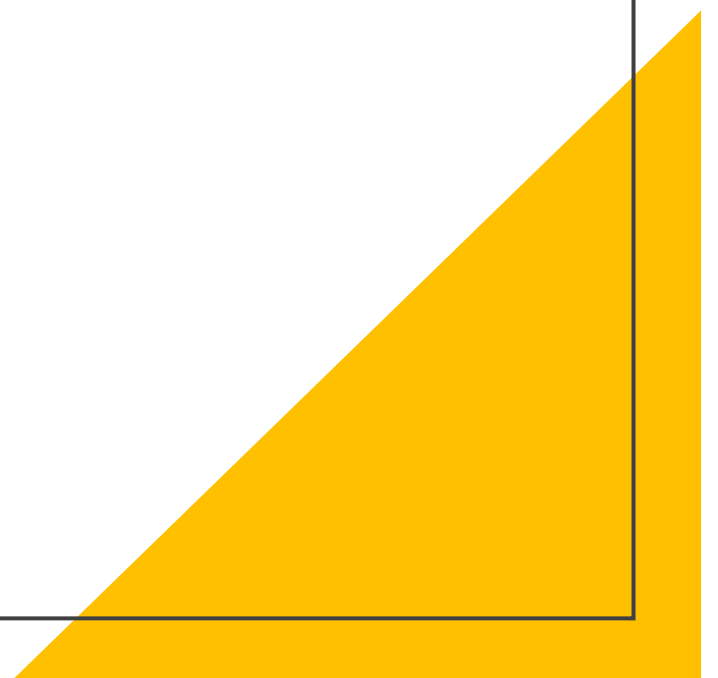


# Crime and Poverty

Project 1

November 21, 2020

Presenters: Sean Galloway, Peter Vlahos, Eliot Chen, Cora Micsunescu



# Explanation of topic

- It is an overall goal in society to reduce poverty and crime. We set out to look at this in Sacramento County as we all live here and care about the welfare of our community.
- This analysis matters as it can inform policy makers and political leaders in how to best allocate resources towards reducing crime and poverty.

# Initial Questions used to formulate null hypothesis

- What are the overall crime rate trends in Sacramento County and Oak Park in particular?
- What are the crime rates for Oak Park for 2014-2018?
- How would a heat map of the crime rate for each zip code in Sacramento County appear?
- How would a heat map of the poverty rate for each zip code in Sacramento County appear?
- What is the crime rate versus the poverty rate for Sacramento County for 2014-2018?
- Is there a correlation between the poverty rate and the crime rate?
- Is there a correlation between median home prices and the crime rate?

