

The George Washington University

School of Business

Police Violence & Racial Equity:

An analysis of police violence by state, demographic and politics

ISTM 6217 - Internet of Things Management

GROUP 10

Yubika Bhattacharai

Getinet Bogale

Nishellie Navaratne

Table of Contents

I.	Executive Summary	3
II.	Data Description	4
III.	Research Questions	6
IV.	Data Cleaning and Preprocessing	8
V.	Methodology	8
VI.	Analyses and Model Results	9
VII.	Conclusion	22
VIII.	References and Citations	24

Executive Summary

The 2020 killing of George Floyd sparked a nationwide protest and a movement to address police violence against African Americans. This analysis aims to understand crime trends in major metropolitan areas and states, and further aims to understand the influence of race on police violence. The police violence statistics data analyzed include socioeconomic factors, demographic data, political leaning and city law enforcement budgeting.

The dataset used in this analysis is part of a multi-series collection of data aimed to understand police violence. Our team used SQL and visualization techniques for this analysis because 1) our dataset consists of mostly numbers 2) we would be utilizing a series of queries to answer the research questions we had.

Through our analysis our team found: 1) a higher number of arrests of black individuals 2) an increase in police funds and revenue did not have a positive impact on crime rate and 3) while African Americans are significantly more impoverished than white individuals in major metropolitan areas no correlation between politics and police violence was found based on this dataset.

Data Description

Police Violence & Racial Equity Dataset

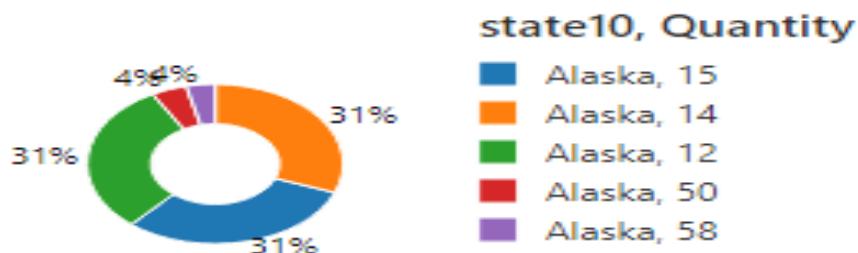
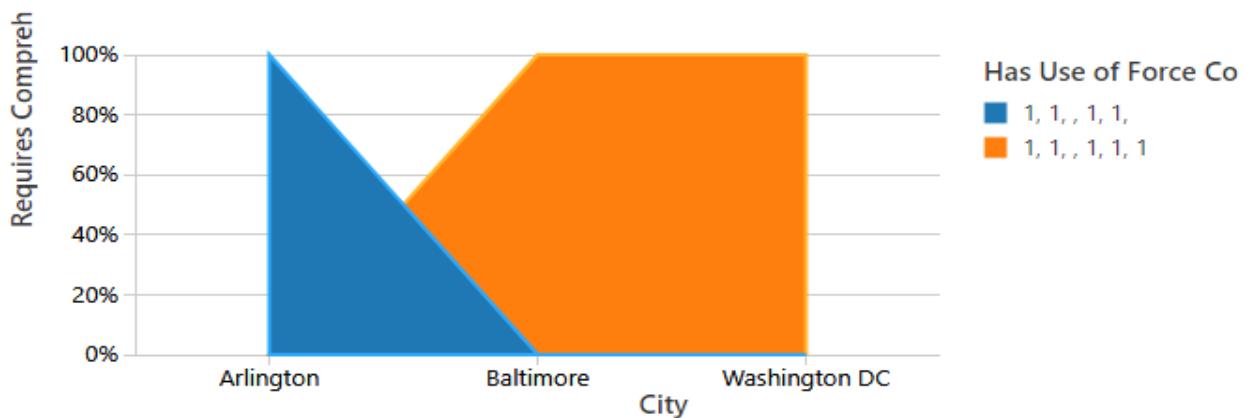
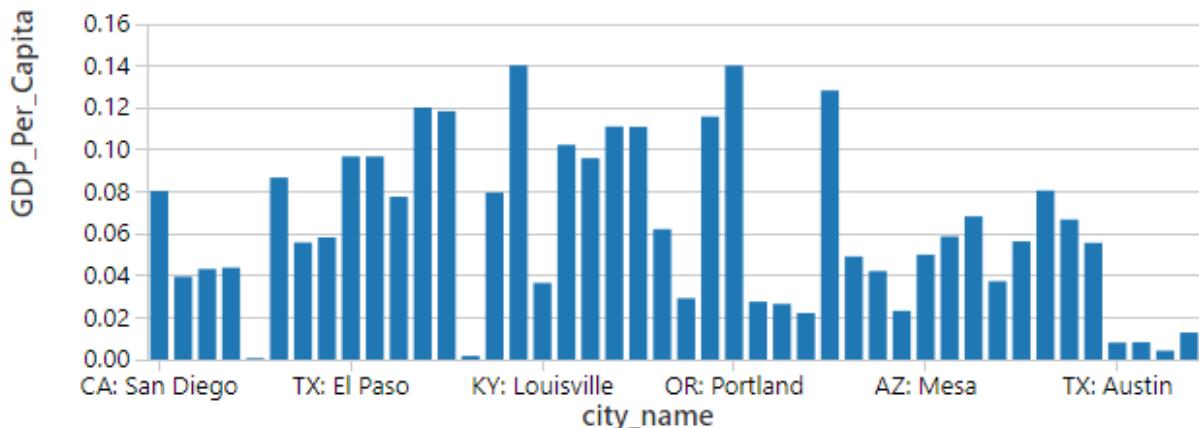
The dataset we decided to use for this project is titled “Police Violence & Racial Equity.” This dataset can be found on Kaggle.com, and the size of the dataset is 3.49 GB. The unique interesting fact about this dataset is that it has three parts that are pulled together from several different sources that explore police violence and racial equity that exists in the United States. Within these three parts, there are between five to six separate CSV files containing the followings:

- Part 1. Consists of citizen deaths, police deaths, and other outcomes
- Part 2. Consists of social and economic data, Political learnings of Citizens, Sales of DoD equipment to law enforcement agencies, city budgets, Police department head counts, Police department policies and contract provisions, Juvenile arrests by type of crime and race, and crimes and arrests for the prime city in the four largest metro areas.
- Part 3. Consists of Protest activity, Police response, Press activity, and Video clips of incidents.

