My title*

My subtitle if needed

Robert Ford

long??

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

2 Data

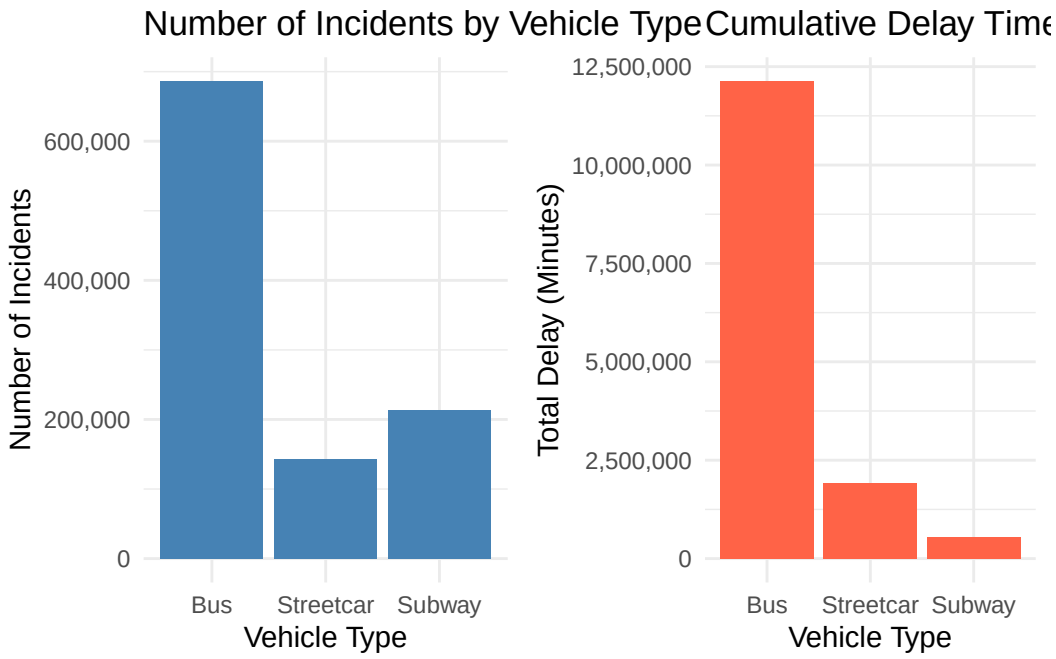

Figure 1: (left) Total Number of Incidents (right) Total Amount of Delay Time

*Code and data are available at: https://github.com/Ford-Robert/STA304_City-Of-Toronto-Data.git

