Final Project: MPD Use of Force

Kaylhan Garcia, Phoebe Lin, Adam Billen

2023-10-17

MPD Use of Force Final Project Proposal

Introduction:

Police have existed in America in some form for nearly its entire history. Police forces themselves have their philosophical origin in Patrick Colquhoun's 1797 "Treatise on the Police of the Metropolis." These ideas first took form in London in 1829, but Colquhoun himself spent his early career as a British agent for cotton manufacturers. His conception of policing grew from his time spent enforcing slave codes and wielding slave patrols in Virginia. The first hiring of police through legislative means was in Boston in 1838, but before then "police" simply referred to slave patrols. North Carolina's first state police force was formed in order to stamp out the publishing of "The Liberator," William Lloyd Garrison's weekly abolitionist newspaper, soon after he was almost killed in a mob attack in Boston in 1835. The modern American police force began and grew from these dark roots of American history.

The existing literature has established that Black, Hispanic and American Indian/Native Americans are disproportionately likely to be killed by police in America, especially at younger ages. One study using data from between 2013 and 2018 found that Black men are ~2.5 times more and Black women ~1.4 times more likely to be killed by police in their lifetime than white men and white women, respectively. This amounts to some 96 out of 100,000 Black men and boys, between 36 and 81 American Indian/Alaskan Native men and boys, and 53 out of 100,000 Latino men and boys being killed by police over their lifetime, as compared to about 39 out of 100,000 white men and boys being killed. The risk is substantially lower for women across all racial groups. Between 2.4 and 5.4 Black women and girls, 2.4 American Indian/Alaskan Native women and girls, and 2 Latino and white women and girls out of 100,000 are expected to be killed by police throughout their lifetime.

In recent years, the Black Lives Matter Movement rose up in this context and caused the public to emerge as a strong force to fight for racial justice and equality in the United States. This movement also sparked the public to think about the appropriation of the use of force from the police to the general public. Also, the topic of how gender and race can act upon the use of force. This causes the public to request the government to provide basic information of the use of force (Schwartz, 2020). Therefore, information about use of force cases becomes more transparent for the general public to access in order to maintain public trust.

Moreover, in the study of Tanya Golash-Boza et al, "Broken windows and order-maintenance policing in gentrifying Washington, DC", she observed that the majority of citizen-police interactions were involved with the Black residents and the Black individuals were even more likely to be subjected to stop(Golash-Boza, T. Et al, 2022). According to her findings, this provided us with valuable insights regarding the demographic group that could be the primary focus of targeting.

For this project, we will be using the dataset about the use of force that was conducted from the Washington D.C. Metropolitan Police Department, in other words, the MPD. MPD is unique in a number of ways. 51% of its officers were Black and 34% white in 2021 compared to the 46.2% of residents who are white and 25% who are Black. This makes it one of the few police forces in the US where white officers are not over represented as compared to the community. This is commonly thought of as a positive for police forces, and studies have found that when police diversity matches a neighborhood, crime rates tend to be lower.

The dataset focused on the incidents that occurred between January 1, 2021 to December 31, 2021. The Use of force being defined as any form of physical effort employed to compel, sway, or convince an individual to adhere to an officer's directive from the MPD. The reportable forces and incidents include "Any use of force resulting in injury or a complaint of injury or pain where the injury or pain is directly associated with a member's use of force" and "The drawing and pointing of a firearm at, or in the direction of, another person when no other force was used" (Metropolitan Department, 2022).

The dataset contain 19 variables, 14 of them are categorical variables and 5 of the others are quantitative. Variables include **"IncidentDate", "IncidentTime", "IncidentDistrict", CaseStatus","DateClose","uof_type","disposition","department_action","officer_id","OfficerAssignment" , "OfficerGender", "OfficerRace", "year_force", "subject_age", "subject_race", "subject_gender", "serious",** and "**CD**"(Appendix A). The categorical variables in this dataset are mainly ordinal categorical variables. And there are 2135 rows in total. The dataset is mostly completed, only a few boxes contain "NA". This data, extracted from authoritative sources, is a rich resource for exploring the patterns, implications, and disparities in how force is applied in the nation's capital.

This project aims to ...

In our case, we will be considering variables of *"IncidentDistrict", "uof_type", "IncidentDate", "Disposition", "department_action", "OfficerGender", "OfficerRace", "subject_age", "subject_race", "subject_gender", "serious",* and *"CD(civil disturbance)"* mainly in our project. Since the rest of the variables wouldn't act significantly towards our project.

Hypotheses:

We hypothesize that Black people/African Americans are overrepresented in use of force cases in Washington, DC. They represented 45% of the DC population in 2021, so if $> 45\%$ of the subjects of uses of force in our data are Black/African American, they would be over represented. We also hypothesize that white police officers are over represented in use of force cases in Washington, DC. As of 2021, 34.55% of officers were white. If the % of white officers in use of force cases $> 34.55\%$, this would indicate over-representation. This question could serve as a case study for how accurate representation of a policed community affects the policing itself.

We are particularly interested in predicting the "disposition," or how the department classified a use of force, based on the variables available to us. We hypothesize that uses of force labeled unjustified may be more common in wards 7 and 8 that are more diverse and poorer. The location variable which is available, "IncidentDistrict," is police district rather than ward, but the police districts are roughly analogous to the wards:

- District 1 = Ward 6
- District 2 = Wards 2 & 3
- District 3 = Ward 1
- District 4 = Ward 4
- District 5 = Ward 5
- District 6 = Ward 7
- District 7 = Ward 8

We could do the same type of prediction but for a binary variable like "serious," which represents the severity of a given use of force.

