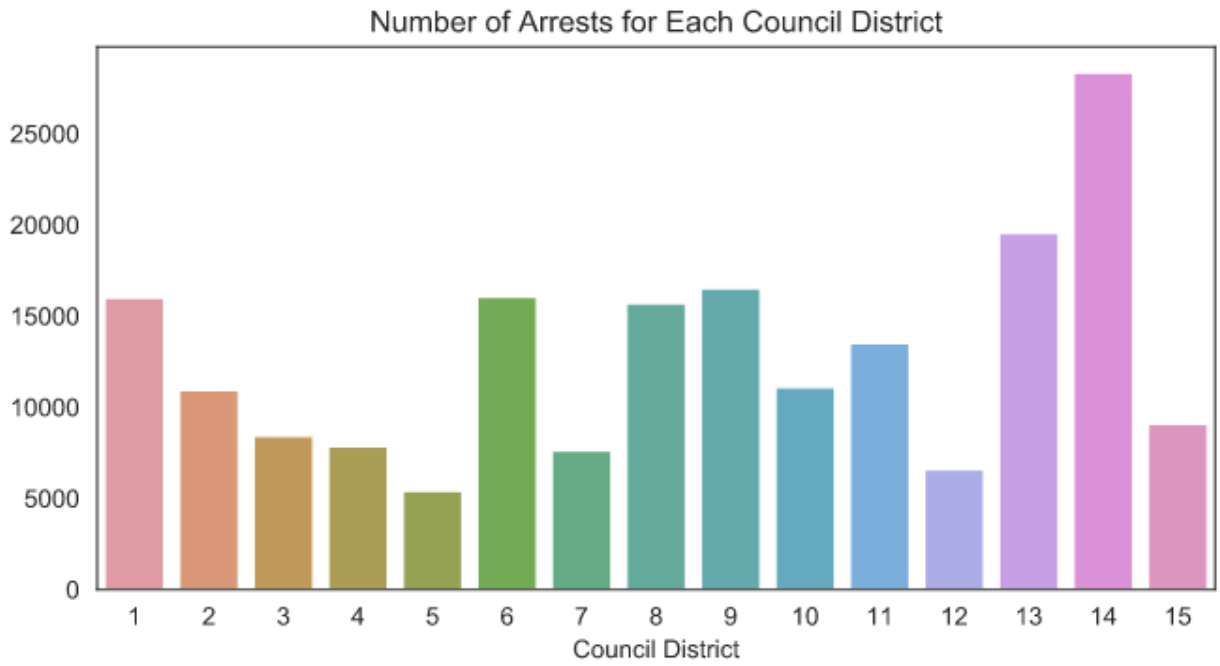
An Analysis of Police Arrests in Los Angeles

By: **Franco Gonzales**

Police killings has been a driving force for many social movements in the modern United States. In the past decade, an accumulation of incidents lead to the police being the hot seat of national news. Recent incidents such as the killing Stephon Clark brought to light what many would say the unfair treatment of minority groups by police. This isn't just a one-off thing either. Police brutality in the states spans back to decades. As a result, I became curious of police activity in general. Where are the majority of arrests occurring? Are the police targeting a specific group? Questions like those came to my head. Fortunately enough, the city of Los Angeles offers their data to the public online. Their Open Data website has a wealth of information available related to the city; ranging from permit information to population demographics. Using this resource, I downloaded arrests and population statistics (specifically income and racial breakdowns) to further explore my questions. Do police really show signs of unfair treatment to minority groups? How does income come into play?

The arrest data has a large amount of information to play with, having 17 unique features. The information that you'd normally think present is available: race, gender, charge, the address, etc. To make this a little easier on my local machine, I filtered out all of the recorded arrests from the beginning of 2018 to now, the end of 2019. The most important feature given is location, given in latitude and longitude. There is a problem though. The population statistics like income or race are given via Council Districts. Council districts are 15 unique sections of Los Angeles used by the city council for voting purposes On the website, if you were to search for a feature like income, it is not given in a 1:1 match with the location of the arrest. Available here is the shapefile which contains the coordinates that cover each district. Using R and its libraries (sf, rgdal, sp, and rgeos), I was able to match each arrest with the specific districts that they occurred (this took forever). We can look at which district has the most arrests for a start.



