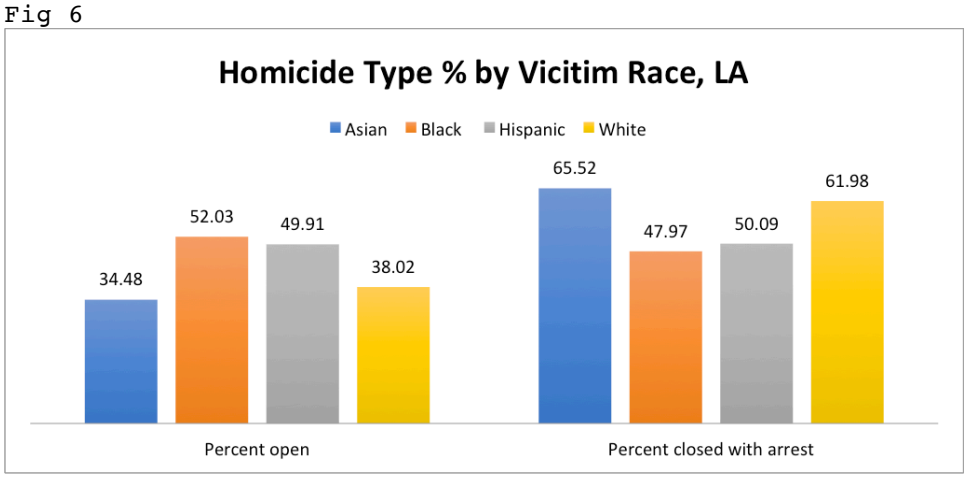
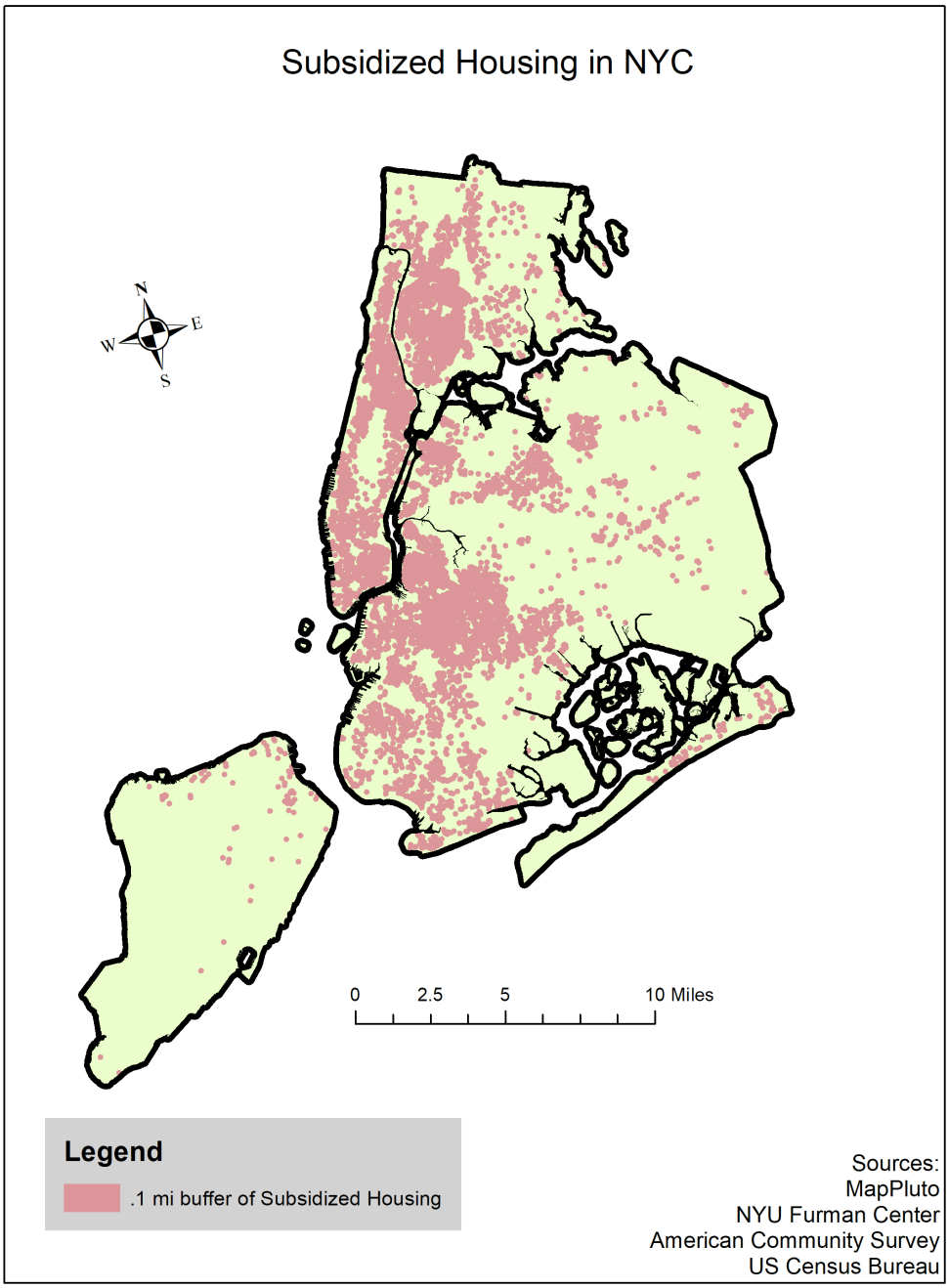


