Analyzing TTC Bus Delay Times and Incident Causes*

Exploring Delay Trends, Daily Patterns, and Potential Improvements in Toronto's Public Transit System

Hanqing Yang

September 27, 2024

This paper analyzes TTC delay patterns throughout the year and week, incident occurrences, and the relationship between delay time and incident type. We found that the first quarter experiences the most delays, with minimal variation across days, mechanical incidents are frequent, but diversions contribute to the longest delays. Our findings emphasize the need for the TTC to focus on mechanical issues and diversions to improve bus delays. The paper is structured into Introduction, Data, Visualization, Results and Discussion sections.

1 Introduction

Mobility affects the daily lives of millions of citizens and the economic productivity and social development of a city. Since it heavily relies on public transportation, increasing the efficiency of transportation is crucial. The Toronto Transit Commission (TTC) serves some 5.5 million people in Toronto, the delay will cause a negative impact for commuters, thus, it is important to study and identify the leading factors of delay.

This paper examines the delay time patterns of the TTC throughout the year and across different days of the week, while also analyzing the frequency of various incidents that lead to delays. Specifically, we explore the relationship between delay times and the types of incidents that occur, aiming to uncover trends that can inform future transit management strategies. Our findings indicate that delays are more frequent in the first quarter of the year, with little variation across the days of the week. While mechanical incidents occur frequently, it is the diversions that cause the most significant delays, underscoring a critical area for improvement in transit operations

*Code and data are available at: https://github.com/Lorina-Y/TTC-Bus-delays.git.

This research is important as it highlights the need for the TTC to prioritize addressing mechanical issues and managing diversions more effectively. By focusing on these areas, the TTC can enhance its service reliability, improve commuter satisfaction, and ultimately foster greater public trust in public transportation.

The paper is organized into four sections: the Introduction provides the background and motivation; the Data section outlines the data sources and methodology; The Visualization section plot the data and relationship between variables; The Results section presents the key findings; and the Discussion interprets the implications of these findings and limitations in analysis

2 Data

2.1 Data Source

Our data of TTC bus delay time in 2024 come from Gelfand (2022), and reference from R Core Team (2023) and Wickham et al. (2019) also be used. The paper study the trend of number of TTC bus daley in 2024, with incident causes, delay duration, direction and location. While similar datasets on TTC bus delay times from previous years exist, I chose to focus on the delays in 2024 because the frequency of incidents can change over time. Analyzing updated data will provide more relevant insights that can better inform future improvements to the transit system.

2.2 Measurement of data

The dataset collects information about each TTC bus delay, including the date, day of the week, incident type, and delay duration in minutes. The date and day are recorded automatically when a delay occurs. The delay duration is measured by clock on each bus. Incidents causing delays are categorized into 12 predefined types and recorded for each delay occurred. These measurement methods ensure the data is accurate and reliable.

2.3 Data variables

The cleaned TTC dataset contains four variables: “Date” which represents the occurrence of each TTC delay, formatted as year-month-day. “Day” which Indicates the day of the week, categorized with six possible values from Monday to Sunday. “Incident” which Identifies the specific incident causing each bus delay, comprising 12 categorical values, including “Vision,” “General Delay,” “Security,” “Operations - Operator,” “Emergency Services,” “Utilized Off

Route,” “Cleaning - Unsanitary,” “Diversion,” “Mechanical,” “Collision - TTC,” “Investigation,” and “Road Blocked - NON-TTC Collision.” and variable “Min_Delay” which Provides the duration of each TTC bus delay in minutes.

2.4 Data table

Below table show a part of the “ttc_data” dataset been used in this paper. The function Zhu (2021) in package Xie (2023) was used for display table.

date	day	incident	min_delay
2024-01-01	Monday	Vision	10
2024-01-01	Monday	General Delay	20
2024-01-01	Monday	General Delay	0
2024-01-01	Monday	Security	0
2024-01-01	Monday	Security	0
2024-01-01	Monday	Operations - Operator	8
2024-01-01	Monday	General Delay	20
2024-01-01	Monday	Emergency Services	21
2024-01-01	Monday	Security	10
2024-01-01	Monday	Utilized Off Route	0

Figure 1: TTC bus delay dataset

3 Visualizations

3.1 Trend of daily TTC Bus delays time

This line graph plot the trend of daily TTC bus delay time in minutes over 2024(Figure 2), from@opendatatoronto.

