

Incidents Leading to Time-Consuming TTC Bus Delays in 2024*

The time period covered is January 2024 to August 2024

Aliza Abbas Mithwani

September 24, 2024

This paper discusses the causes for TTC bus delays in 2024 from January until August. We find that although mechanical incidents cause the most frequent delays, diversions cause the most time-consuming delays. People seem to be frustrated by the TTC delays and this is reducing their consumer satisfaction. Uncovering the leading causes for these delays can allow the Toronto Transit Commission (TTC) to target their investigation and solutions.

1 Introduction

In 2019, the Global News reported that although the majority of TTC riders said they are satisfied with the company's service, half have complained of regular delays when using the service. This problem has led a reduction in satisfaction for these people. Riders at rush hour reported being more affected by unanticipated TTC delays. This data was collected by a poll conducted in May 2019, which sampled almost 2,000 riders in Toronto, Calgary, Edmonton, Vancouver and Montreal. (Vuong (2019))

The goal of this research paper is to discover which incidents are causing the most time-consuming TTC bus delays. The data set used in this paper is "TTC Streetcar Delay Data" and specifically for 2024 (the file name is "ttc-streetcar_delay_data-2024.xlsx". This data set was taken from City of Toronto's Open Data Portal and published by the Toronto Transit Commission (TTC). For the sake of simplicity, this paper only focuses on TTC buses as opposed and does not include or draw conclusions about TTC subways and streetcars. Type of incident is not featured in the TTC Subway Delay Data and TTC Streetcar Delay Data included a couple of additional incident types.

*Code and data are available at: <https://github.com/alizamithwani/TTC-Delays-Analysis.git>

The rest of this paper is structured such that we will discuss the data (Section 2) next. More specifically, we will go over where this data comes from (the source), a more comprehensive context for this data, and a description/summary of the data itself.

2 Data

As mentioned earlier, the data set was found on City of Toronto’s Open Data Portal and titled “TTC Bus Delay Data”. When simulating the data we used tidyverse (Wickham et al. (2019)), when downloading the data, we used opendatatoronto and tidyverse (Gelfand (2022) and Wickham et al. (2019)), and when cleaning and testing the data we used tidyverse (Wickham et al. (2019)). Furthermore, we used the statistical programming language R (R Core Team (2023)) to write any code associated with this paper. The data set obtained contained 11 data sets, each corresponding to one of the last 11 years. Each data set has 10 variables: Report Date, Route, Time, Day, Location, Incident, Min Delay, Min Gap, Direction and Vehicle. One specific variable of interest is Incident, which provides a description of the incident that caused the delay. The second variable of interest is Min Delay, which is the delay to the schedule for the next bus (in minutes).

Only the 2024 data set (January-August) was used to analyze the most up-to-date situation of the TTC bus delays. There might be other contributing factors in older years, which would result in reduced applicability of conclusions to the current situation if those contributing factors have changed. The data set was published by the Toronto Transit Commission (TTC), which is the company that provides the bus service in discussion. The Global News reported that half of the riders they sampled said they were dissatisfied by the TTC delays (Vuong (2019)), sparking an interest in why these delays occur and is there a certain incident that is more responsible for these delays compared to rest the rest. Finding an answer to this question can lead to more focused investigation into the problem and more tailored solutions.

Analyzing how frequently each incident has occurred over the last year (Figure 1), we were able to see which incidents occurred more than others. Although this table tells us a lot about which incidents the TTC buses are facing the most, the feelings of dissatisfaction from TTC delays likely have more to do with the amount of time wasted or the lateness of the rider in reaching their destination. Therefore, although it is useful to know that mechanical incidents were the most common cause for TTC bus delays and that the TTC can target working to improve in that department, it does not reveal the most time-consuming cause of TTC bus delays.

To investigate which incidents are causing the most time-costing bus delays, we find the total delay time due to each type of incident as displayed in Figure 2. Here we used the variable Incident for the x-axis and the variable Min Delay for the y-axis, however we converted minutes to hours by dividing by 60. This makes it easier to fit the data in the graph as the sum of total delay time in minutes for each incident was too large. As we can clearly see, this visual tells a different story from Figure 1. In Figure 2 we can see that diversions are, in fact, the

| Incident Type | Number of Incidents |
|----------------------------------|---------------------|
| Cleaning - Unsanitary | 1422 |
| Collision - TTC | 2385 |
| Diversion | 2380 |
| Emergency Services | 1607 |
| General Delay | 2411 |
| Investigation | 613 |
| Mechanical | 12041 |
| Operations - Operator | 5972 |
| Road Blocked - NON-TTC Collision | 117 |
| Security | 2550 |
| Utilized Off Route | 1703 |
| Vision | 1037 |

Figure 1: Total number of incidents in each category

most time-consuming incident and have resulted in the most delay time overall in 2024 so far. Interestingly, the Transit app has had an update recently that allows riders access to live updates regarding TTC detours and diversions (Wilson (2024)). Since this update took place in September and this data set only contains data from January till August, it will be interesting to see if this Transit app update will result in a better experience for riders as they will be able to become aware of diversions more efficiently and can avoid using the TTC bus service if they are in a hurry.

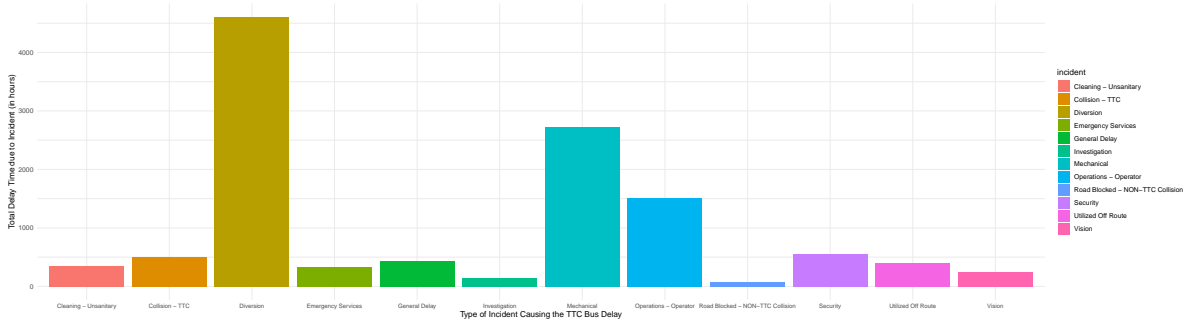


Figure 2: Total Delay Time for TTC Buses by Incident Type (2024 Year-to-Date)

3 Model

3.1 Model set-up

3.1.1 Model justification

4 Results

5 Discussion

5.1 First discussion point

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

B.2 Diagnostics

C References

R Core Team (2023)

Wickham et al. (2019)

Gelfand (2022)

Wickham (2016)

Xie (2022)

Vuong (2019)

Wilson (2024)

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.

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

Vuong, Oriana. 2019. *Most Transit Commuters Satisfied with TTC, but Many Affected by Unexpected Delays: Poll*. <https://globalnews.ca/news/5338014/ttc-service-toronto-transit-commuters-poll/>.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. New York, USA: Springer-Verlag. <https://ggplot2.tidyverse.org>.

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

Wilson, Jermaine. 2024. *Starting Today, TTC Riders Are Getting a New Way to Avoid Commuting Headaches*. https://www.thestar.com/news/gta/starting-today-ttc-riders-are-getting-a-new-way-to-avoid-commuting-headaches/article_9a9cc6f6-6627-11ef-978a-ef7f079ba1ea.html.

Xie, Yihui. 2022. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.