Fig 6

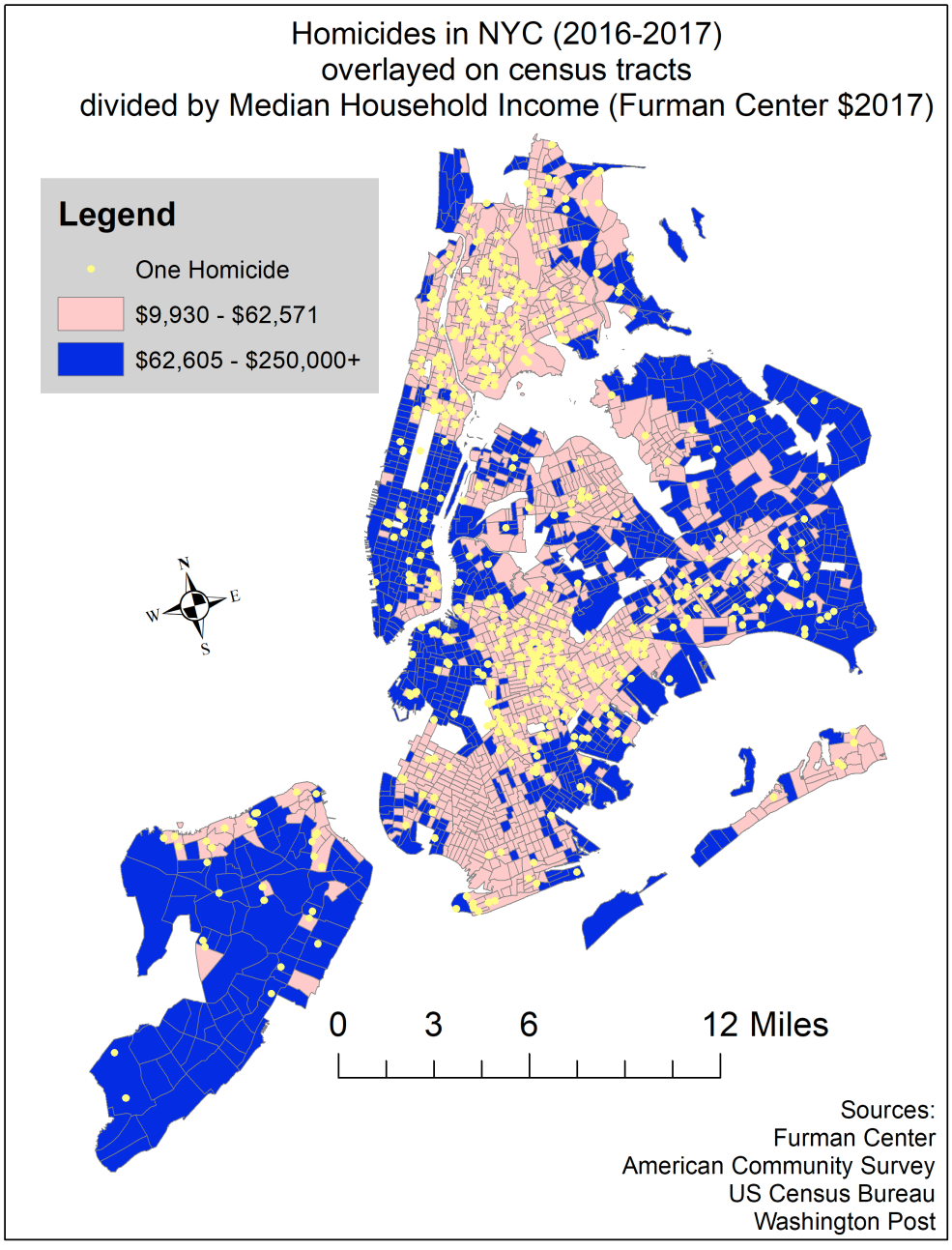


(Sources: US Census Bureau, LACounty.gov, American Community Survey, Washington Post)