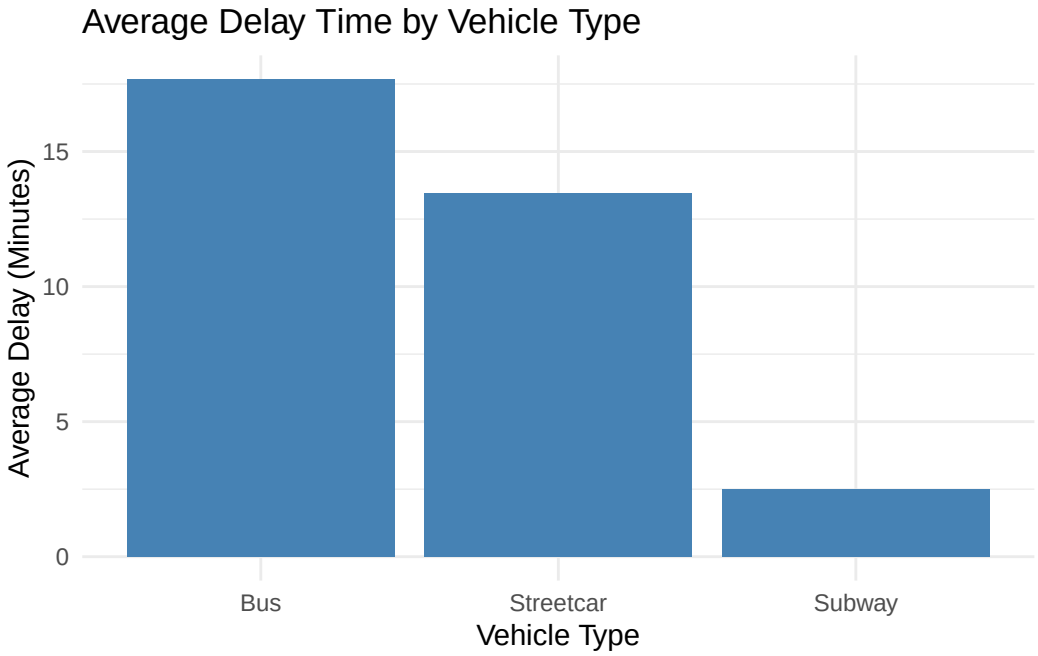

Figure 2: Average Amount of Delay Time by Vehicle Type

```r
# Step 1: Calculate total delay time for each incident type
data_incident_total_delay <- data %>%
  group_by(Incident) %>%
  summarise(total_delay = sum(Delay, na.rm = TRUE), .groups = 'drop') %>%
  mutate(Incident = fct_reorder(Incident, total_delay, .desc = TRUE))  # Reorder by total_del

# Step 2: Create the bar chart for total delay time by incident type
ggplot(data_incident_total_delay, aes(x = Incident, y = total_delay)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  labs(title = "Total Delay Time by Incident Type",
       x = "Incident Type", y = "Total Delay Time (Minutes)") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1, vjust = 1, size = 10))
```