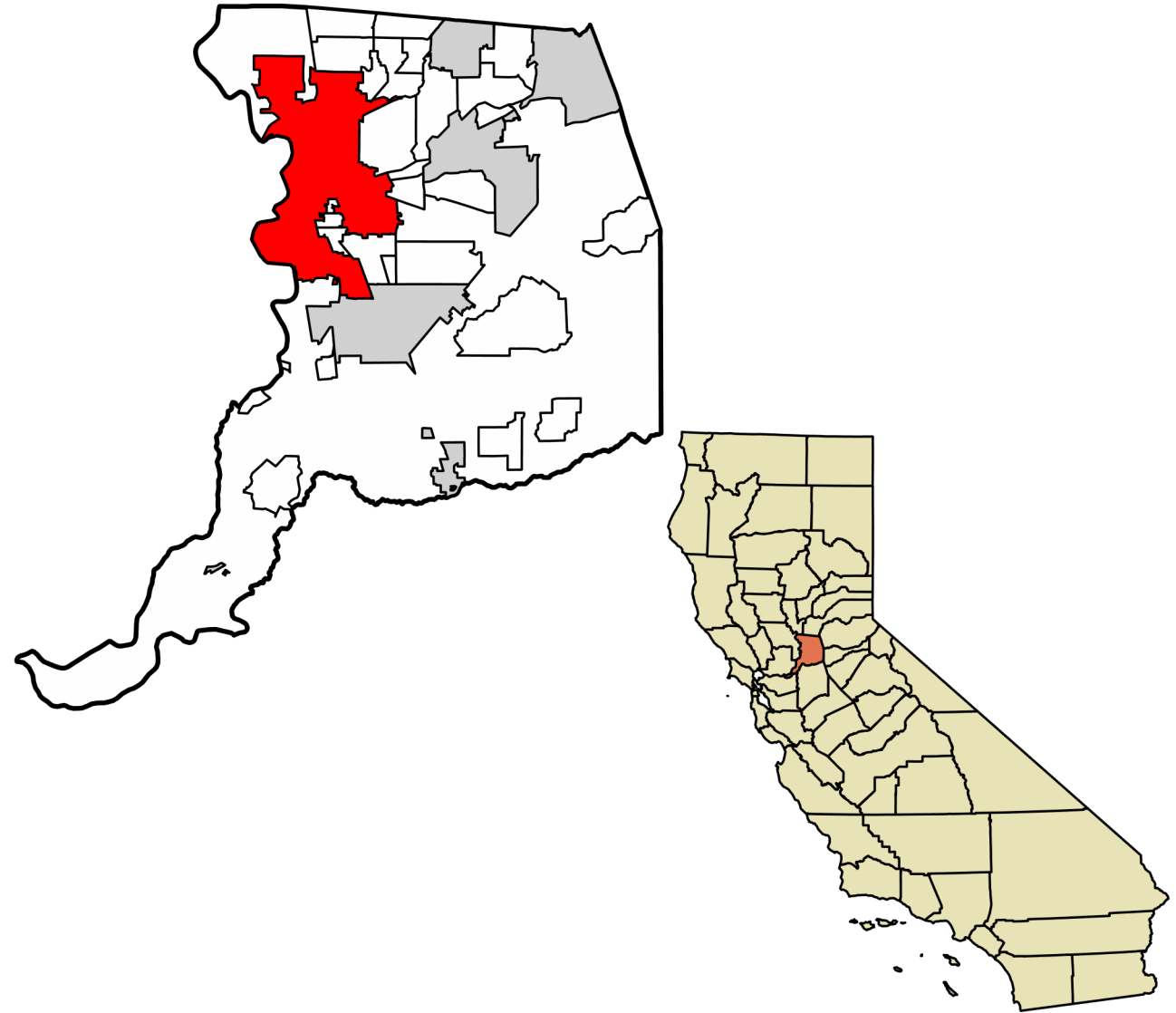
# Null Hypothesis

- For many years Oak Park has been notorious for high crime rates and high poverty rates. Recently it has experienced rising median home prices due to gentrification. Simultaneously, it has been observed that crime rates and poverty rates have decreased. Does the data reflect these observations? Does Oak Park follow the same trends found in rest of Sacramento County?
- A decrease in the poverty rate is correlated with a decrease in the crime rate.
- An increase in the median home price correlates to a decrease in crime.
- Sacramento County and Oak Park both follow the same trends with respect to the relationships between poverty, crime, and median home values.
- Confidence level of 95%

# Definitions of Terms Used

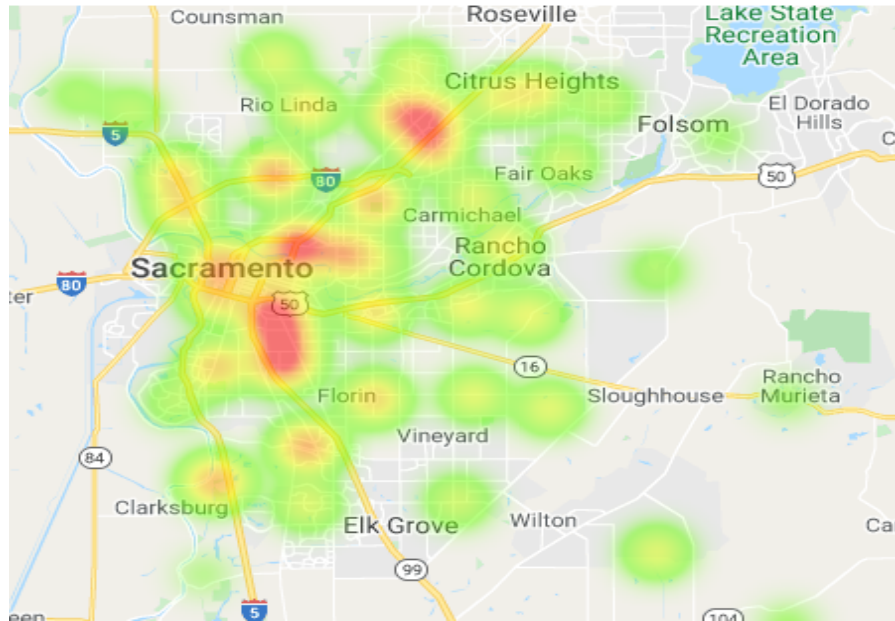
- “PC Personal”: refers to California Penal Code Crimes Against the Person which include Homicide, Mayhem, Kidnapping, Hostages, Robbery, Attempts to Kill, Assaults, False Imprisonment and Human Trafficking, and Assault and Battery.
- “PC All”: refers to the California Penal Code definition of all the types of crime including crimes against property, crimes against public justice, crimes against the person, criminal threats, and miscellaneous crimes.
- Additional information can be found here:  
<https://leginfo.legislature.ca.gov/>