The x-axis represents the date variable, covering the period from January 1 to August 31, 2024, while the y-axis shows the total duration of bus delays in minutes for each day. The line graph clearly illustrates the trend in delay times over this period, indicating that delays are more pronounced in the first quarter of the year compared to the subsequent months; however, the difference is not substantial. Consequently, we can conclude that there is not much variation in bus delay times over the three-quarters of 2024 analyzed.

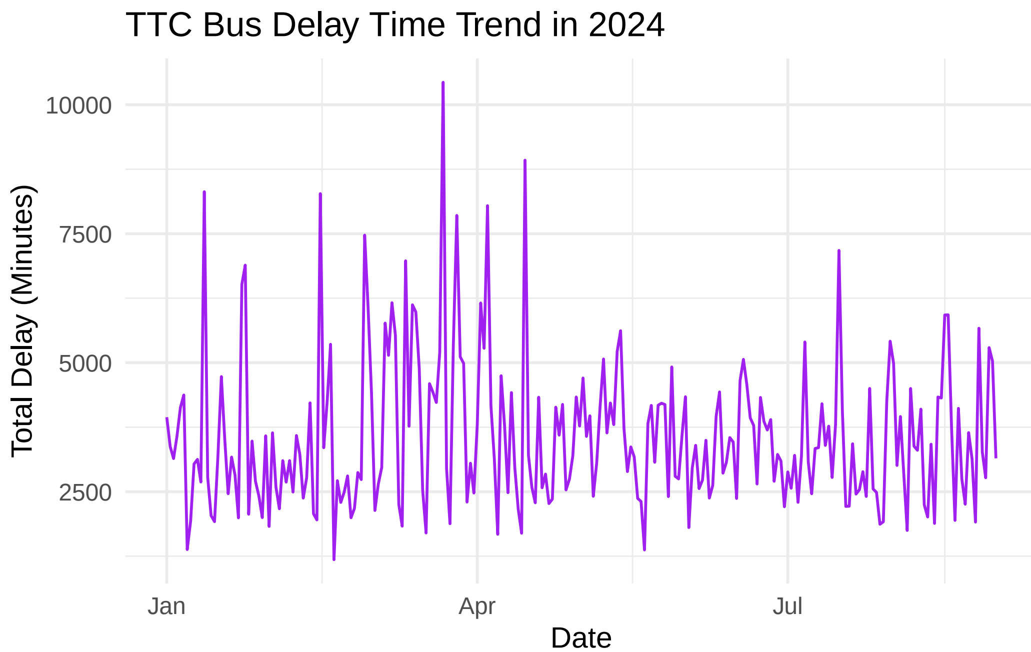

Figure 2: Trend of daily TTC Bus delays time in minutes over 2024

3.2 Pattern of TTC bus delay time across day of the week

This bar chart show the average TTC bus delay time by days of the week over the three-quarters of 2024(Figure 3), data from@opendatatoronto.

In this bar chart(Figure 3), the x-axis represents the days of the week, with each rectangular bar corresponding to a specific day. The y-axis indicates the average delay time for each day of the week throughout 2024. Different colors are used to fill each rectangular bar, clearly illustrating the variations in delay times for each day. However, the average delay times show minimal differences across the days, indicating that TTC delays are not significantly related to the day of the week.

From the two visualization above, We can conclude that the TTC bus delay time not significantly related to time over the year nor days of weeks.

3.3 Occurance of different type of incidents over 2024

Also, to understand the potential causation of bus delay, we need to study the incidents lead to each delay. The bar chart(Figure 4) illustrate the number of occurence of different kind of incidents over 2024, data from@opendatatoronto.