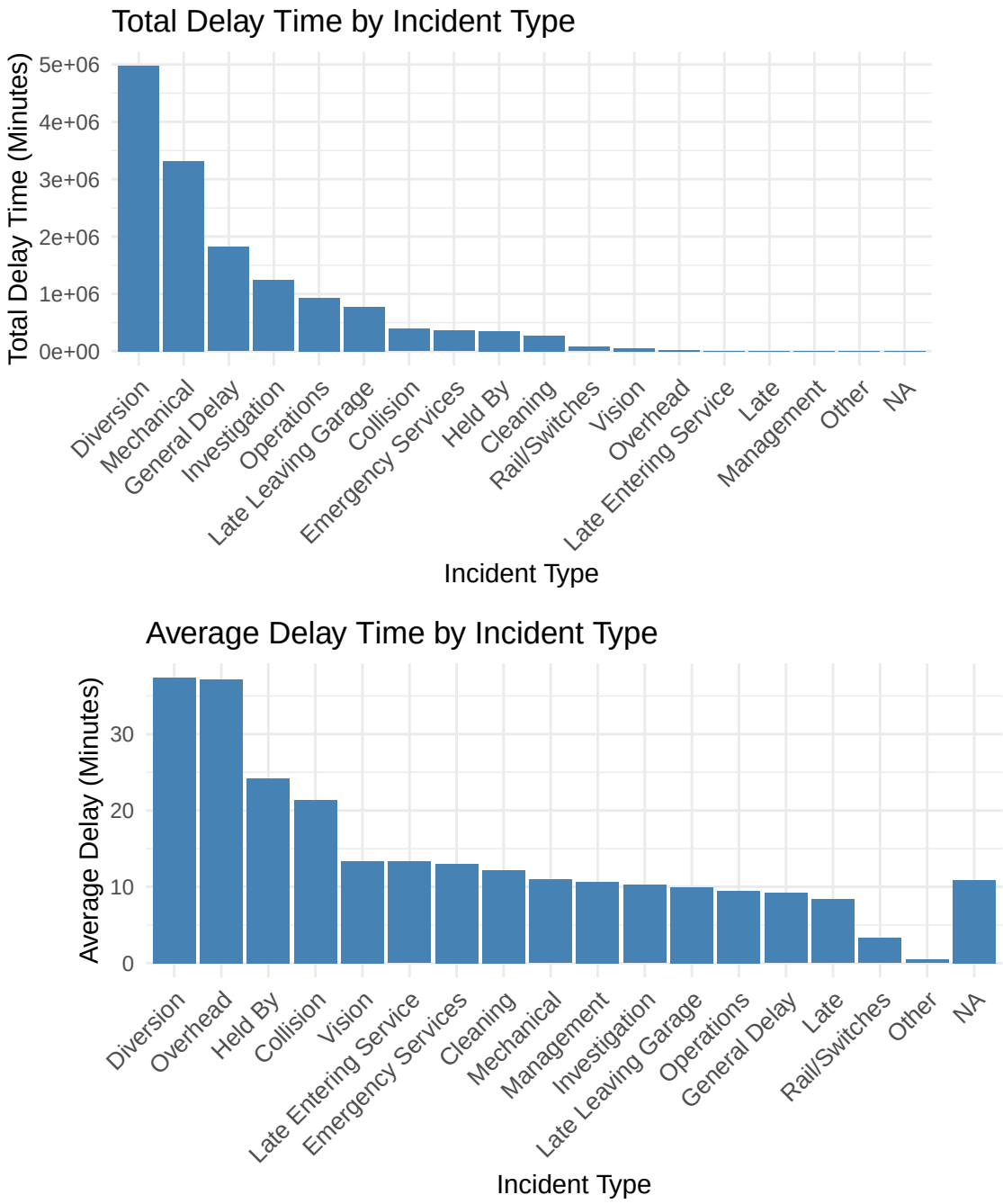

Figure 3: Average Amount of Delay Time by Incident Type

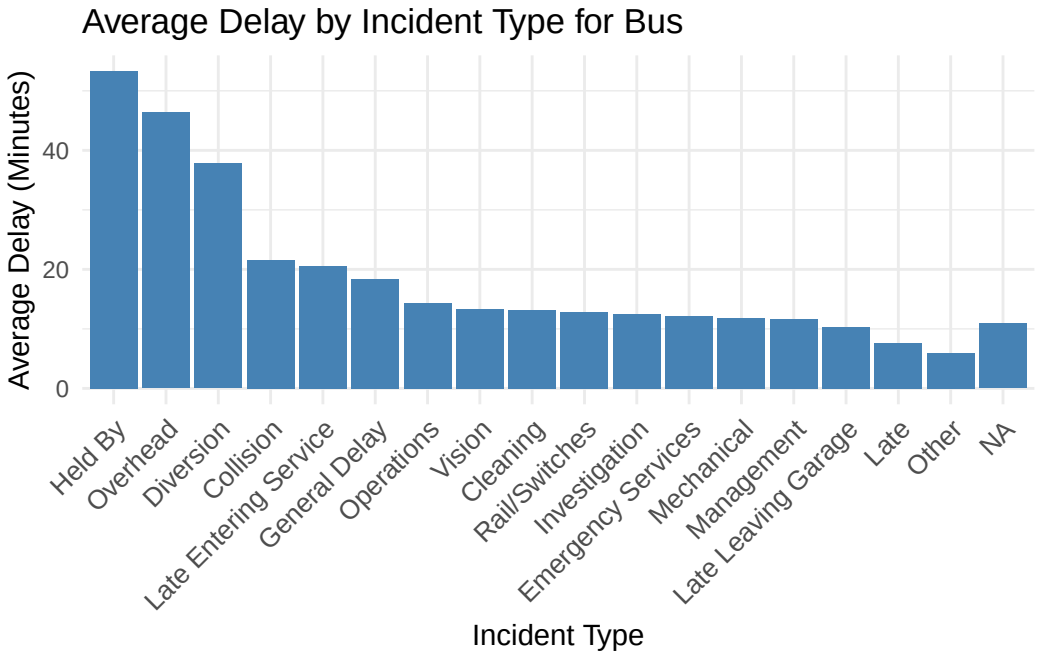

Figure 4: Average Amount of Delay Time for Buses by, Incident Type

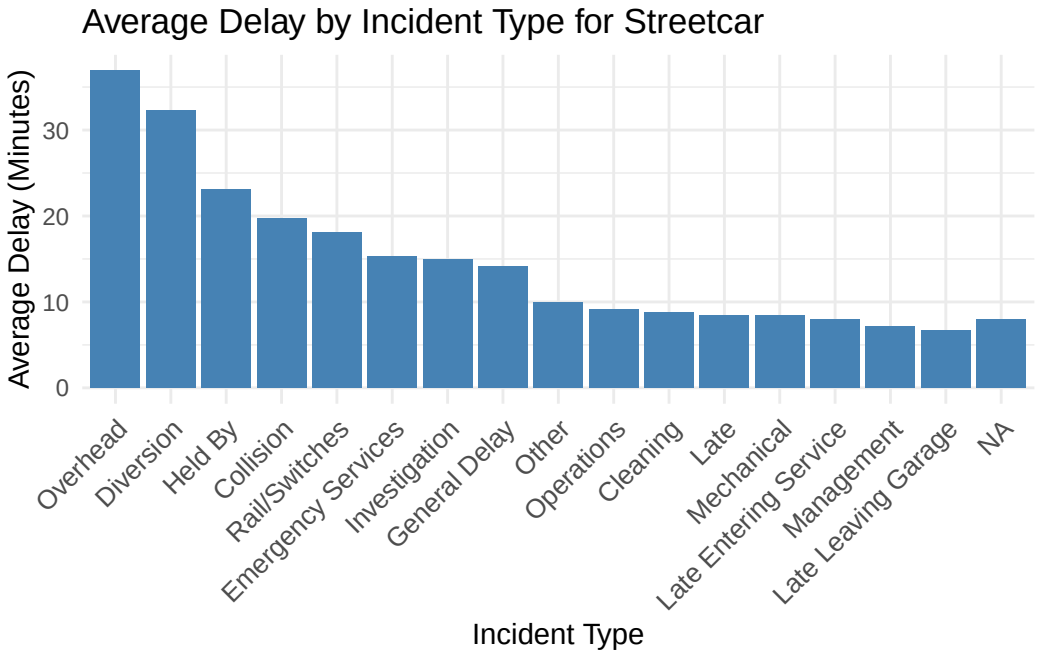

Figure 5: Average Amount of Delay Time for Streetcars, by Incident Type

```r
# Calculate total delay by vehicle and incident
data_incident_delay <- data %>%
  group_by(vehicle, Incident) %>%
  summarise(total_delay = sum(Delay, na.rm = TRUE)) %>%
  ungroup()
```

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

```r
# Calculate proportions for each incident by vehicle
data_incident_delay <- data_incident_delay %>%
  group_by(vehicle) %>%
  mutate(total_vehicle_delay = sum(total_delay),
         proportion = total_delay / total_vehicle_delay * 100) %>%
  ungroup()

# Keep top 5 incidents per vehicle and group the rest as "Other"
data_top5_incidents <- data_incident_delay %>%
  group_by(vehicle) %>%
  arrange(vehicle, desc(proportion)) %>%
  mutate(rank = row_number()) %>%
  mutate(Incident = ifelse(rank > 5, "Other", Incident)) %>%
  group_by(vehicle, Incident) %>%
  summarise(total_delay = sum(total_delay),
            proportion = sum(proportion)) %>%
  ungroup()
```

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

```r
incident_colors <- c(
  "Cleaning" = "#1f77b4",
  "Collision" = "#ff7f0e",
  "Diversion" = "#2ca02c",
  "Emergency Services" = "#d62728",
  "General Delay" = "#9467bd",
  "Held By" = "#8c564b",