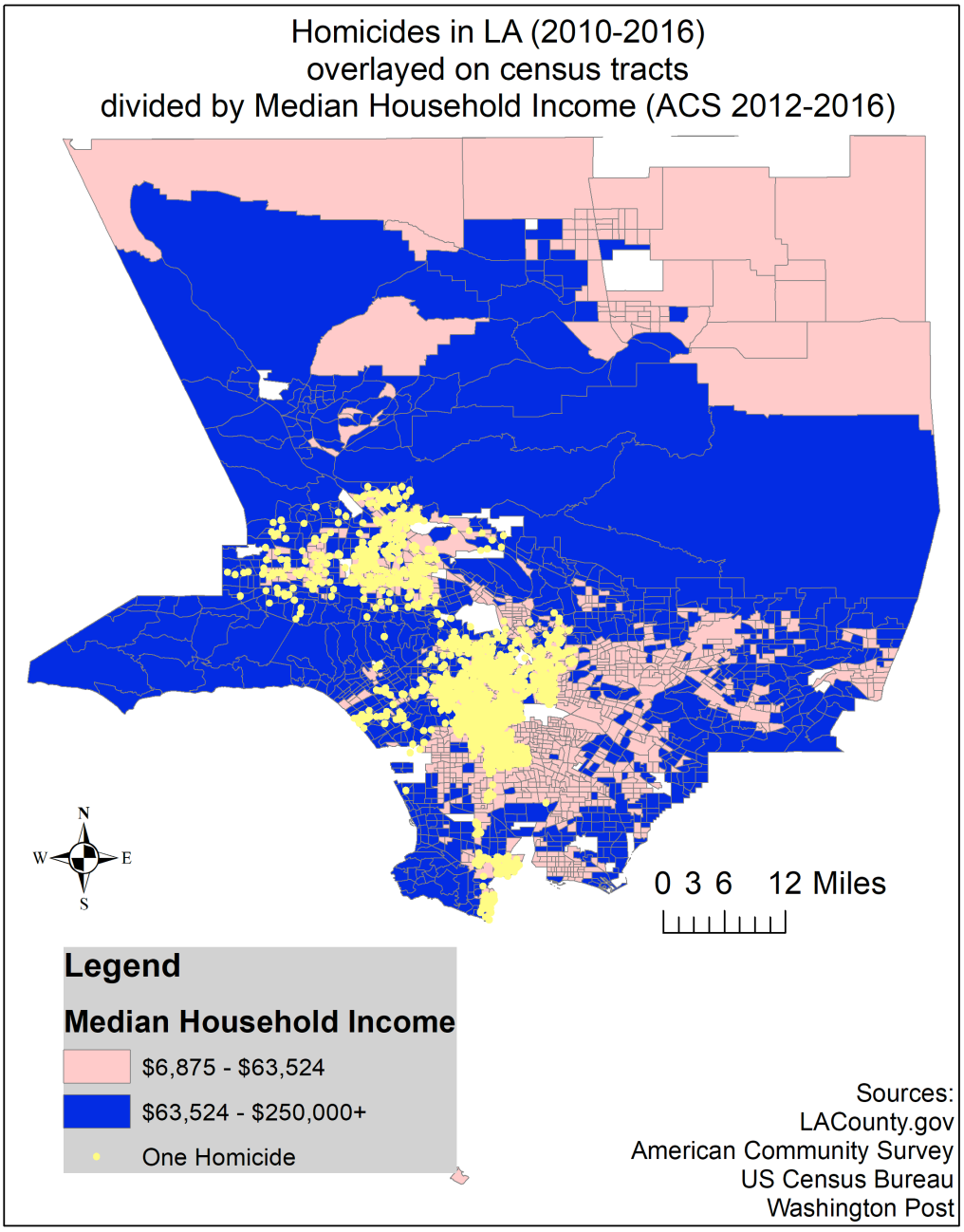
Map 1



Map 2


Homicides in NYC (2016-2017)
overlayed on census tracts
divided by Median Household Income (Furman Center $2017)
Legend
One Homicide
$9,930 - $62,571
$62,605 - $250,000+
N
E
S
W
0
3
6
12 Miles
Sources:
Furman Center
American Community Survey
US Census Bureau
Washington Post

Map 3


Homicides in LA (2010-2016)
overlayed on census tracts
divided by Median Household Income (ACS 2012-2016)
N
W
E
S
0 3 6 12 Miles
Legend
Median Household Income
$6,875 - $63,524
$63,524 - $250,000+
One Homicide
Sources:
LACounty.gov
American Community Survey
US Census Bureau
Washington Post

Technical Appendix

Data Sources

The datasets used in this analysis include:

- Cartographic Boundary File by Census Tract from the US Census Bureau for New York State and California[4]
- New York City Borough Boundaries, clipped to shoreline[5]
- MapPluto shapefiles from 2017, including data on every BBL in NYC, separated by borough[6]
- Homicide data from a Github account corresponding to a Washington Post story about unsolved murders[7]
- LA Country Boundary Line File [8]
- MEDIAN INCOME IN THE PAST 12 MONTHS (IN 2016 INFLATION-ADJUSTED DOLLARS) from 2012-2016 American Community Survey 5-Year Estimates for California by Census Tract[9]
- LA public housing site points, in report referred to as 'subsidized housing'[10]
- Median Household Income (2017$) for NYC from the Furman Center[11]
- Subsidized Housing points for NYC from the Furman Center[12]

[4] https://www.census.gov/geo/maps-data/data/cbf/cbf_tracts.html

[5] https://www1.nyc.gov/site/planning/data-maps/open-data/districts-download-metadata.page

[6] https://www1.nyc.gov/site/planning/data-maps/open-data/dwn-pluto-mappluto.page

[7] https://github.com/washingtonpost/data-homicides

[8] https://egis3.lacounty.gov/dataportal/drp_county_boundary/

[9] https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml

[10] https://egis3.lacounty.gov/dataportal/2015/10/27/cdc-public-housing-sites-points/