Sacramento  
County,  
California

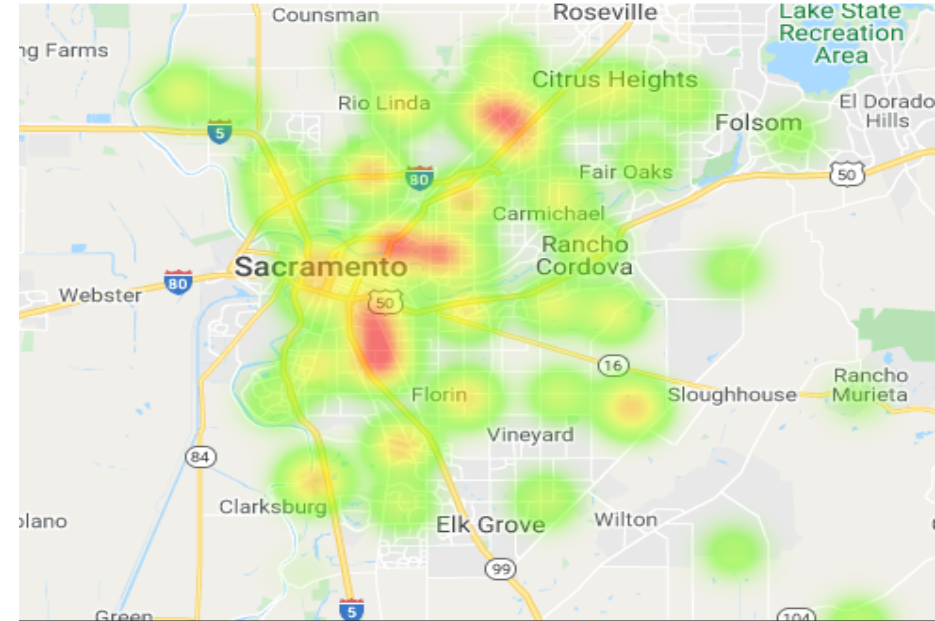


# Overall poverty rate was flat

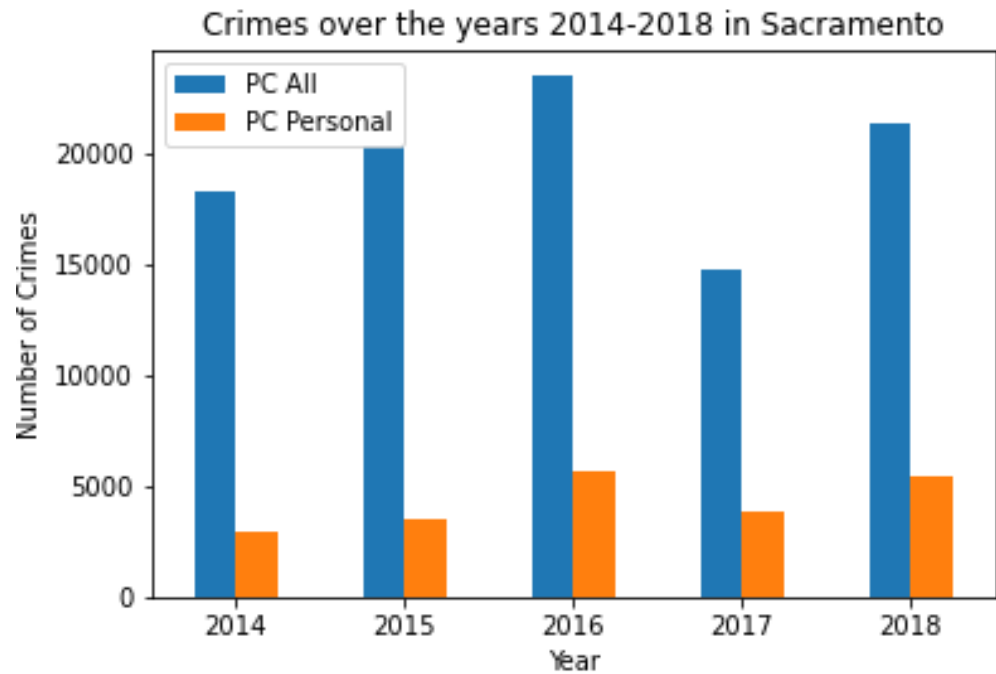
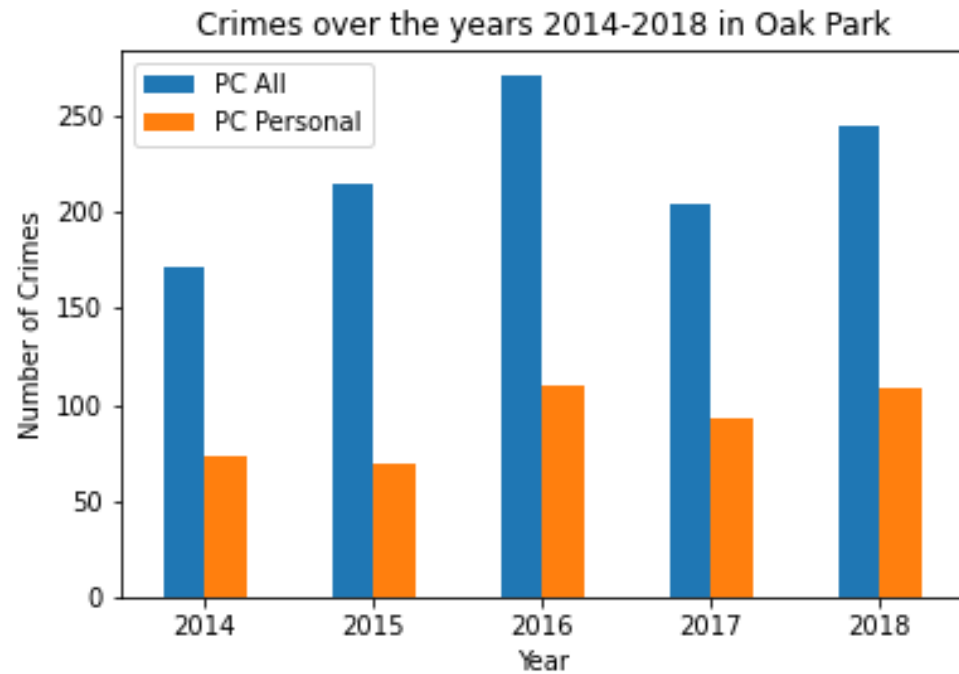
2014