  "Investigation" = "#d62728",
  "Late Entering Service" = "#7f7f7f",
  "Late Leaving Garage" = "#bcbd22",
```

```
  "Management" = "#17becf",
  "Mechanical" = "#1f77b4",
  "Operations" = "#ff7f0e",
  "Overhead" = "#2ca02c",
  "Rail/Switches" = "#2ca02c",
  "Security" = "#d62728",
  "Vision" = "#8c564b",
  "Other" = "#7f7f7f",
  "N/A" = "#7f7f7f"
)

# Function to create a pie chart for a given vehicle
create_pie_chart <- function(data, vehicle_type) {
  data_vehicle <- data %>% filter(vehicle == vehicle_type)

  ggplot(data_vehicle, aes(x = "", y = proportion, fill = Incident)) +
    geom_bar(width = 1, stat = "identity", color = "white") +
    coord_polar("y") +
    labs(title = vehicle_type) +
    theme_void() +  # Removes background and axis
    theme(legend.position = "none",  # Removes the legend
          plot.title = element_text(hjust = 0.5)) +
    geom_text(aes(label = paste0(Incident, ": ", round(proportion, 1), "%")),
              position = position_stack(vjust = 0.5), size = 3) +
    scale_fill_manual(values = incident_colors)  # Apply the consistent color palette
}

# Create pie charts for each vehicle type
bus_pie <- create_pie_chart(data_top5_incidents, "Bus")
streetcar_pie <- create_pie_chart(data_top5_incidents, "Streetcar")
subway_pie <- create_pie_chart(data_top5_incidents, "Subway")

# Arrange the pie charts side by side
(bus_pie | streetcar_pie | subway_pie) +
  plot_annotation(title = "Proportion of Total Delay Time by Incident (Top 5 + Other)")
```