Methods:

Our predictors will be the police district in which the use of force took place, the date of the incident, the gender and race of the officer and the age, race, and gender of the subject. The target will be the type

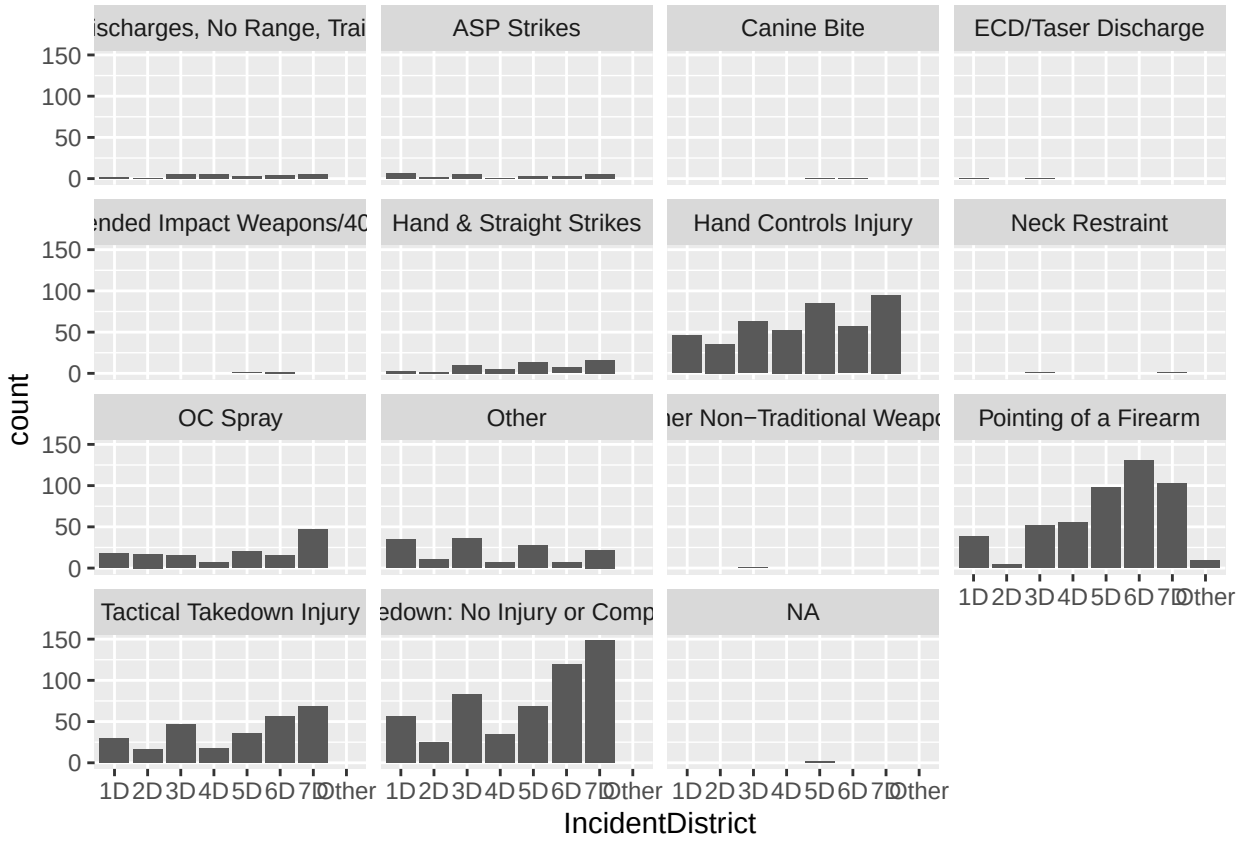
of use of force, the disposition, the department action, or the severity of the force. We won't include the other potential targets as predictors for one another, because severity, for example, will likely be a very high predictor of the type of use of force, muddling our models. We will not include variables such as case status, when a case closed, the anonymized officer IDs, or officer assignments as predictors. We will remove all uses of force against animals, as we are only interested in cases involving humans.

```
library(conflicted)
library(dplyr)
library(tidyverse)
```

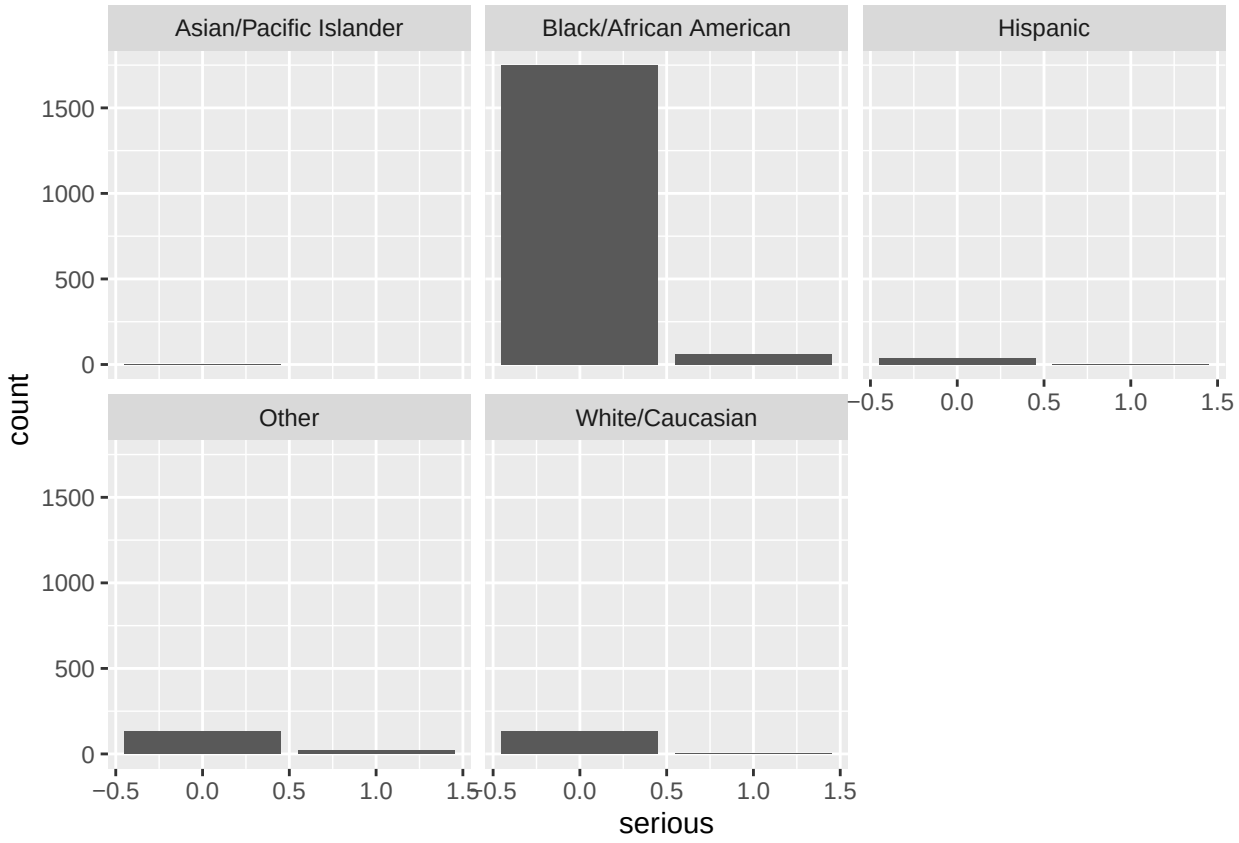
```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v forcats   1.0.0     v readr     2.1.4
## v ggplot2   3.4.3     v stringr   1.5.0
## v lubridate 1.9.2     v tibble    3.2.1
## v purrr     1.0.2     v tidyr     1.3.0
```

```
library(ggthemes)
library(ggplot2)
library(tibble)
# https://mpdc.dc.gov/node/1635896
df <- read.csv("UoF_mpd_2021_public.csv")
#df
```