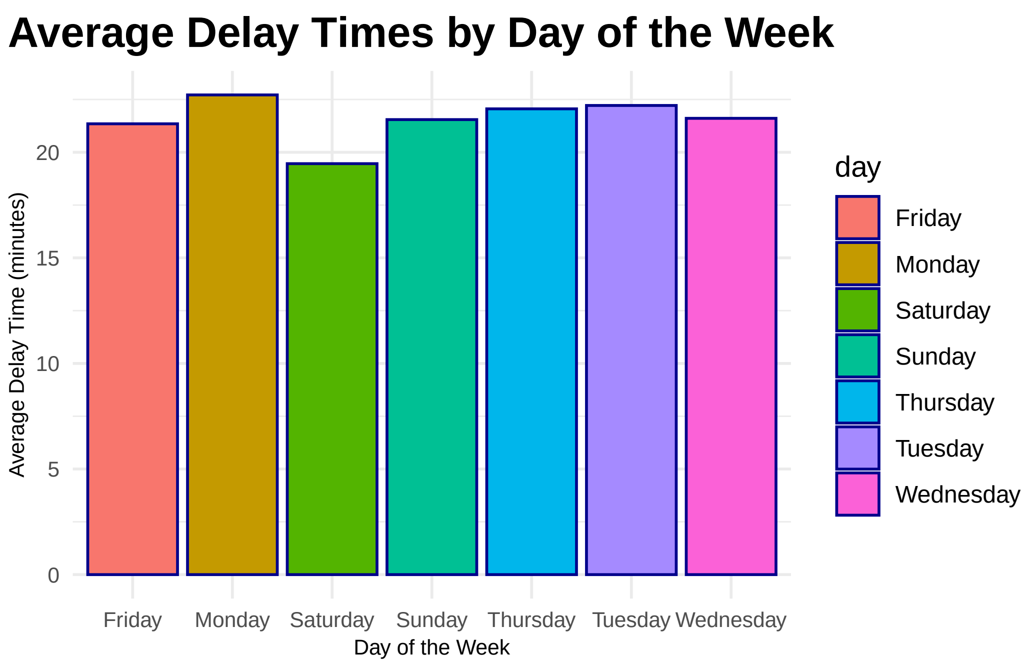

Figure 3: pattern of TTC bus delay time across day of the week.

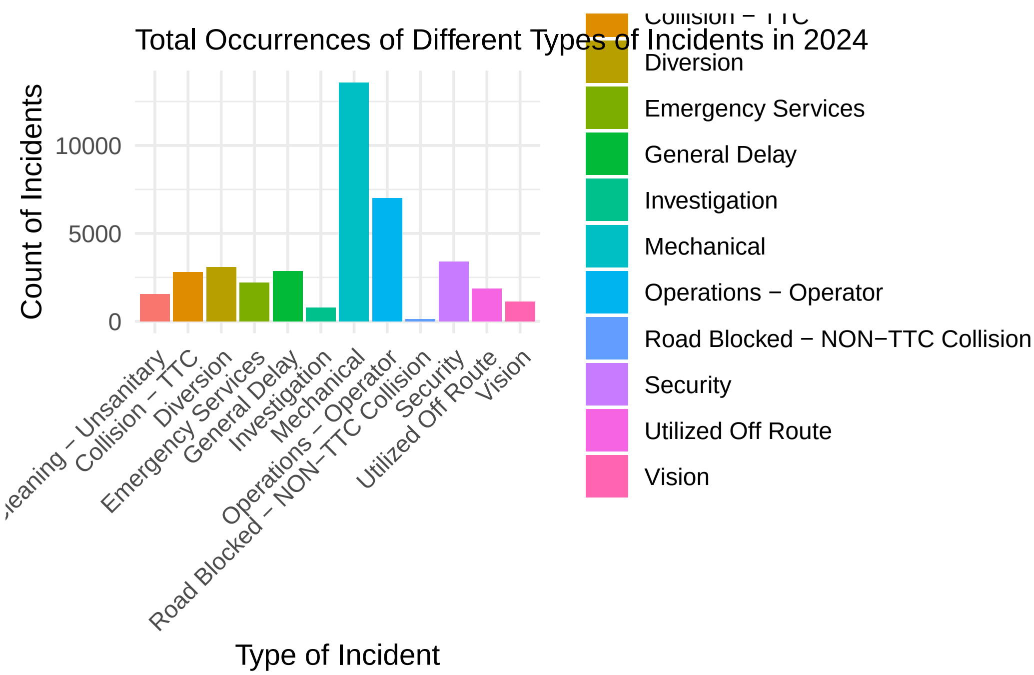

Figure 4: Occurance of different type of incidents over 2024

In (Figure 4), We present all 12 types of incidents that contribute to TTC bus delays throughout the year. The x-axis represents the types of incidents, while the y-axis shows the total count of occurrences for each incident from January 1 to August 31, 2024. The graph illustrates that mechanical incidents occur most frequently, with around 13,000 occurrences during this period, followed by operational issues. In contrast, road blockage incidents are the least likely to occur.

3.4 Relationship between incident and TTC bus delay time

Then, the (Figure 5) illustrate the relationship between incidents and TTC delay time.