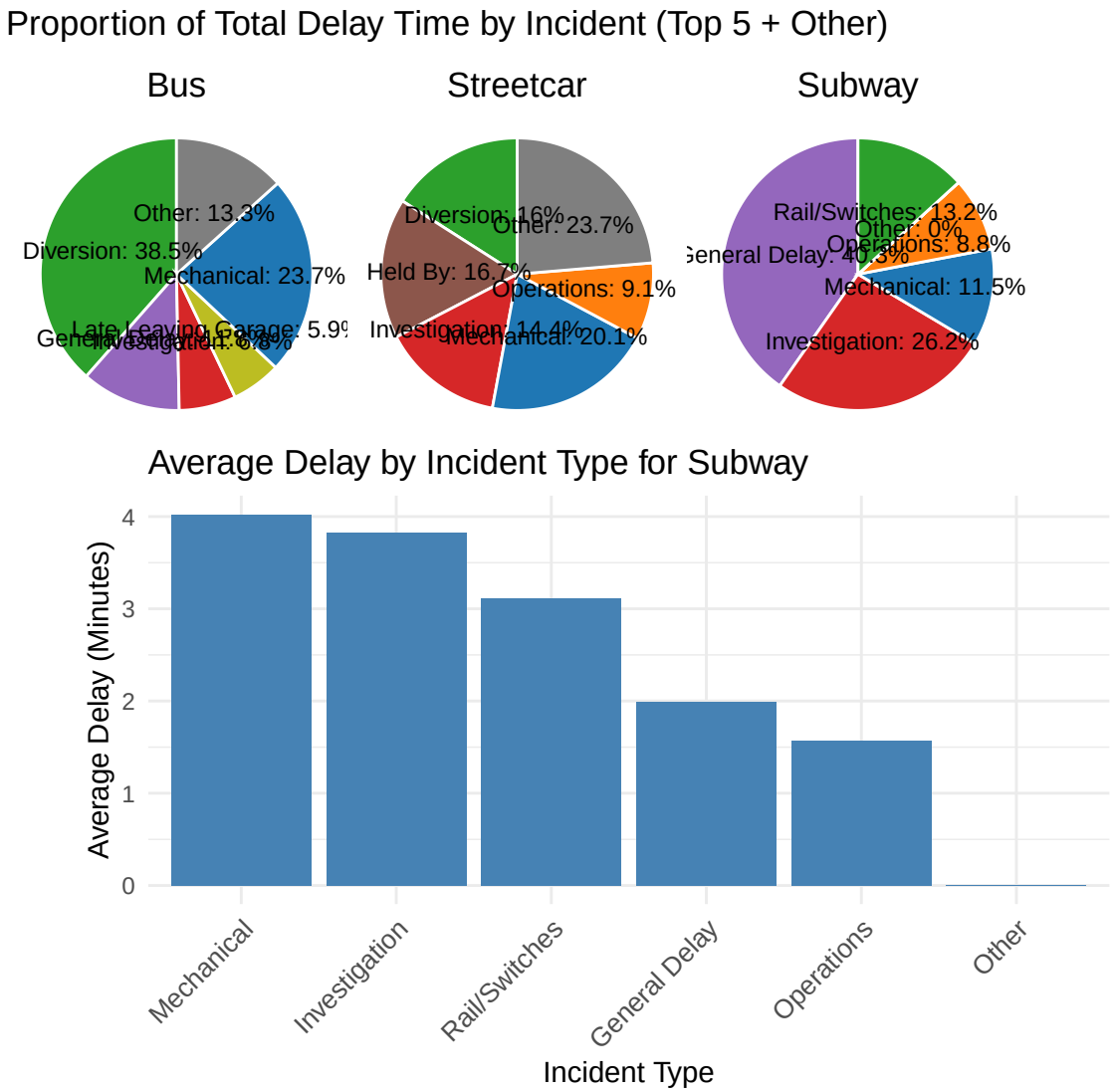

Figure 6: Average Amount of Delay Time for Subways, by Incident Type

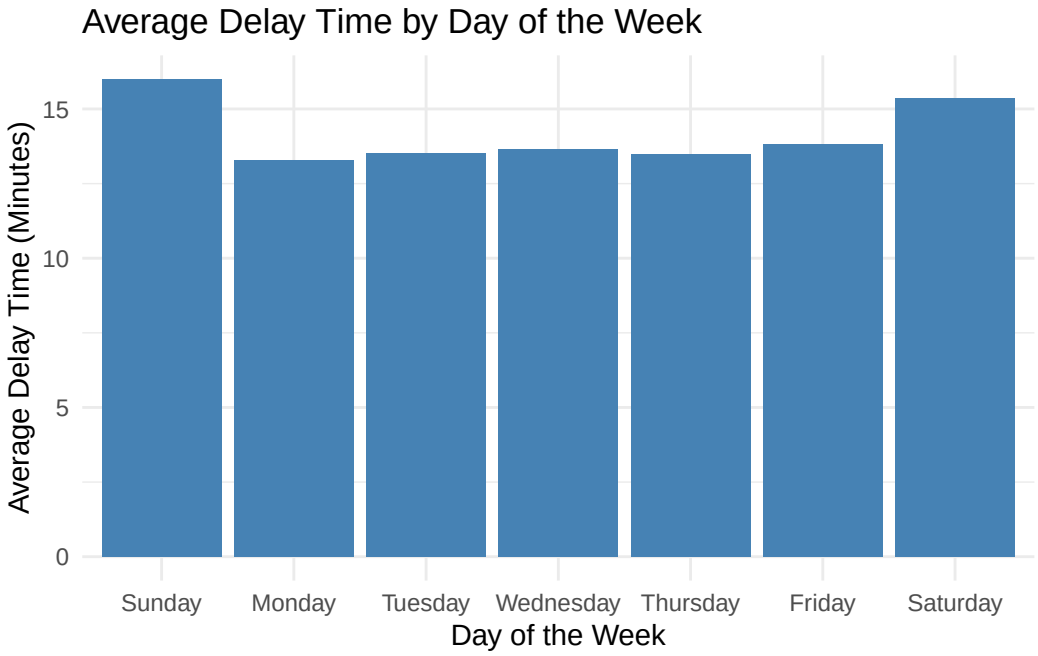

Figure 7: Average Amount of Delay Time by Day

