Checkpoint 1: SQL Analytics Findings

The Freedom Deer: Tianchang Li, Hualiang Qin, Qingwei Lan

October 14, 2021

The main goal of our research is to investigate the conditions under which police officers tend to use force on civilians. Specifically we believe that race and environmental conditions play a critical role in this, therefore in this report, we utilized SQL to analyze and answer some important questions.

The questions are split into sections and each section contains multiple questions.

Note that all `.sql` files can be run with `psql cpdb cpdb < filename.sql`.

1. Information about Victims and Officers

We would like to first gather some background information about the officers and victims in the use of force cases, such as race distribution. The data in this section is taken from the `trr_trr` and `data_officer` datasets.

1a. What is the racial distribution of the victims involved in cases of use of force?

Database Query Results:

```
          subject_race          | count
--------------------------------+-------
 Black                          | 49747
 Hispanic                       |  9369
 White                          |  6540
 Asian/Pacific Islander         |   431
 Native American/Alaskan Native |    54
```

From the results, we can see the race distribution of the subjects involved in the use of force cases. The table shows that black subjects are the dominant race which has contributed to cases more than the total of all other races combined. However, as we know that black people do not dominate the total population in Chicago, this indicates a worth-of-investigating racial discrimination involved in the police use of force.

1b. What is the racial distribution of police officers involved in these cases?

Database Query Results:

```
          officer_race         | count
-------------------------------+-------
 White                         | 38731
 Hispanic                      | 15064
 Black                         | 10599
 Asian/Pacific                 |  2028
 Native American/Alaskan Native |   310
```

From the results, we can see the race distribution of police officers involved in use of force cases. At the top is white police officers whose count exceeds the total of the rest of the races, especially in contrast to the number of Asian/Pacific and Native American/Alaskan Native police officers. This potentially leads to a white-dominated culture in police department.

1c. What portion of the total use of force cases involves an officer that is of a different race than that of the victim (cross-race use of force)?

Database Query Results:

```
 cross_race_percentage
-----------------------
              0.732046
```

Based on the results, we can see that cross-race use of force cases make up 73.2% of total use of force cases, which is less surprising considering the fact that victims are dominated by black people and police offers are dominated with white. Nevertheless, 73.2% is high enough to raise follow-up questions about the different dynamics between police and victims. Specifically, we will look into the racial composition of the cross-race cases.

1d. What portion of the cases in use of force containing firearm usage.

Database Query Results:

```
 firearm_used_percentage
-------------------------
                  0.0153
```

Based on the results, we can see that only 1.5% use of force cases involved the usage of firearms. This indicates a less number of firearm usage than what we expected, despite large media coverage on this topic.

1e. What are the percentages of use of force cases grouped by officer race and subject race? (i.e. what is the percentage of white officers using force on black subjects)

Database Query Results:

percentage	officer_race	subject_race
0.416733	White	Black
0.157269	Hispanic	Black
0.141796	Black	Black
0.083096	White	Hispanic
0.065161	White	White
0.043256	Hispanic	Hispanic
0.020457	Hispanic	White
0.019950	Asian/Pacific Islander	Black
0.008147	White	
0.006923	Black	White
0.006849	Black	Hispanic
0.005536	Asian/Pacific Islander	Hispanic
0.004238	White	Asian/Pacific Islander
0.003969	Asian/Pacific Islander	White
0.003357	Native American/Alaskan Native	Black
0.002552	Hispanic	
0.001805	Black	
0.001044	Hispanic	Asian/Pacific Islander
0.000731	Black	Asian/Pacific Islander
0.000627	Native American/Alaskan Native	Hispanic
0.000537	White	Native American/Alaskan Native
0.000463	Native American/Alaskan Native	White
0.000418	Asian/Pacific Islander	Asian/Pacific Islander
0.000358	Asian/Pacific Islander	
0.000194	Hispanic	Native American/Alaskan Native
0.000179	Native American/Alaskan Native	
0.000045	Black	Native American/Alaskan Native
0.000030	Asian/Pacific Islander	Native American/Alaskan Native

The results provides a more detailed view of the racial components of the subjects and police officers in all the use of force cases. A cursive scan shows us that 41% of all cases come from white police officers' use of force on black subjects. Further analysis shows that cases with black subjects make up 71.58% of all use of force cases. This indicates that the black population are more prone to police's use of force.

2. Environmental Factors That May Affect an Officer's Decision to Use Force

Next we want to investigate the influence of environmental factors on a police officer's decision to use force. The following questions will looking into some of these factors.

2a. What portion of the use of force happened under different lighting conditions?

