Exploring TTC Delays*

A Multimodal Analysis

Luca Carnegie

January 25, 2024

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

Table of contents

*Code and data are available at: https://github.com/lcarnegie/OpenDataToronto. Thank you to Nescafé coffee, Rajan Maghera, Hannah Yu, and Sehar Bajwa for your love and support - could not have done it without you. No thanks go to Rohan in particular for not pointing out all the R he wanted us to learn that was in his textbook… sigh. I will try the appendix cocktails someday, though. :)

1 Introduction

A key area of improvement in Canadian sustainability is in increasing the use of public transportation in big cities like Toronto. This makes sense, since cars account for about 80% of transportation emissions (Dia et al. 2019). A major drawback preventing the widespread use of public transportation is the existence of delays within transit systems. These delays push citizens toward using less sustainable alternatives like private cars, leading to congestion on highways and consequent increased emissions. The lack of timely transit services have a demonstrated impact on traffic. One study by Anderson et al. estimated a whopping 47 percent increase in highway delay when transit service ceased completely in a big city such as Los Angeles (Anderson 2014). Learning from this example, it is imperative that policy makers understand and fix these delays to improve sustainability in Toronto.

The Toronto Transit Commission (TTC) has been Toronto's transit system since 1921, serving the city population with a variety of transportation methods, particularly subways, streetcars, and buses. The openly available data surrounding the delays of these various modes of transport offers the opportunity to investigate the root causes of these delays and prompt discussion as how to rectify them. Understanding the common causes of a delay, when a delay is typically caused and where it was caused gives the TTC more insight into how to deliver their services to the public more efficiently and effectively, setting an improved example of what an effective transit agency looks like.

Patterns of public transit delay time are found across various genres of transportation. First, the data source and analysis employed are covered in Section Two. Section Three then critically examines the data, delving deeper to arrive at various conclusions, as well as discussing implications and proposing new areas of exploration for this data. The difference between bus, streetcar, and subway delays is shown to be stark. This analysis highlights the need to focus on especially reducing bus delays on several key routes.