For this project, we agreed to use partial datasets from part 1 and part 2. Some of the variables from part 1 on the deaths_arrests.csv file is *state, city, PD, Black People Killed, Hispanics Killed, Native Americans Killed, Asians Killed, Pacific Islanders Killed, White People Killed, and Unknown Race Killed*. The next file we will use is Police_Killings_MPV.CSV, and this file includes variables such as *Victims name, Victims age, Victims gender, Victims race, Date of incident, Street Address of Incident, City, State, and Zipcode*.

The datasets from part 2 focuses on Crime_data. The first CSV file is arrests_national_adult.csv. The variables used are *State_abbr, year, offense_code, offense_name, agencis, population, total_male, and total_female*. The Demographics folder includes CSV files like education_census_bureau.csv, housing.csv, Politics_538.csv, and poverty_census_bureau.csv. The next folder we will use in our project is police_info, and this folder includes CSV files such as, budgets.csv, dod_equipment_purchases.csv, police_contracts.csv,

police_employment_fbi.csv, and police_policies.csv. These datasets we will use for our project are interesting because we can further examine and analyze the problems that exist within certain police departments in the US, as well as police violence and racial unfairness due to certain factors that play a part within society. Below a visual of GDP per capita, how often police force is used, and the quantity of equipment purchased are shown to help us better understand the data description in a visual form.



Research Questions

Throughout the United States, especially in 2020 and even earlier years, there has been an uprising in police violence and struggles with racial equity. This data does not represent the entire United States, nor does it show the social dysfunction at question. The purpose of this data is to highlight trends and statistics in the US's four largest major metropolitan areas in regards to social and economic information.

Overall, our goal is to utilize the data provided in order to analyze crime trends in four different metropolitan areas in order to understand how each area is affected by police violence and racial equity.

The dataset specifically includes information about:

- Social and economic data
- Political leanings of citizens
- Sales of DoD equipment to law enforcement agencies
- City budgets
- Police department headcounts
- Police department policies and contract provisions
- Juvenile arrests by type of crime and race
- Adult arrests by type of crime, age, gender and race
- Crimes and arrests for the prime city in the four largest metro areas.

With this in mind, our group wanted to narrow our research to focus on specifics when it came to police brutality in cities with higher funded police budgets, and a more diverse population in terms of gender, race, and political affiliations.

Through our project we will be looking into three main questions as well as a model:

- 1) Is there a relationship between police department policies, purchases, poverty rates and political leanings of citizens in the four metropolitan areas?
 - a) This will focus on policies, equipment purchases, poverty rates and politics within the four metropolitan areas (Chicago, Dallas, New York, & LA)
- 2) What is the total revenue generated by each state according to each criminal offense?
 - b) This will focus on the type of crime in each state, how much revenue is generated.
- 3) What are common denominators seen with the arrests made across the entire United States?
 - c) This will focus on national arrests and the most common type of arrests made along with the perpetrators' characteristics such as their age, sex and race. Additionally, we will take a look into states' local government spending as a percent of total personal income to see if a correlation exists between the number of arrests along with state welfare spending.

Upon reviewing the data, we feel that it is most important to look over crime statistics and trends in all four metropolitan areas provided. This is important to us because we believe it will provide more information on which areas struggle the most with police brutality and violence, alongside racial equity. Because we are focusing on all four areas along with national statistics, for purposes of comparison as well as contrast, it is critical for us to see if crime trends and arrests are based on racial demographics in order to consider the societal issues of race inequality and biased violence.

In 2021 alone, over 1,200 individuals have been killed by police. Black people alone account for 27 percent of this number. According to Mapping Police Violence (MPV), a nonprofit organization that tracks fatal encounters with police, black individuals are three times more likely to be killed by police, yet 1.3 times more likely to be unarmed compared to white people. It is likely that personal and political biases may influence factors such as police department policies, city budgets, police contract provisions. Looking into the data, we must keep these biases in mind to cultivate a factual analysis.

Data Cleaning and Preprocessing

Our questions cover a variety of topics within the dataset. Overall, the data includes information about politics, socioeconomic data, police budgets, sales of DoD equipment, and overview of crimes in the four largest metropolitan areas. Because our research questions are focusing on political leanings in comparison to state DoD equipment purchases, city revenue correlations to city crimes, and most common arrests and arrestee characteristics, our data comes from different csv files. Overall, our approach to answering our research questions revolves around investigating 6 different files through data cleaning via SQL.

In SQL, in order to clean our data, we eliminated columns from each table we did not feel were appropriate for our research and used specialized SQL case statements to cut out unnecessary rows. Furthermore, we queried specific variables for each of the three questions our group came up with. As all questions were different and required different datasets to pull information from, all questions were related to the entire data group which overall proves trends in the existence of police brutality and racial equity.

Methodology

To be able to organize and clean our data, we used SQL commands to query the proper tables for observations including joining and combining datasets for further observations. Using SQL was very helpful in order to analyze arrests, political leaning, socioeconomic data, budgets and financial data. Multiple SQL queries were conducted in order to answer our research questions followed by figures visualizing some of the findings.

We determined SQL to be the best methodology for our analysis based on the type of research questions we were answering. Our research questions involved a lot of queries to determine factors such as: political leanings, Census Bureau Division contributions, state generated revenue by criminal offense, state spending and offense type by race. Furthermore, for more advanced queries and data some of the methods we used include case statements and functions including pipeline and stringIndexer.

Our team determined SQL would be the method utilized due to the nature of our dataset which consists of mostly numbers. Our team found that the dataset we are working with did not require TF-IDF, Collaborative Filtering and other methodologies because our dataset and research questions did not involve text mining or filtering. Therefore, SQL statements and graphs were found to be the most appropriate methodology for this analysis.

Analyses and Model Results

Question 1: Is there a relationship between police department policies, equipment purchases, poverty rates and political leanings of citizens in the four metropolitan areas?

