



Los Angeles, CA

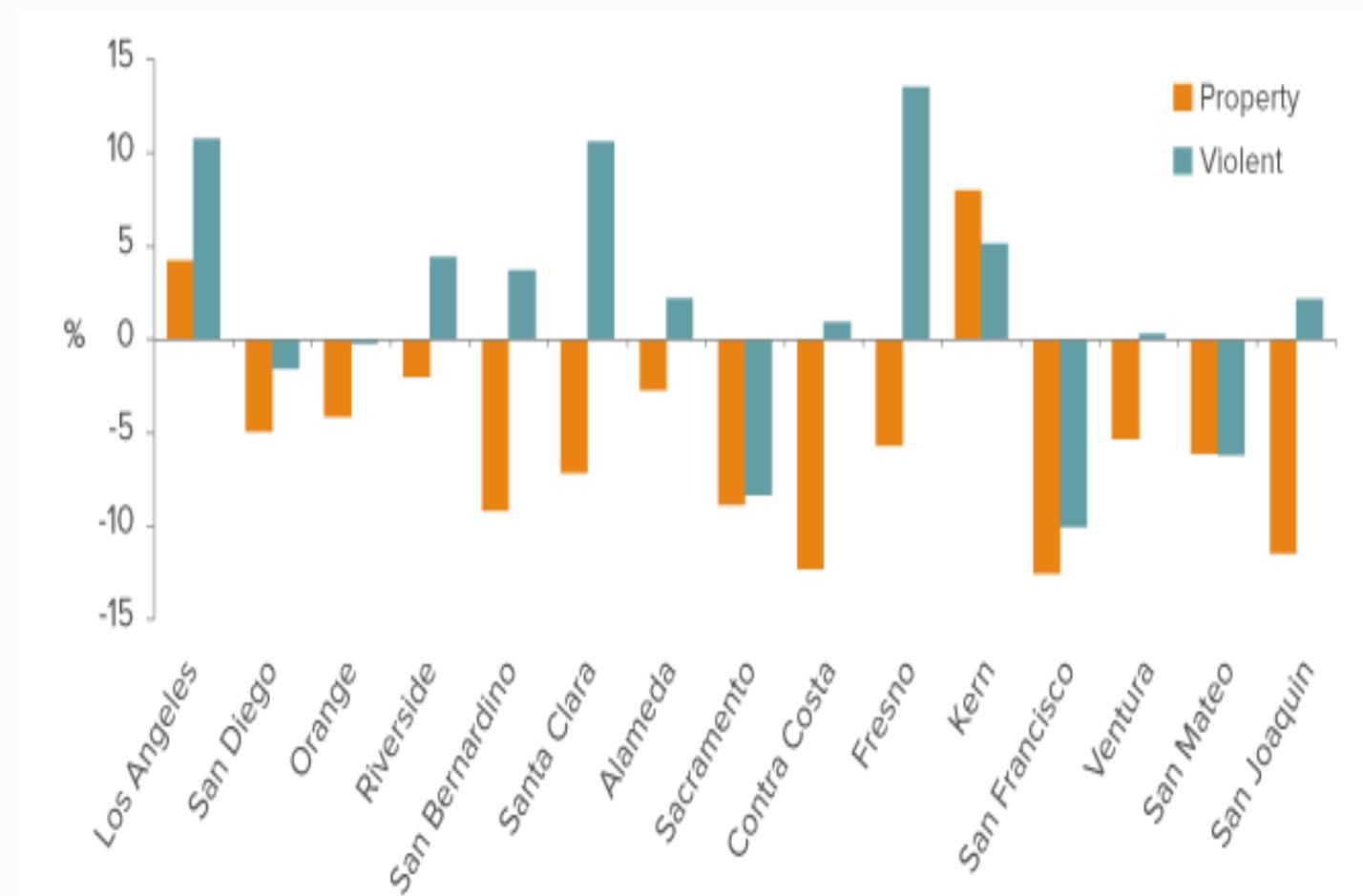
ANALYSIS OF VIOLENT CRIME IN LOS ANGELES, CA (2010 – 2017)

Jon Heston, Eleanora Trittipo, Shrawantee Saha

MOTIVATION

Crime Trends in California*

- California's violent crime rate rose in 2016—but it remains historically low.
- **Violent crime increased in a majority of counties ...**



*Public Policy Institute of California publication (November 2017): <http://www.ppic.org/publication/crime-trends-in-california/>

WHAT ARE WE PRESENTING?

Focus

Los Angeles City
2010 to 2017

Analyze
Violent
Crime

1. By Month and Years
2. Identify safe versus unsafe neighborhoods
3. Do demographics matter?

Tools

1. Graphs and Charts
2. Simple Regressions
3. Informative Map

REST OF THE
PRESENTATION

1. Data Sources

2. Data Retrieval and Cleaning

3. Description

4. Results

5. Conclusions

6. Future Work

I. DATA SOURCES

Incidence of Crime Data: Los Angeles Open Data:
[\(https://data.lacity.org/A-Safe-City/Crime-Data-from-2010-to-Present/y8tr-7khq\)](https://data.lacity.org/A-Safe-City/Crime-Data-from-2010-to-Present/y8tr-7khq)

Demographic and Socio-Economic Data: Los Angeles Census Reporter:
[\(https://censusreporter.org/profiles/16000US0644000-los-angeles-cal\)](https://censusreporter.org/profiles/16000US0644000-los-angeles-cal)

I. DATA SOURCES: DESCRIPTION

I. Incidence of Crime Data: Los Angeles Open Data:

- ✓ Reflects incidents of crime in the City of Los Angeles dating back to 2010.
- ✓ Data provided by Los Angeles Police Department. Last updated: August 14, 2018
- ✓ Transcribed from original crime reports that were typed on paper and therefore there may be some inaccuracies within the data. Address fields are only provided to the nearest hundred block in order to maintain privacy.

II. Demographic and Socio-Economic Data: Los Angeles Census Reporter

- ✓ The American Community Survey is at the heart of Census Reporter.
- ✓ This survey is conducted continuously, and data is released annually in two forms.
- ✓ We use the 5-year release, which is available for almost all Census geographies, block group and larger.
- ✓ It provides information on age, sex, race and housing, education, income, occupation, veteran status, ancestry, and other interesting topics.

2. DATA RETRIEVAL AND CLEANING

I. We used Python bindings for the Socrata Open Data API to retrieve our raw dataset.

12702 lines (12701 sloc) | 598 KB

<> | ⌂ | Raw | Blame | History | ⬕ | ⌁ | ⏙

```
In [1]: import pandas as pd
from uszipcode import ZipcodeSearchEngine
pd.set_option('display.max_rows', 500)
pd.set_option('display.max_columns', 500)

from datetime import datetime
import numpy as np
import matplotlib.pyplot as plt
```

```
In [4]: from sodapy import Socrata

# Unauthenticated client only works with public data sets. Note 'None'
# in place of application token, and no username or password:

# Note that you'll have to paste our API key here. I haven't included it since this will go to the Github repo
client = Socrata("data.lacity.org", "LLaYxNjJnTyIyES7N63sJAUDv")

#Only getting the first 1000 rows to ease computational requirements but when we have all of our codes written
#we should grab the full dataset.
result_list = client.get("7fvc-faax", limit=2000000)

# Convert to pandas DataFrame
crime_df = pd.DataFrame.from_records(result_list)
len(crime_df)
#crime_df.to_csv("RawCrimeData.csv")
```

```
In [7]: # Converting "date_occ" row to a datetime object
```

In [8]:

```
1 lat = []
2 lon = []
3 split_lat_longs = []

4
5 # For each row in a variable,
6 for x in range(len(crime_df["location"])):
7     # Try to,
8     try:
9         # Split the row by comma and append
10        # everything before the comma to lat
11        location = crime_df["location_1"][x]["coordinates"]
12        split_lat_longs.append(location)
13        lon.append(location[0])
14        lat.append(location[1])
15        # Split the row by comma and append
16        # everything after the comma to lon
17        #lon.append(location.split(',')[1])
18    # But if you get an error
19    except:
20        # append a missing value to lat
21        lat.append(np.NaN)
22        # append a missing value to lon
23        lon.append(np.NaN)

24
25 crime_df["Latitude"] = lat
26 crime_df["Longitude"] = lon
27 crime_df.head()
```

2. DATA RETRIEVAL AND CLEANING

2. One of the challenging parts of the data cleaning part was to split the latitude and longitude information from a Point Datatype (a Location datatype) in Socrata.

(<https://dev.socrata.com/docs/datatypes/point.html#2.1>)

```
{ "type": "Point", "coordinates": [ -87.653274, 41.936172 ] }
```

3. Data Retrieval and Cleaning

In []:

```
1 from uszipcode import ZipcodeSearchEngine
2 search = ZipcodeSearchEngine()
3 zipcodes = []
4 counter = 0
5 counter_2 = 0
6 for index, row in crime_df.iterrows():
7
8     #Get zipcodes based on lat-longs
9     try:
10         crime_lat = row["Latitude"]
11         crime_lon = row["Longitude"]
12         zipcode = search.by_coordinate(crime_lat, crime_lon, returns=1)
13         zipcodes.append(zipcode[0]["Zipcode"])
14         counter_2 = counter_2+1
15     except:
16         zipcodes.append("NaN")
17         counter = counter + 1
18         print(f'{counter} exceptions identified on index {counter_2}')
19         counter_2 = counter_2+1
20
21 crime_df["Zip Code"] = zipcodes
```

3. The second challenging and time consuming part of the data retrieval and cleaning process was to convert the Latitude and Longitude information into Zip-codes.