Database Query Results:

lighting_condition	percentage
GOOD ARTIFICIAL	0.395291
DAYLIGHT	0.293887

```
NIGHT              |   0.118295
POOR ARTIFICIAL    |   0.111580
                   |   0.054283
DUSK               |   0.021113
DAWN               |   0.005551
```

This table shows that the use of force cases are dominated by the scenarios with good lighting (either with daylight of good artificial light). This, to some extend, suggests that lack of visibility is not a risk for someone to experience use of force, which is opposite to our hypothesis.

2b. What portion of the use of force happened indoors against outdoors?

Database Query Results:

```
 indoor_or_outdoor | percentage
-------------------+------------
 Outdoor           |   0.705248
 Indoor            |   0.240962
                   |   0.053791
```

This shows that most cases occurred outdoors with still descent amount of cases happening indoors. Civilians are usually under protection in their residential homes unless under criminal investigation. This leads to the problem of unwarranted search and would be interesting to look into the conditions and the allegation outcomes of these cases.

2c. What portion of the use of force happened under different weather conditions?

Database Query Results:

```
 weather_condition | percentage
-------------------+------------
 CLEAR             |   0.810218
 RAIN              |   0.060058
                   |   0.056685
 OTHER             |   0.038959
 SNOW              |   0.027977
 FOG/SMOKE/HAZE    |   0.003312
 SEVERE CROSS WIND |   0.001477
 SLEET/HAIL        |   0.001313
```

This indicates that adverse weather conditions might not be an attribute for use of force which is opposite to our hypothesis. This would allow us to eliminate the influence of weather conditions from our future research.

2d. What portion of the use of force happened under different locations?

Database Query Results:

location_recode	percentage
Street	0.278100
Sidewalk	0.219997
Residence	0.063683
Apartment	0.057417
Police Facility/Veh Parking Lot	0.057014
Alley	0.056178
Residence Porch/Hallway	0.043943
Residential Yard (Front/Back)	0.031573
Parking Lot/Garage (Non-Residential)	0.026037
Other	0.021904
Chicago Housing Authority Property	0.019412
Gas Station	0.010863
Park Property	0.009729

..... truncated .....

This shows that police's use of force more likely to occur on street/sidewalks but not by much. This doesn't give us much information at the moment.

2e. Under what combinations of different conditions (lighting, indoor or outdoor, weather, location) is a police officer more likely to use force?

Database Query Results:

weather	indoor_or_outdoor	lighting_condition	location_recode	count
CLEAR	Outdoor	GOOD ARTIFICIAL	Street	5870
CLEAR	Outdoor	DAYLIGHT	Street	5587
CLEAR	Outdoor	DAYLIGHT	Sidewalk	4621
CLEAR	Outdoor	GOOD ARTIFICIAL	Sidewalk	4290
CLEAR	Outdoor	NIGHT	Street	2278
CLEAR	Outdoor	NIGHT	Sidewalk	1891
CLEAR	Indoor	GOOD ARTIFICIAL	Police Facility/Veh Parking Lot	1836
CLEAR	Indoor	GOOD ARTIFICIAL	Apartment	1498
CLEAR	Indoor	GOOD ARTIFICIAL	Residence	1357
CLEAR	Outdoor	POOR ARTIFICIAL	Street	1313
CLEAR	Outdoor	DAYLIGHT	Alley	1294
CLEAR	Outdoor	POOR ARTIFICIAL	Sidewalk	1198

..... truncated .....

This table provides a more comprehensive view of the narratives for the most common scenarios of police's use of force. Cases often time happened with good lighting on the street.

3. How does the influence of the top 10 combinations of conditions vary from race to race?

Race: Black

Database Query Results:

weather	indoor_outdoor	lighting	location	pct
CLEAR	Outdoor	DAYLIGHT	Street	0.0891
CLEAR	Outdoor	GOOD ARTIFICIAL	Street	0.0818
CLEAR	Outdoor	DAYLIGHT	Sidewalk	0.0774
CLEAR	Outdoor	GOOD ARTIFICIAL	Sidewalk	0.0612
CLEAR	Outdoor	NIGHT	Street	0.0336
CLEAR	Outdoor	NIGHT	Sidewalk	0.0277
CLEAR	Indoor	GOOD ARTIFICIAL	Police Facility/Veh Parking Lot	0.0269
CLEAR	Indoor	GOOD ARTIFICIAL	Apartment	0.0214
CLEAR	Outdoor	DAYLIGHT	Alley	0.0205
CLEAR	Outdoor	POOR ARTIFICIAL	Street	0.0204