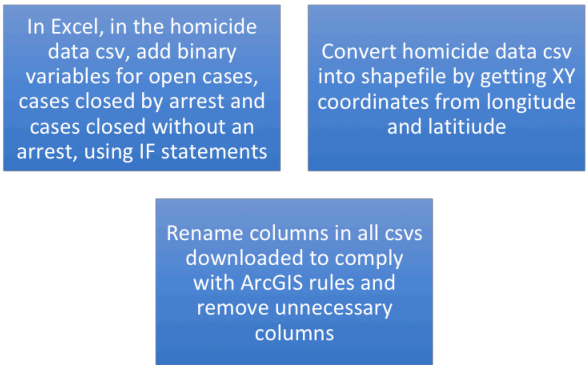
[11] http://app.coredata.nyc

[12] http://app.coredata.nyc

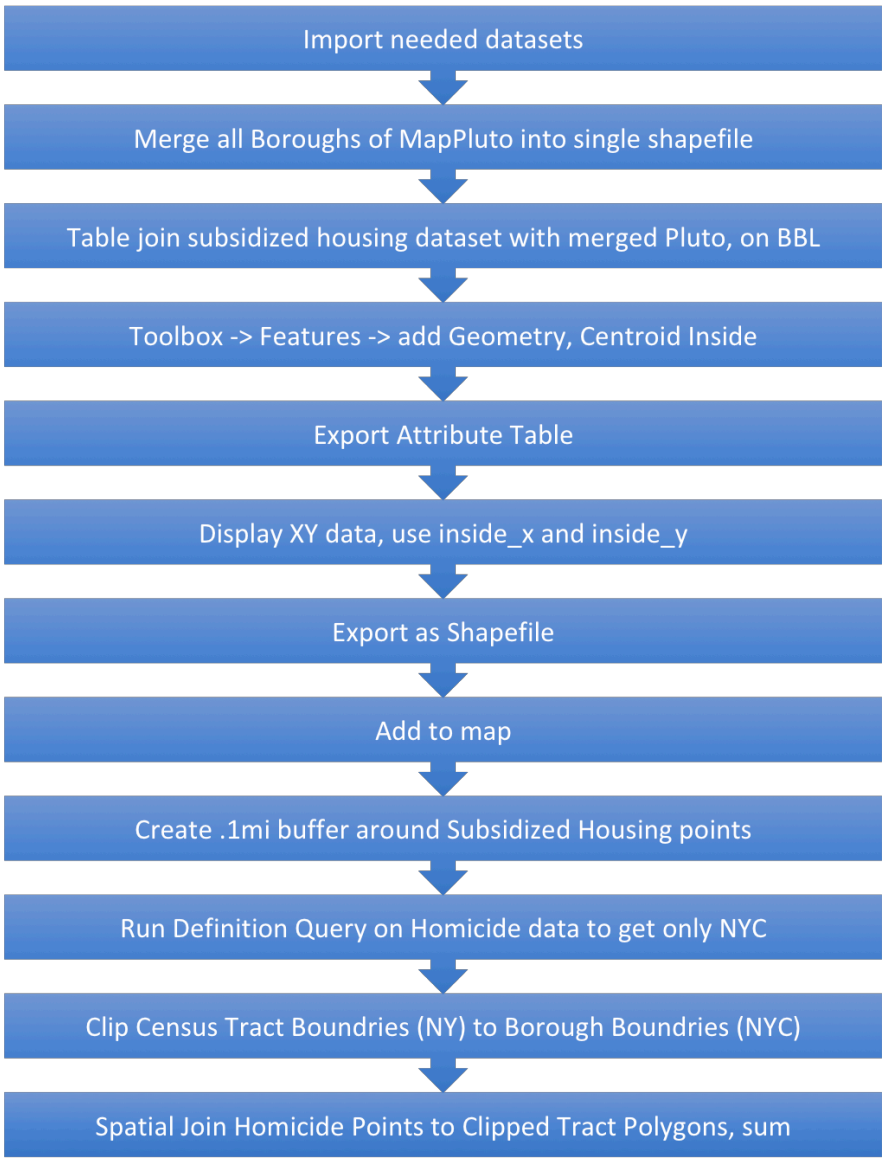
Flow Charts (to be done in order)