Note that “uszipcode” is a powerful and easy-to-use programmable zip-code database, and is also a search engine in Python

3. DESCRIPTION

- Number of observations (raw dataset) : 1,800,985
- Years (raw dataset) : 2010 to 2018 (August)
- Years (included in our analysis): Dropped 2018

3. DATA DESCRIPTION

CLASSIFICATION OF CRIME (FYI)

- US Department of Justice classifies crime in two major groups:
 - Violent Crime
 - Property Crime
- **Violent crime includes**
 - Homicide
 - Rape
 - Robbery
 - Assault
 - Purse snatching and pocket picking

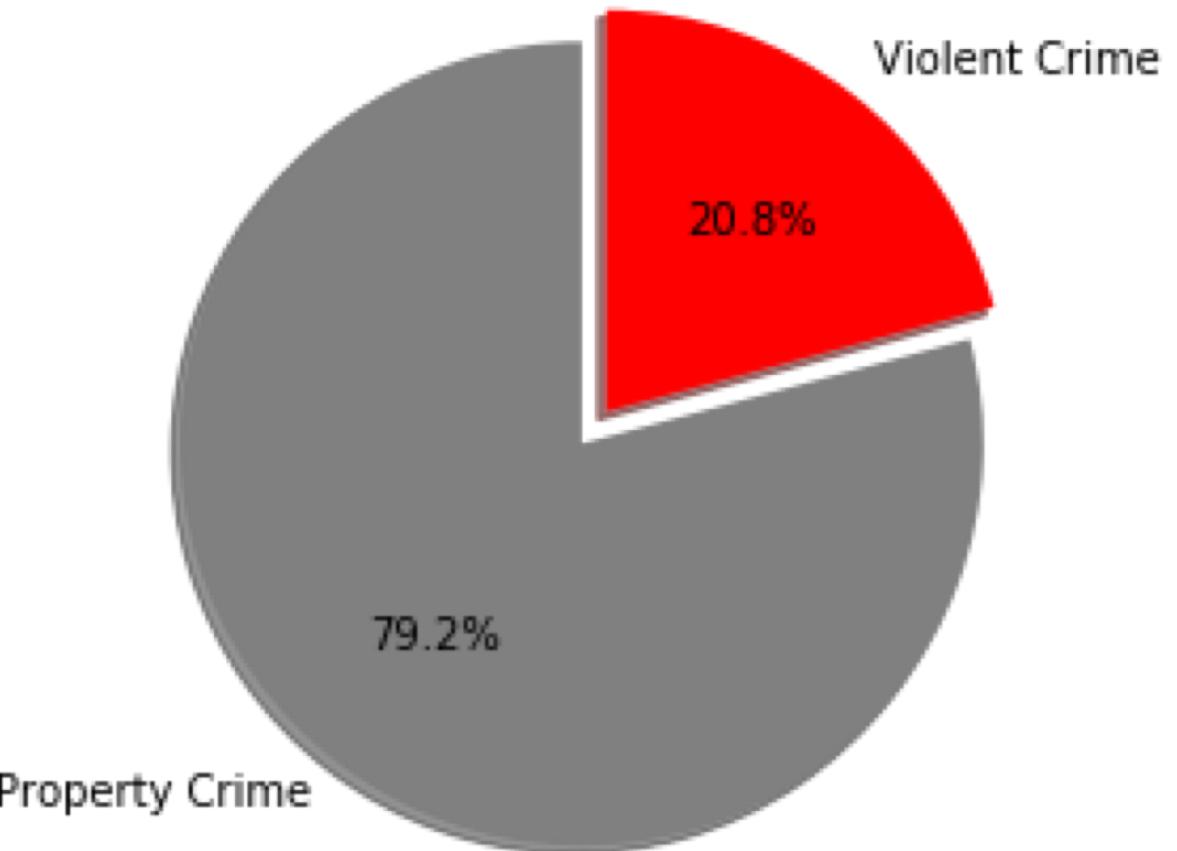
*For more information: <https://www.bjs.gov/index.cfm?ty=tp&tid=93> (as viewed on 18th August 2018)

VIOLENT CRIME IN LA CITY: A SNAP SHOT

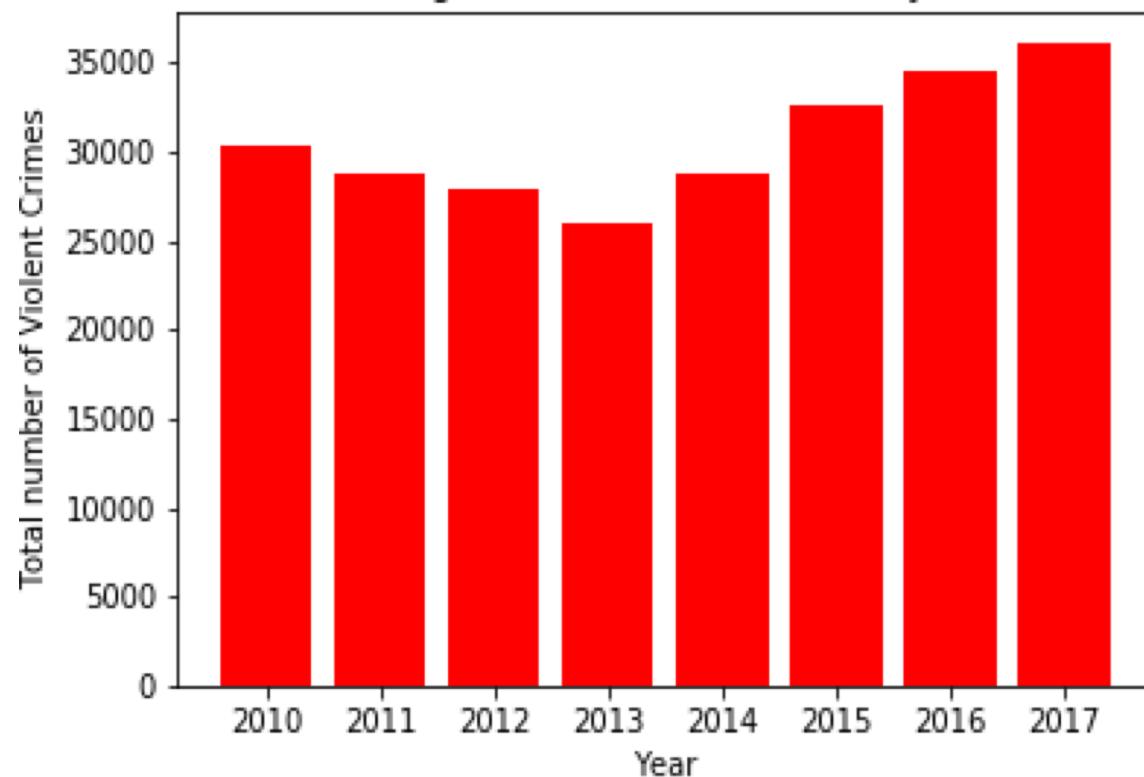
"BATTERY - SIMPLE ASSAULT", "ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT", "ROBBERY", "THEFT PERSON","INTIMATE PARTNER - AGGRAVATED ASSAULT", "BURGLARY, ATTEMPTED", "BATTERY WITH SEXUAL CONTACT", "RAPE, FORCIBLE", "CHILD ABUSE (PHYSICAL) - SIMPLE ASSAULT", "CRM AGNST CHLD (13 OR UNTER) (14-15 & SUSP 10 YRS OLDER)0060", "CHILD NEGLECT", "BATTERY POLICE (SIMPLE)", "SEX, UNLAWFUL", "DISCHARGE FIREARMS/SHOTS FIRED", "ARSON", "OTHER ASSAULT", "CRIMINAL HOMICIDE", "SHOTS FIRED AT INHABITED DWELLING", "EXTORTION", "KIDNAPPING", "ORAL COPULATION", "SEXUAL PENETRATION WITH A FOREIGN OBJECT", "CHILD ABUSE (PHYSICAL) - AGGRAVATED ASSAULT", "ASSAULT WITH DEADLY WEAPON ON POLICE OFFICER", "THROWING OBJECT AT MOVING VEHICLE", "LEWD CONDUCT", "SODOMY/SECUAL CONTACT B/W PENIS OF ONE PERS TO ANUS OTH 0007=02", "CHILD STEALING", "RAPE, ATTEMPTED", "FALSE IMPRISONMENT", "KIDNAPPING - GRAND ATTEMPT", "SEXUAL PENETRATION W/FOREIGN OBJECT", "PIMPING", "CRIM AGNST CHLD (13 OR UNDER) (14-15 & SUSP 10 YRS OLDER)", "SEX,UNLAWFUL(INC MUTUAL CONSENT, PENETRATION W/ FRGN OBJ0059", "BATTERY ON A FIREFIGHTER", "THEFT FROM PERSON - ATTEMPT", "SHOTS FIRED AT MOVING VEHICLE, TRAIN OR AIRCRAFT", "SEX,UNLAWFUL(INC MUTUAL CONSENT, PENETRATION W/ FRGN OBJ", "HUMAN TRAFFICKING - COMMERCIAL SEX ACTS", "SODOMY/SEXUAL CONTACT B/W PENIS OF ONE PERS TO ANUS OTH", "LEWD/LASCIVIOUS ACTS WITH CHILD", "CHILD PORNOGRAPHY", "LYNCHING", "DRUGS, TO A MINOR", "HUMAN TRAFFICKING - INVOLUNTARY SERVITUDE", "LYNCHING - ATTEMPTED", "INCITING A RIOT", "INCEST (SEXUAL ACTS BETWEEN BLOOD RELATIVES)", "ABORTION/ILLEGAL", "MANSLAUGHTER, NEGLIGENT", "TRAIN WRECKING"]