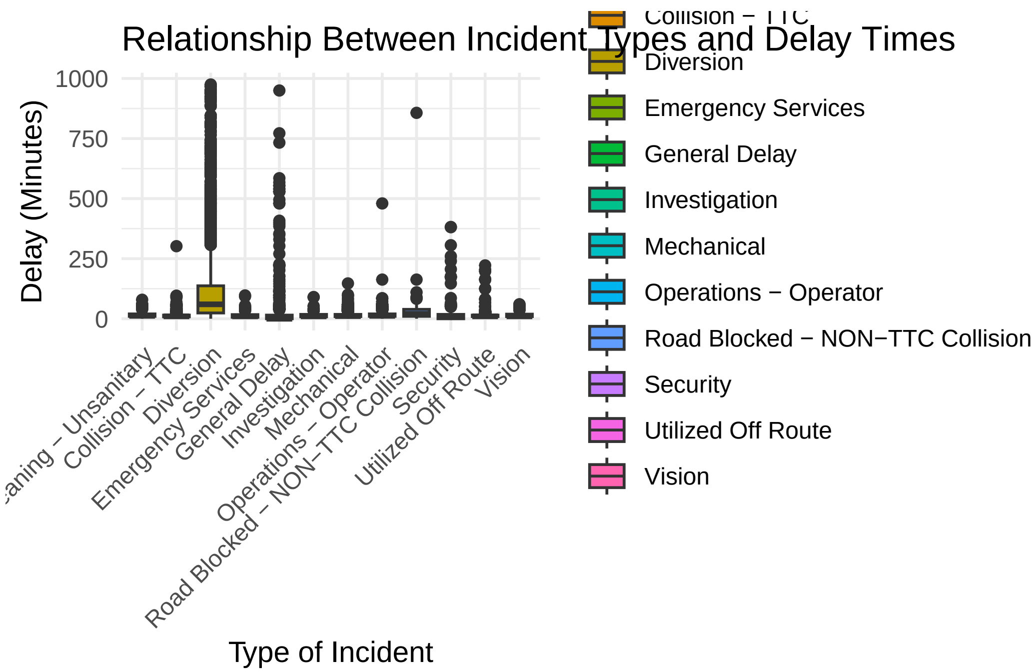

Figure 5: Relationship between incident and minutes of TTC bus delay

The box plot(Figure 5) is utilized to examine which types of incidents are most likely to cause TTC bus delays or lead to longer delay times. A box plot was chosen because it effectively displays the distribution of the dataset, including potential outliers. The x-axis represents the types of incidents, while the y-axis indicates the delay times. From the plot, we observe that diversions exhibit a greater spread and variance in delay times, suggesting that they are more likely to cause bus delays compared to other incidents.

4 Results:

The analysis of TTC bus delays throughout 2024 reveals 4 key findings. First, there is little variation in bus delay times across the three-quarters analyzed, indicating a relatively consistent delay pattern. Second, the delays do not appear to be significantly influenced by the day of the week, suggesting that other factors are more critical in determining service reliability. Third, mechanical incidents are the most common cause of delays, followed by operational issues, while road blockage incidents are the least frequent. Lastly, diversions are associated with longer delays, highlighting their impact on overall service efficiency.

5 Discussion

5.1 Suggested Improvement can be made for reduce bus delay:

Based on the result, the following strategies are recommended for enhance TTC bus service and reduce delays:

1.Focus on Mechanical Maintenance: Since mechanical incidents are the most frequent cause of delays, the TTC should prioritize routine maintenance and implement a proactive maintenance schedule, which could help to identify and address potential issues before they lead to delays.

2.Optimize Operational Procedures: Since operational issues are a significant contributor to delays, the TTC should review and refine its operational guideline. Training programs for staff and clear communication can help to reduce these incidents.

3.Enhance Diversion Management: Since diversions are linked to longer delays, the TTC could develop better diversion management strategies. This includes timely communication with passengers about expected delays and alternative routes, as well as collaborating with local authorities to minimize the need for diversions whenever possible.

4.Keep trace the data and continue evaluate delay pattern regularly and adjust operations accordingly.

5.2 Weaknesses and next steps

One limitation of this paper is that in the analysis for incidents type of time delay(Figure 5) does not demonstrate a causal relationship between incident type and delay time. While the data indicates a correlation suggesting that diversions are linked to longer delays, we cannot definitively conclude that diversions are the cause of bus delays. Further analysis is necessary to validate this point and provide a clearer understanding of the nature of these relationships.

6 References

Gelfand (2022), Müller and Wickham (2022), Xie (2023), Kassambara (2023), Zhu (2021), Firke (2023) R Core Team (2023), Wickham et al. (2019)

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://sharlagelfand.github.io/opendatatoronto/.

Kassambara, Alboukadel. 2023. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. https://rpkgs.datanovia.com/ggpubr/.

Müller, Kirill, and Hadley Wickham. 2022. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.