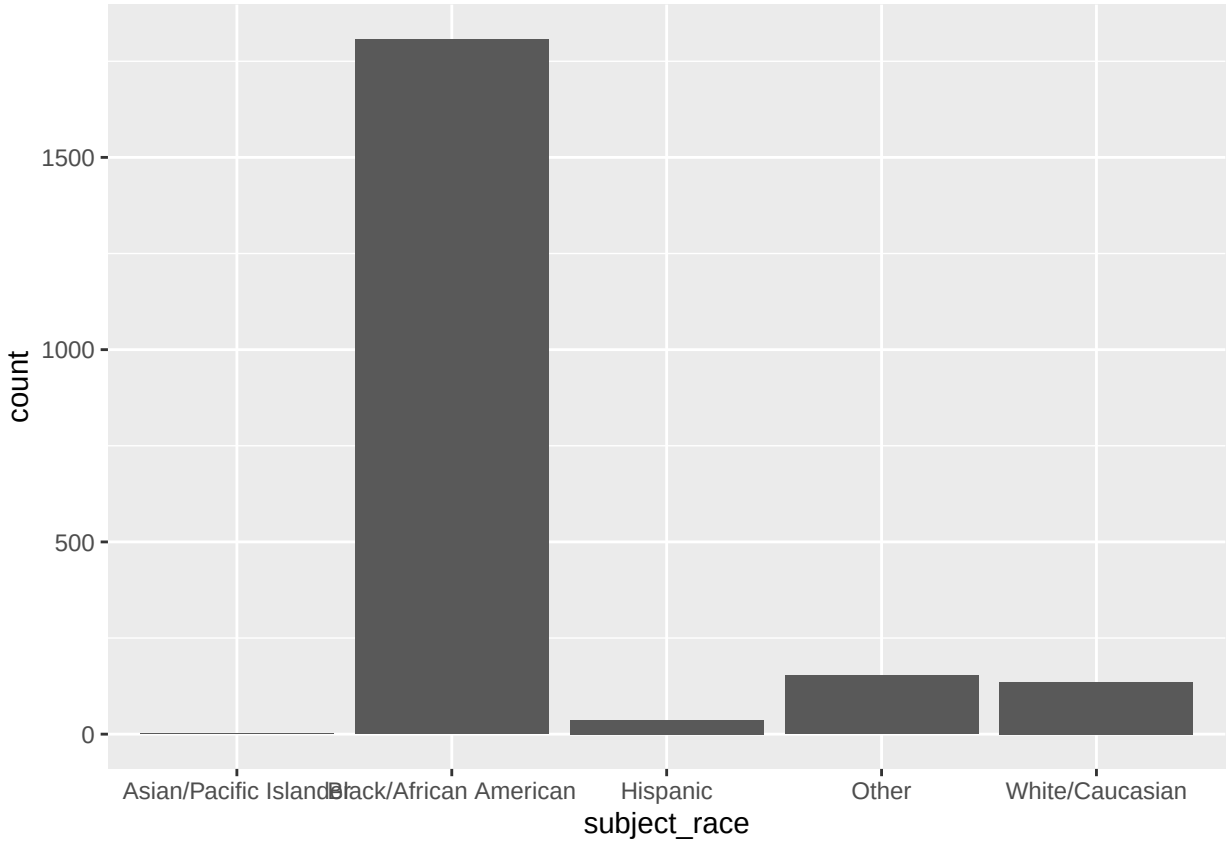
```
mpd_dat <- read.csv('UoF_mpd_2021_public.csv')
#head(mpd_dat)
#View(mpd_dat)

ggplot(data = mpd_dat, mapping = aes(x = IncidentDistrict)) +
  geom_bar() +
    facet_wrap(~uof_type)
```



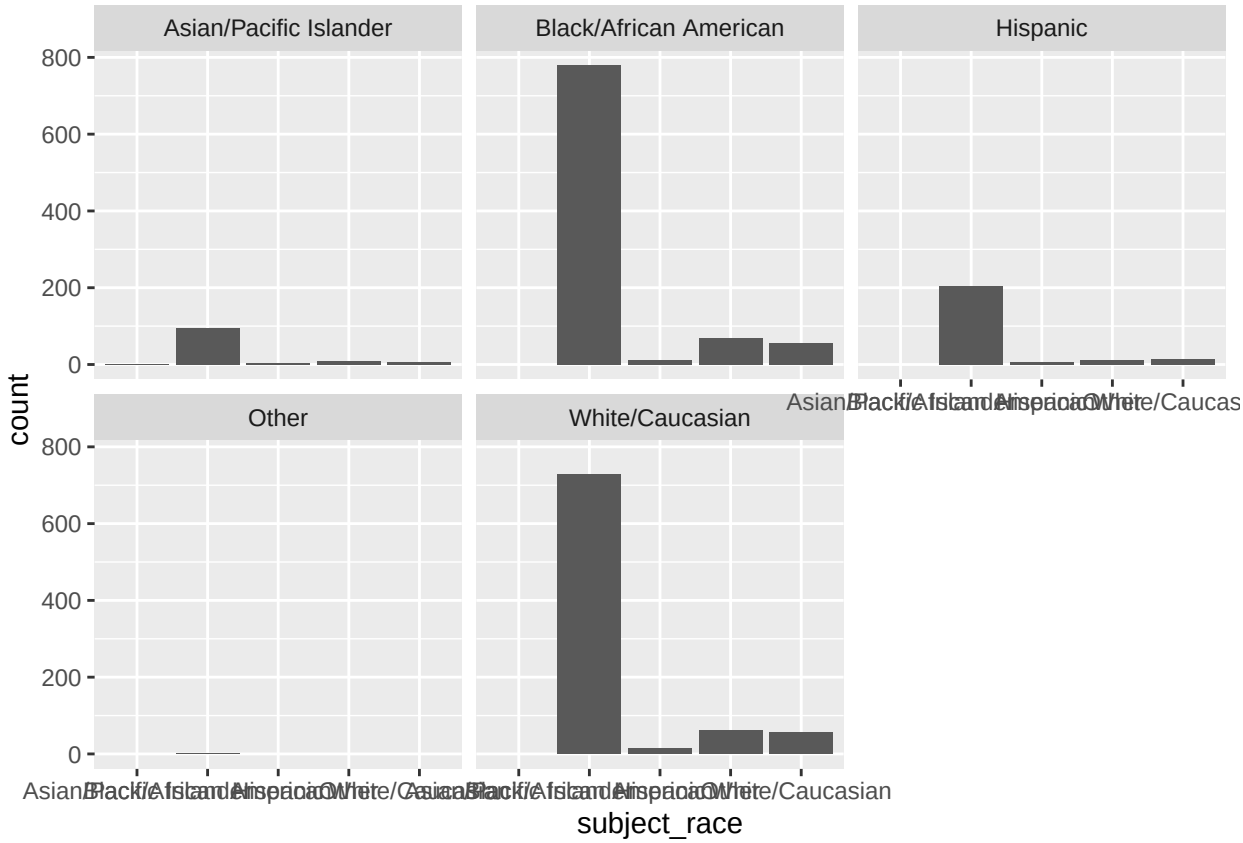

```
ggplot(data = mpd_dat, mapping = aes(x = serious)) +
  geom_bar() +
    facet_wrap(~subject_race)
```




```
ggplot(data = mpd_dat, mapping = aes(x = subject_race)) +
  geom_bar()
```