2018



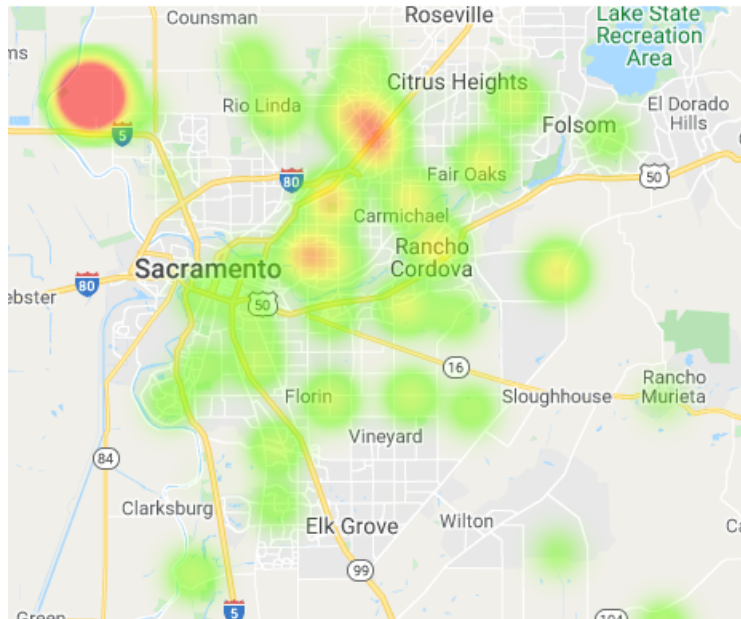
# Overall crime rate was flat



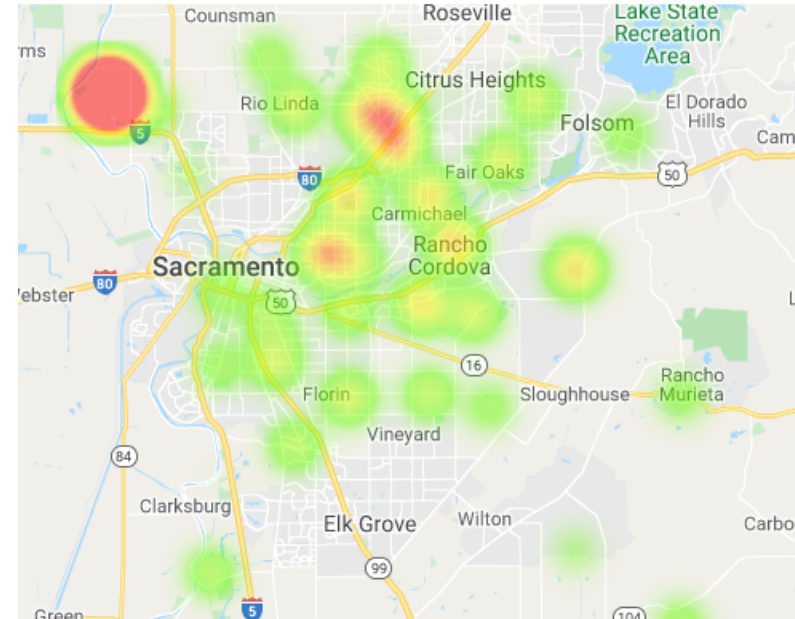


A heat map of the crime rates reflects the same trend

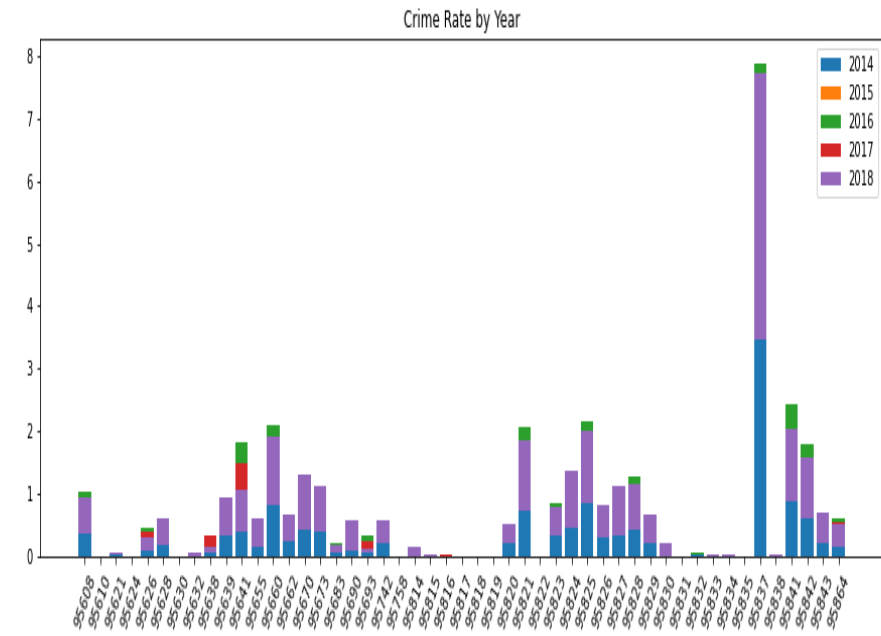
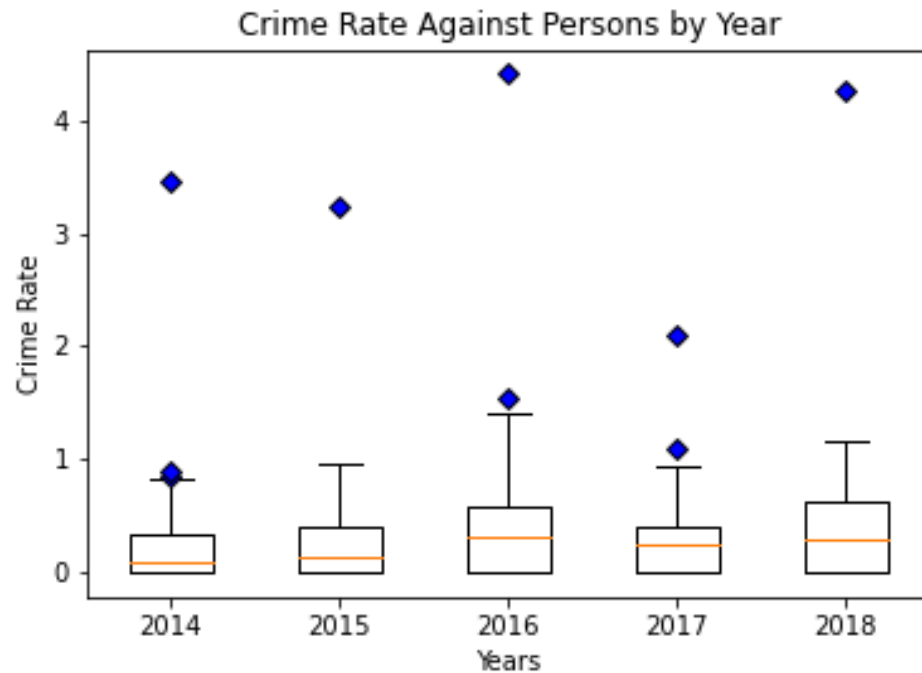
2014



2018



# An outlier in the data was identified, additional analysis confirmed it



The bar graph above clearly