3. DESCRIPTION: VIOLENT VS PROPERTY

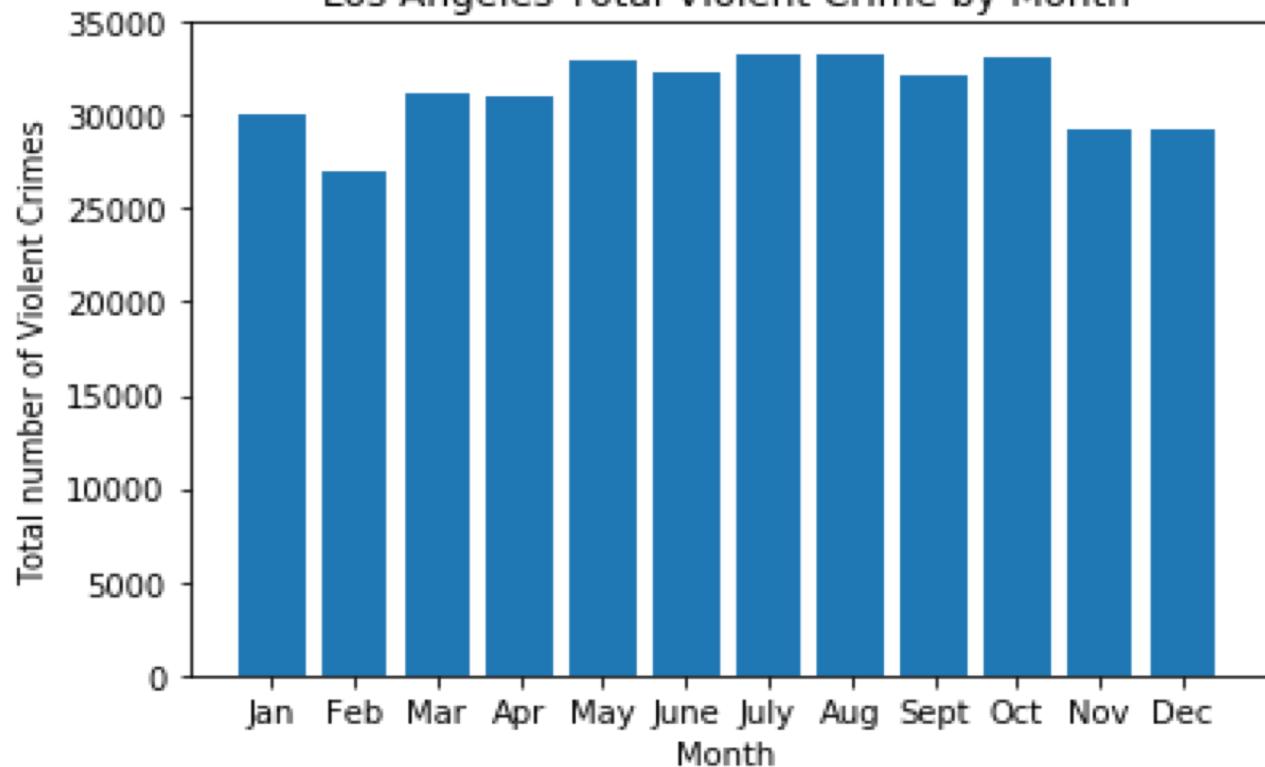
Total LA Crime



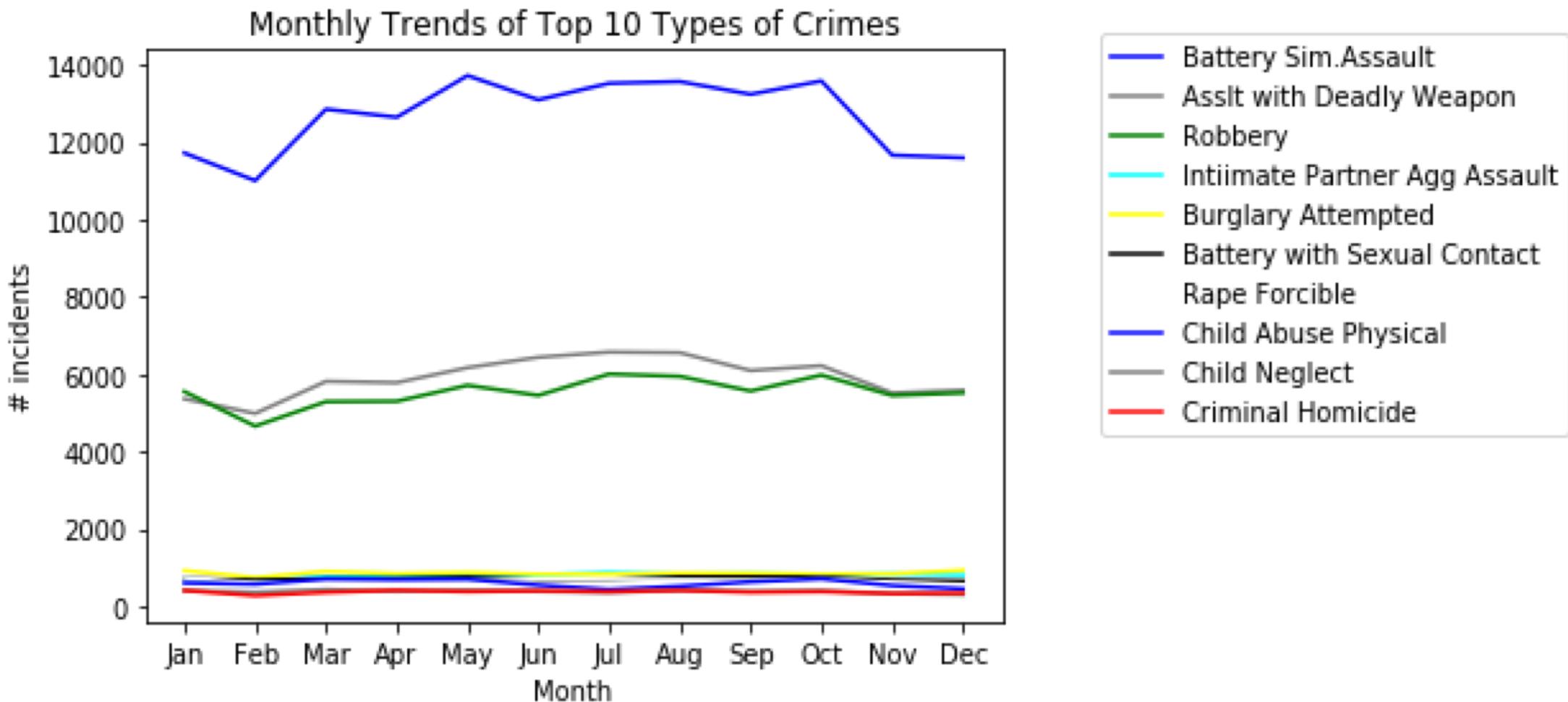
Los Angeles Total Violent Crime by Year



Los Angeles Total Violent Crime by Month

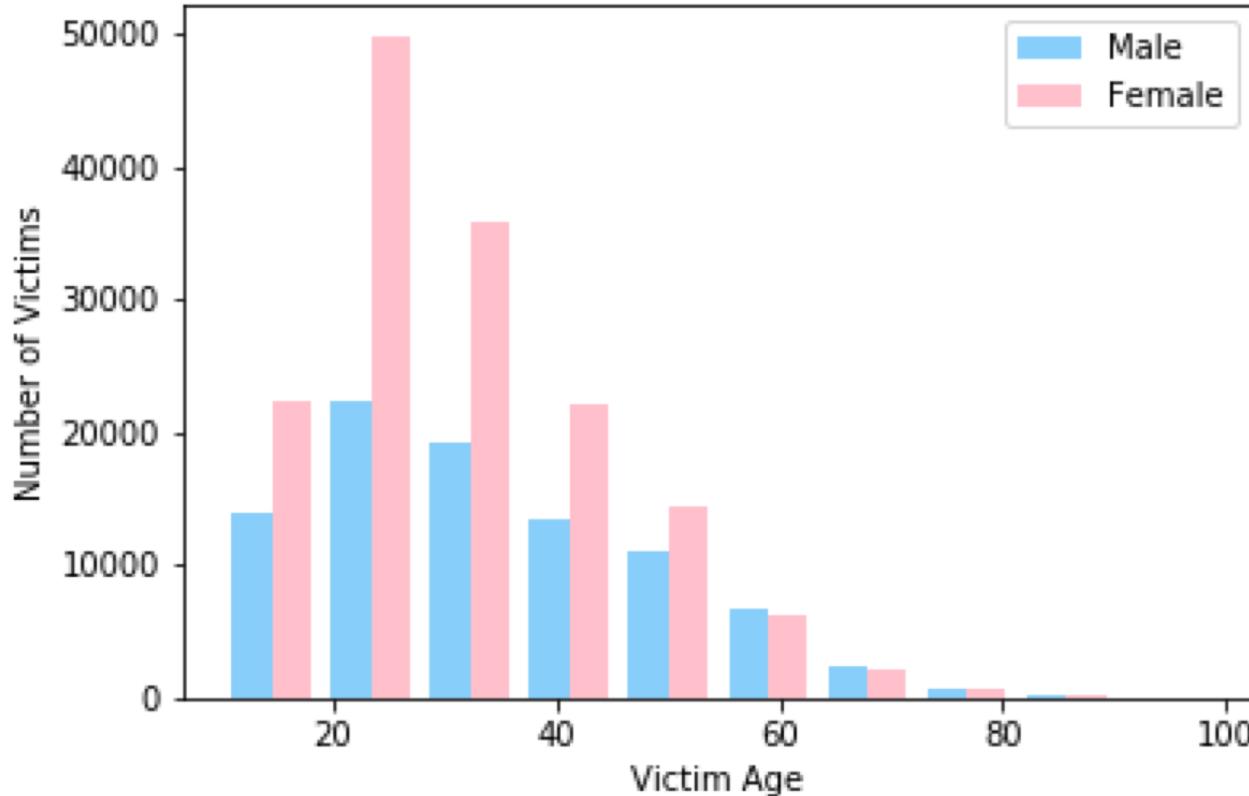