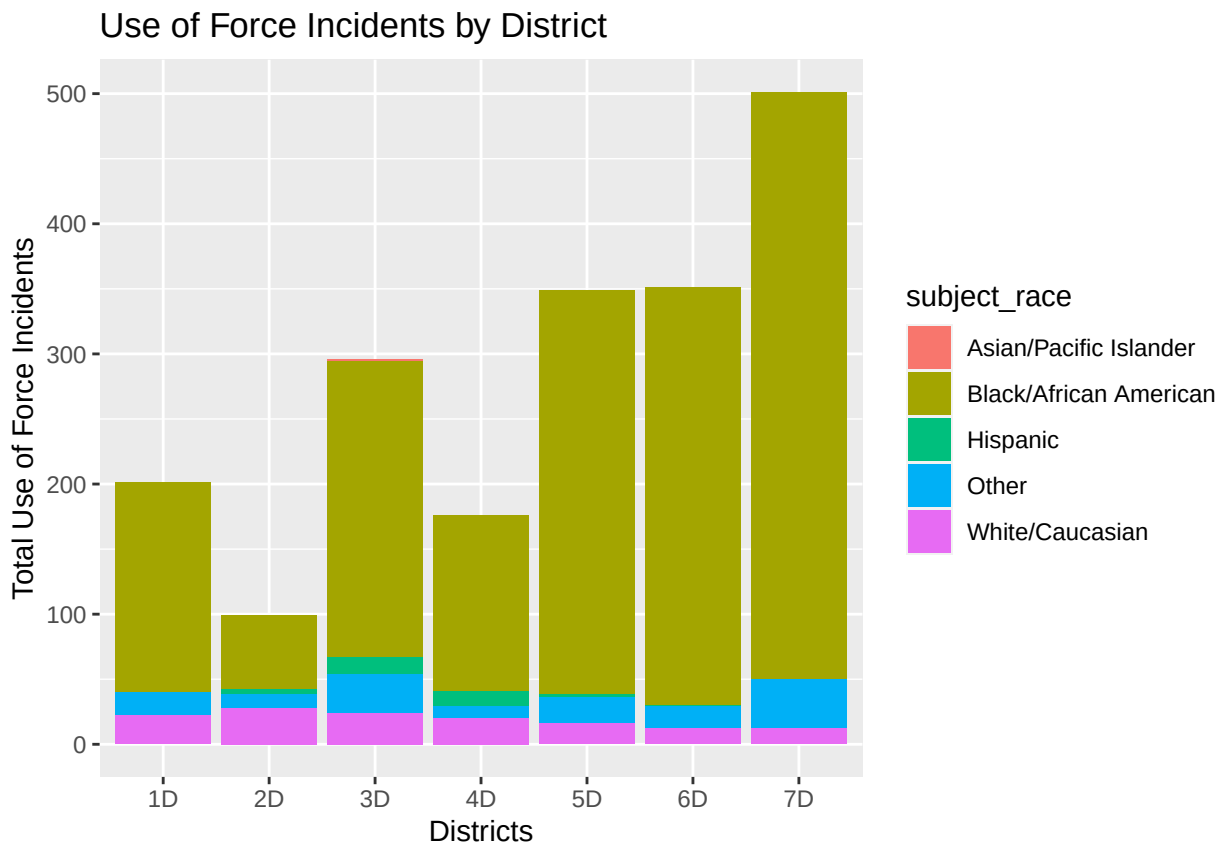

```
ggplot(data = mpd_dat, mapping = aes(x = subject_race)) +
  geom_bar() +
    facet_wrap(~OfficerRace)
```




```
force_by_district_race <- mpd_dat %>%
  group_by(OfficerAssignment, subject_race) %>%
  summarize(Total_Use_of_Force = n()) #total num of cases of use of force
```

