

NYP Shootings

Fredy Paz

NYPD Shootings Analysis

Overview

The data set used in this analysis is from the City of New York. It contains detailed records of shooting incidents.

The intent of this document is to conduct analysis and gather insights. We will look at a breakdown of incidents per boroughs, date, and time of day.

Importing data into variable “nypddata” from <https://data.cityofnewyork.us/api/views/833y-fsy8/rows.csv?accessType=DOWNLOAD>

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.5.1      v tibble    3.2.1
## v lubridate  1.9.4      v tidyr     1.3.1
## v purrr      1.0.4
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
nypddata <-
read.csv("https://data.cityofnewyork.us/api/views/833y-fsy8/rows.csv?accessType=DOWNLOAD")
```

Summary of data

```
summary(nypddata)
```

```
##  INCIDENT_KEY      OCCUR_DATE      OCCUR_TIME      BORO
##  Min.   : 9953245    Length:28562    Length:28562    Length:28562
##  1st Qu.: 65439914   Class :character Class :character Class :character
##  Median : 92711254   Mode  :character Mode  :character Mode  :character
##  Mean   :127405824
##  3rd Qu.:203131993
##  Max.   :279758069
##
##  LOC_OF_OCCUR_DESC  PRECINCT  JURISDICTION_CODE LOC_CLASSFCTN_DESC
```

```
## Length:28562      Min.   : 1.0   Min.   :0.0000   Length:28562
## Class :character  1st Qu.: 44.0   1st Qu.:0.0000   Class :character
## Mode :character   Median : 67.0   Median :0.0000   Mode :character
##                  Mean    : 65.5   Mean    :0.3219
##                  3rd Qu.: 81.0   3rd Qu.:0.0000
##                  Max.    :123.0   Max.    :2.0000
##                  NA's    :2
## LOCATION_DESC     STATISTICAL_MURDER_FLAG PERP_AGE_GROUP
## Length:28562      Length:28562      Length:28562
## Class :character   Class :character   Class :character
## Mode :character    Mode :character    Mode :character
##
##
##
## PERP_SEX          PERP_RACE          VIC_AGE_GROUP      VIC_SEX
## Length:28562      Length:28562      Length:28562      Length:28562
## Class :character   Class :character   Class :character   Class :character
## Mode :character    Mode :character    Mode :character    Mode :character
##
##
##
## VIC_RACE          X_COORD_CD          Y_COORD_CD          Latitude
## Length:28562      Min.   : 914928   Min.   :125757   Min.   :40.51
## Class :character   1st Qu.:1000068   1st Qu.:182912   1st Qu.:40.67
## Mode :character    Median :1007772   Median :194901   Median :40.70
##                  Mean    :1009424   Mean    :208380   Mean    :40.74
##                  3rd Qu.:1016807   3rd Qu.:239814   3rd Qu.:40.82
##                  Max.    :1066815   Max.    :271128   Max.    :40.91
##                  NA's    :59
## Longitude         Lon_Lat
## Min.   : -74.25   Length:28562
## 1st Qu.: -73.94   Class :character
## Median : -73.92   Mode :character
## Mean    : -73.91
## 3rd Qu.: -73.88
## Max.    : -73.70
## NA's    : 59
```

Data Cleaning

Removing columns PRECINT, JURISDICTION_CODE

```
nypddatatidy <- nypddata %>% select(-JURISDICTION_CODE, -PRECINCT)
```

Changing OCCUR_DATE to Date format