This research question aims to understand any correlation or trends between the political leaning of citizens in the four metropolitan areas (Los Angeles, Chicago, New York City, and Dallas) and factors such as police department policies, equipment purchases, and poverty rates. First, relevant datasets were loaded to PySpark, then an SQL table was created and used to query a series of tables that would be analyzed to answer this research question.

POLITICAL LEANING:

The first query conducted for this research was on political parties of the four metropolitan areas.

```
1 # Create a view or table
2 temp_table_name = "Politics"
3 df.createOrReplaceTempView(temp_table_name)
```

```
1 %sql
2 SELECT *
3 FROM Politics
4 WHERE Area LIKE '%Dallas%' OR Area LIKE '%Chicago%' OR Area LIKE '%New York City%' OR Area LIKE '%Los Angeles%';
```

	Area	Republican Vote Share	Partisan Segregation
1	Greater New York City area	0.333	0.47
2	Greater Chicago area	0.296	0.4
3	Greater Los Angeles area	0.274	0.34
4	Greater Dallas area	0.472	0.32

We found that Dallas had the largest percentage of population leaning Republican at 47.2%, while Los Angeles had the smallest percentage of Republicans at 27.4%. This query was followed by a subsequent query that gave us the total population percent below poverty level in the four metropolitan areas.

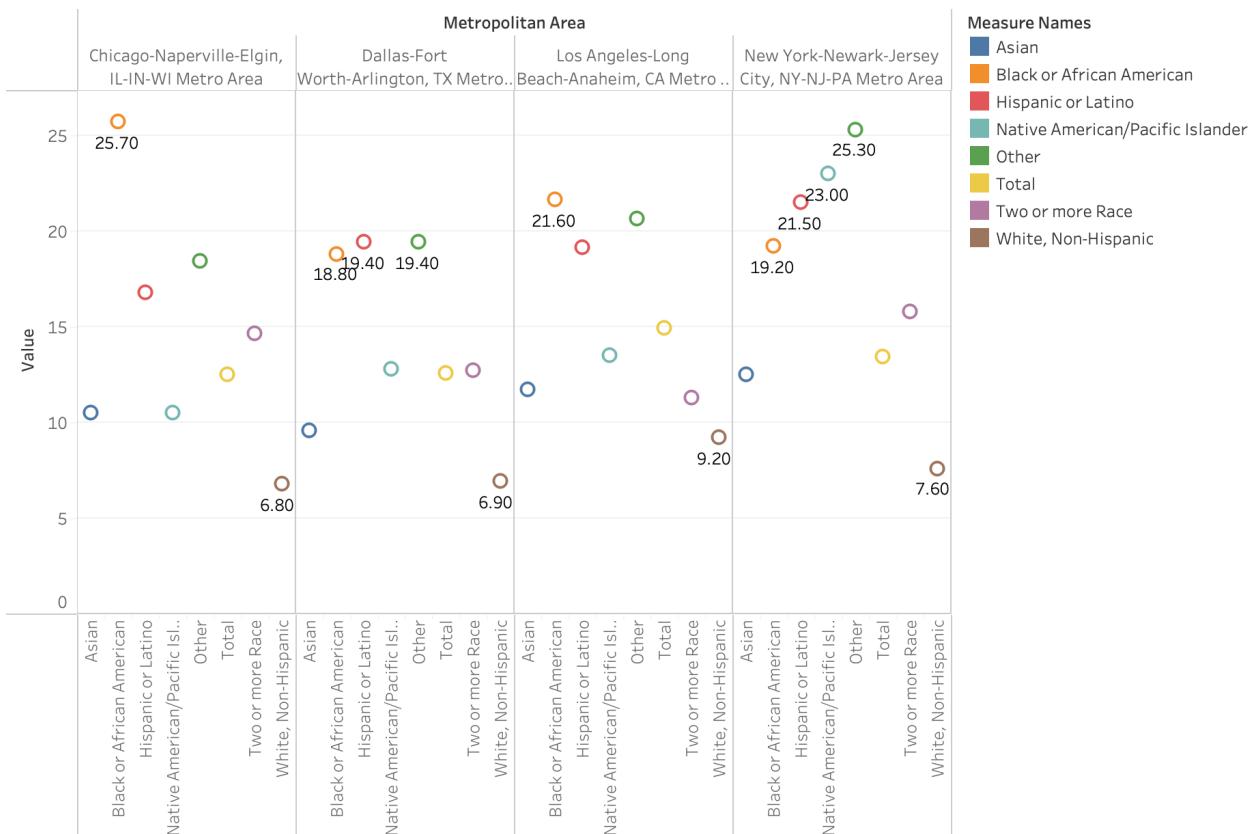
POPULATION BELOW POVERTY LEVEL:

```
1 %sql
2 SELECT Geographic_Area_Name, Total_Population_Percent_Below_Poverty_Level
3 FROM Poverty
4 WHERE Geographic_Area_Name LIKE '%Dallas%' OR Geographic_Area_Name LIKE '%Chicago%' OR
5 Geographic_Area_Name LIKE '%New York%' OR Geographic_Area_Name LIKE '%Los Angeles%';
```

	Geographic_Area_Name	Total_Population_Percent_Below_Poverty_Level
1	Chicago-Naperville-Elgin, IL-IN-WI Metro Area	12.5
2	Dallas-Fort Worth-Arlington, TX Metro Area	12.6
3	Los Angeles-Long Beach-Anaheim, CA Metro Area	14.9
4	New York-Newark-Jersey City, NY-NJ-PA Metro Area	13.4

Based on total population percentage level no strong correlation is seen between Republican leaning cities and non-republican leaning cities. Chicago had the second smallest percentage of Republicans (29.6%) while Dallas had the largest percentage of Republicans (47.2%), yet they both had similar population percentages below poverty level. Chicago at 12.5% and Dallas at 12.6%.

To find out if there are any large differences in poverty levels between races, an SQL query was done and the below figure created. The figure shows a breakdown of the percentage of population below poverty level in each of the four metropolitan areas by race.