3. DESCRIPTION BY YEAR AND MONTH

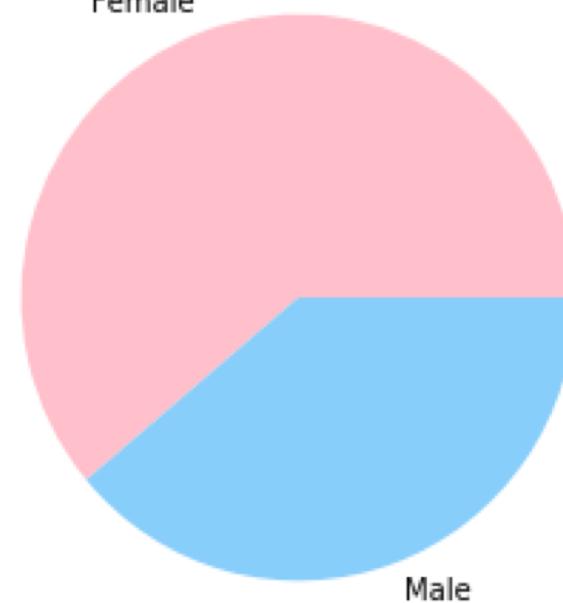


3.DESCRIPTION MONTHLY TRENDS FOR TOP 10 CRIMES

Number of Victims by Sex and Age



Breakdown of Victim Sex



3. DESCRIPTION:
VICTIMS BY SEX & AGE

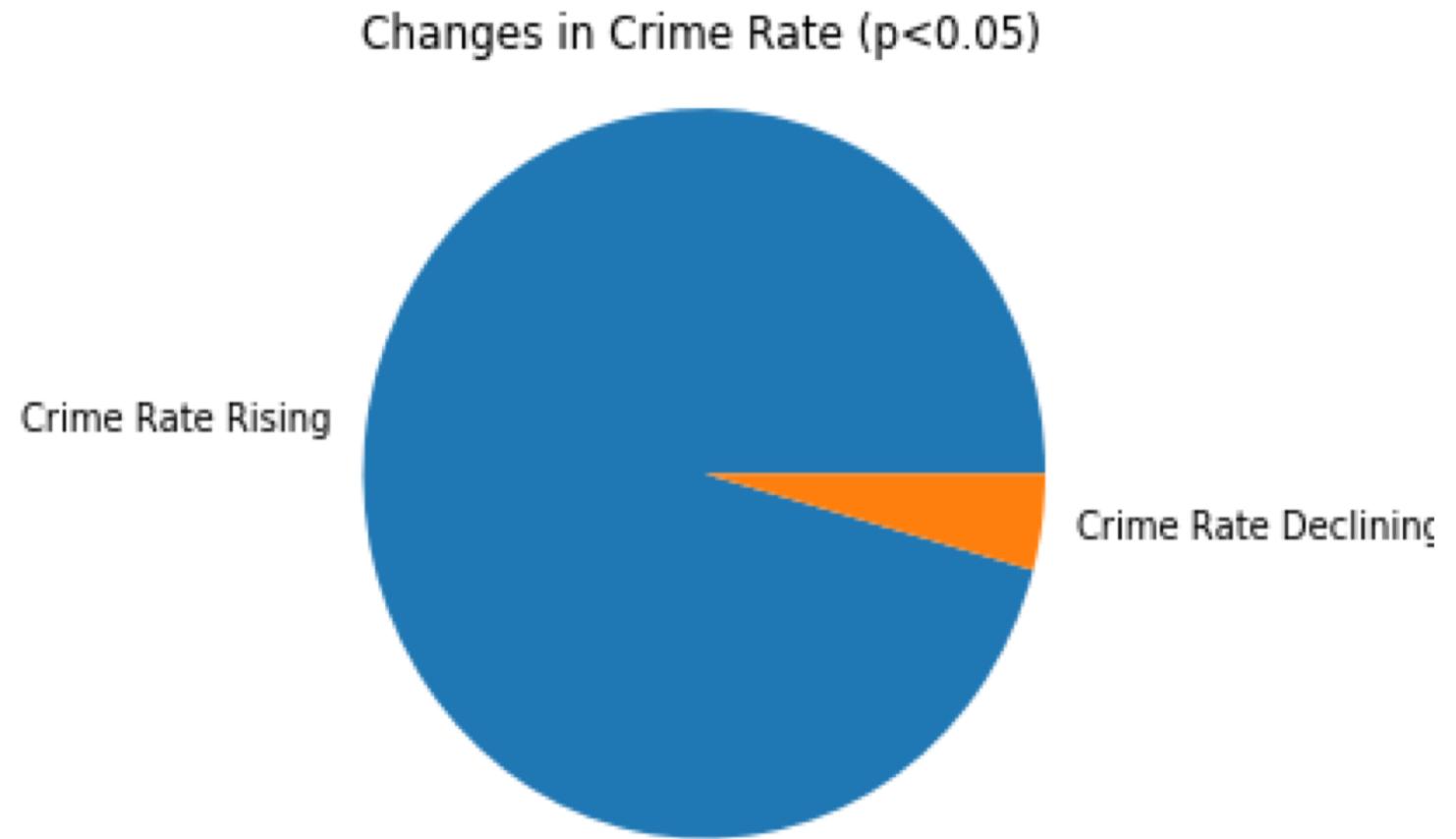
WHAT'S NEXT?