illustrates we have an outlier, 95837. The zip code, which is sparsely populated and isn't indicative of the demographics of Sacramento County, was removed from further analysis as it threw off subsequent calculations.

Is there a correlation between the poverty rate and the crime rate?

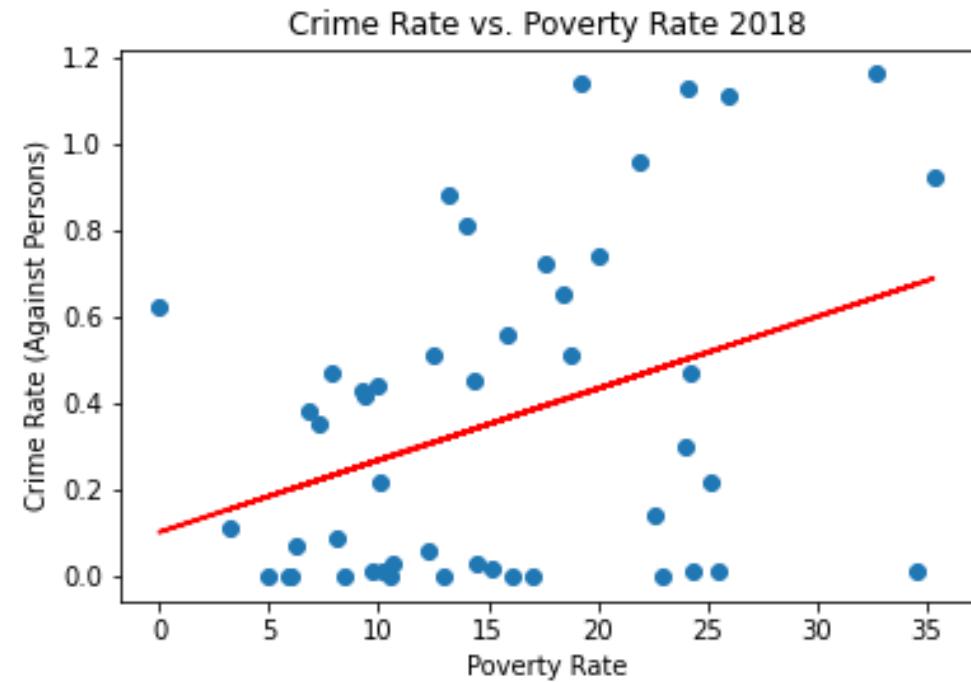
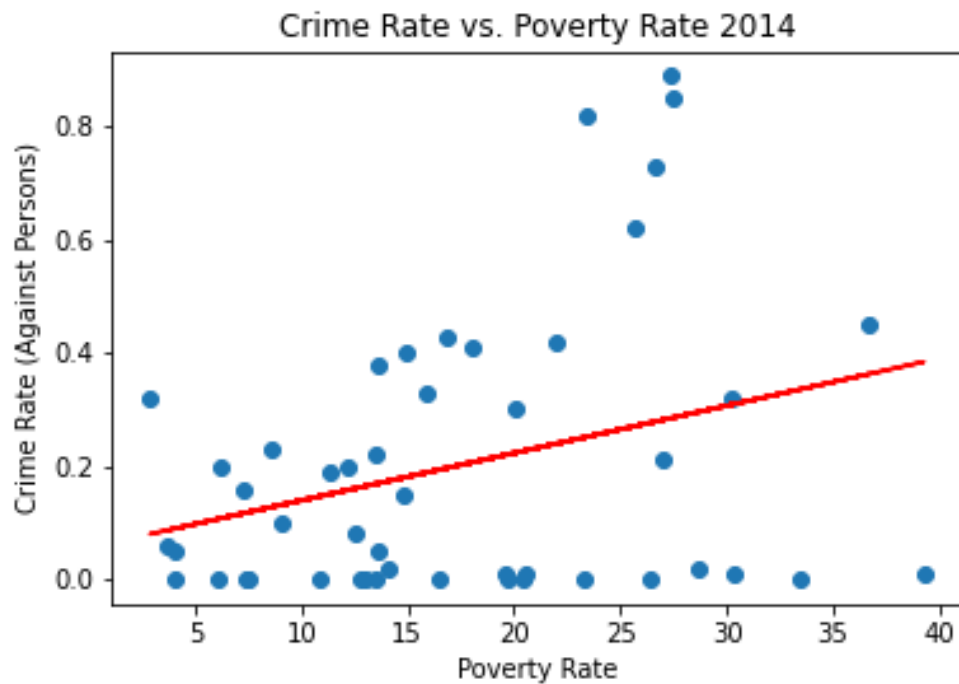
---