The findings of the above figure suggests that African American and Hispanic have a higher percentage of population below poverty level in comparison to Non Hispanic Whites. This was found to be true for all four metropolitan areas of this study. The percentage of African American below poverty level in all four metropolitan areas was between 18.8% - 25.7%, in comparison, the percentage of Non-Hispanic White below poverty level was significantly lower between 6.8% - 9.2%. However, because there are no major differences within the same race in the four metropolitan areas analyzed, no correlation between political leaning and poverty level by race can be concluded.

PURCHASES OF POLICE EQUIPMENT:

To determine if there were any patterns between political leaning and purchase of police equipment from DOD, the following query was done.

```

1 %sql
2 SELECT (CASE WHEN Station_Name_LEA LIKE '%DALLAS%' THEN 'Dallas'
3             WHEN Station_Name_LEA LIKE '%CHICAGO%' THEN 'Chicago'
4             WHEN Station_Name_LEA LIKE '%NEW YORK CITY%' THEN 'New York City'
5             WHEN Station_Name_LEA LIKE '%LOS ANGELES%' THEN 'Los Angeles'
6         END) as Station_Name_LEA, COUNT(Quantity) AS Equipment_quantity
7 FROM Equipments
8 GROUP by (CASE WHEN Station_Name_LEA LIKE '%DALLAS%' THEN 'Dallas'
9                 WHEN Station_Name_LEA LIKE '%CHICAGO%' THEN 'Chicago'
10                WHEN Station_Name_LEA LIKE '%NEW YORK CITY%' THEN 'New York City'
11                WHEN Station_Name_LEA LIKE '%LOS ANGELES%' THEN 'Los Angeles'
12            END);

```

	Station_Name_LEA	Equipment_quantity
1	Dallas	145
2	Los Angeles	2071
3	null	152159
4	Chicago	403
5	New York City	2

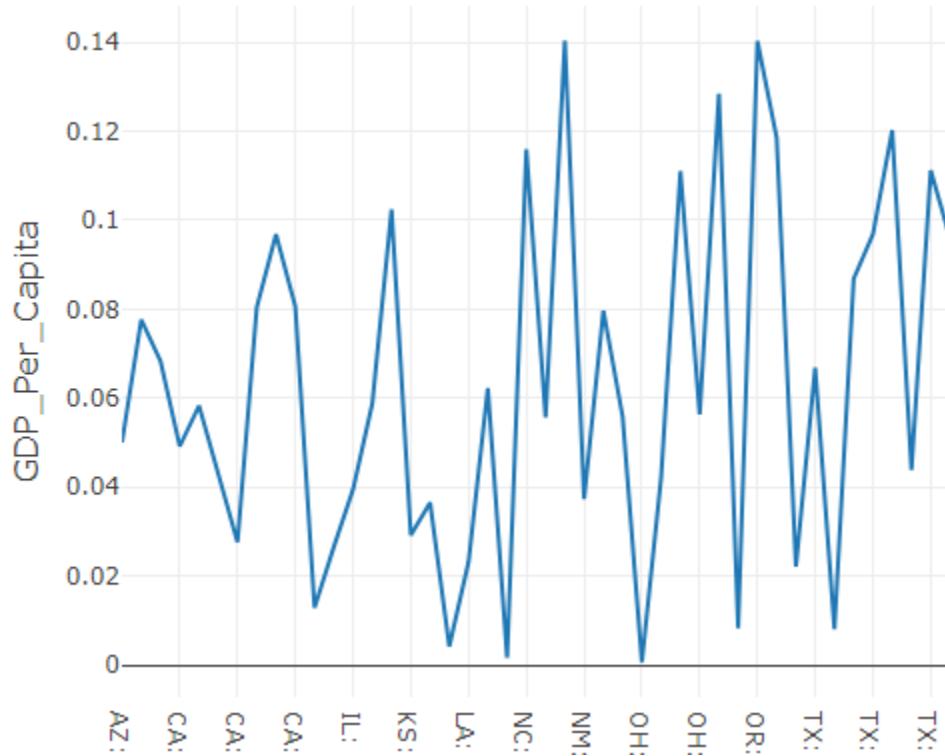
The result of the query shows the volume of police equipment purchased from the Department of Defense in the four Metropolitan Areas. No significant trends were found between the purchases and political leaning. According to this dataset, Chicago had the highest volume of equipment purchased while New York had the lowest. A better dataset needs to be obtained to understand police equipment purchasing behavior between the four metropolitan areas and determine if any correlation exists between that and political leaning.

Finally, a query and analysis of police policies on types of criminal intervention in the four metropolitan areas had many null values and thus generated inconclusive results and was excluded from this analysis.

Question 2: What is the total revenue generated by each state according to each criminal offense?

Before answering the main question of how much money is generated based on each crime, we should take a look at where each State and City stands economically. In most cases, a City that

does not produce economically is more prone to display a higher unemployment rate, and poverty lifestyle to its residents. In order to determine where the City/State stands economically we first need to determine its GDP per capita, as shown by the line graph below we can determine that there is heterogeneity variation amongst each state on their GDP per capita production. Based on the analysis, Omaha, Nebraska produces 0.14% GDP per capita on average. That is the highest amongst the other States, and Cincinnati, Ohio is the lowest producer of GDP per capita by 0.0063% on average according to our analysis.