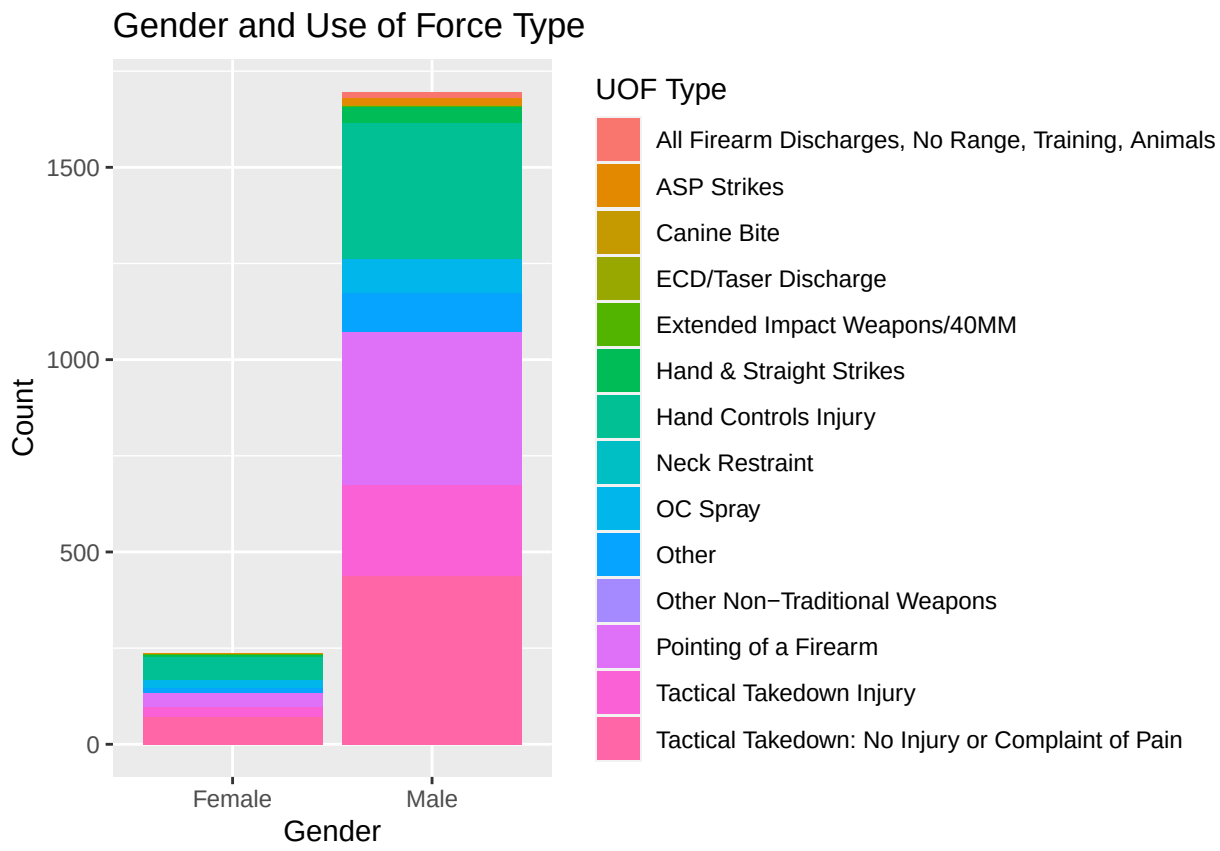
```
## `summarise()` has grouped output by 'OfficerAssignment'. You can override using
## the `.groups` argument.
```

```
ggplot(subset(force_by_district_race, OfficerAssignment %in% c("1D", "2D", "3D", "4D", "5D", "6D", "7D"
       aes(x = OfficerAssignment, y = Total_Use_of_Force, fill = subject_race)) +
  geom_bar(stat = "identity") +
  labs(title = "Use of Force Incidents by District") +
  xlab("Districts") +
  ylab("Total Use of Force Incidents")
```




```
#what dispositions are we most interested in knowing more about?
```

```
#graph for gender and uof type
mpd_dat <- mpd_dat[complete.cases(mpd_dat), ] #removed nulls

ggplot(mpd_dat, aes(x = subject_gender, fill = uof_type)) +
  geom_bar() +
  labs(title = "Gender and Use of Force Type",
       x = "Gender",
       y = "Count") +
  scale_fill_discrete(name = "UOF Type")
```




```
unique(mpd_dat$disposition)
```

```
## [1] "Justified - Tactical Improvement Opportunity"
## [2] "Tracking Only"
## [3] "Justified - Within Department Policy"
## [4] "Not Justified - Not Within Departmental Policy"
```

```
#rename
#Justified - tactical improvement opp -> jtip
#Justified - within department policy -> jwdp
#not justified - not within departmental policy -> njnwdp

mpd_dat$disposition <- recode(mpd_dat$disposition,
  "Justified - Tactical Improvement Opportunity" = "JTIP",
  "Tracking Only" = "Tracking",
  "Justified - Within Department Policy" = "JWDP",
  "Not Justified - Not Within Departmental Policy" = "NJNWDP",
  "NA" = "Unknown"
)
```

```
unique(mpd_dat$disposition)
```

```
## [1] "JTIP"     "Tracking" "JWDP"     "NJNWDP"
```

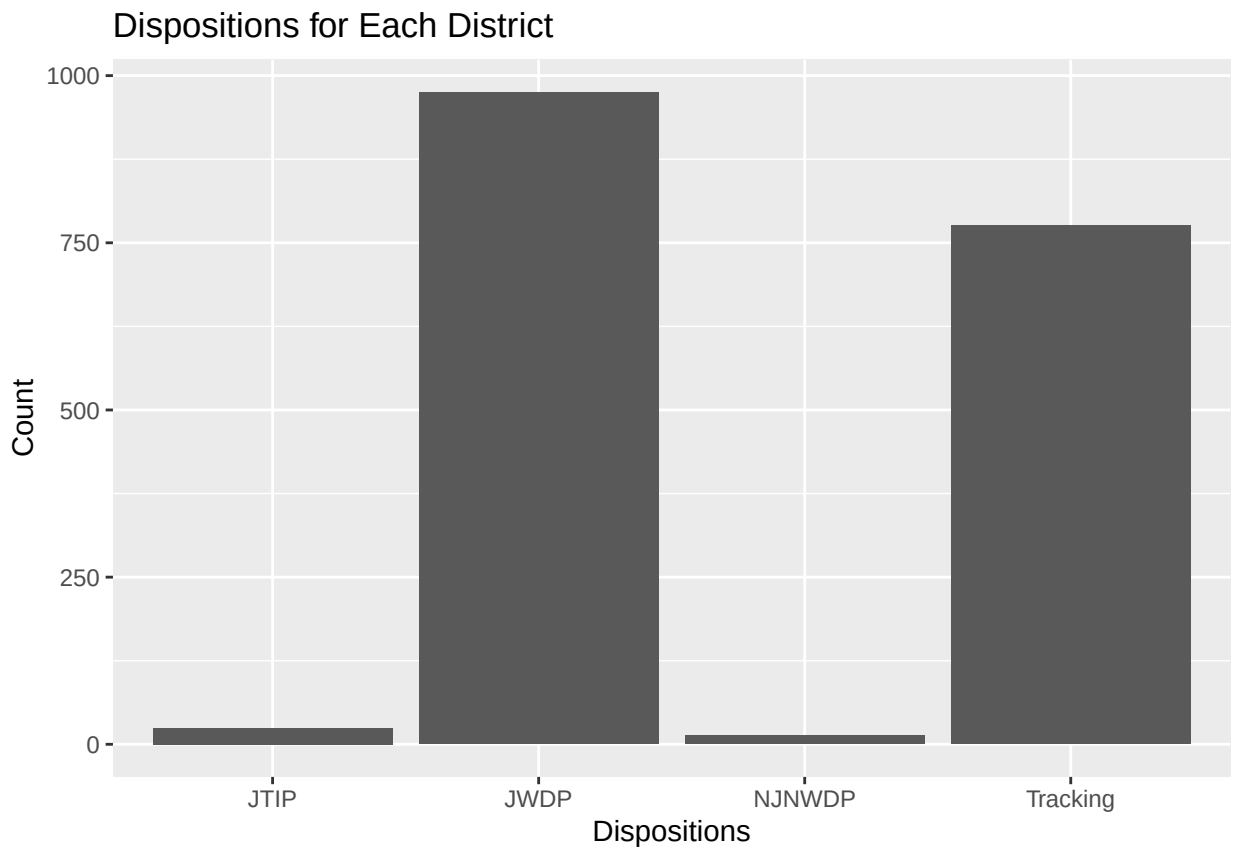
```r
unique(mpd_dat$uof_type)
```

```
##  [1] "Hand Controls Injury"
##  [2] "Hand & Straight Strikes"
##  [3] "Tactical Takedown: No Injury or Complaint of Pain"
##  [4] "Tactical Takedown Injury"
##  [5] "All Firearm Discharges, No Range, Training, Animals"
##  [6] "Pointing of a Firearm"
##  [7] "OC Spray"
##  [8] "ASP Strikes"
##  [9] "Other"
## [10] "Extended Impact Weapons/40MM"
## [11] "Canine Bite"
## [12] "ECD/Taser Discharge"
## [13] "Other Non-Traditional Weapons"
## [14] "Neck Restraint"
```

```r
unique(mpd_dat$OfficerAssignment)
```

```
##  [1] "5D"
##  [2] "6D"
##  [3] "4D"
##  [4] "1D"
##  [5] "3D"
##  [6] "7D"
##  [7] "2D"
##  [8] "School Safety Division"
##  [9] "Narcotics and Special Investigations Division/Violent Crime Suppression Division"
## [10] "Special Operations Division"
## [11] "Metropolitan Police Academy Division"
## [12] "Criminal Investigations Division"
## [13] "Strategic Change Division"
## [14] "Youth and Family Services Division"
## [15] "Administration"
```

```r
districts <- c("1D", "2D", "3D", "4D", "5D", "6D", "7D")
mpd_dat_filtered <- mpd_dat[mpd_dat$OfficerAssignment %in% districts, ]


ggplot(mpd_dat_filtered, aes(x = disposition)) +
  geom_bar() +
  labs(title = "Dispositions for Each District",
       x = "Dispositions",
       y = "Count")
```