- Our hypothesis is that as the poverty rate goes up, the crime rate increases.



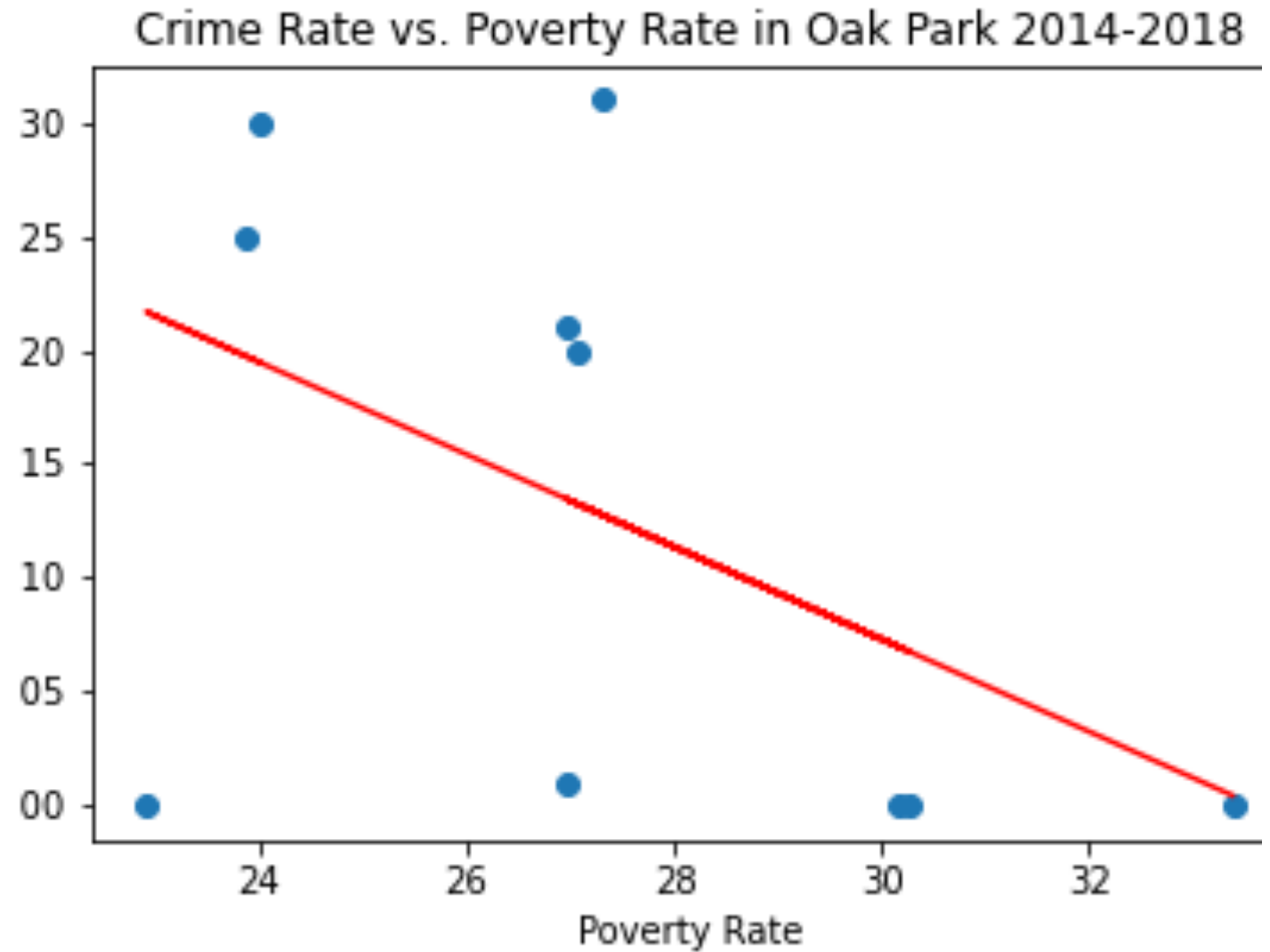
# Sacramento County

- The average p-value is 0.02 and the average r-square value is 0.12. Given these values and the diagrams' appearance, our hypothesis is confirmed.