Let's look at some of the more noticeable districts.

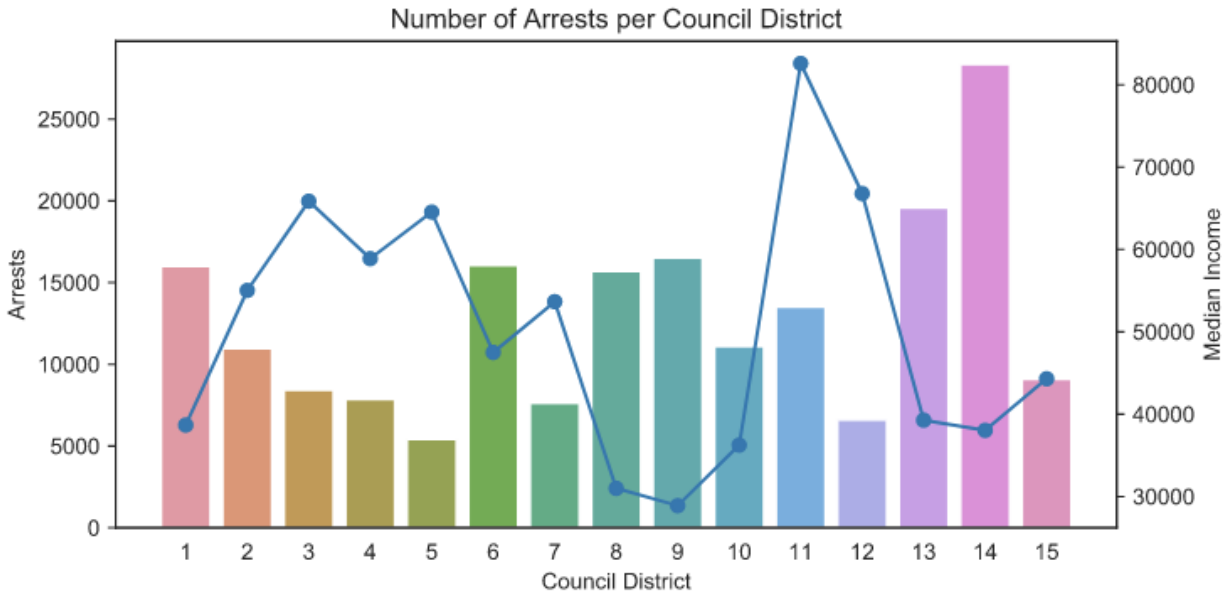
- District 14, the highest amount of arrests: Boyle Heights and Northeast Los Angeles. Northeast Los Angeles consists of Eagle Rock, Highland Park, and Glassell Park. Skid Row is a notable landmark in 14.

- District 13, the second most number of arrests: Central Los Angeles, covering parts of Hollywood, East Hollywood, Atwater Village, and Silver Lake (directly south of Glendale).

- District 5, the lowest number of arrests: Westside Los Angeles, Central-Eastern Santa Monica Mountains, and central-southern San Fernando Valley. Notable landmarks: UCLA, Fairfax, Melrose, portion of Beverly Hills.