```
nypddatatidy <- nypddatatidy %>% mutate(OCCUR_DATE = mdy(OCCUR_DATE))
```

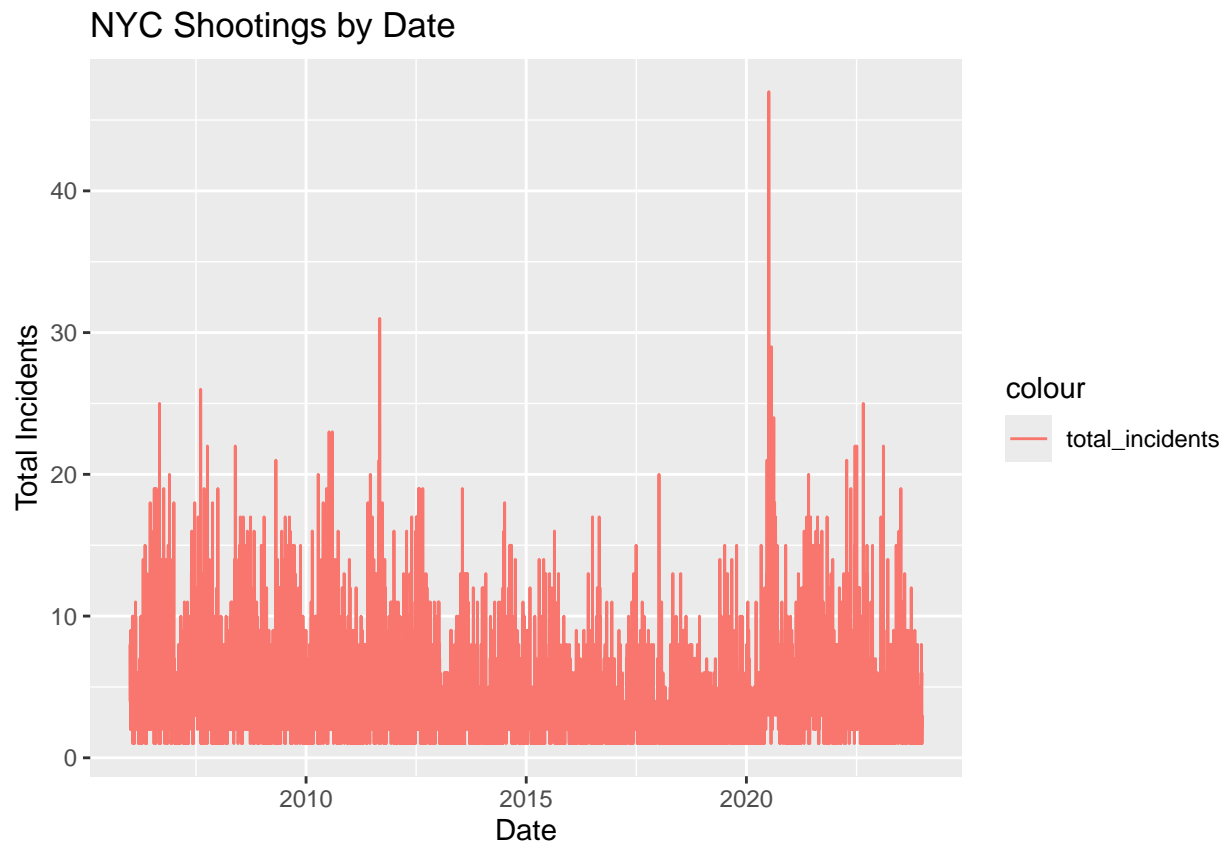
Data Analysis

Group shootings by date, and provide a count of shooting by date

```
nypd_bydate <- nypddatatidy %>% group_by(OCCUR_DATE) %>% summarize(total_incidents = n())
```

Adding a plot to view shootings by date