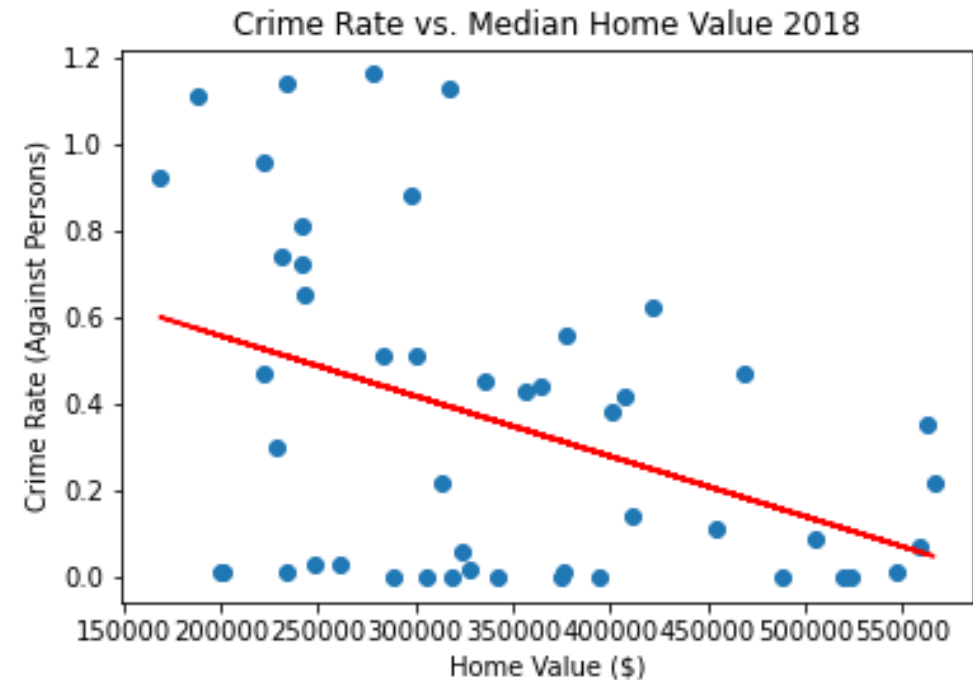
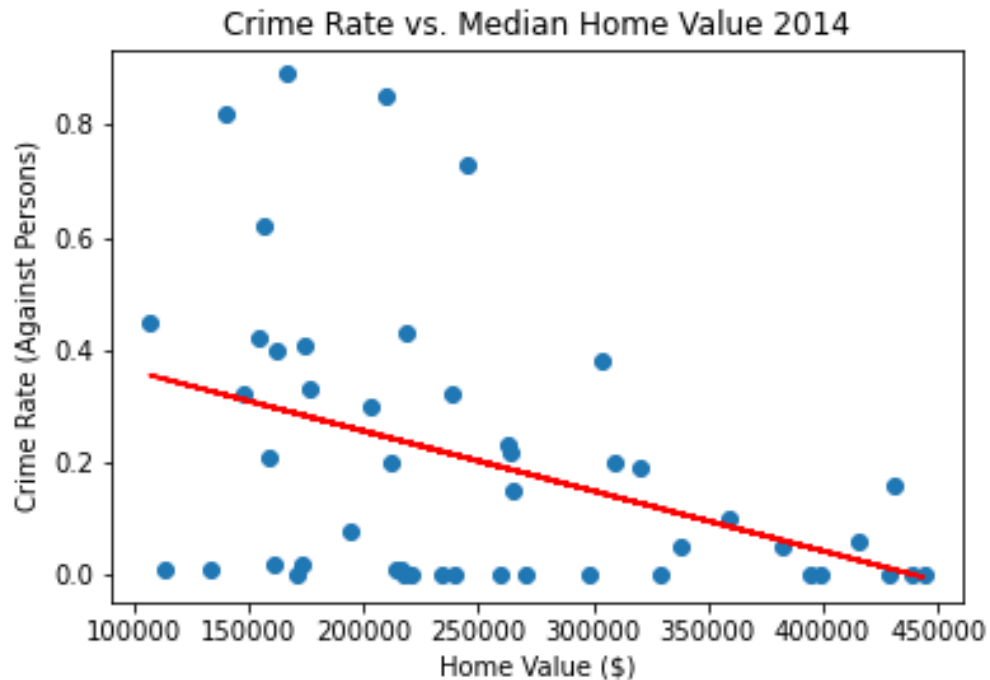
## Oak Park

- The p-value is 0.23 and the r-square value is 0.16. Given these values and the diagram's appearance, we conclude that Oak Park goes against the overall trend in Sacramento County.