Prep

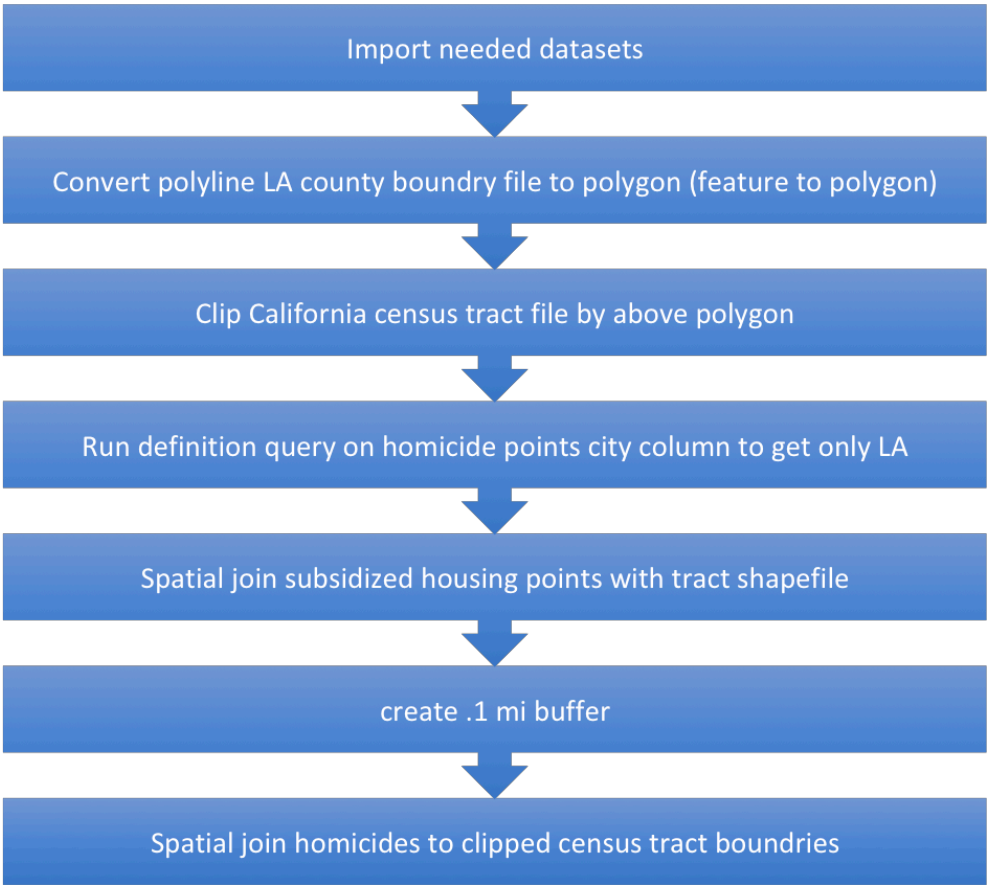
(note, rows with null data for variables of interest are removed)