Results

Conclusions

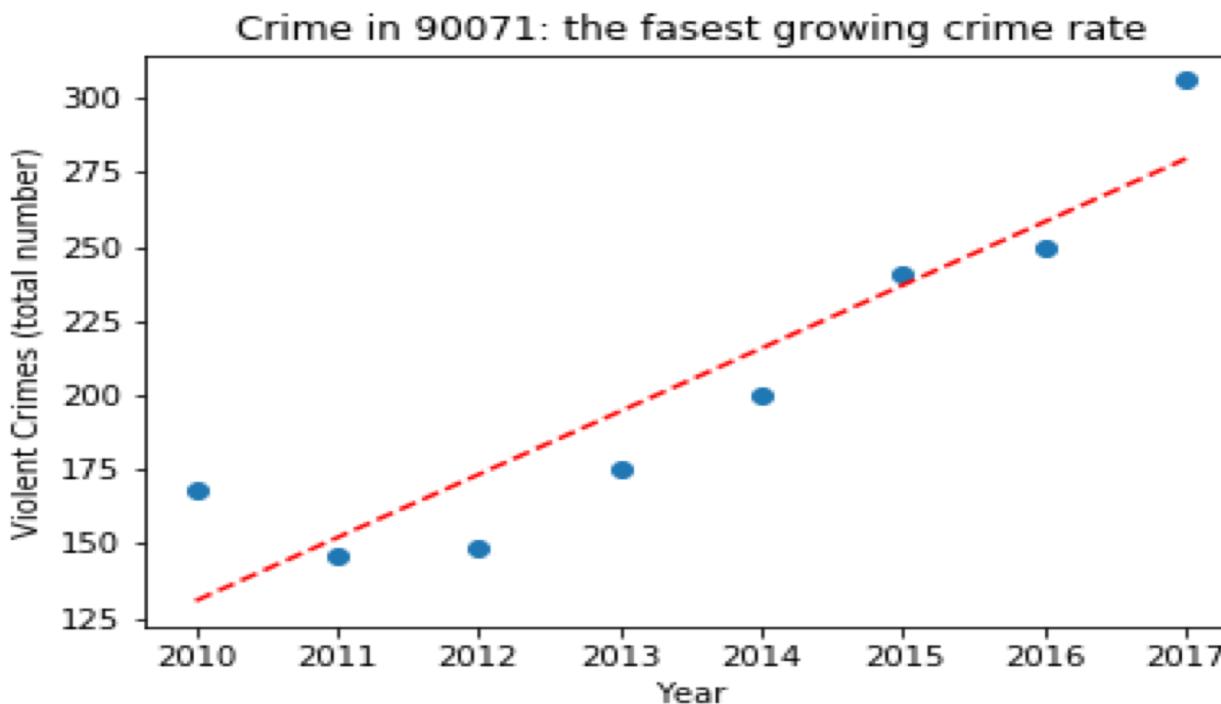
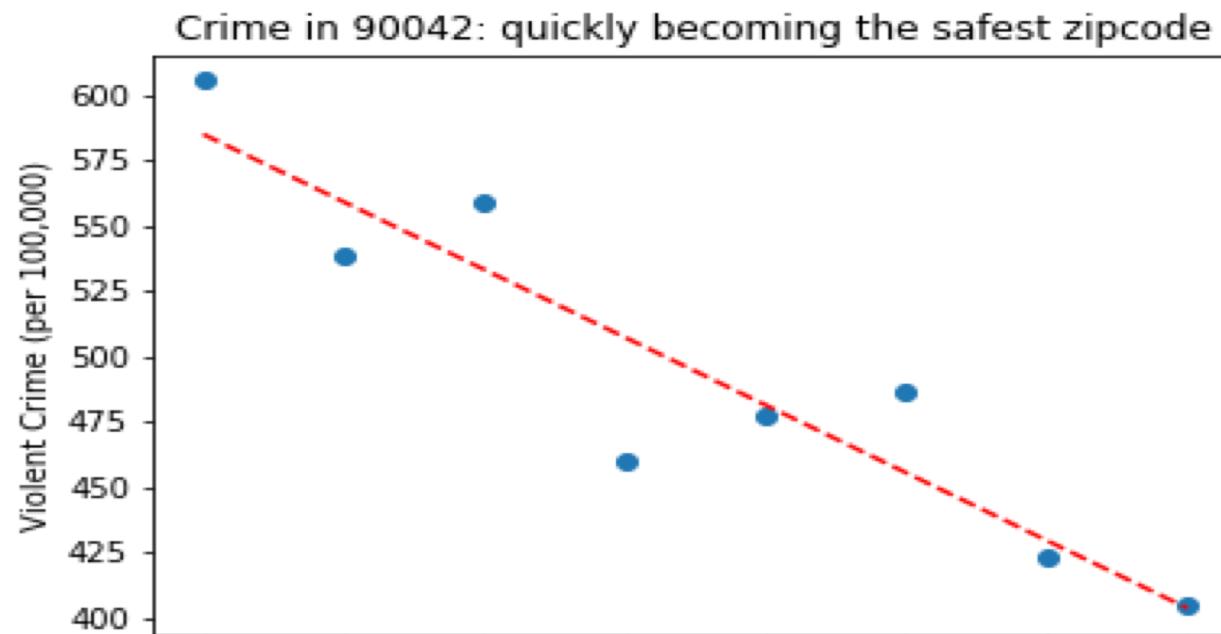
Future Work

4. RESULTS:



4. RESULTS:

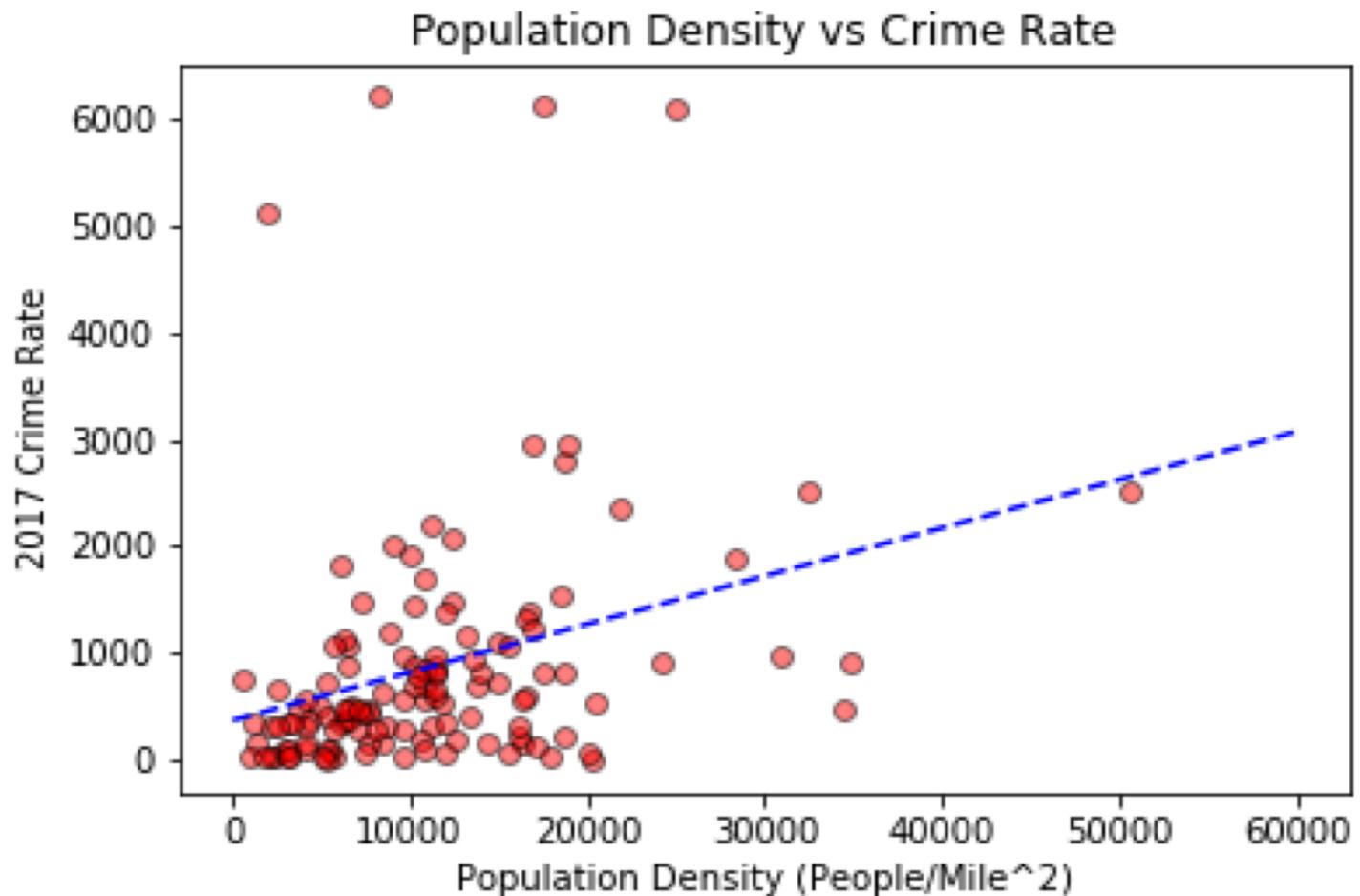
SAFE
VS.
“NOT-SO-SAFE” ZIP-CODE



4. RESULTS:

REGRESSION WITH CRIME RATE:

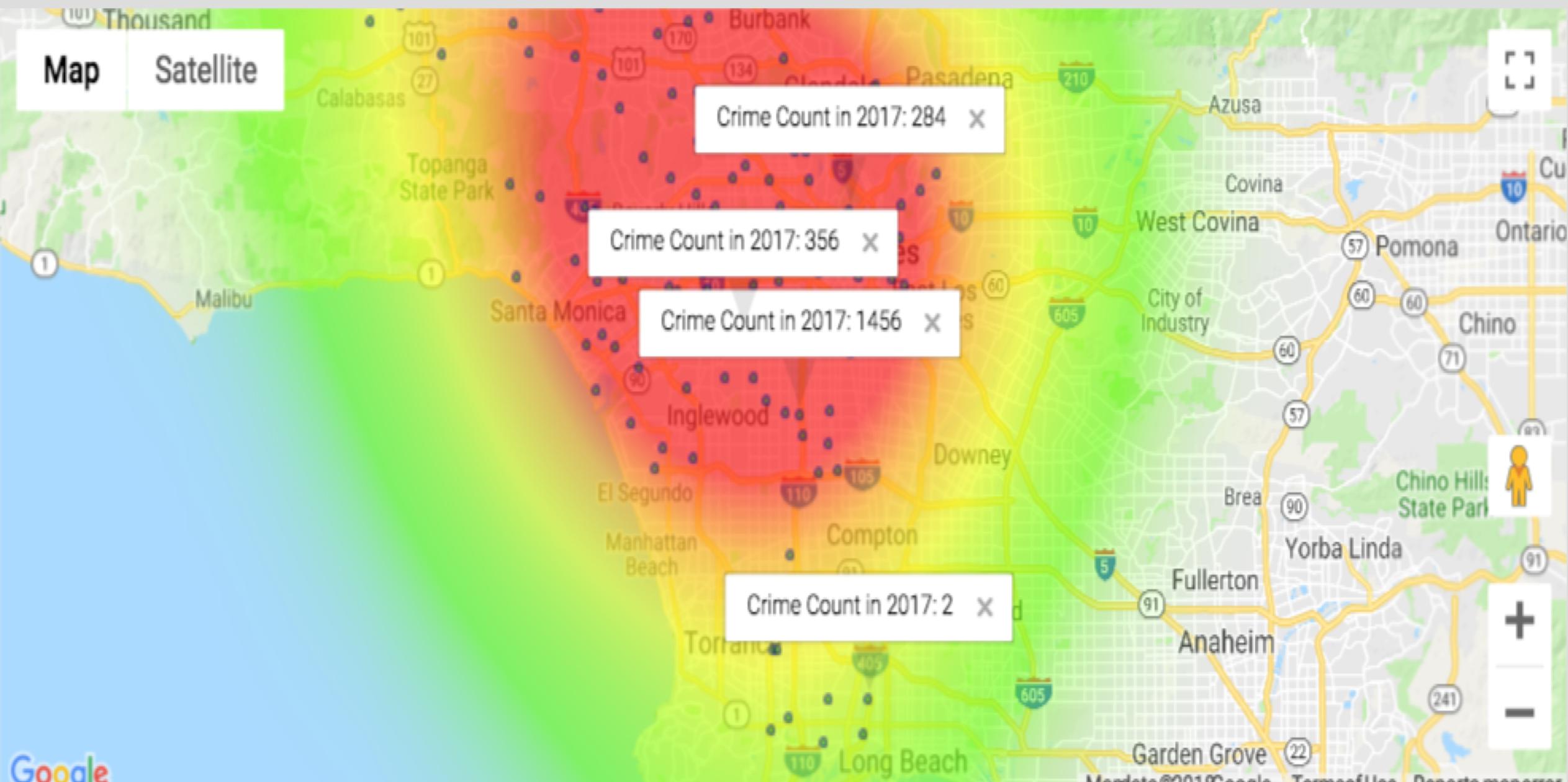
Does Population Density Matters?



4. RESULTS:

HEAT MAP: DOES DENSITY MATTER?

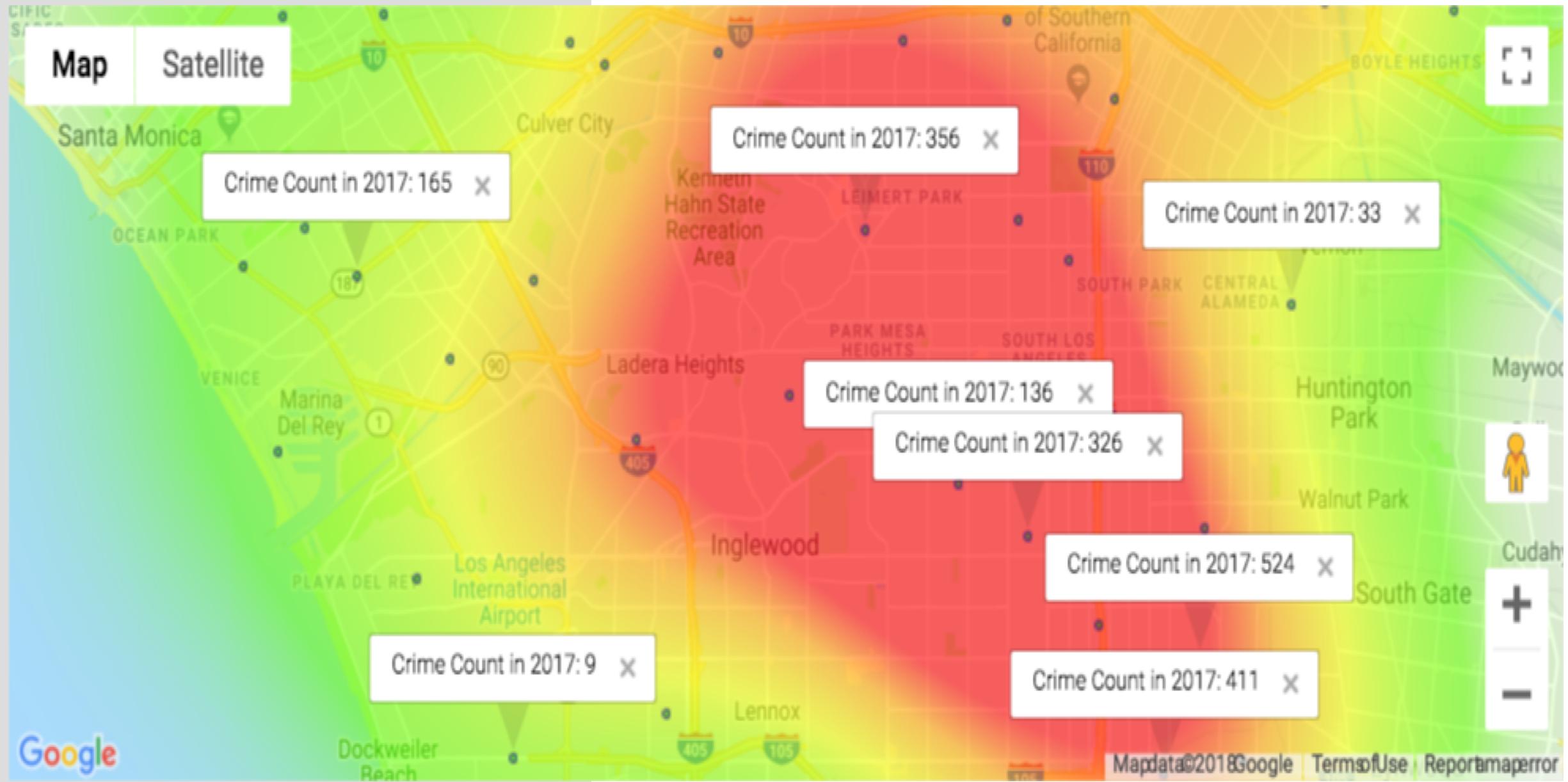
CRIME COUNT VS DENSITY



4. RESULTS:

HEAT MAP: DOES RACE MATTER?