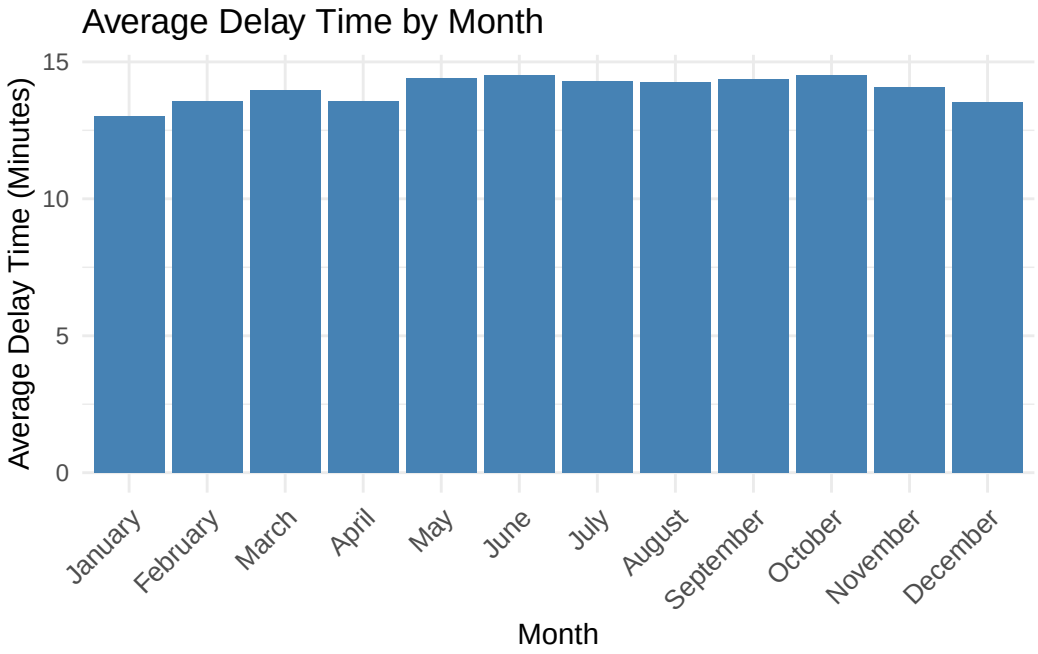

Figure 8: Average Amount of Delay Time by Month

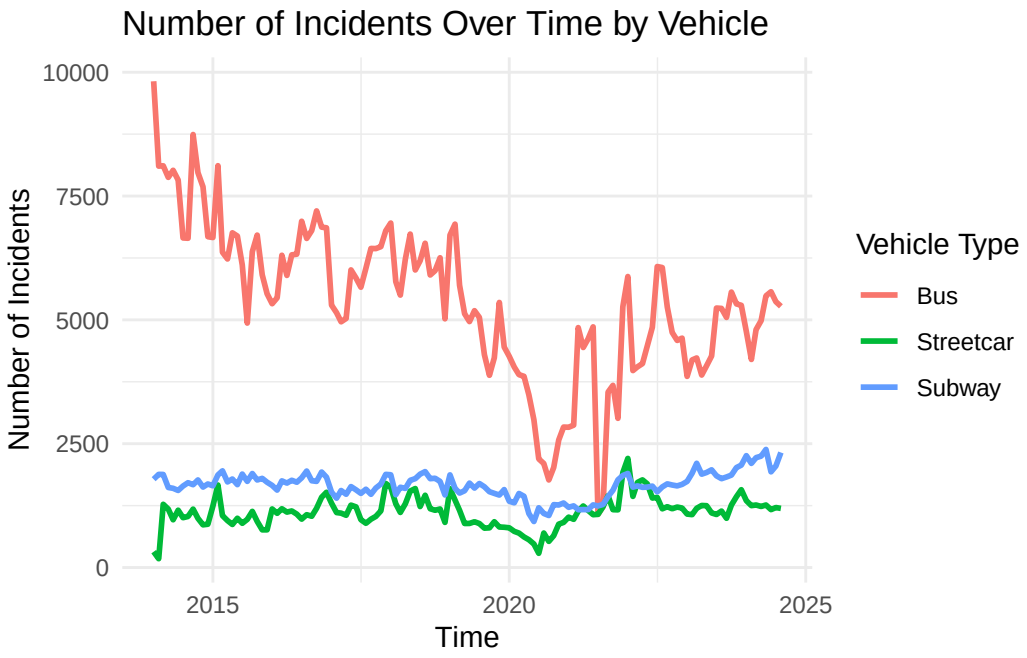

Figure 9: Number of Incidents Over Time

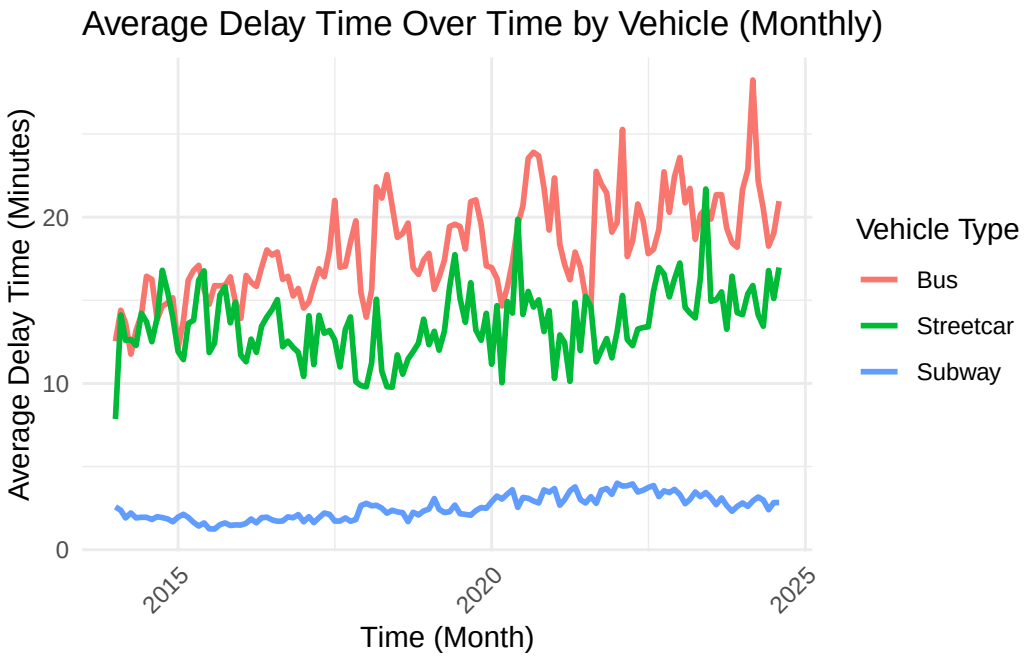

Figure 10: Delay Time from 2014 to 2024

3 Discussion

3.1 First discussion point

3.2 Second discussion point

3.3 Third discussion point

3.4 Weaknesses and next steps

4 References