After analyzing the GDP per capita, we looked at the number of law enforcement branches that receive contribution or donation from the Census Bureau Division (CBD), based on the East South Central division branch. The City of Alaska, Fairbanks received the most contribution by the Census Bureau Division. We were able to do this analysis by joining two table together based on the population, analyzing which cities gets the highest contributions will help us determine how funds are being distributed amongst these branches of law enforcements and help us understand that the reason for these funds and equipment programs are designed to incentivize harmful policing and lessen the effectiveness of local and state policing to a certain extent. To be more specific, the more funding these law enforcement agencies receive the less they work for

the community, because either way they are receiving funds. Therefore, Fairbanks Alaska has the poorest score when it comes to the quality of policing guideline to serve the community.

```
%sql
SELECT budgets.city_name, budgets.city_types, employment_fbi.division_name, employment_fbi.agency_type_name
FROM budgets, employment_fbi
WHERE budgets.city_population = employment_fbi.population
AND division_name = 'East South Central'
ORDER BY city_name
LIMIT 5
```

(1) Spark Jobs

	city_name	city_types	division_name	agency_type_name	
1	AK: Fairbanks	core	East South Central	County	
2	AK: Fairbanks	core	East South Central	County	
3	AK: Fairbanks	core	East South Central	County	
4	AK: Fairbanks	core	East South Central	City	
5	AK: Fairbanks	core	East South Central	County	

To answer the question of which state generates the most revenue based on the types of criminal offense, let's look at the sql statement below which shows the amount of revenue that was accumulated based on each crime offense and it is significantly high. The query shows city_name, city_population, year, crime offense type, and the total city revenue that is greater than 5000.

```
%sql
SELECT budgets.city_name, budgets.city_population, national_arrests.year,
national_arrests.offense_name, budgets.rev_total_city
FROM budgets, national_arrests
WHERE budgets.rev_total_city > 5000
ORDER BY city_name
```

1) Spark Jobs

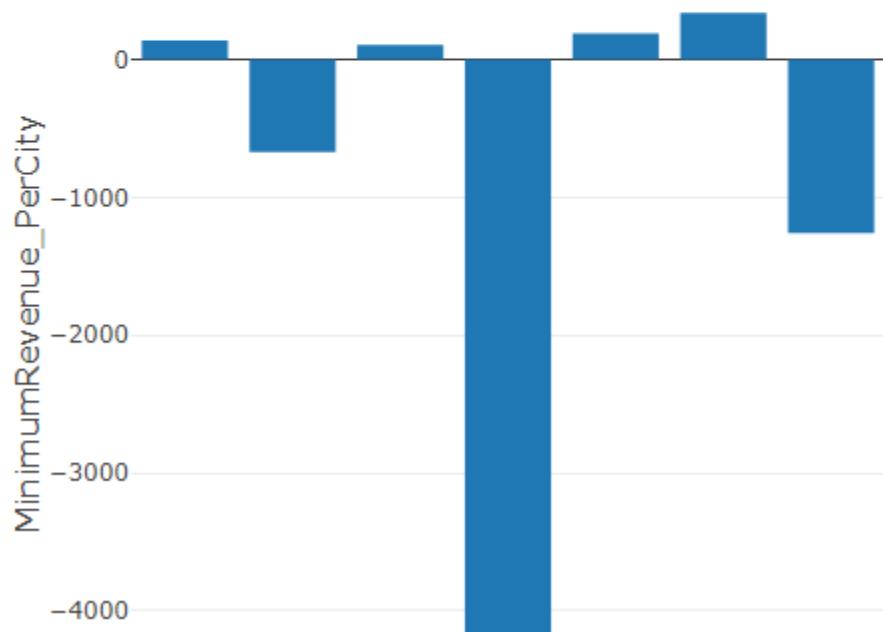
	city_name	city_population	year	offense_name	rev_total_city
1	AK: Anchorage	177000	2004	Prostitution and Commercialized Vice	5948.99
2	AK: Anchorage	174500	1998	Drunkenness	5342.24
3	AK: Anchorage	177000	2004	All Other Offenses	5948.99
4	AK: Anchorage	174500	2007	Simple Assault	5342.24
5	AK: Anchorage	174500	1998	Drug Abuse Violations	5342.24
6	AK: Anchorage	177000	2012	Offenses Against the Family and Children	5948.99
7	AK: Anchorage	177000	2004	Motor Vehicle Theft	5948.99

cated results showing first 1000 rows.

According to the sql statement and the chart below, we can see the minimum total revenue that was generated by each city. The lowest revenue generated according to that bar chart is by the city of Harrisburg, Pennsylvania at -4000. Followed by the City Milwaukee, Wisconsin at about -1258.38

```
%sql
SELECT year, city_name, MIN(rev_total_city) AS MinimumRevenue_PerCity
FROM budgets
WHERE rev_total_city < 500
GROUP BY year, city_name
HAVING MIN(rev_total_city)
```

