

## Visualization- Task1

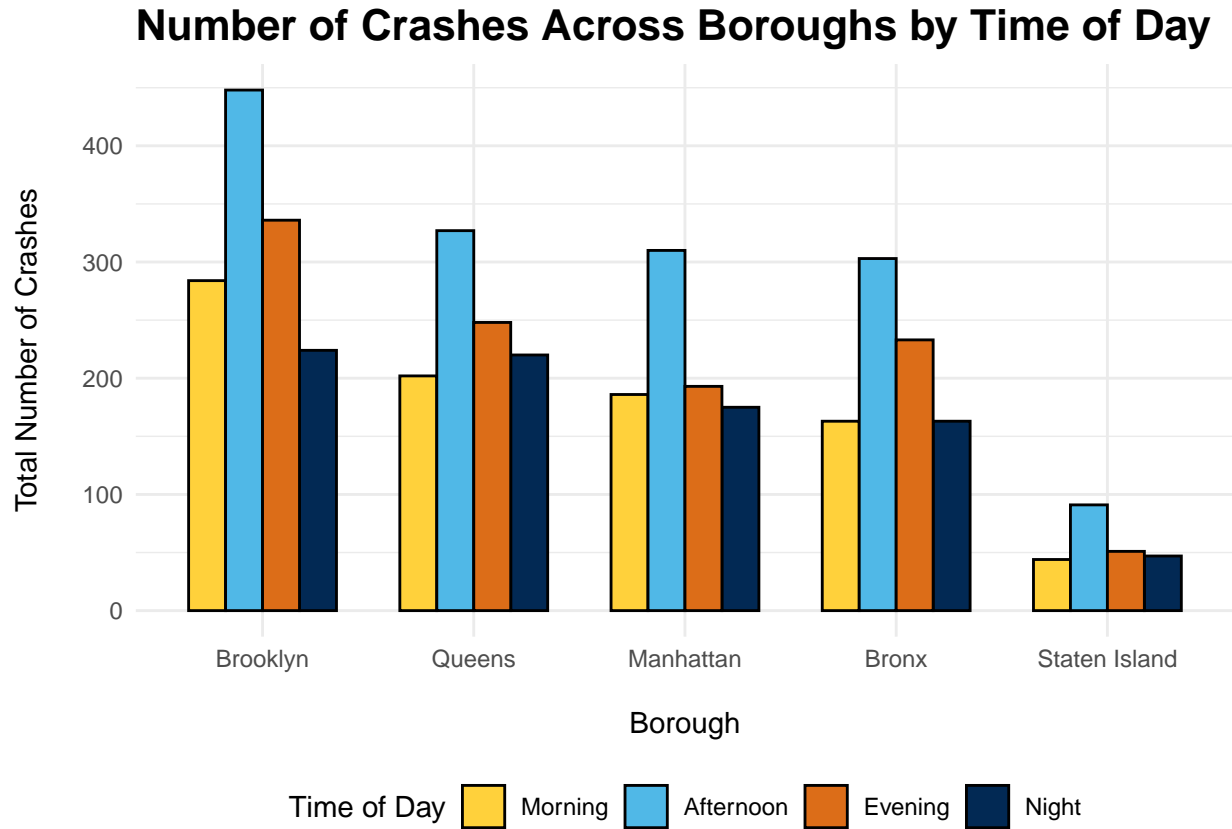
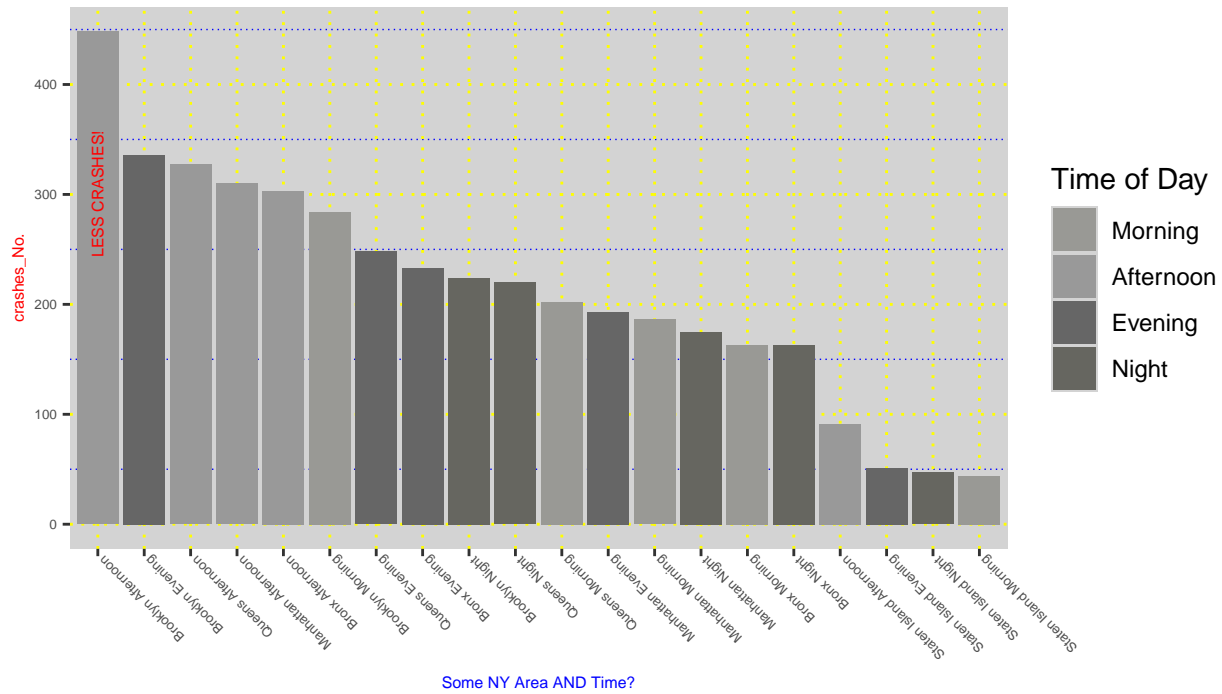