```
ggplot(nypd_bydate, aes(x = OCCUR_DATE, y = total_incidents)) + geom_line(aes(color = "total_incidents"))
```



Getting exact date with most shootings

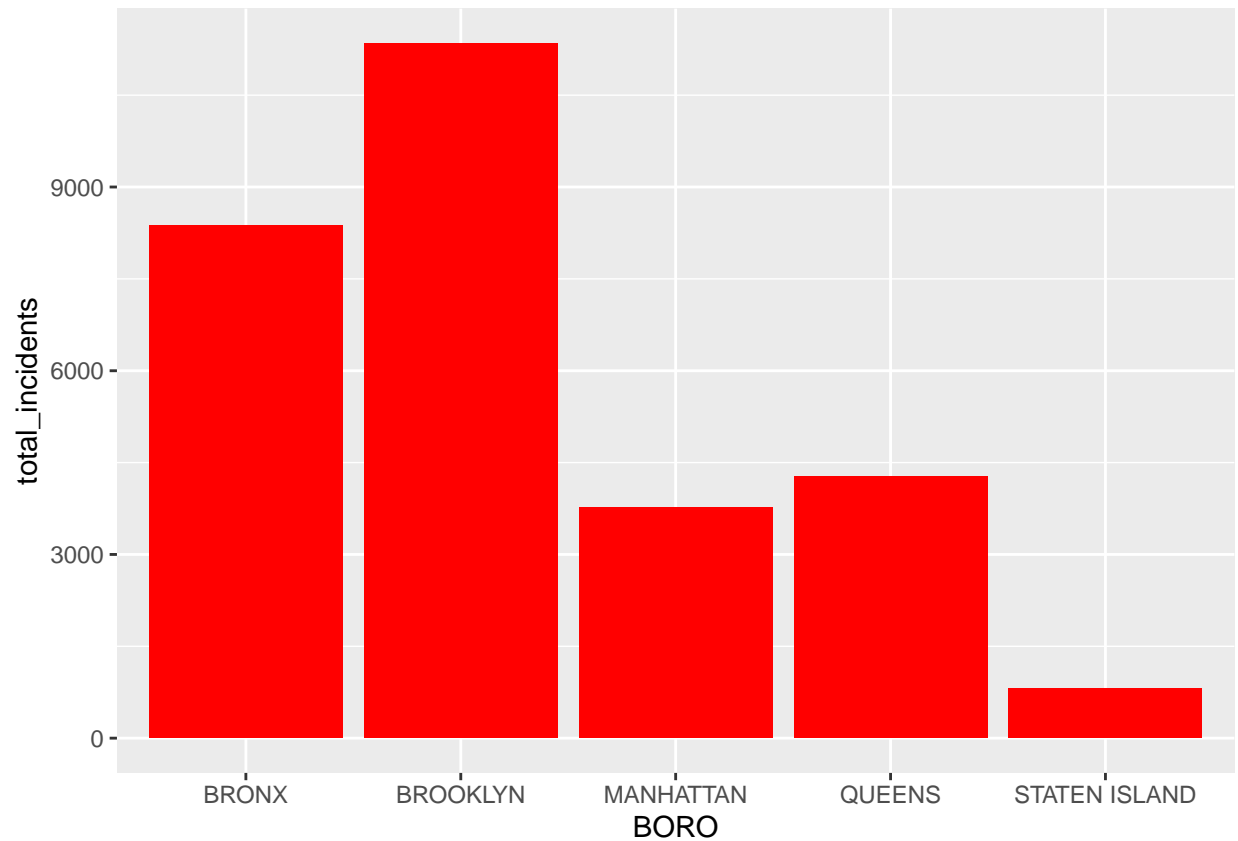
```
maxdate <- nypd_bydate %>% filter(total_incidents == max(total_incidents))
```

Group shooting by Boroughs

```
nypd_byboro <- nypddatatidy %>% group_by(BORO) %>% summarize(total_incidents = n())
```

Adding a bar chart to show shootings by Boroughs