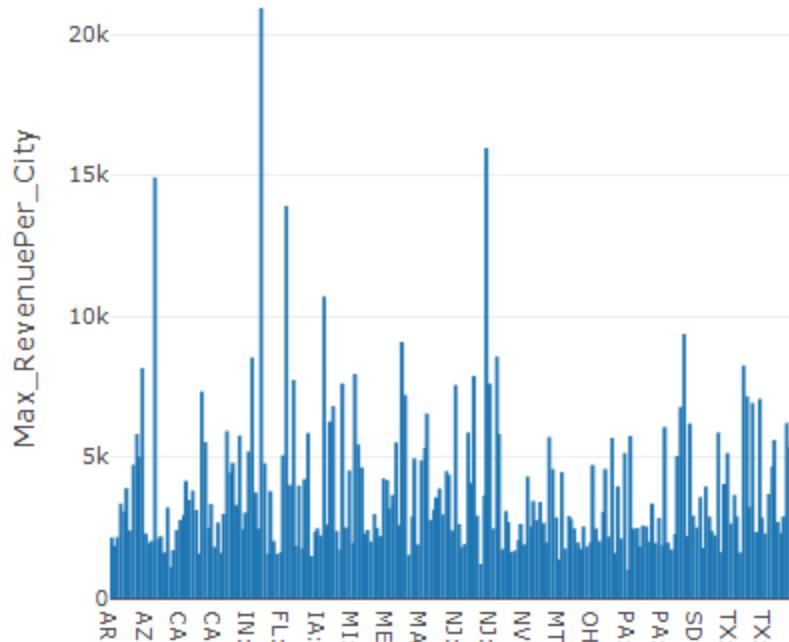
(2) Spark Jobs



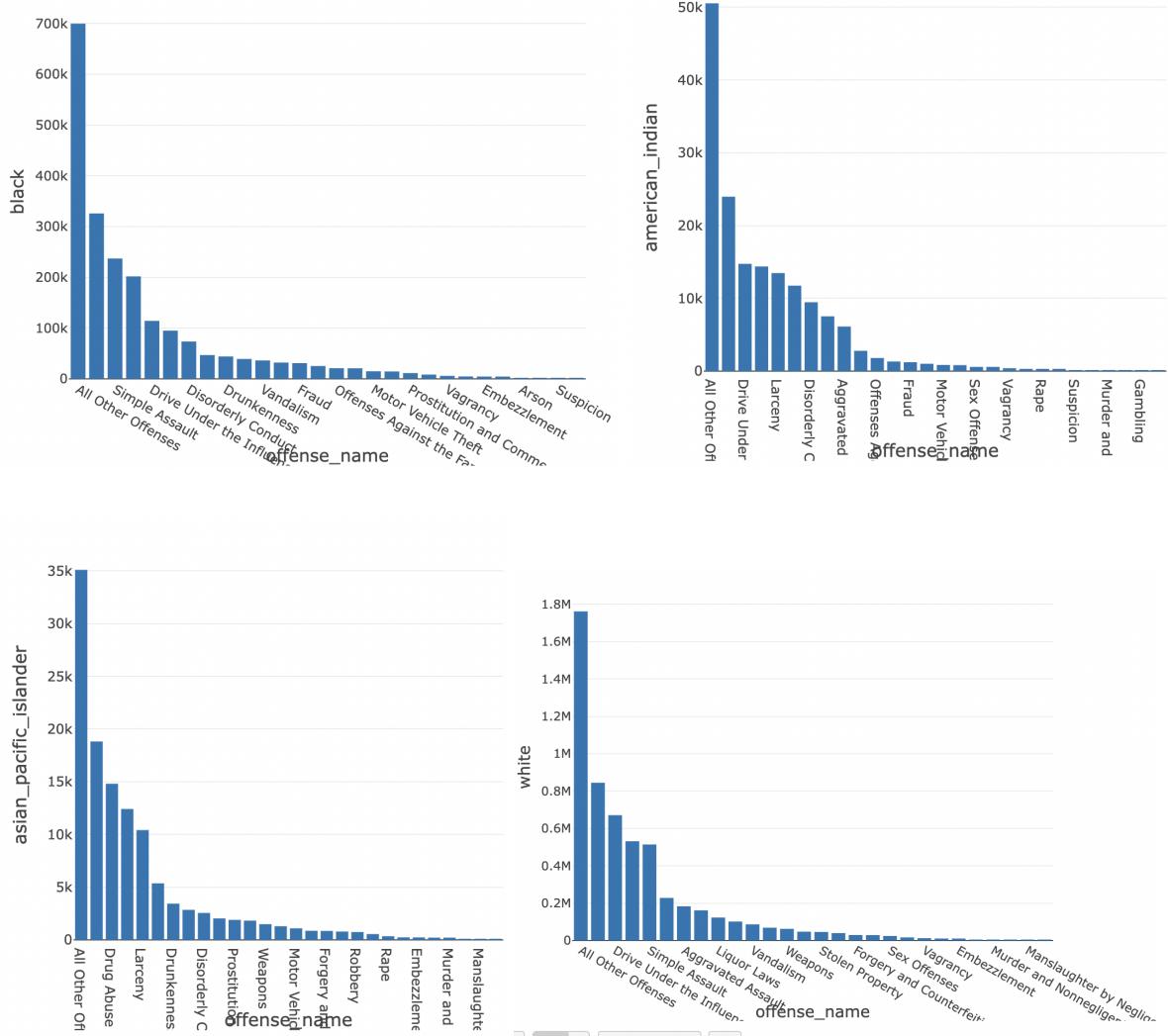
The city that generated the highest or the most revenue based on each crime offense was Washington, DC at about 20,934.83. Followed by the second highest total revenue, which was New York city at about 15,962.39, and then San Francisco comes in at 3rd for the highest total revenue collected at about 14,916.28. The bar chart below depicts an accurate result of the maximum amount of total revenue amongst each State.

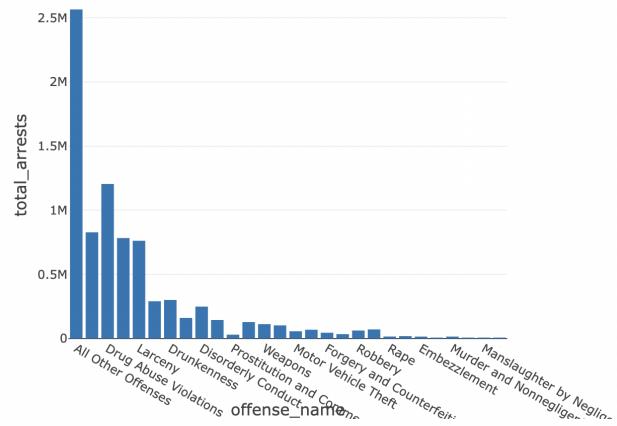
```
%sql  
SELECT city_name, MAX(rev_total_city) AS MaximumRevenue_PerCity  
FROM budgets  
GROUP BY city_name
```