Figure 1. This visualization presents a breakdown of vehicular crashes across New York City boroughs, segmented by the time of day. Each colored bar represents a distinct time interval: morning, afternoon, evening, and night. The x-axis denotes the different NYC boroughs, while the y-axis shows the total number of crashes. The data provides insights into periods and locations of heightened crash activity, which can guide city planning and safety initiatives. The information was sourced from the Serious Injury Response, Tracking & Analysis Program (SIRTA) dataset.”

Figure2 !  
 Some crashes... Figure below gives some crashes in SF ?  
 The x-axis is borough + time (both) and y axis is crashes\_no.

Data from somewhere...Not known.



Important data source:

Another graph. This plot maybe shows a view or count of CRASHES! across NY areas. The x-axis is borough + time (both) and y axis is crashes\_no. The data is about city stuff. The data was sourced from SIRTa.

The questions to be answered via the plots:

- Which New York City boroughs experience the highest number of vehicular crashes?
- Are there any specific times of the day when certain boroughs experience a surge in vehicular crashes?

#### DATA DESCRIPTION-

The dataset originates from the Serious Injury Response, Tracking & Analysis Program (SIRTA) maintained by the New York City Department of Transportation (NYC DOT). It offers a comprehensive overview of severe vehicular crashes within New York City, capturing key details such as the date and time of the crash, precise location coordinates, the borough in which the incident occurred, and metrics on the severity of the crash, including non-severe injuries, severe injuries, and fatalities. Additional data points include street information, nearby landmarks, zip codes, and historical data on similar incidents in the past three years at the same locations. This dataset is instrumental for urban traffic safety analysis and infrastructure decisions.

#### DATA SOURCE:

New York City Department of Transportation. Last Updated August 7, 2023.

Serious Injury Response, Tracking & Analysis Program (SIRTA) Data. NYC Open Data.

Data link:

<https://data.cityofnewyork.us/Transportation/Serious-Injury-Response-Tracking-Analysis-Program-/xeqp-qz8h>

Some of the errors in Figure 2:

- Indistinct Colors: The shades of gray used for different times of the day are very similar, making it very difficult to distinguish between the categories at once
- Incorrect Data Presentation: The data is presented in a random order, making it challenging to find patterns or make comparisons.
- Redundant and misleading label: The “LESS CRASHES!” label in bright red color is distracting and incorrect.
- Incomplete & Incorrect title: The title and figure caption is oddly spaced, making it difficult to comprehend. It also gives out incomplete and wrong information.
- The use of red for the y-axis label and blue for x-axis label is not appropriate as it distracts from the data.
- The x-axis values are titled and are very difficult to read at once.
- X and Y axis labeling: The y-axis is labeled “crashes\_No.” which is an incorrect and unclear way of labeling. The x-axis label “Some NY Area AND time?” is vague and uninformative.
- Unnecessary Invisible Grid lines: The yellow and blue dotted grid lines are almost invisible against the gray background
- Distracting Background: The dark gray background makes the plot look gloomy and further worsens the interpretation of already indistinct bars.