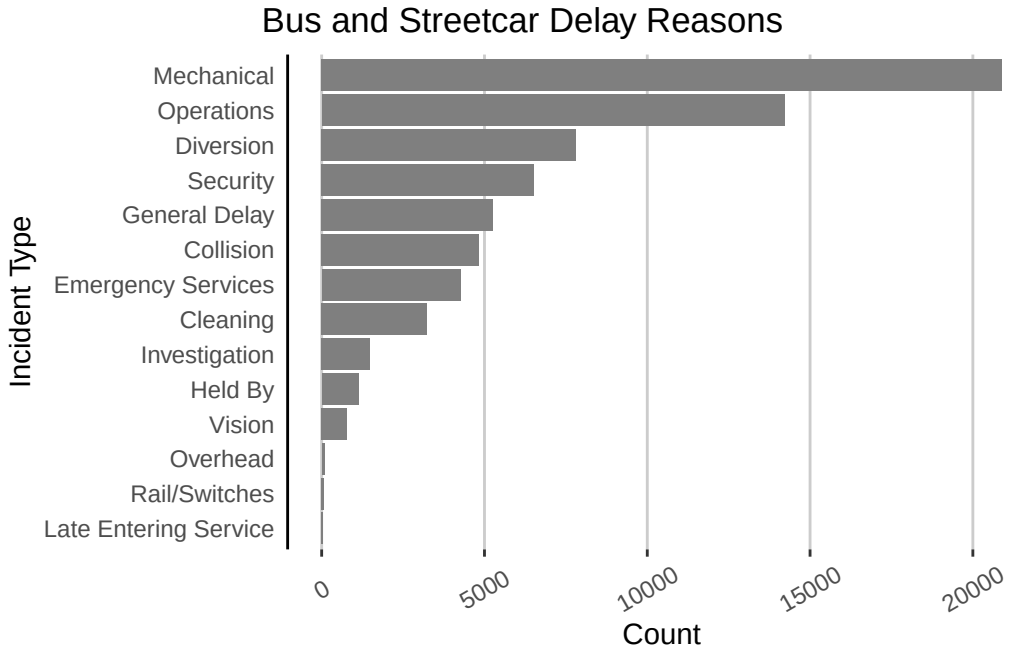
2 Data

To investigate transit delays in Toronto, data on bus (TTC 2024a), subway (TTC 2024c), and streetcar (TTC 2024b) delay incidents for the year 2023 were downloaded using the OpenDataToronto R package (Gelfand 2022) from the Toronto Open Data Catalogue. The data was then cleaned using R (R Core Team 2022), tidyverse (Wickham et al. 2019) and it's associated packages. After combining the datasets together, variables common between datasets and relevant to the analysis were selected, creating a dataset with 93,569 observations. For this analysis, the date (year, month, day), time, day of week, vehicle type, route/line, incident location, cause of delay, and time delay (in minutes) were considered.

Table 1: Sample of Cleaned Delay Data

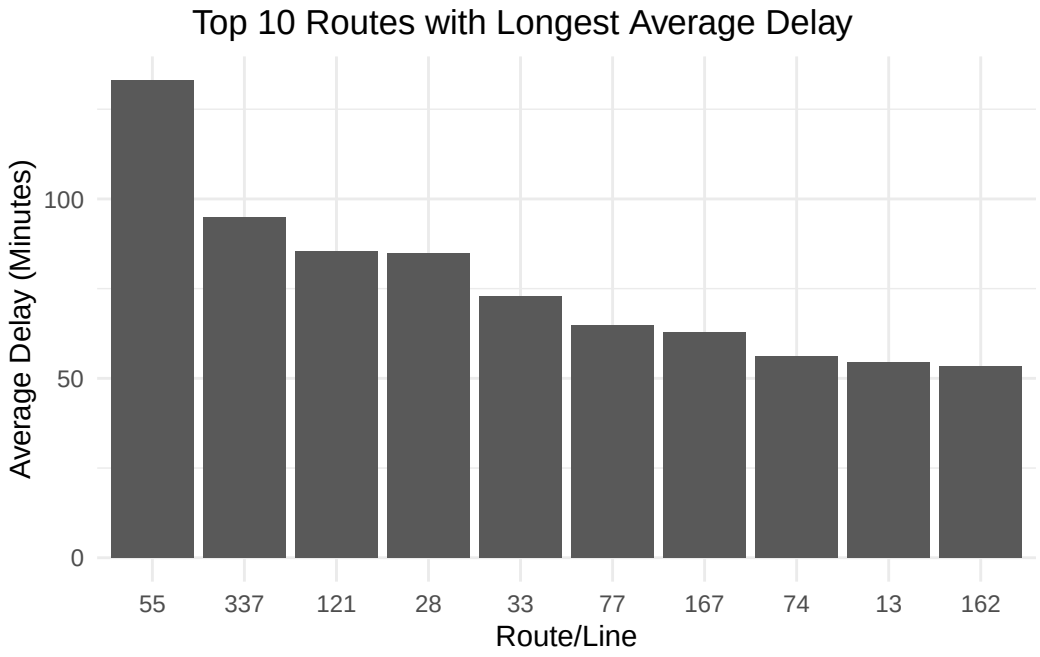
Date	Time	Day	Vehicle	Route/Line	Location	Reason	Delay (Minutes)
2023-01-01	02:30:00	Sunday	Bus	91	WOODBINE AND MORTIMER	Diversion	81
2023-01-01	02:34:00	Sunday	Bus	69	WARDEN STATION	Security	22
2023-01-01	03:06:00	Sunday	Bus	35	JANE STATION	Cleaning	30
2023-01-01	03:14:00	Sunday	Bus	900	KIPLING STATION	Security	17
2023-01-01	03:43:00	Sunday	Bus	85	MEADOWALE LOOP	Security	1