(2) Spark Jobs



Question 3: What are common denominators seen with the arrests made across the entire United States?





```

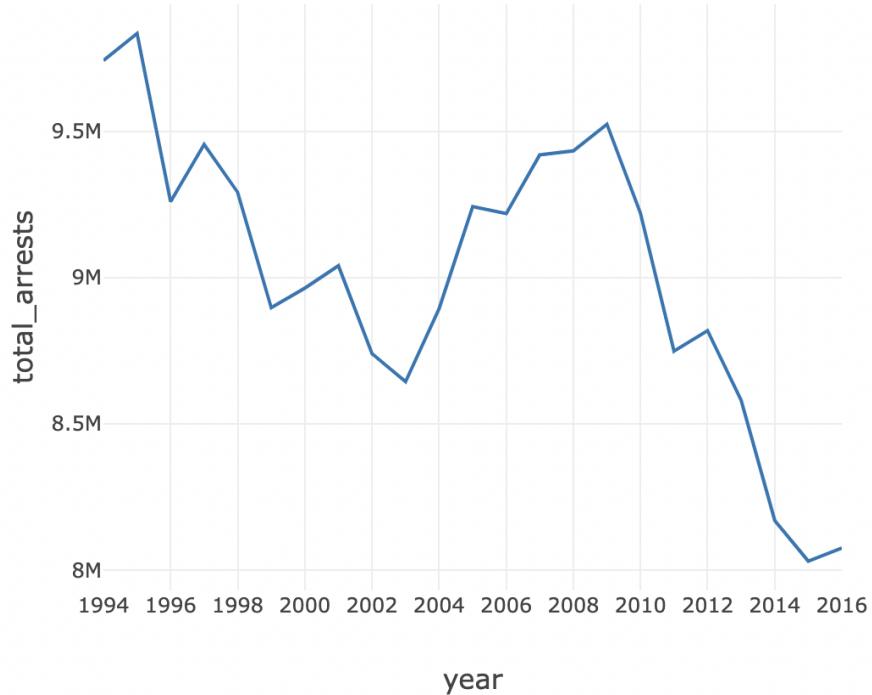
1 %sql
2 select
3 year,
4 offense_code,
5 offense_name,
6 total_male + total_female as total_arrests,
7 white,
8 black,
9 asian_pacific_islander,
10 american_indian
11 from `arrests_national_adults__1__csv`
12 where year = 2016
13 order by black desc

```

RACE:

The five bar charts shown above display Black, White, Asian & Pacific Islander along with American Indian individuals arrested across the United States during 2016. One bar chart specifically shows the number of total arrests during that year along with the offenses of the crimes committed. Based on the arrests for each of the different races, we can see which races were arrested more often than others for specific crimes. An example of this can be seen where aside from “all other offenses” the offense White individuals were mostly arrested for in 2016 was driving under the influence. On the other hand, Black individuals in 2016 had the top offense of simple assault. Additionally, according to the graphs more white individuals were arrested (1.76m) for “all other offenses” whereas under 700k Black individuals were arrested for “all other offenses”. If we look at the United States census in 2016, there were about 40.2 million Black Americans and 233.6 million White Americans. If we take these arrests as a snippet of the total population of individuals to exist in the United States, the percentage of Black individuals

arrested would be 1.74% compared to 0.75% of white individuals arrested for the same level of offenses. This is where we see an issue of racial bias existing when arrests are made across the United States.

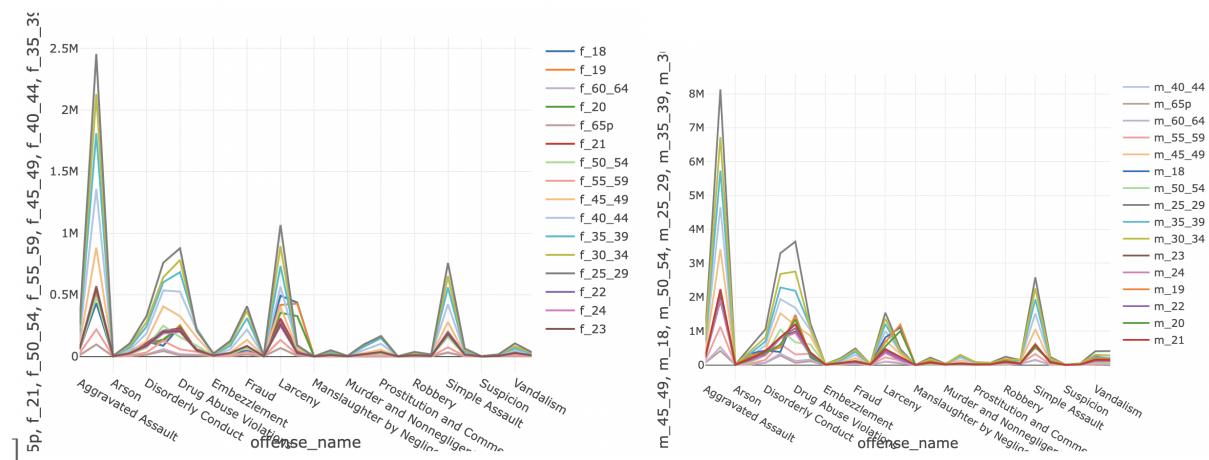


```
1 %sql
2 select
3 year,
4 total_male + total_female as total_arrests,
5 offense_name
6 from `arrests_national_adults__1__csv`
```

YEAR:

In this line chart, the total number of arrests made by the year is displayed. In the dataset used, the data ranged from 1994 to 2016, including arrests of 4 different races (Black, White, Asian, and American Indian) along with a range of ages from 18 years old to 65 and above. This line chart demonstrates the movement of how arrests have changed through three different decades. We can see that the highest peak of this line exists in 1995 reaching close to 10 million arrests

within the year. A huge decrease is seen exactly 20 years later in 2015 with the total number of arrests being a little over 8 million. When thinking about this drastic change in arrests, although a hump exists with another peak in 2009, most likely due to the distressing time of the United States financial crisis, we can assume that less individuals have been arrested due to advancements in technology that can prove a suspect's connection to a crime in order to better enforce accurate arrests. In the 1990s, police departments did not have the same type of technology that they do today, along with similar ways to store criminal records. It is likely that many arrests in the 1990s could have been wrongly done whereas even though there are innocent individuals arrested, and even killed, today for crimes they did not commit, the number of arrests made is significantly lower which may have been influenced by police forces making more appropriate and accurate decisions in terms of arrests.