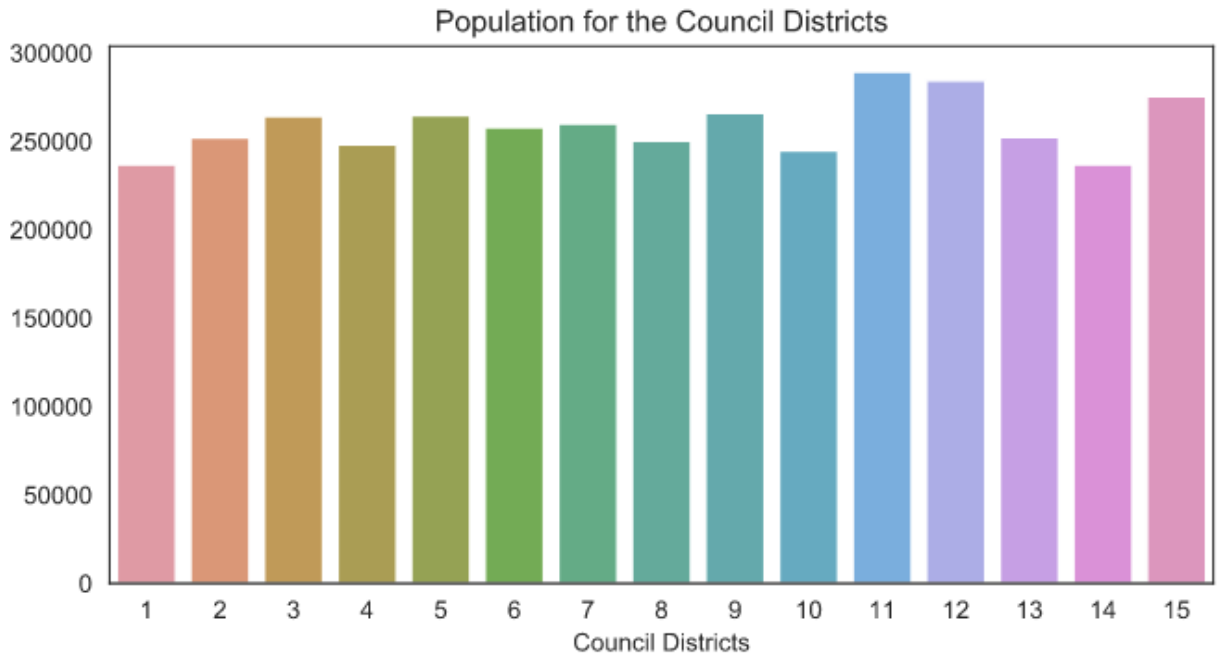
- District 12, the second lowest number of arrests: Notably the *only* Republican majority district. Encompasses the west and northwestern sections of the San Fernando Valley. To put some names: Chatsworth, Granada Hills, Northridge, Porter Ranch, West Hills, and lastly Sherwood Forest.

Going through each district and naming specific landmarks is a little unnecessary, but it is nice to be able to see where the low and high activity happens by name. With the income information that we were given, we could superimpose the median income and number of arrests for each district.