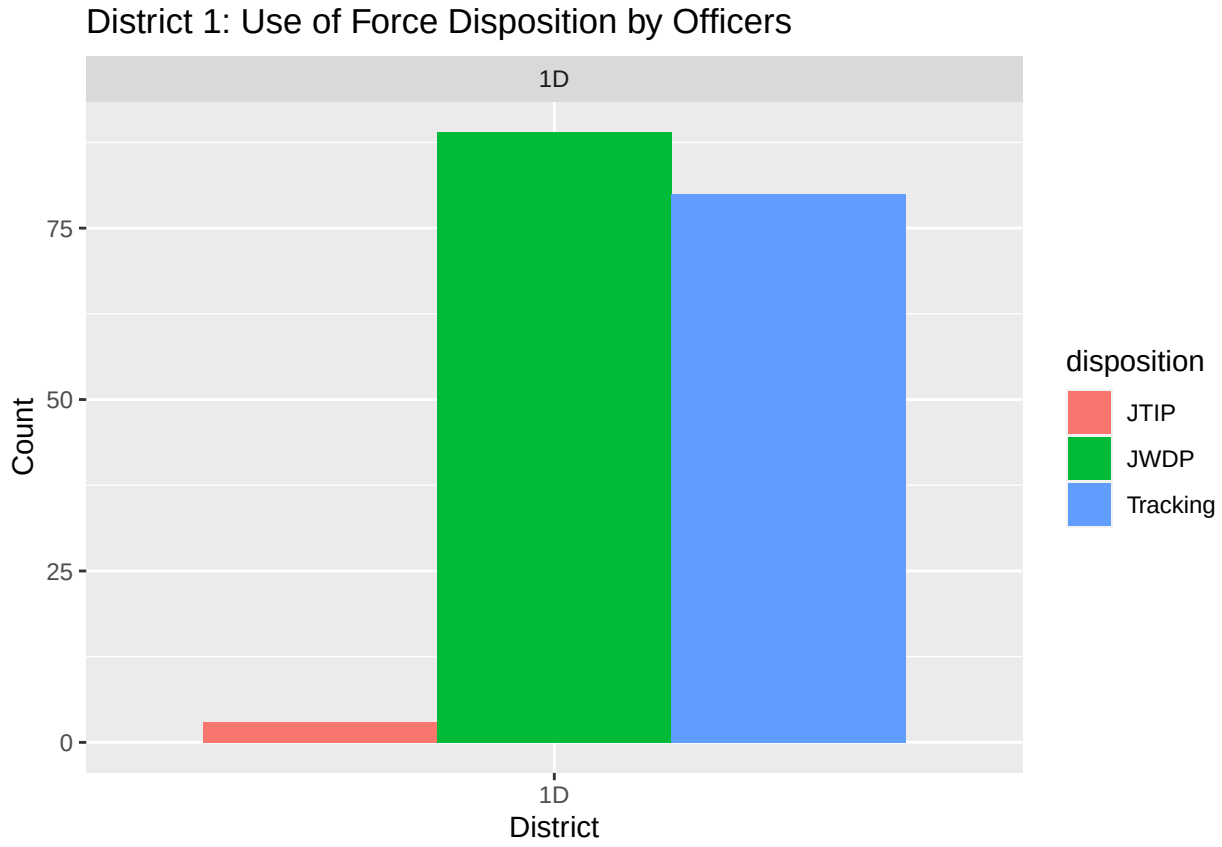



```
#df with officer assignment, disposition, officer race
officer_by_district <- mpd_dat %>%
  group_by(OfficerAssignment, disposition, OfficerRace)
```

```
summary_data <- officer_by_district %>%
  filter(OfficerAssignment %in% c("1D")) %>%
  group_by(OfficerAssignment, disposition) %>%
  summarise(count = n())
```

```
## `summarise()` has grouped output by 'OfficerAssignment'. You can override using
## the `.groups` argument.
```