## Code Used To Make These Plots

Below are the code chunks used to make these plots.

```
# Load necessary libraries
library(ggplot2)

# Load the dataset
data <- read.csv('/Users/riyanshibohra/Desktop/Coursework- UofA/Data Viz assignments/Data Viz6/SIRTA.csv')

# Convert Crash Time into a categorical variable for each time of the day
data$Time_Interval <- cut(as.numeric(
  format(as.POSIXct(data$Crash.Time, format="%H:%M"), "%H")),
  breaks = c(-1, 6, 12, 18, 24),
  labels = c("Night", "Morning", "Afternoon", "Evening"))

# Calculating data by Crash Borough and Time_Interval for number of crashes
data_crash_count <- aggregate(Crash.Date ~ Crash.Borough + Time_Interval,
  data=data, FUN=length)
names(data_crash_count)[3] <- "Total_Crashes"

# Setting Time_Interval
data_crash_count$Time_Interval <- factor(data_crash_count$Time_Interval,
  levels = c("Morning", "Afternoon", "Evening", "Night"))

# Pretty plot color palette
custom_colors <- c("Morning" = "#ffd13b",
  "Afternoon" = "#50b8e7",
  "Evening" = "#DC6d18",
  "Night" = "#022954")

# Reorder the Crash Borough
data_crash_count$Crash.Borough <- with(data_crash_count,
  reorder(Crash.Borough, -Total_Crashes))

# Create the "pretty" bar plot
pretty_plot <- ggplot(data_crash_count,
  aes(x = Crash.Borough, y = Total_Crashes, fill = Time_Interval)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.7), width = 0.7,
    color="black") +
  labs(title = "Number of Crashes Across Boroughs by Time of Day",
    x = "Borough",
    y = "Total Number of Crashes") +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    plot.title = element_text(face = "bold", size = 16),
    axis.title.x = element_text(margin = margin(t = 15, r = 0, b = 0, l = 0)),
    axis.title.y = element_text(margin = margin(t = 0, r = 15, b = 0, l = 0))
  ) +
  scale_fill_manual(values = custom_colors, guide = guide_legend(title = "Time of Day"))

print(pretty_plot)
```

```

# Ugly plot color palette
gray_shades <- c("Morning" = "#999995",
                 "Afternoon" = "#999999",
                 "Evening" = "#666666",
                 "Night" = "#666660")

# Combining borough and time interval for the x-axis
data_crash_count$Borough_Time <- paste(data_crash_count$Crash.Borough,
                                       data_crash_count$Time_Interval)

# Create the "ugly" bar plot
ugly_plot <- ggplot(data_crash_count, aes(x = reorder(Borough_Time, -Total_Crashes),
                                           y = Total_Crashes, fill = Time_Interval)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Figure2 !
Some crashes... Figure below gives some crashes in SF ?
      The x-axis is borough + time (both) and y axis is crashes_no.",
       subtitle = "Data from somewhere...Not known.",
       caption = "Important data source:",
       x = "Some NY Area AND Time?",
       y = "crashes_No.") +
  theme_gray() +
  theme(
    legend.position = "right",
    plot.title = element_text(size = 8, color = "red"),
    plot.subtitle = element_text(size = 7, color = "blue"),
    axis.title.x = element_text(size = 6, color = "blue"),
    axis.title.y = element_text(size = 6, color = "red"),
    axis.text.x = element_text(size = 5, angle = 135, hjust = 1),
    axis.text.y = element_text(size = 5),
    panel.background = element_rect(fill = "lightgray"),
    panel.grid.major = element_line(color = "yellow", linetype = "dotted"),
    panel.grid.minor = element_line(color = "blue", linetype = "dotted")
  ) +
  scale_fill_manual(values = gray_shades, guide = guide_legend(title = "Time of Day"))+
  annotate("text", x = 1, y = max(data_crash_count$Total_Crashes) * 0.8,
         label = "LESS CRASHES!", color = "red", size = 2, angle = 90, hjust = 1)

print(ugly_plot)

```