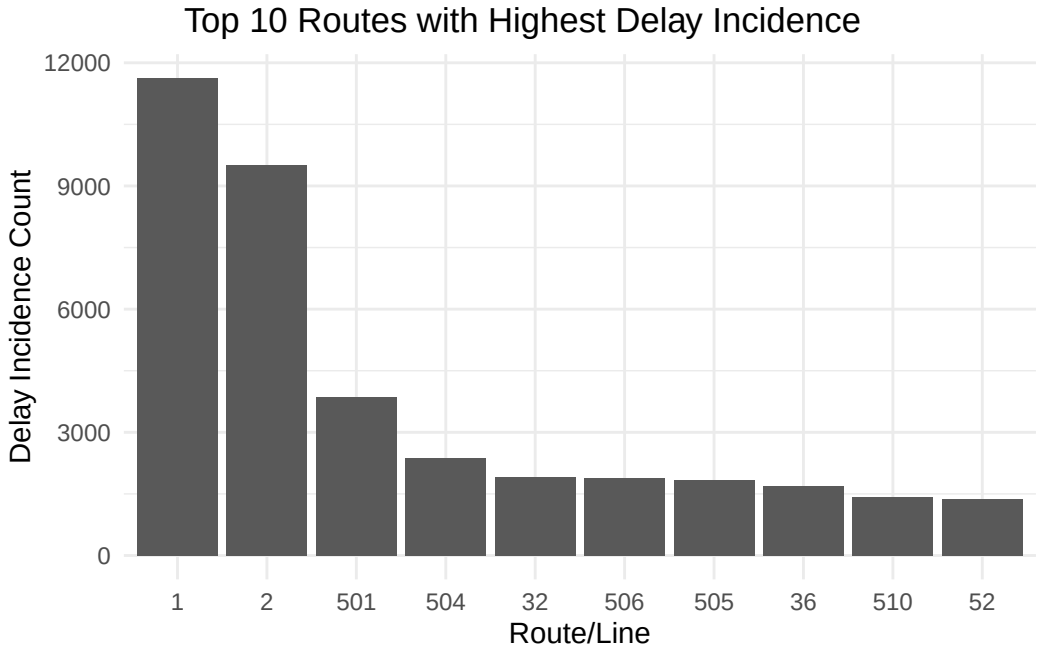
2.1 Incident Type



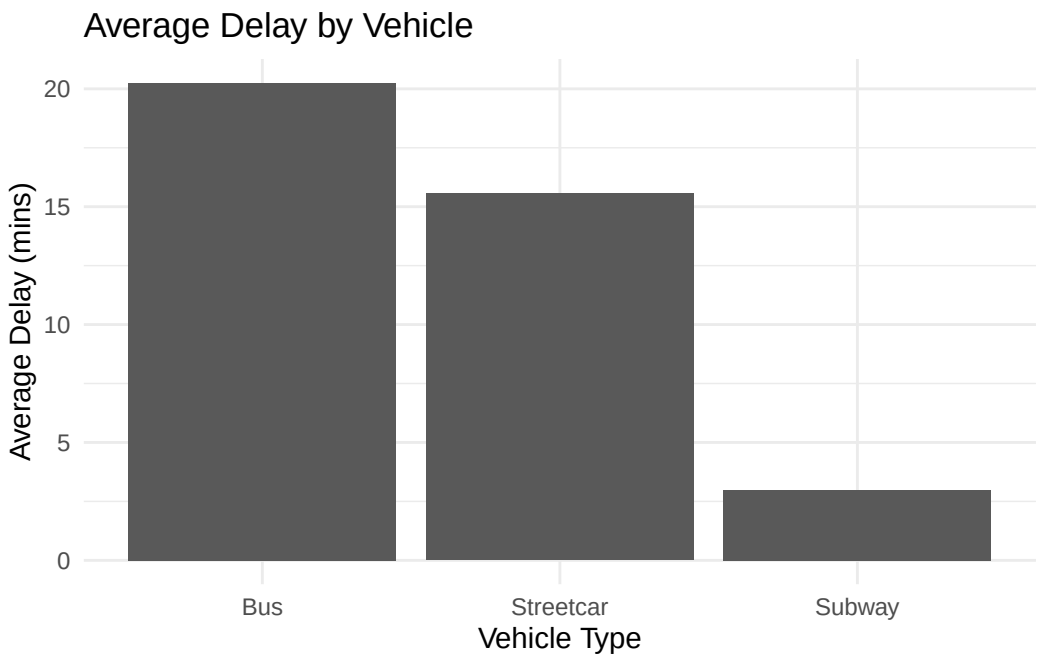
Subways didn't have any reason for delay, yet have the most incidence of delay.