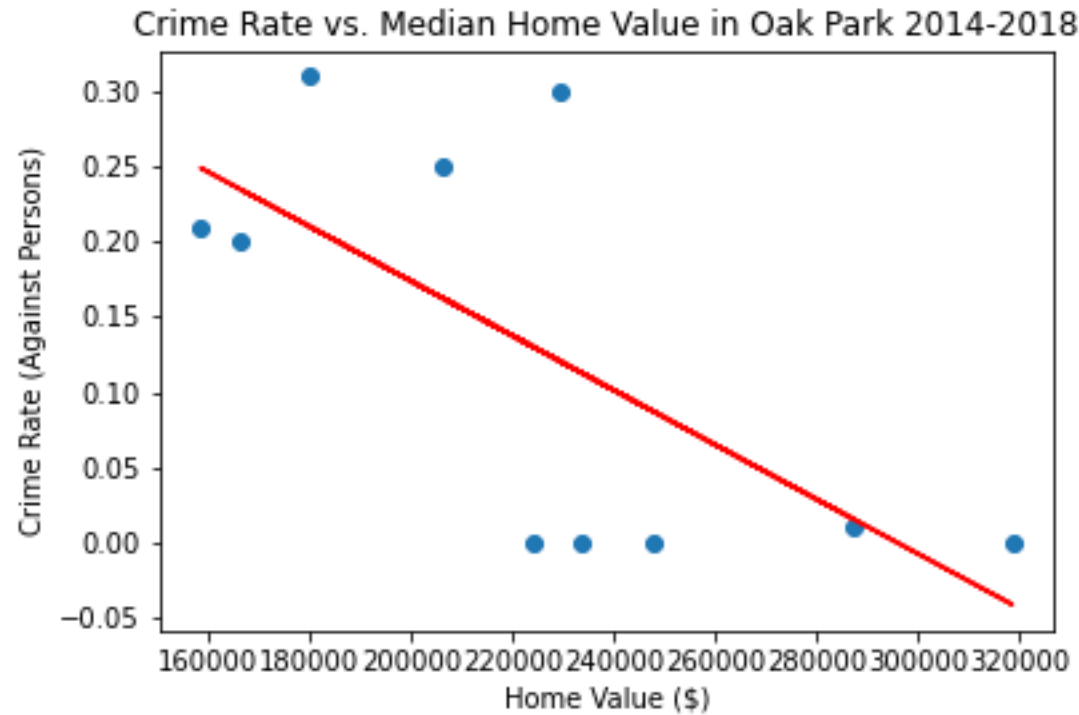


## Crime Rate vs. Median Home Values in Sacramento County for 2014-2018

- The average p-value is 0.01 and the average r-square value is 0.15. Given these values and the diagrams' appearance, our hypothesis, which is that as median home values increase the crime rate decreases, is confirmed.



# Crime Rate vs. Median Home Values in Oak Park for 2014-2018



The p-value is 0.03 and the r-square value is 0.46. The chart demonstrates that Oak Park mirrors the same trend as Sacramento County.

# Overall Conclusions

We had two main hypotheses:

- Rising median home prices correlate with lower crime rates
- Rising poverty rates correlate with higher crime rates

Overall the data for Sacramento County confirmed these hypotheses. When we looked exclusively at Oak Park, results from the data were mixed. Oak Park had a decreasing crime rate as median home prices went up, as expected. In contrast, Oak Park had a decreasing crime rate even as the poverty rate went up, which did not confirm our hypotheses. More research into Oak Park is needed to understand this.



# Data Sources

- Sacramento County Datasets: [https://data.saccounty.net/datasets/9a7f2df25a584ff9b55db274704ad7c9\\_0/geoservice](https://data.saccounty.net/datasets/9a7f2df25a584ff9b55db274704ad7c9_0/geoservice).
- Capitol Impact list of Sacramento County Zip Codes: [http://www.ciclt.net/sn/clt/capitolimpact/gw\\_ziplist.aspx?FIPS=06067](http://www.ciclt.net/sn/clt/capitolimpact/gw_ziplist.aspx?FIPS=06067)
- Census API: <https://api.census.gov/data.html>
- Google Maps API: <https://cloud.google.com/maps-platform>

# Data Cleaning & EDA

- Census data was problematic and required two types of API calls, the census.acs5 path and get requests/JSON as we could not get all the years in our data set.
  - ACS5 only went up to 2018 as that is the latest year with available data.
  - Census data needed extensive clean-up due to incomplete data and highly irregular data on the housing prices, i.e. negative home values.
  - Had to build a flat file from the data dump from Census.
  - Used this dataset for our poverty rate and median home values
- The dataset from Sacramento County provided the crime rate statistics and came in a flat file format.
  - Removed all rows with no data on zip codes or where the zip codes were incomplete.
  - Renamed columns and added new columns to include the Penal Code Violation as a float.
  - Filtered out data to only include the years 2014-2018, which was the latest five-year period of data available.
- Google Maps used in the creation of the heat maps.
  - It required try/except blocks to weed out bad or incomplete information.