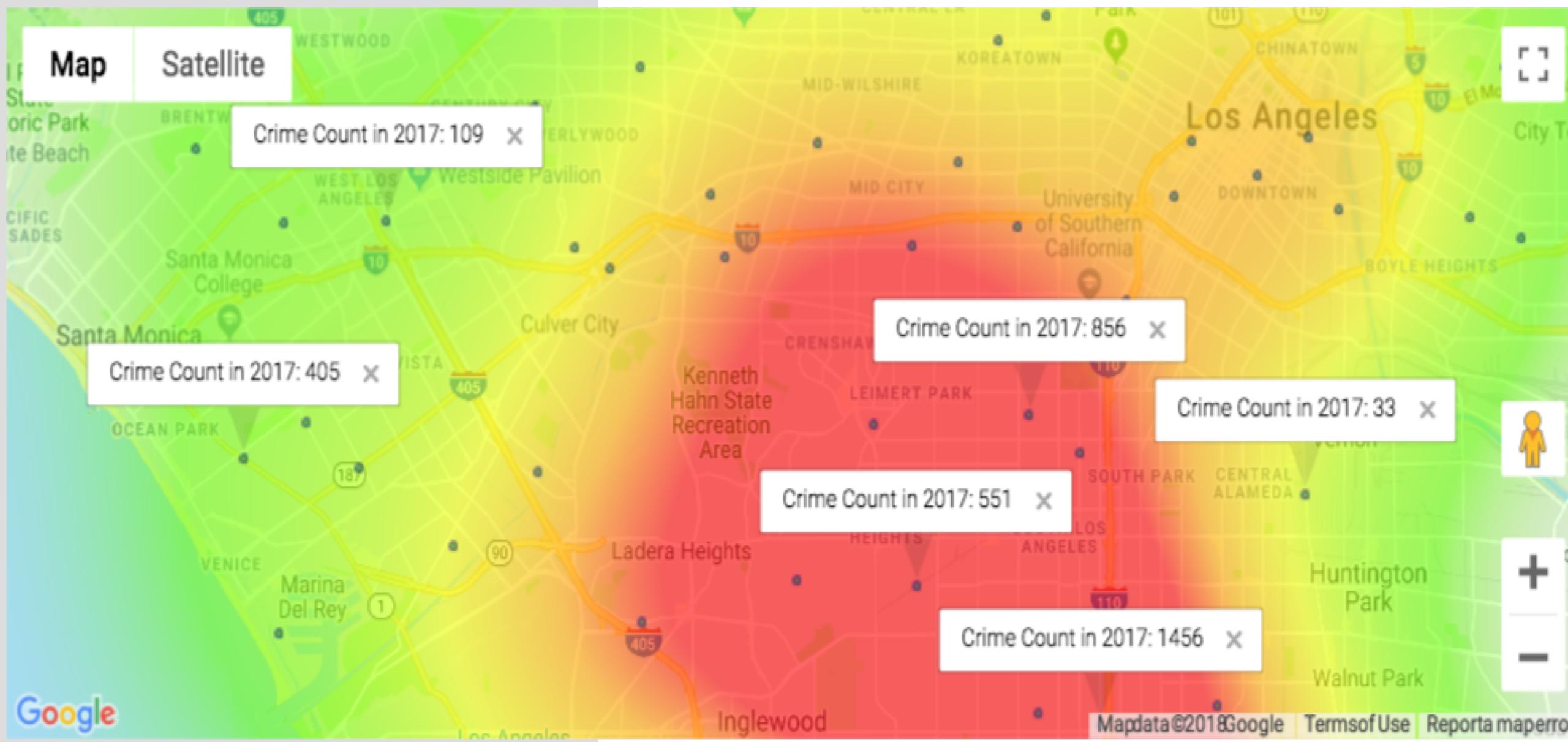
CRIME COUNT VS BLACK



4. RESULTS:

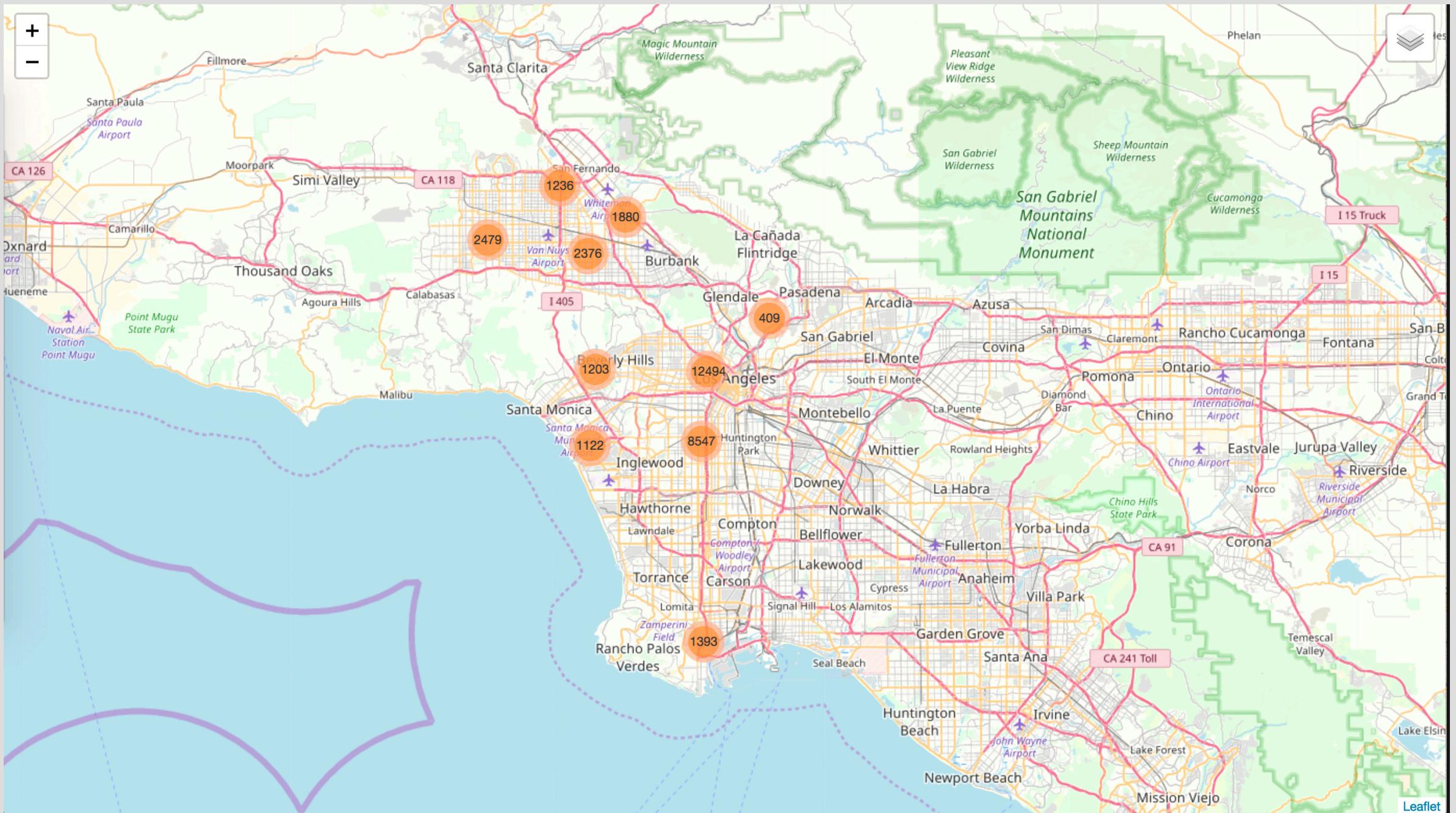
HEAT MAP: DOES RACE MATTER?