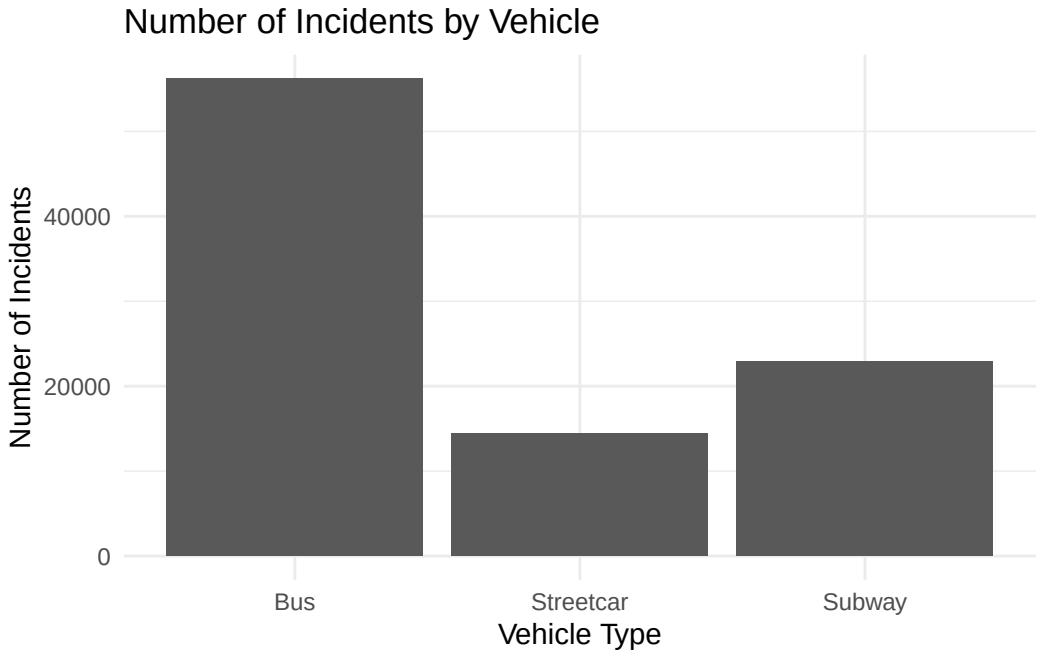
2.2 Line/Route



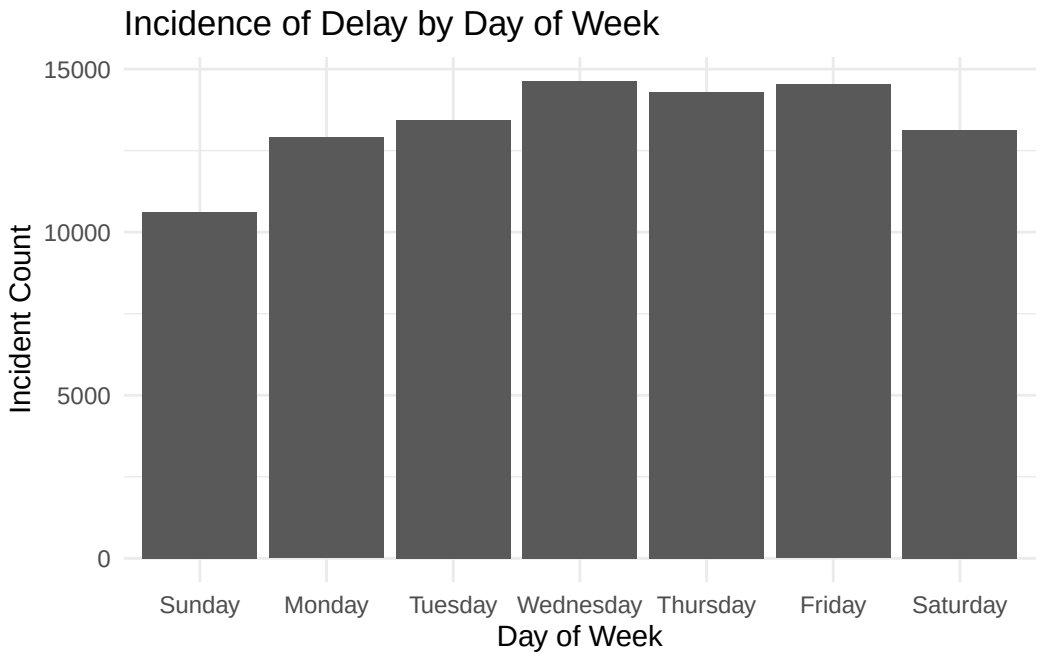


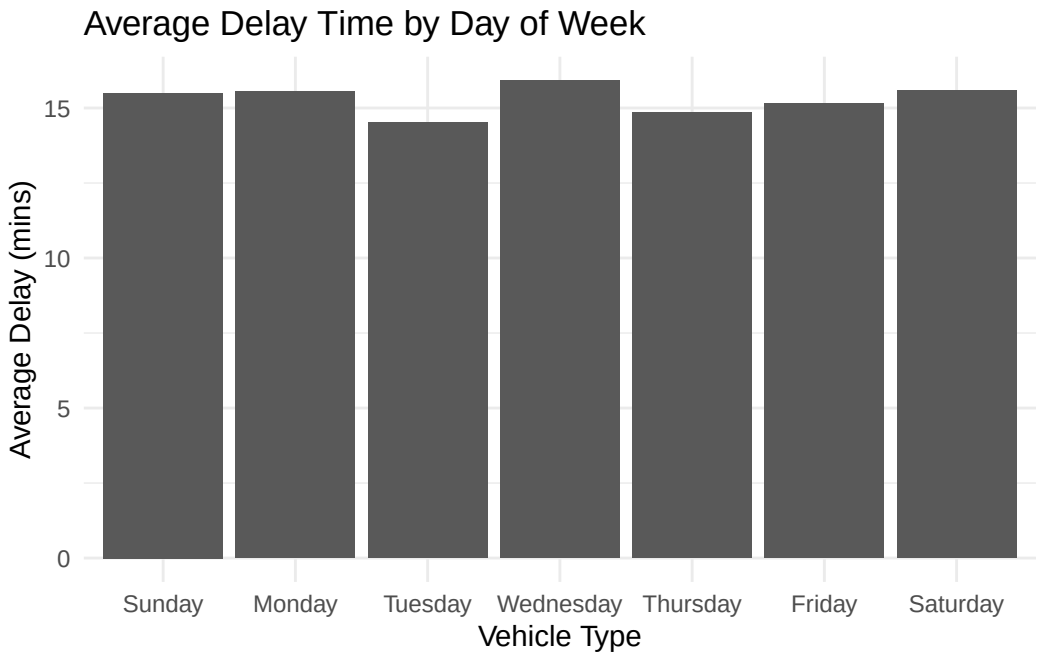
2.3 Vehicle Type



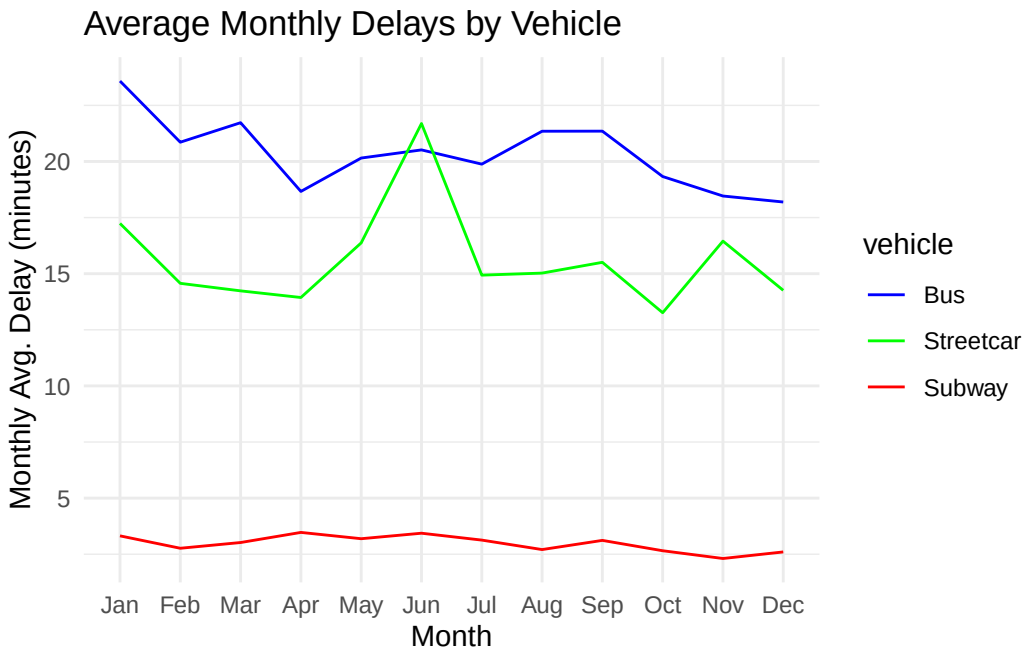


2.4 Day of Week





2.5 Time of Year



3 Discussion

3.1 Results

3.2 Implications

3.3 New Areas of Exploration

4 Conclusion

5 References

Appendix: TTC Route Nomenclature

Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review* 104 (9): 2763–96. https://doi.org/10.1257/aer.104.9.2763.

Dia, Hussein, Michael Taylor, John Stone, Sekhar Somenahalli, and Stephen Cook. 2019. "Low Carbon Urban Mobility." In, 259–85. https://doi.org/10.1007/978-981-13-7940-6_14.

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://CRAN.R-project.org/package=opendatatoronto.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

Martin, Layla, Michael Wittmann, and Xinyu Li. 2021. "The Influence of Public Transport Delays on Mobility on Demand Services." *Electronics* 10 (4). https://www.mdpi.com/2079-9292/10/4/379.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

TTC. 2024a. "TTC Bus Delay Data." https://open.toronto.ca/dataset/ttc-bus-delay-data/.

———. 2024b. "TTC Streetcar Delay Data." https://open.toronto.ca/dataset/ttc-streetcar-delay-data/.

———. 2024c. "TTC Subway Delay Data." https://open.toronto.ca/dataset/ttc-subway-delay-data/.

Wickham, Hadley. 2011. "Testthat: Get Started with Testing." *The R Journal* 3: 5–10. https://journal.r-project.org/archive/2011-1/RJournal_2011-1_Wickham.pdf.

———. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.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.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.