AGE AND SEX:

Another denominator looked at for national arrests made, were individual age groups a suspect may belong to. The two line graphs above display 17 different age groups for men and women with ages ranging from 18 years old to over 65. M indicates men and F represents women. In the chart, the most arrests from 1994 to 2016 were aggravated assault for men were 8.1 million individuals in the 25 to 29 age groups. On the other hand, women additionally had the same

offense for the most arrests and the age group was the same, however there were 2.5 million arrests. The significance of the difference between the number of female and male arrests with the same age range can be for a multitude of reasons. However, the age range for men is interesting to note because it corresponds with ages and cases of many individuals who have endured police brutality. These cases include Daunte Wright, Breonna Taylor, Stephon Clark and many more just to name a few. Although 8.1 million men were arrested from 1994 to 2016, it is possible that not all arrests were accurately made, especially since 5.6 million more men were arrested. It is possible younger age groups committed more crimes due to poverty, lack of basic education, and poor state funding. We can look into this next.

```

1 from pyspark.ml.feature import StringIndexer
2 from pyspark.ml import Pipeline
3 from pyspark.sql.functions import col
4 indexer = [StringIndexer(inputCol=column, outputCol=column+"_index") for column in list(set(nd.columns)-set(['overall']))]
5 pipeline = Pipeline(stages=indexer)
6 transformed = pipeline.fit(nd).transform(nd)
7 transformed.show()

```

	State	Basic Education	Public Welfare	Police	Personal Income (\$M)	State_index	Basic Education_index	Public Welfare_index	Police_index	Personal Income (\$M)_index
[Connecticut.....]	3.6	1.5	0.5	257,714	6.0	10.0	12.0	2.0	16.0	
[Maine.....]	4.1	5.2	0.5	62,060	19.0	15.0	23.0	2.0	42.0	
[Massachusetts.....]	3.5	4.8	0.6	463,931	21.0	9.0	21.0	0.0	34.0	
[New Hampshire.....]	3.7	3.0	0.5	80,122	29.0	4.0	7.0	2.0	48.0	
[Rhode Island.....]	4.5	5.5	0.8	55,934	39.0	2.0	5.0	3.0	40.0	
[Vermont.....]	5.1	5.7	0.7	32,570	45.0	21.0	24.0	1.0	24.0	
[Delaware.....]	4.3	5.3	0.8	47,782	7.0	1.0	4.0	3.0	37.0	
[District of Colum...]	5.0	7.1	1.1	55,510	8.0	20.0	29.0	6.0		

AREA:

Above is a table which displays top states with spending on basic education, public welfare, police, and personal income. The states shown are Connecticut, Maine, Massachusetts, New Hampshire, Vermont, Delaware, and District of Columbia, which isn't a state but is included in the data. Although the number of arrests doesn't involve states perpetrators are from, we can note that most of the states listed in the data shown above are a part of the United States known as New England. As of 2021, Black Americans compromised about 7% of the region's population. This area of the United States is also known to be notoriously wealthy and primarily

white. As this does not specifically correlate with arrests made, a connection can be made where a higher percentage of Black individuals were arrested than White individuals. This is across the United States, but noting that New England has more spending in its states for things like police, basic education and welfare, an assumption can be made that the individuals who live in these areas do not take part in much crime. It is likely because there are more resources and access to resources, many individuals do not resort to making criminal decisions. We can use the denominator of an area as a factor as to why more arrests are made. Areas include factors like state government spending on basic education and welfare which can significantly impact an individual's life and why they make the decisions that they do, or put them in positions where they are stereotyped and exist in a racially biased environment.

Conclusion

Overall, the data our group used to research and analyze our questions for this project will show that police brutality and racial bias exist in the United States. This has been a problem the country has faced for centuries, and has become more of a recognizable problem and discussion topic with recent events of innocent killings of Black individuals and attacks on Asian individuals. In this section of our project, we would like to conclude what we have found through our individual research questions.

Question 1:

An analysis on political leaning of the four metropolitan areas and its comparison to factors such as: police department policies, equipment purchases, and poverty rates gave a lot of inconclusive results. While the analysis did determine that the four metropolitan areas had different political leaning. No correlation was found between total population below poverty level and party leaning. Dallas and Chicago had significantly different political leaning but were found to have very numerically similar population percentages below poverty level. When the percentage below poverty level was further analyzed by race, the analysis found a strong disparity between percentages of Whites and African American below poverty level, with African American in all four major metropolitan areas being significantly more impoverished. However, further analysis needs to be conducted to determine the impact of police violence on poverty levels within

different races. Furthermore, an analysis of a dataset containing statistics on police violence victims by race may be useful. A comparison of police violence victims by race and poverty level by race could shed some light on any patterns between police violence, political leaning and economic disparities. The volume of police equipment purchased was also analyzed as part of this research question and compared to political leaning and no significant correlations were found. Further analysis must be conducted to adequately determine any pattern between political leaning and policies, purchase and poverty of the four metropolitan areas.

Question 2:

After performing multiple analysis on the budget dataset, organizing the data to fit the analysis, and running multiple sql queries to help us predict and find the answer to our question. It is safe to say that police budget and fund allocations do not help enhance or maintain a good police approach when it comes to avoiding police brutality or police injustice. We have noticed from our analysis that, even if law enforcement are generating revenue or receiving funds as a form of donation, the crime rate is still increasing and it does not show a sign of slowing down. As mentioned earlier, Washington, DC generates the highest total revenue capital than any other state, but the fact of the matter is, DC has an overall of 45% police violence rate, 63% deadly force used during arrests, and the department receives 51% Misconduct complaints on average. Based on the analysis, It seems to be the higher revenue generated, the poor quality police work and the lack of unjustified accountability performance that is displayed to the public.

Question 3:

While looking at the data, there was noticeable information such as a large population of White Americans, a larger percentage of Black Americans arrested, and a decrease in the number of arrests amongst other factors. The variables that were most impactful were race, year, age and sex, and area. All these variables have affected the number of arrests along with who is arrested. Today, we live in an era where police brutality exists, especially to individuals of color. A racial bias also exists which impacts how many individuals are able to live across the United States. It is noticeable that individuals in wealthier areas like New England, where the majority population is white, are not affected by poverty the way that areas where the majority of the population is Black or another race. While there has been a significant decrease in the amount of arrests across

the nation, there is still a higher number of arrests of Black individuals over any other race with respect to the nation's entire Black population. This in turn, supports the social dysfunction that exists in this country where police brutality and racial bias exist.

References

<https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/2016/>

<https://statisticalatlas.com/division/New-England/Race-and-Ethnicity>

<https://www.nytimes.com/interactive/2020/07/03/us/george-floyd-protests-crowd-size.html>

<https://www.vera.org/downloads/publications/for-the-record-unjust-burden-racial-disparities.pdf>