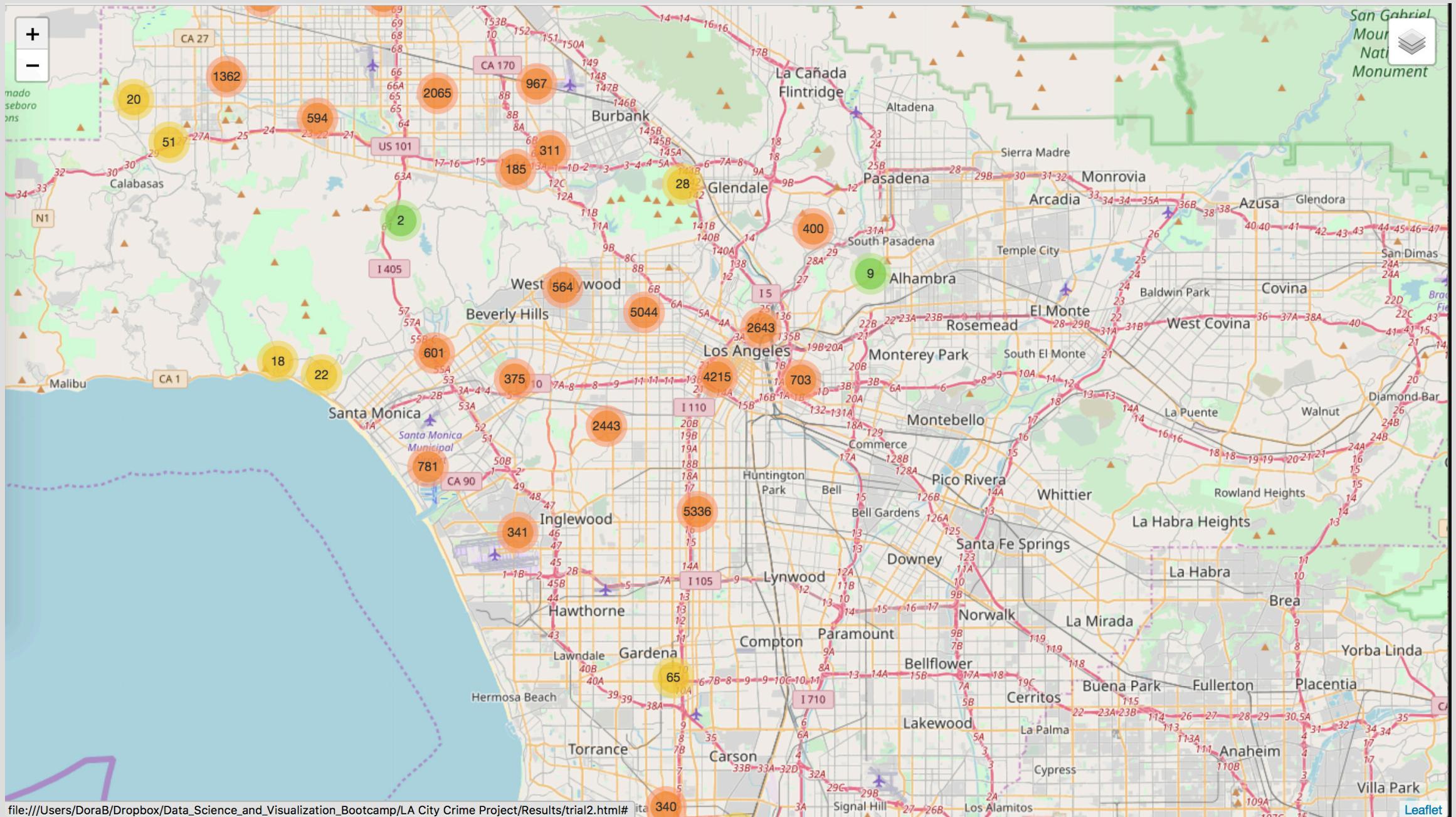
CRIME COUNT VS HISPANIC

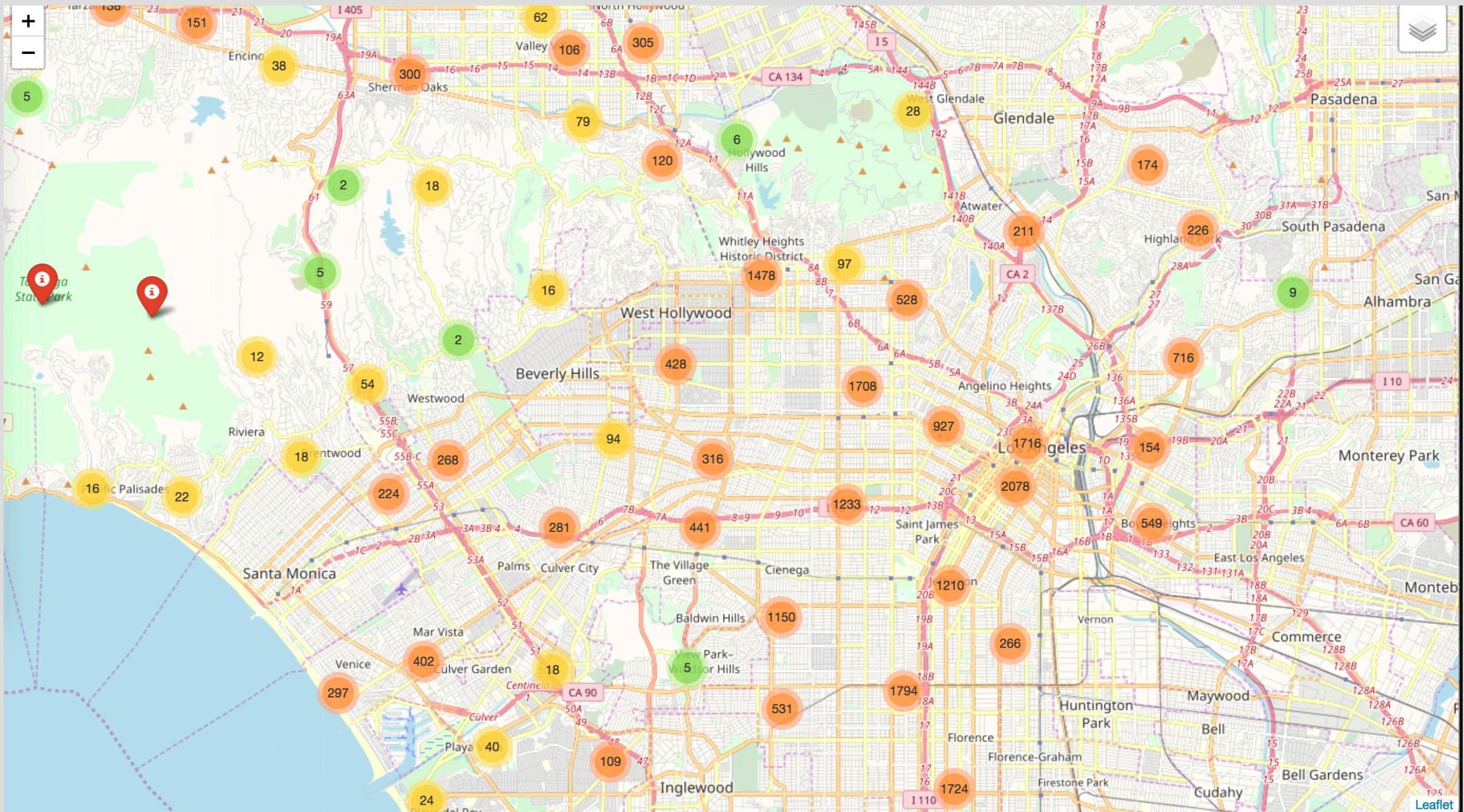


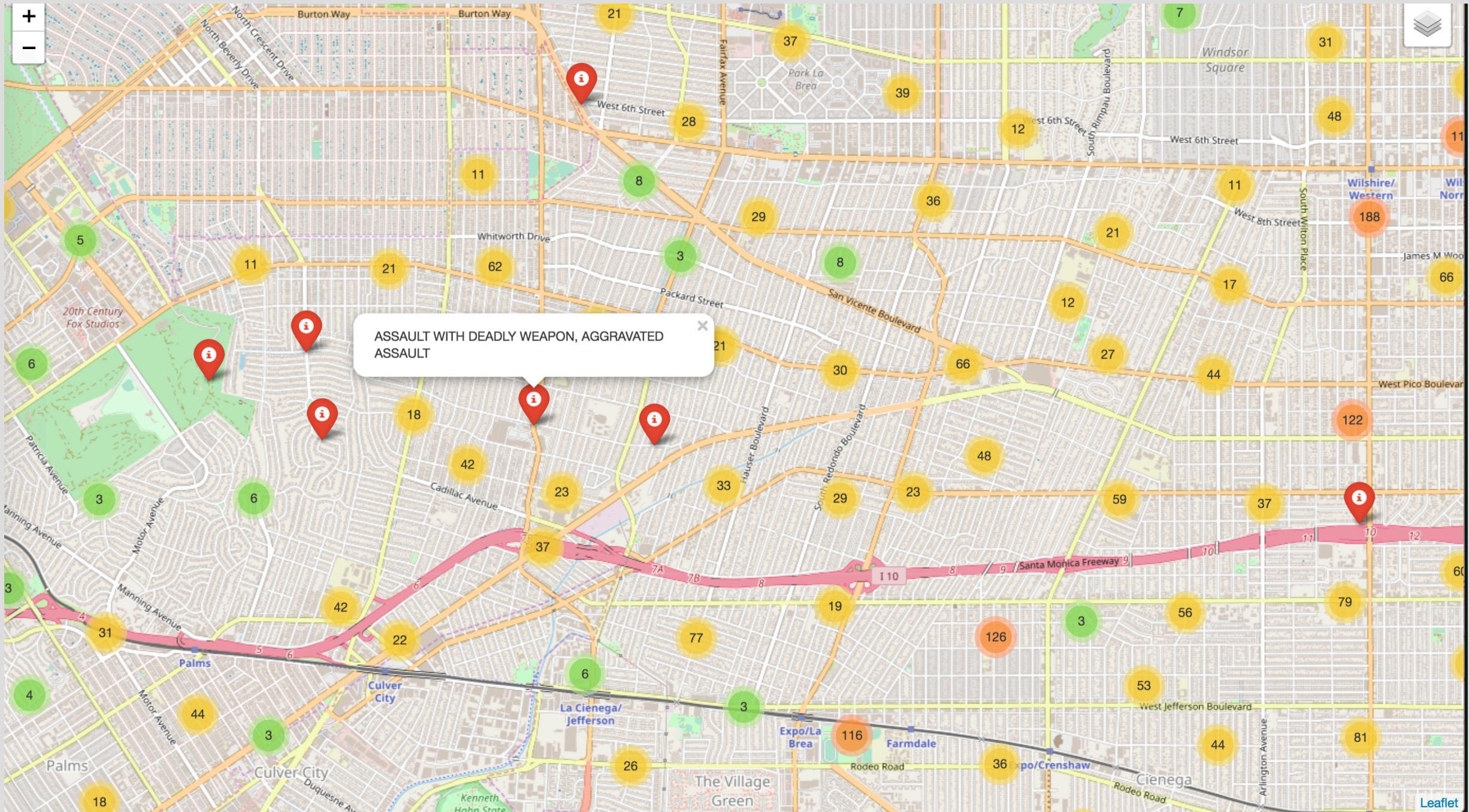
FUN STUFF: INTERACTIVE MAP

[Click here!](#)









TAKE AWAYS.

- Are certain months or years more crime prone?
 - Over the years, 2010 to 2017, violent crimes have increased in Los Angeles.
 - With the exception of February, November and December, there seems to be no significant monthly variations in the incidence of violent crimes.
- Can we identify safe versus not-so-far neighborhoods?
 - In general violent crime rate (as measured by the number of incidences per 100,000 individuals) has increased in the city of Los Angeles for the time period 2010 to 2017, at the zip-code level.
 - There are, however, some zip-codes that have experienced significant decreases in crime rate over this time period, while others have become more unsafe.
- Do demographics matter in the incidence of crime?
 - Population Density (persons per square miles) has a positive correlation with crime rate.
 - Minority Groups: Anecdotal evidence using heatmaps show that zip-codes with a higher percentage of non-Hispanic Black and Hispanic population are correlated with higher crime counts.

FUTURE WORK

- One of the trends that we did not explore was property crime which is reported to have decreased in most counties in CA.
- Work on the regressions – Run multivariate regressions.
- Create more interactive maps → Choropleth maps with zip-codes shaded for the percentage of minorities.

CITATIONS

- U.S. Census Bureau (2016). *American Community Survey 1-year estimates*. Retrieved from *Census Reporter Profile page for Los Angeles, CA*.
- <https://www.bjs.gov/index.cfm?ty=tp&tid=93>
- *Folium 0.6.0* (<https://pypi.org/project/folium/>)
- <http://python-visualization.github.io/folium/docs-v0.5.0/modules.html#module-folium.map>
- <https://blog.dominodatalab.com/creating-interactive-crime-maps-with-folium/>

WE ARE READY TO TAKE
QUESTIONS!

Thanks to Sandip, Jake and Chris!