

# EDA

## Exploratory Data Analysis

### Imports

```
library(tidyverse)

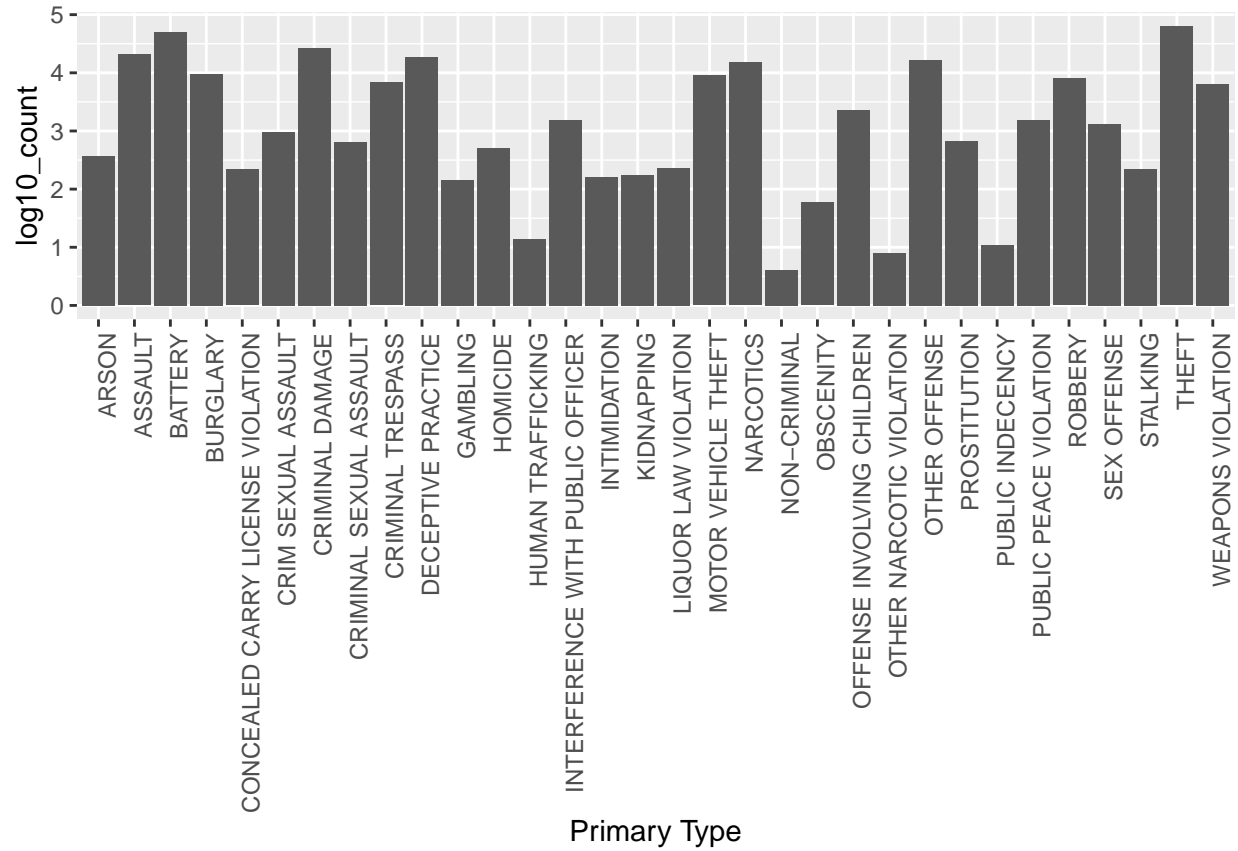
## -- Attaching packages ----- tidyverse 1.3.0 --
## v ggplot2 3.3.2      v purrr  0.3.4
## v tibble  3.0.4      v dplyr  1.0.2
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

### Load data

```
df = read_csv("./data/crime-2019.csv", col_types = cols())

df %>%
  group_by(`Primary Type`) %>%
  tally(name = "count") %>%
  mutate(log10_count = log10(count)) %>%
  ggplot() +
  geom_col(aes(x=`Primary Type`, y=log10_count)) +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```



Sp