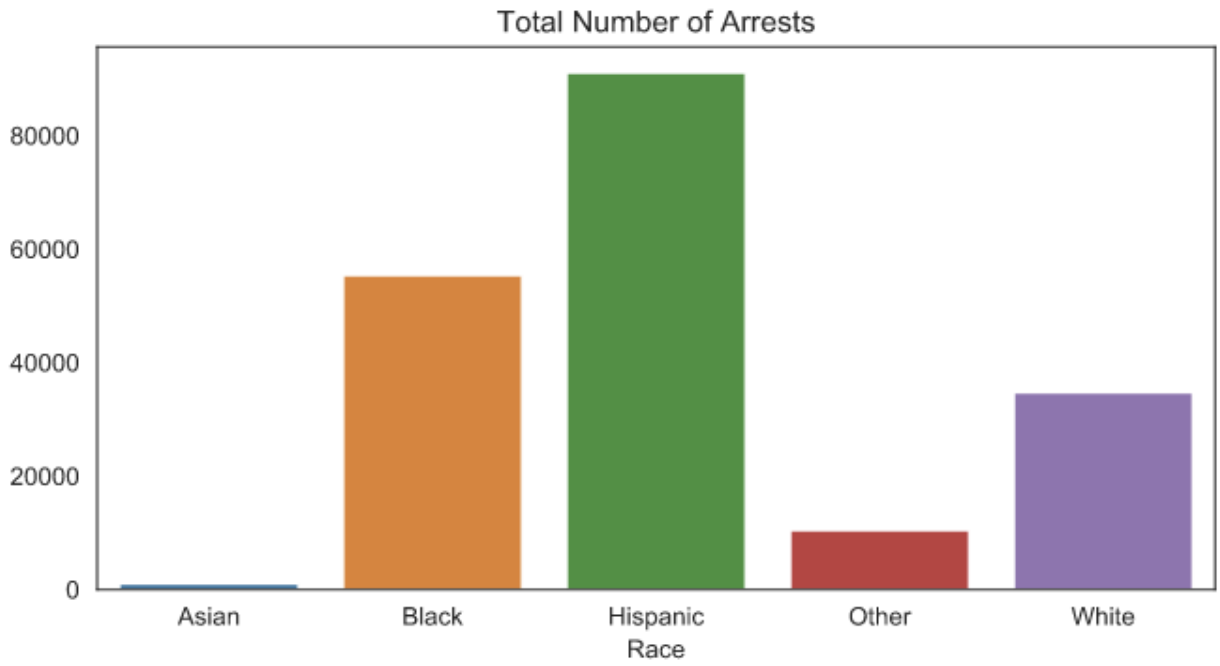
If we were to single out the districts that we had pointed out earlier, it would seem as if income and crime do have a relationship. Districts 14 and 13, where the most crime occurred, had lower incomes relative to the median value, around $49,000. The lowest incomes were districts 9 and 8 respectively. So while there is a correlation between income and occurrence of crime, the relationship isn't as binary as we'd think. Instead of using it as an absolute measure, it should be used as a guideline. Take a look at district 11, the highest income out of all the other districts. It doesn't necessarily have the lowest amount of crime, but it is low when you compare it to the other regions.

I wanted to see if the total population was different, leading to higher arrests in one area. In a vacuum, the higher number of residents, the higher the arrests. Below is the population counts for each district.