```
ggplot(nypd_byboro, aes(x = BORO, y = total_incidents)) + geom_col(fill = "red")
```



Adding a variable to change OCCUR_TIME to time format hms

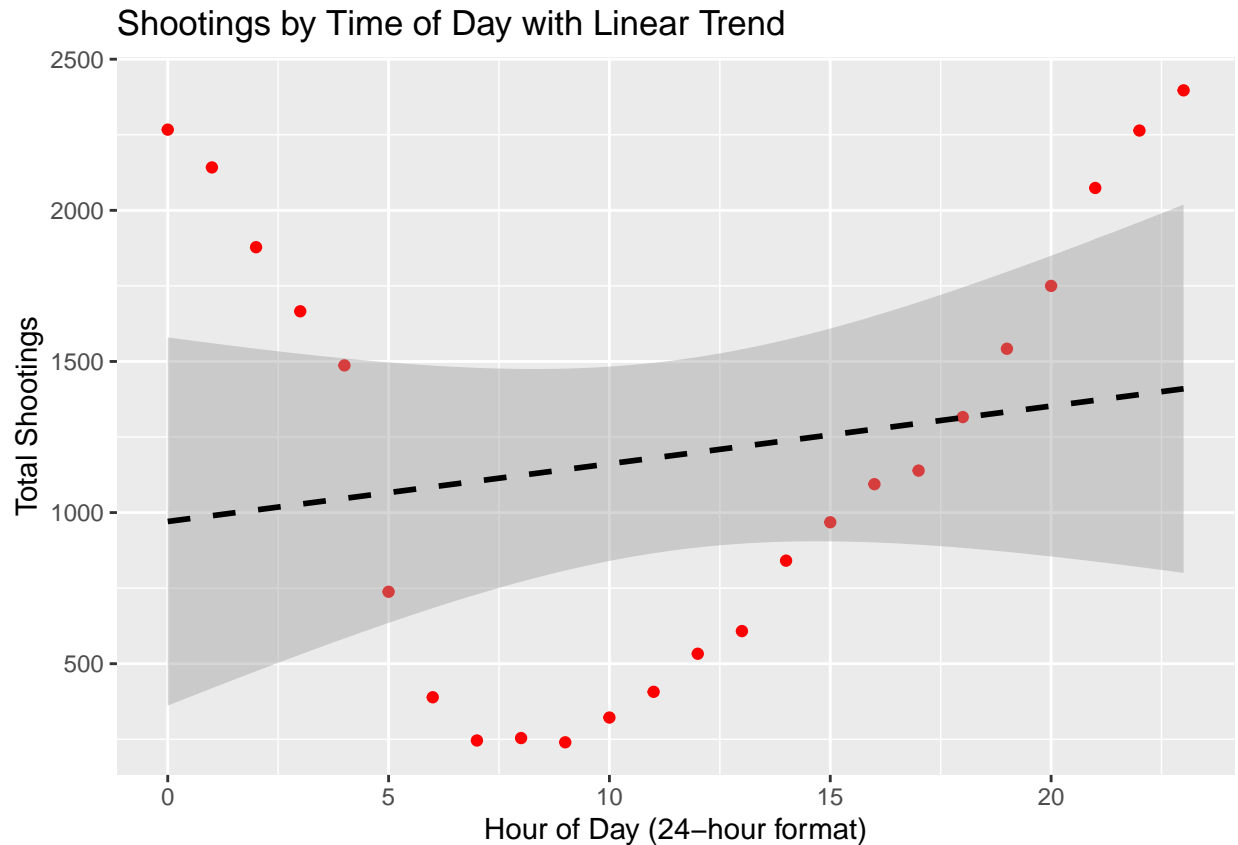
```
nypddatatidy <- nypddatatidy %>% mutate(time_hour = hour(hms(OCCUR_TIME)))
```

Grouping by hour of day and count shootings

```
nypd_byhour <- nypddatatidy %>% group_by(time_hour) %>% summarise(total_incidents = n())
```

Adding a plot with linear model showing shootings by time of day

```
ggplot(nypd_byhour, aes(x = time_hour, y = total_incidents)) +
  geom_point(color = "red") +
  geom_smooth(formula = y ~ x, method = "lm", se = TRUE, color = "black", linetype = "dashed") +
  labs(title = "Shootings by Time of Day with Linear Trend",
       x = "Hour of Day (24-hour format)",
       y = "Total Shootings")
```



The borough with the most shootings over the history of the data is Brooklyn with 11,346 total incidents.

```
maxboro <- nypd_byboro %>% filter(total_incidents == max(total_incidents))
maxboro
```

```
## # A tibble: 1 x 2
##   BORO      total_incidents
##   <chr>          <int>
## 1 BROOKLYN      11346
```

The date with the most shootings was July 5th 2020 with 47 shootings

```
maxdate
```

```
## # A tibble: 1 x 2
##   OCCUR_DATE total_incidents
##   <date>          <int>
## 1 2020-07-05         47
```

The bias that I identified was assuming that the most shootings occurred during peak covid times and that they would be mostly located in neighborhoods that we have historically associated with violence and poverty. I assumed Manhattan would not be at the top.

I reset my mindset to not make assumptions, relied on the data, and corroborated my findings by searching news for the specific date with most deaths. See resources below.

I found a few news articles mentioning the spike in violence but there was no single explanation or cause. It seems to be a combination of July 4th activity coupled with covid restrictions easing.

Resources:

<https://www.nbcnewyork.com/news/local/bullet-strikes-nypd-patrol-vehicle-misses-officers-sitting-inside/2500243/>

<https://abc7ny.com/nyc-shootings-2020-last-night-this-week-in/6299513/>