Bring More Buses: Transit Delays in the Toronto Transit Commission*

An Exploratory Analysis with Open Data Toronto

Luca Carnegie

January 26, 2024

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

Table of contents

*Code and data are available at: https://github.com/lcarnegie/TTCTransitAnalysis. Thank you to Nescafé coffee, Rajan Maghera, Hannah Yu, and Sehar Bajwa for your love and support - I could not have done it without you. Special thanks go to Rohan Alexander for the immense amount of R he taught me in such a short time and his extremely easy-to-navigate textbook (that has a cocktail appendix, fortunately).

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). However, a major drawback preventing the widespread use of public transportation is the existence of delays within the transit system. 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 ceases completely in Los Angeles (Anderson 2014), a large city like Toronto. It is then imperative that policy makers understand and fix these delays to improve sustainability.

The TTC, or Toronto Transit Commission, has been Toronto's transit system since 1921, serving the city population with mainly subways, streetcars, buses, along with smaller-scale transport offerings. The openly available data surrounding the delays of these various modes of transport offers opportunity to investigate the root causes of the longer-lasting delays. Delays are likely the best way to assess a transit agency's performance from both the rider and management perspective, since riders' experience is typically determined through how low waiting and travel times are, while performance of staff is assessed "in terms of on-time performance standards (or adherence to schedules)" (Diab, Badami, and El-Geneidy 2015). Both of these metrics are highly correlated with delay time. Therefore, understanding what caused a delay, when it was caused and where it was caused could give the TTC more insight into how to deliver their services to the public more efficiently and effectively.

This paper finds relationships and patterns of public transit delay time across various genres of transportation. The paper is structured in that Section Two covers the data source, and analysis employed. Section Three critically examines the data, comes to several conclusions and proposes new areas of exploration of this dataset. The difference between bus, streetcar, and subway delays is shown to be stark. This analysis highlights the need to focus on especially on reducing bus delays on several key routes to improve efficiency and effectiveness of service.

2 Data Analysis

To investigate transit delays in the TTC, recorded data on bus (TTC 2024b) , streetcar (TTC 2024c), and subway (TTC 2024d) delay incidents for the year 2023 were obtained using the OpenDataToronto R package (Gelfand 2022), providing data from the Toronto Open Data Catalogue. All the datasets were current as of January 18th, 2024.

All the data was then cleaned using R (R Core Team 2022) and it's tidyverse (Wickham et al. 2019), janitor (Firke 2023), and lubridate (Grolemund and Wickham 2011) packages. Variables common between datasets and relevant to this analysis were selected, creating a combined dataset with 93,569 observations made over the year. For this analysis, the date (year, month, day), time, day of week, vehicle type, route/line, incident location, cause of delay, and time delay gap (in minutes) were made of interest. A sample of the cleaned dataset is shown in Table 1. In the analysis itself, the knitr (Xie 2023) and ggplot2 (Wickham 2016) R packages were used to create the tables and graphs.

Table 1: Delay Incidence by Month

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

The dataset, primarily categorical, was analyzed on a variable-by-variable basis. Each variable, barring 'delay reason', was examined in two ways: through its relationship with delay occurrences and its impact on the average delay duration, measured in minutes.

2.1 Reason of Delay

First, we assess the various recorded reasons of a particular delay on the TTC and their impacts. In Figure 1, we observe several key groupings of delay reasons for buses and streetcars. Before analyzing further, however, it is of important note that the reasons for subway delay were not recorded during data collection, subways must be omitted from this particular chart.

In Figure 1, Mechanical Issues, Operational Issues, and Route Diversion emerge as the top three causes of delays in streetcars and buses across the TTC. The incidence of these delays varies significantly. In particular, there 6682 more mechanical delays than operational delays and 13083 more mechanical delays than diversion delays.

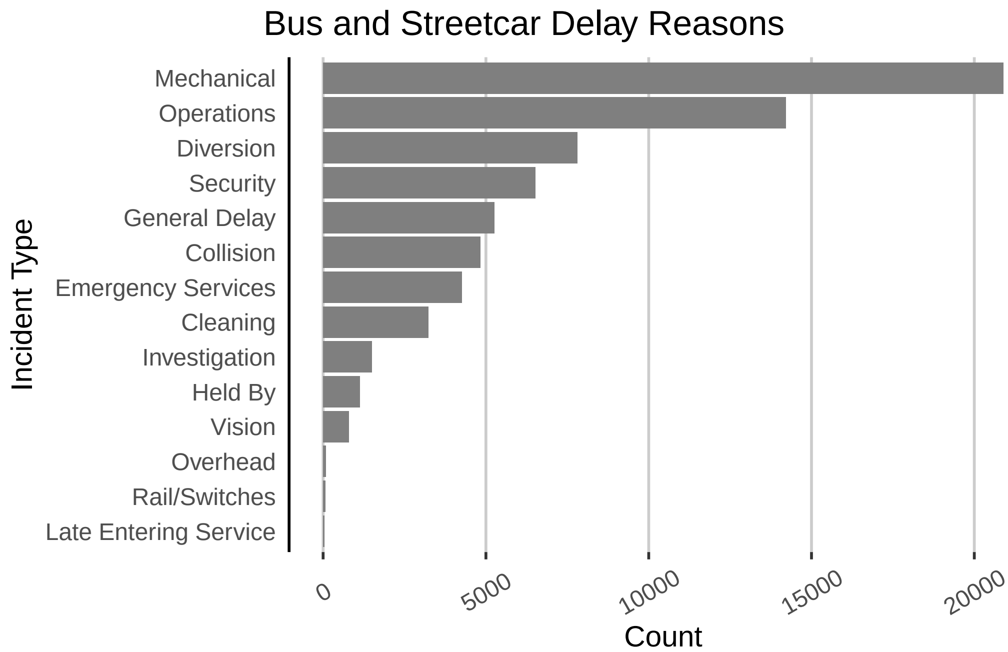


Figure 1: Distribution of Bus and Streetcar Delay Reasons

2.2 Day of Week

Next, the variation of delay incidents by day of the week was analyzed. Figure 2a shows the number of delay incidents for each day.