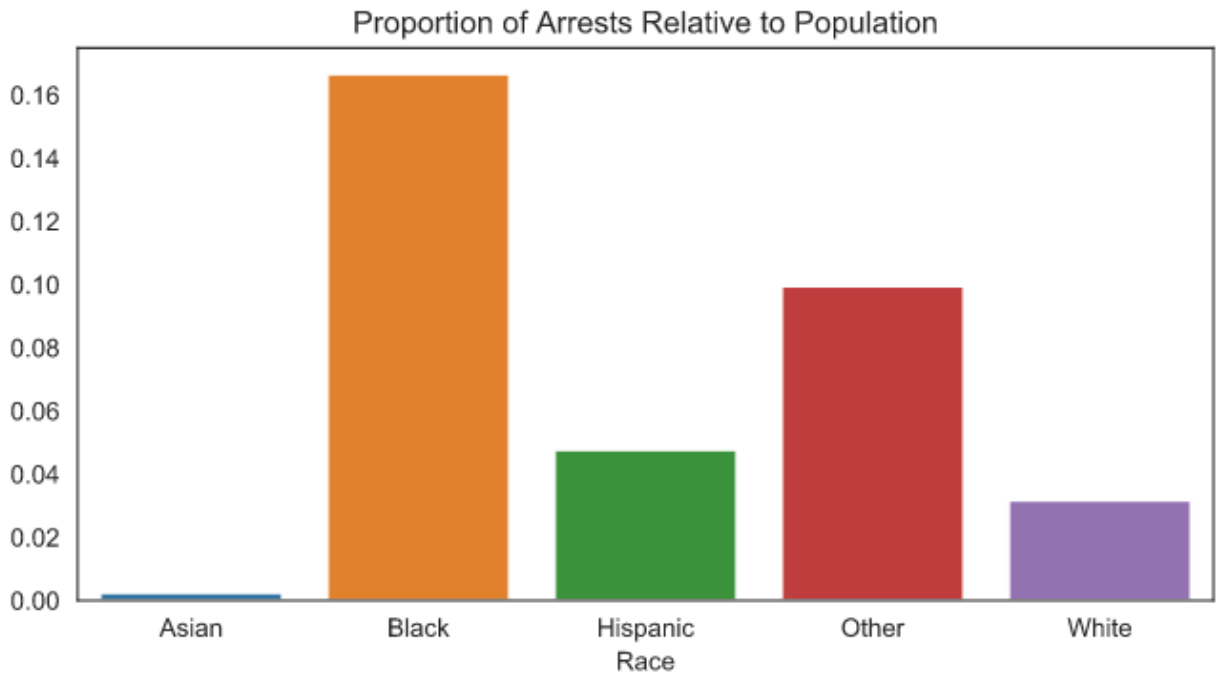


So as we can see, the differences between each district is so small that population size shouldn't have affected our plot comparing income and arrests above. If anything, look at 14, the highest arrest region. It actually has one of the lower counts.

Now let's look at race. The arrest data has a feature called Descent Code, which contains a mix of different ethnicities. Due to the wide variety of codes, I tried to make generalize the races into groups: Hispanic, Black, Asian, White, and Other. This took some playing around with as the original data set gave variations of our group. For example, Filipino and Laotian had their own categories which I placed into Asian. So after rearranging them into our umbrella terms, we get the following plot which shows the number of arrests for each race.



To be honest this was about what I expected. The only real surprising fact here is how low the number of arrests are for Asians. The bar is barely visible compared to the other races. They still *are* considered a minority, although they are portrayed differently by the media. That's a topic for another conversation, but that could easily explain the plot above. In contrast, the race with the most arrests is Hispanic, with Black directly following it. Unfortunately, this information is useless without context. We can't just look at the total number of arrests for the races and make conclusions. To explore a little more, I went and grabbed the population breakdown by race, and found the proportion (total arrests for the race/ population for the race). The results, while a little shocking, isn't really surprising.



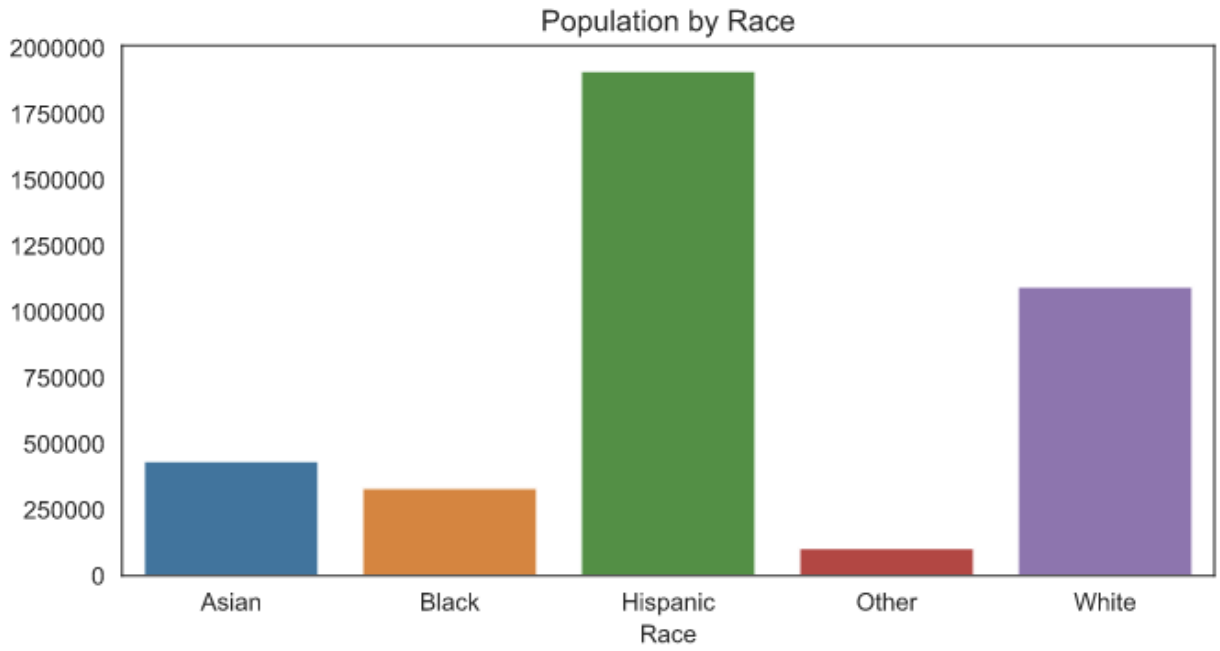
This answers the questions that I asked in the beginning. What the plot above shows is that around 16% of the black population in Los Angeles will have had been arrested by the police. While this could be due to a multitude of other reasons, the topic of unfair targeting of Black people by the police can make some cases here. The population of Blacks living in Los Angeles is *significantly* smaller than the other groups, but they are arrested exponentially greater than other groups. Why exactly? Well, we'd need to delve deep into history and other facts to get an answer, which isn't the main point here. Along with Black, the proportion of Other arrests also shot-up, which is already a very vague description. Looking at the table for the Descent Code feature, they covered essentially the vast majority of ethnicities. So Other could refer to mixed race individuals or those that are hard to identify. Another point of interest is while there are a significant amount of arrests of Hispanics, the low proportion says that there are just a higher amount of Hispanic residents in Los Angeles. What *is* interesting is why are Asians low for both proportions and total arrests? This could segue into the topic of the Model Minority, which is also a touchy subject in its own right.

