Checkpoint 1: SQL Analytics Findings

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

October 13, 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 information about the officers and victims in the use of force cases. 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?

SQL file: `1a.sql`

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 use of force cases. As we can see, black subjects are most common.

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

SQL file: `1b.sql`

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.

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)?

SQL file: `1c.sql`

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.

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

SQL file: `1d.sql`

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.

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)

SQL file: `1e.sql`

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 give us a high level overview of the racial distribution of both 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.

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 address some of these factors.

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

SQL file: `2a.sql`

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
```

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

SQL file: `2b.sql`

Results:

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

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

SQL file: `2c.sql`

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
```

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

SQL file: `2d.sql`

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
Cta Platform	0.008326
Restaurant	0.008013
Hospital Building/Grounds	0.007968
Jail / Lock-Up Facility	0.007565
Vacant Property	0.005148
Vacant Lot/Land	0.004476
Bar Or Tavern	0.004417
Small Retail Store	0.004372
Government Building/Property	0.004163
Vehicle Non-Commercial	0.003954
Public School	0.003298
Grocery/Food Store	0.003178
Hotel/Motel	0.003133
Cta Garage / Other Property	0.003029
Residence-Garage	0.002104
Cta Bus	0.002074
Department Store	0.002059
School, Public, Grounds	0.002044
Driveway - Residential	0.001850
Highway/Expressway	0.001656
Convenience Store	0.001641
Airport	0.001492
Cta Train	0.001447
Tavern/Liquor Store	0.001313
Nursing Home/Retirement Home	0.001104
Airport/Aircraft	0.001074
Drug Store	0.000985
Cta Station	0.000955
Church/Synagogue/Place Of Worship	0.000910
Abandoned Building	0.000910
Other Railroad Prop / Train Depot	0.000851
Cta Bus Stop	0.000821
Sports Arena/Stadium	0.000746
Private School	0.000612
Commercial / Business Office	0.000612
Library	0.000537
Bowling Alley	0.000522
Bank	0.000418
Warehouse	0.000373
Barbershop	0.000343
Currency Exchange	0.000328
College/University	0.000328
Movie House/Theater	0.000269
Bridge	0.000254
Vehicle-Commercial	0.000239
Lakefront/Waterfront/Riverbank	0.000239
Cleaning Store	0.000209
Medical/Dental Office	0.000209
Other Commercial Transportation	0.000194

Construction Site	0.000194
Factory/Manufacturing Building	0.000194
Fire Station	0.000179
Athletic Club	0.000164
Taxicab	0.000134
Forest Preserve	0.000090
Appliance Store	0.000075
Cemetary	0.000075
Coin Operated Machine	0.000060
Car Wash	0.000045
Aircraft	0.000045
Vehicle - Other Ride Service	0.000030
Animal Hospital	0.000030
Day Care Center	0.000030
Airport Transportation System (Ats)	0.000015
Pool Room	0.000015
Cta Tracks - Right Of Way	0.000015

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

SQL file: `2e.sql`

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