Race: White

Database Query Results:

weather	indoor_outdoor	lighting	location	pct
CLEAR	Outdoor	GOOD ARTIFICIAL	Street	0.1057
CLEAR	Outdoor	GOOD ARTIFICIAL	Sidewalk	0.0699
CLEAR	Outdoor	DAYLIGHT	Street	0.0645
CLEAR	Outdoor	DAYLIGHT	Sidewalk	0.0433
CLEAR	Outdoor	NIGHT	Street	0.0336
CLEAR	Indoor	GOOD ARTIFICIAL	Police Facility/Veh Parking Lot	0.0300
CLEAR	Outdoor	NIGHT	Sidewalk	0.0235
CLEAR	Indoor	GOOD ARTIFICIAL	Residence	0.0234
CLEAR	Indoor	GOOD ARTIFICIAL	Apartment	0.0205
CLEAR	Outdoor	POOR ARTIFICIAL	Sidewalk	0.0180

Race: Hispanic

Database Query Results:

weather	indoor_outdoor	lighting	location	pct
CLEAR	Outdoor	GOOD ARTIFICIAL	Street	0.1075
CLEAR	Outdoor	GOOD ARTIFICIAL	Sidewalk	0.0748
CLEAR	Outdoor	DAYLIGHT	Street	0.0630
CLEAR	Outdoor	DAYLIGHT	Sidewalk	0.0457
CLEAR	Outdoor	NIGHT	Street	0.0361

CLEAR	Outdoor	NIGHT	Sidewalk	0.0334
CLEAR	Indoor	GOOD ARTIFICIAL	Apartment	0.0285
CLEAR	Indoor	GOOD ARTIFICIAL	Police Facility/Veh Parking Lot	0.0284
CLEAR	Indoor	GOOD ARTIFICIAL	Residence	0.0244
CLEAR	Outdoor	POOR ARTIFICIAL	Sidewalk	0.0222

Race: Asian/Pacific Islander

Database Query Results:

weather	indoor_outdoor	lighting	location	pct
CLEAR	Outdoor	GOOD ARTIFICIAL	Street	0.0928
CLEAR	Outdoor	DAYLIGHT	Street	0.0905
CLEAR	Outdoor	GOOD ARTIFICIAL	Sidewalk	0.0696
CLEAR	Outdoor	NIGHT	Sidewalk	0.0487
CLEAR	Outdoor	DAYLIGHT	Sidewalk	0.0418
CLEAR	Indoor	GOOD ARTIFICIAL	Apartment	0.0348
CLEAR	Outdoor	NIGHT	Street	0.0325
CLEAR	Indoor	GOOD ARTIFICIAL	Residence	0.0325
CLEAR	Outdoor	DAYLIGHT	Alley	0.0302
CLEAR	Indoor	GOOD ARTIFICIAL	Police Facility/Veh Parking Lot	0.0278

Race: Native American/Alaskan Native

Database Query Results:

weather	indoor_outdoor	lighting	location	pct
CLEAR	Outdoor	GOOD ARTIFICIAL	Sidewalk	0.0926
CLEAR	Outdoor	DAYLIGHT	Street	0.0741
CLEAR	Outdoor	DAYLIGHT	Sidewalk	0.0741
			Street	0.0741
CLEAR	Indoor	GOOD ARTIFICIAL	Apartment	0.0741
CLEAR	Outdoor	GOOD ARTIFICIAL	Street	0.0741
CLEAR	Outdoor	DAYLIGHT	Alley	0.0370
CLEAR	Outdoor	NIGHT	Alley	0.0370
CLEAR	Indoor	GOOD ARTIFICIAL	Residence Porch/Hallway	0.0370
CLEAR	Outdoor	GOOD ARTIFICIAL	Parking Lot/Garage (Non-Residential)	0.0370

Based on the results for each race, we can see that most use of force cases happen outdoors in clear weather under good lighting conditions. Further analysis shows that these results do not differ much between different races. The interesting part here is that most use of force cases don't happen under bad conditions as we hypothesized. This probably indicates that most police activity happens under favorable conditions and that police activity is not common when conditions are bad. Further research should be conducted to dig deeper into each of these conditions.

Conclusion and Future Research

Based on the race distributions of victims and police officers, we can see that the black population is the dominant race in the victims of police use of force. We also saw that cross-race use of force is common and made up 73% of all the use of force cases. Furthermore we also saw that a white police officer is more likely to use force on a black subject. These results suggest that the black population is more prone to police use of force and we would like to dig deeper into racial issues.

We also investigated how environmental conditions affected a police officer's decision to use force and concluded that most police use of force cases happened in favorable environmental condtions, such as good lighting and good weather. Since these environmental conditions didn't seem to influence an officer's decision that much, we would like to further investigate the physical location in which the case occurred, such as the neighborhood's socioeconomic status, median income, and crime rate. We believe that this will provide us with more information going forward.