Charge description is also given in the arrest data. Using the value counts and subsetting the specific columns for each race, I created a table which showed the three highest charges for each of the groups. For all of them, the most frequent was "Miscellaneous Other Violations", which is vague, so I decided to exclude it all together.

Race	Charge 1	Charge 2	Charge 3
Asian	Prostitution/Allied	Liquor Laws	Driving Under Influence
Black	Narcotic Drug Laws	Aggravated Assault	Other Assaults
Hispanic	Driving Under Influence	Narcotic Drug Laws	Aggravated Assault
Other	Driving Under Influence	Aggravated Assault	Narcotic Drug Laws
White	Narcotic Drug Laws	Driving Under Influence	Other Assaults

To be quite honest, the only real surprising thing here is the charges that Asians were given. I was thinking that assault or drugs would be on the list, but they aren't. Contrast that with the other races, where drugs along with assault are a common theme. Another thing to point out is that DUIs are not in the top charges for Blacks. In fact, larceny (theft) and prostitution are more prominent than DUIs.

To summarize, I'd say that income and race does have some sort of influence on police activity. Even if we all tried to be as objective as possible, our biases will always have some sort of effect. This is called implicit bias. The way the media and our peers portrays certain racial groups or "that old building" will be taken in subconsciously regardless of how we actually feel. So going through this analysis, it was interesting how it panned out. Income does seem to have some sort of correlation with police activity. The higher the income, the lower the number of arrests, but we shouldn't use it as a 1:1 measure. Race though, is a bit trickier . Initially we looked at just the total number of arrests and saw that Hispanics with the most by a *significant* margin. Looking at the count plot for populations, it does make sense.



The more present, the more arrests are gonna happen just going off of basic logic. The topic that should be investigated is why the proportion of Blacks are notable larger than the other groups. In comparison, Whites and Asians having the lowest arrests compared to their population is also another talking point. If we were to just look at strictly the information we have presented, we'd say that Black people (out of all minorities) are treated most unfairly by police. Whether or not it's true in the real world, where context isn't just numbers is up for debate.

Links:

- Code
  - Linking shapefiles and districts via R:
    - https://github.com/francogonzales/la_arrests/blob/master/findcouncilID.R
  - The analysis, done via Python:
    - https://github.com/francogonzales/la_arrests/blob/master/analysis.py
- Data
  - https://data.lacity.org/browse