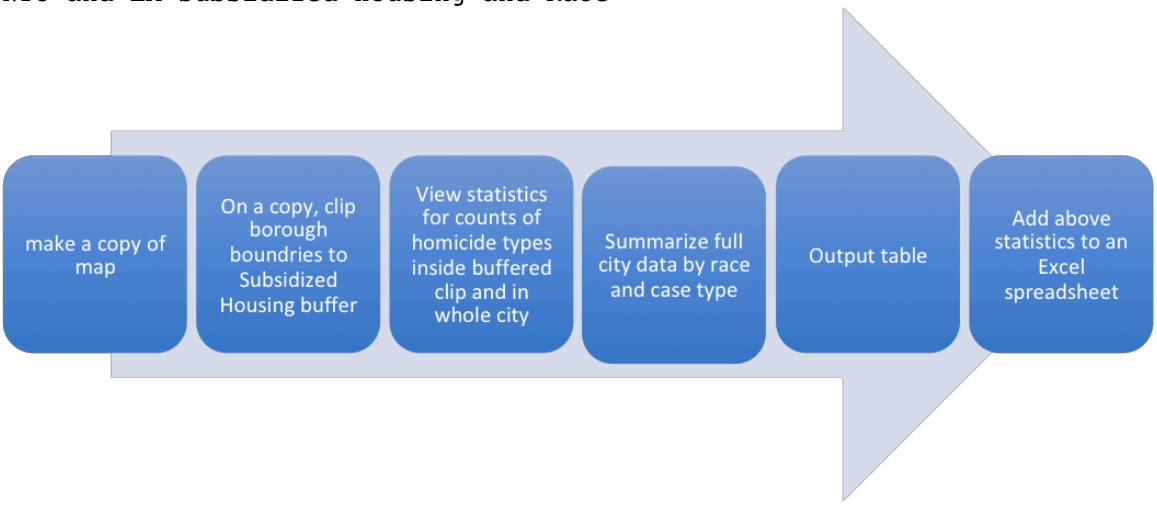
NYC Subsidized Housing and Race



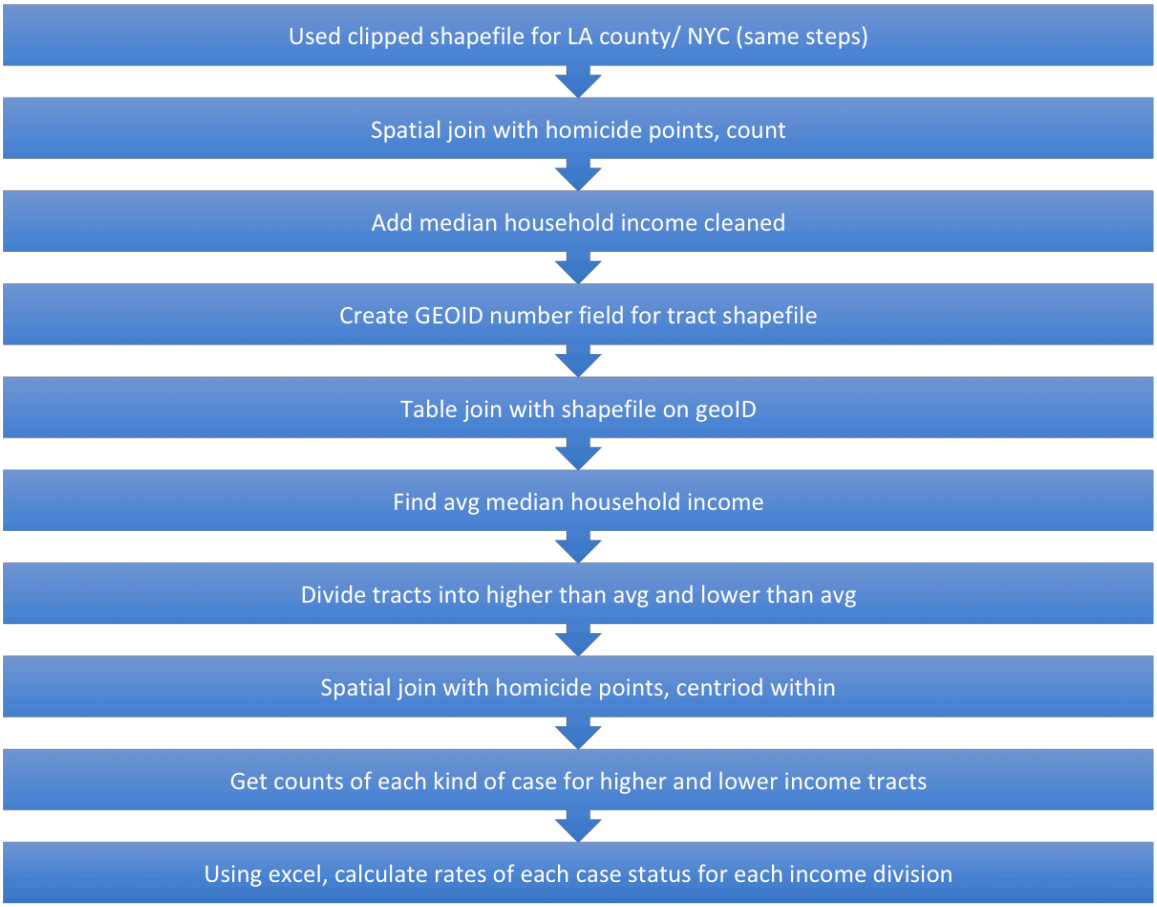
LA Subsidized Housing and Race



NYC and LA Subsidized Housing and Race



NYC and LA Income


Used clipped shapefile for LA county/ NYC (same steps)
Spatial join with homicide points, count
Add median household income cleaned
Create GEOID number field for tract shapefile
Table join with shapefile on geoID
Find avg median household income
Divide tracts into higher than avg and lower than avg
Spatial join with homicide points, centriod within
Get counts of each kind of case for higher and lower income tracts
Using excel, calculate rates of each case status for each income division