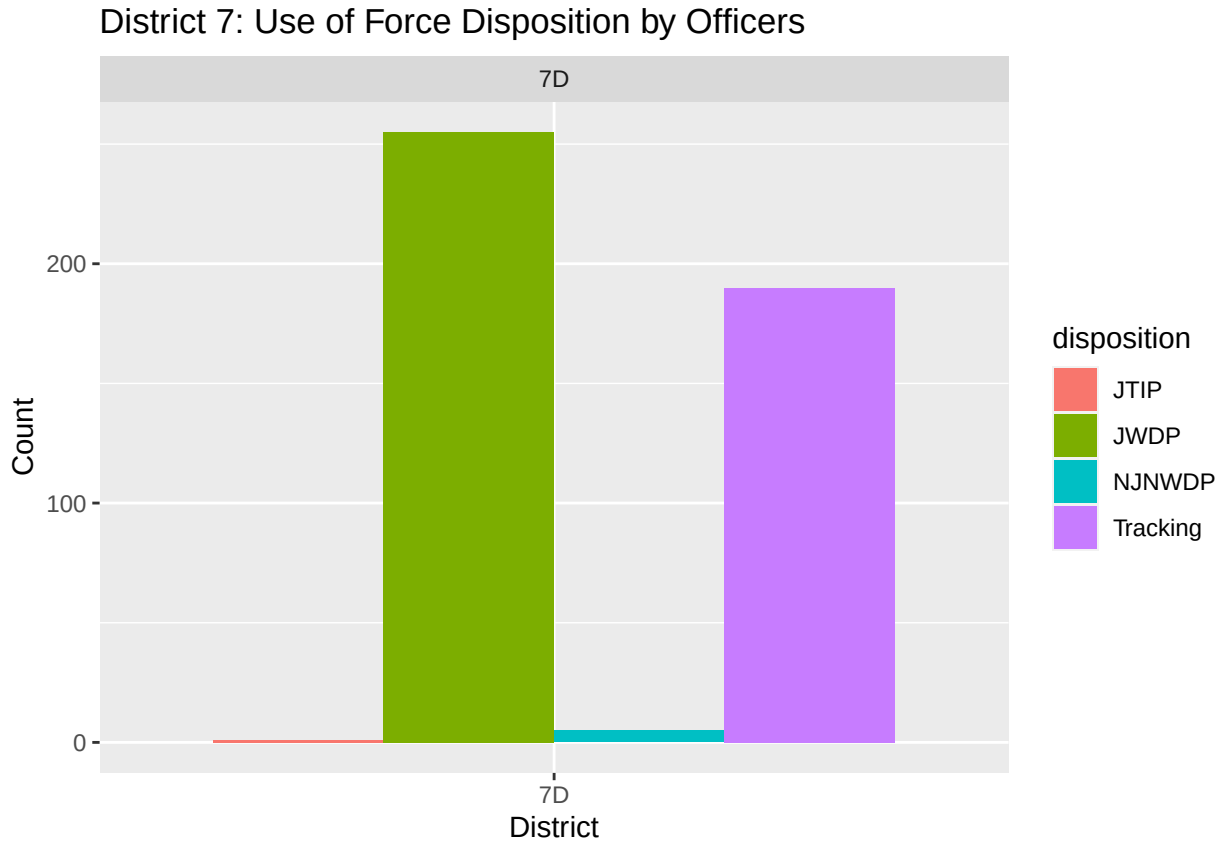
```
ggplot(summary_data, aes(x = OfficerAssignment, fill = disposition, y = count)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "District 1: Use of Force Disposition by Officers",
       x = "District",
       y = "Count") +
  facet_wrap(~OfficerAssignment, ncol = 2)
```



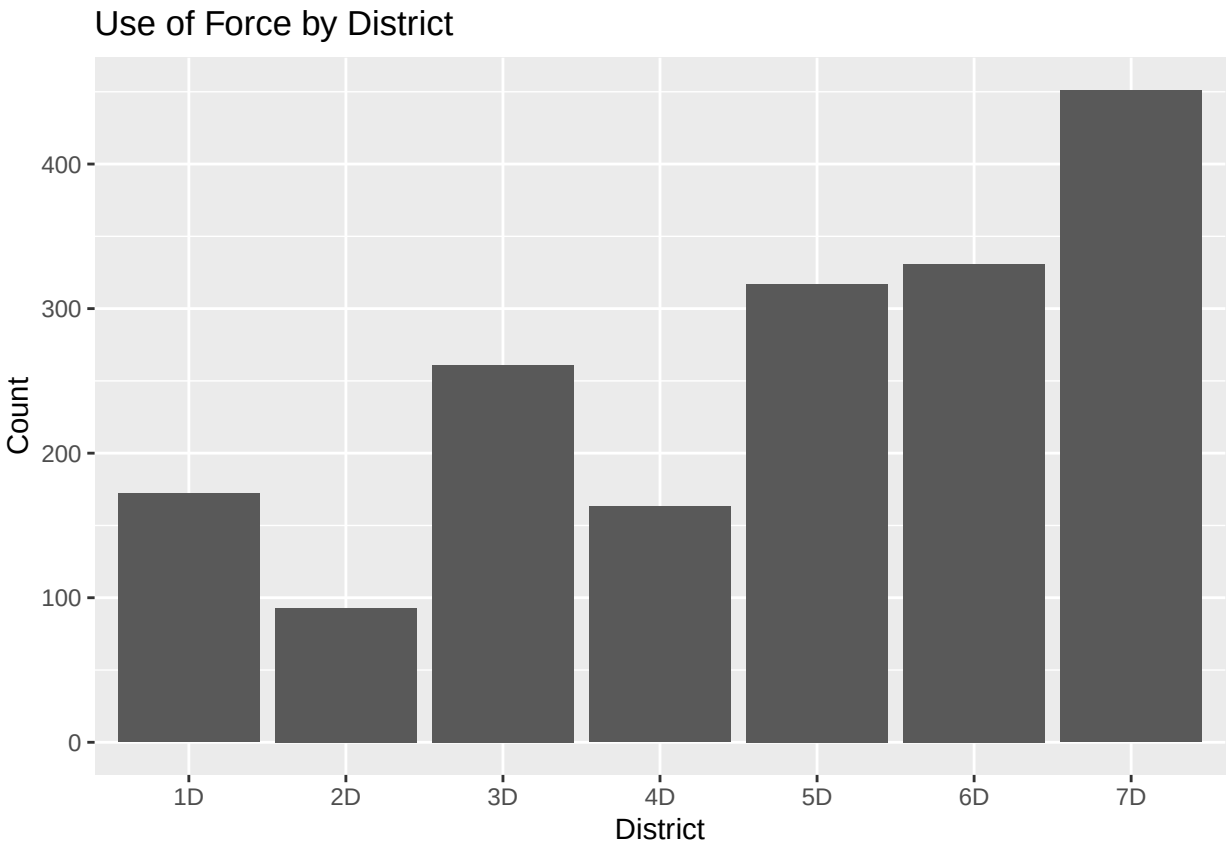

```
summary_data <- officer_by_district %>%
  filter(OfficerAssignment %in% c("7D")) %>%
  group_by(OfficerAssignment, disposition) %>%
  summarise(count = n())
```

```
## `summarise()` has grouped output by 'OfficerAssignment'. You can override using
## the `.groups` argument.
```

```
ggplot(summary_data, aes(x = OfficerAssignment, fill = disposition, y = count)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "District 7: Use of Force Disposition by Officers",
       x = "District",
       y = "Count") +
  facet_wrap(~OfficerAssignment, ncol = 2)
```



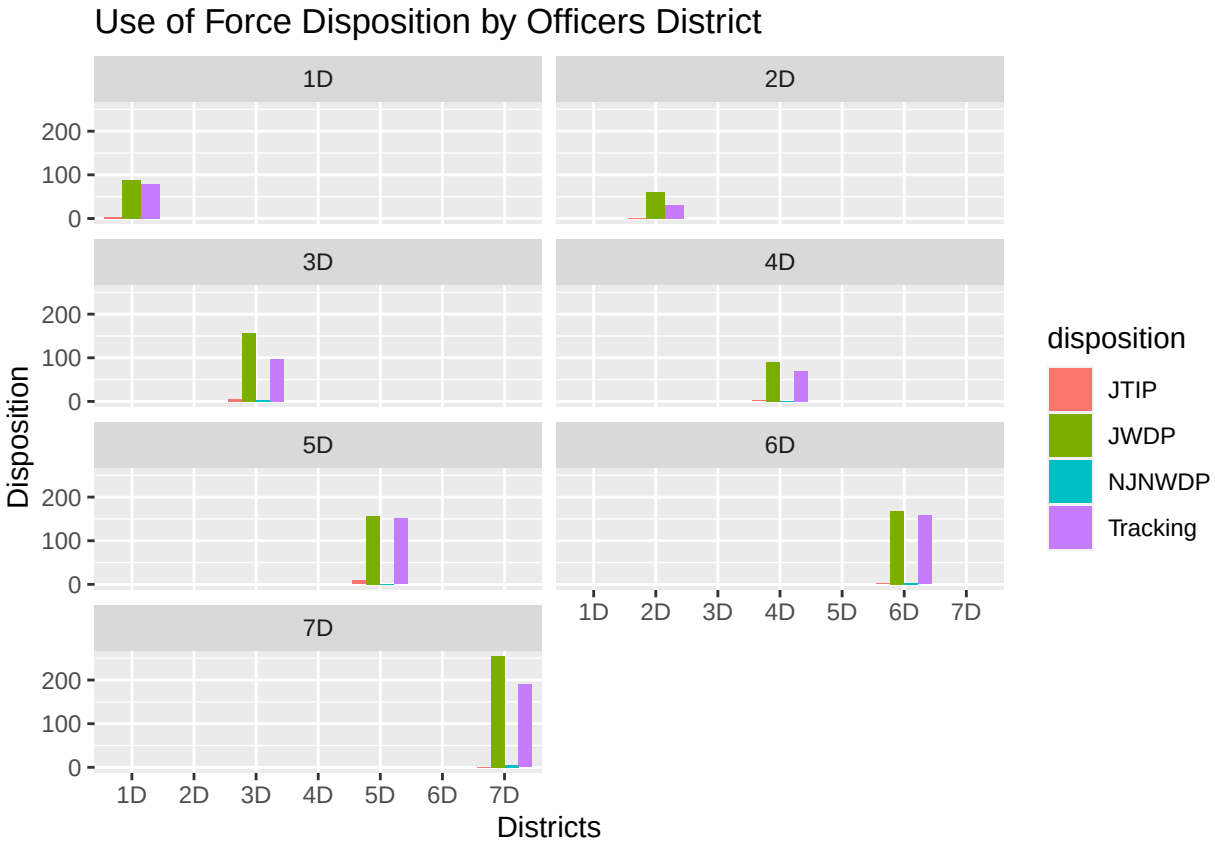

```
districts <- c("1D", "2D", "3D", "4D", "5D", "6D", "7D")
mpd_dat_filtered <- mpd_dat[mpd_dat$OfficerAssignment %in% districts, ]

uof_counts <- mpd_dat_filtered %>%
  group_by(OfficerAssignment) %>%
  summarise(Count = n())

ggplot(uof_counts, aes(x = OfficerAssignment, y = Count)) +
  geom_bar(stat = "identity") +
  labs(title = "Use of Force by District",
       x = "District",
       y = "Count")
```




```
summary_data <- officer_by_district %>%
  filter(OfficerAssignment %in% c("1D", "2D", "3D", "4D", "5D", "6D", "7D")) %>%
  group_by(OfficerAssignment, disposition) %>%
  summarise(count = n())
```

```
## `summarise()` has grouped output by 'OfficerAssignment'. You can override using
## the `.groups` argument.
```

```
ggplot(summary_data, aes(x = OfficerAssignment, fill = disposition, y = count)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Use of Force Disposition by Officers District",
       x = "Districts",
       y = "Disposition") +
  facet_wrap(~OfficerAssignment, ncol = 2)
```



Citation:

Schwartz G., (2020). Mapping fatal police violence across U.S. metropolitan areas: Overall

rates and racial/ ethnic inequities, 2013-2017. EBSCO.

https://web.p.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=6ba04cd8-6484-4883-8e3e-7e0e405082c8%40redis

Golash-Boza, T. Et al, (2022). Broken windows and order-maintenance policing in gentrifying Washington, DC. Taylor & Francis Online.

https://doi.org/10.1080/10439463.2022.2085268

Metropolitan Department. (2022). 2020 and 2021 UoF_Explanatory Notesv2.

Edwards, F., Lee, H., & Esposito, M. (2019). Risk of being killed by police use of force in the United States by age, race–ethnicity, and sex. Proceedings of the National Academy of Sciences, 116(34), 16793–16798. https://doi.org/10.1073/pnas.1821204116

Report reveals over 1,000 uses of forces by DC Police in 2021 | wusa9.com. (n.d.). Retrieved October 17, 2023, from https://www.wusa9.com/article/news/local/dc/annual-use-of-force-report-dc-police-released-more-than-800-complaints-filed/65-9bc43d2b-059d-43bd-a785-3bed9d6ca858

Lepore, J. (2020, July 13). The New Yorker. https://www.newyorker.com/magazine/2020/07/20/the-invention-of-the-police#:~:text=It%20is%20also%20often%20said,the%20history%20of%20the%20police