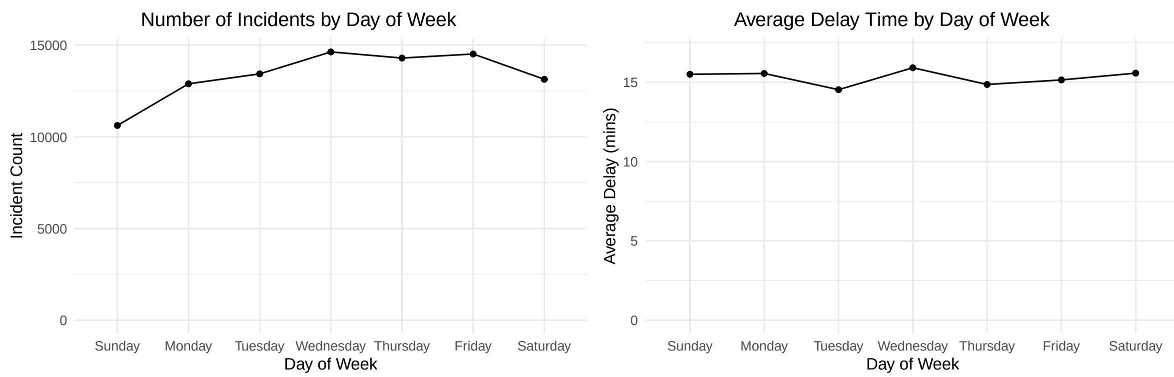
The frequency of delays climbs steadily from Sunday to Friday before peaking on Saturday. However, there does not appear to be a clear upward or downward trend in delays over the course of the week. Further analysis on a weekly basis would be needed to determine if there is a statistically significant pattern in delays based on the day. However, the data indicates overall that weekends, particularly Saturdays, experience more delay incidents than weekdays.

Next, we looked at average delay duration by day. Figure 2b displays the average delay time in minutes for each day. The average remains relatively consistent across the week, hovering around 15 minutes per delay incident.

2.3 Vehicle Type

Next, we graph the data to check for noticeable differences across different genres of transportation. We see a noticeable difference in delays by vehicle in Figure 3

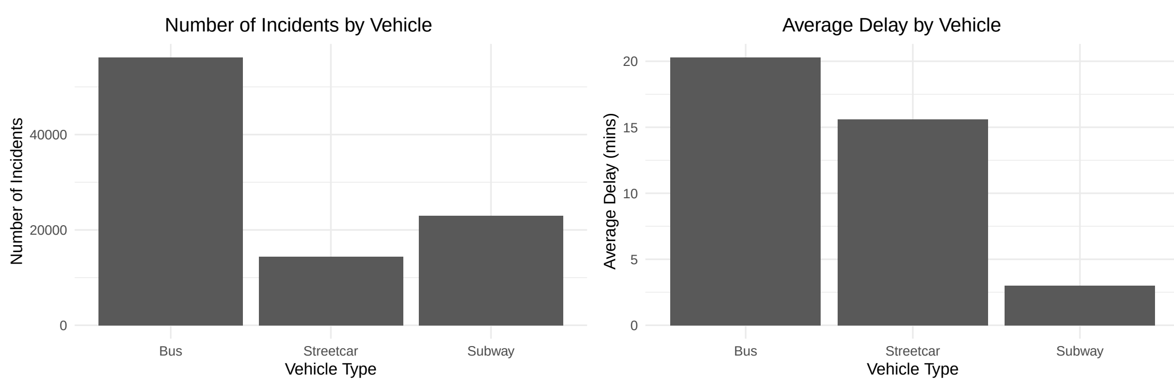
Figure 3a shows that in 2023, buses had 56207 recorded delay incidents in 2023. This is more than twice as many as subways and close to three times more than streetcars.



(a) Incidences

(b) Average Delay Time (mins)

Figure 2: Delays, aggregated by Day of Week



(a) Number of Incidences

(b) Average Delay Time

Figure 3: Delays for each mode of transportation

In Figure 3b, however, we do not see as large a difference in average delay times across different genres of transportation, however. Buses had the longest average delay times, at around 20 minutes, but subway had only 3 minutes of delay time on average, making for a fairly sizeable gap delay times. Interestingly, although there was less incidence of streetcar delay than subway delay, streetcar routes tended to be delayed longer on average throughout 2023.

2.4 Time of Year

To investigate how TTC delays change throughout the year, Figure 4 and Figure 5 were constructed. The graphs show monthly changes in delay count and average delay time over 2023. Buses, streetcars, and subways are represented by blue, green, and red lines, respectively.

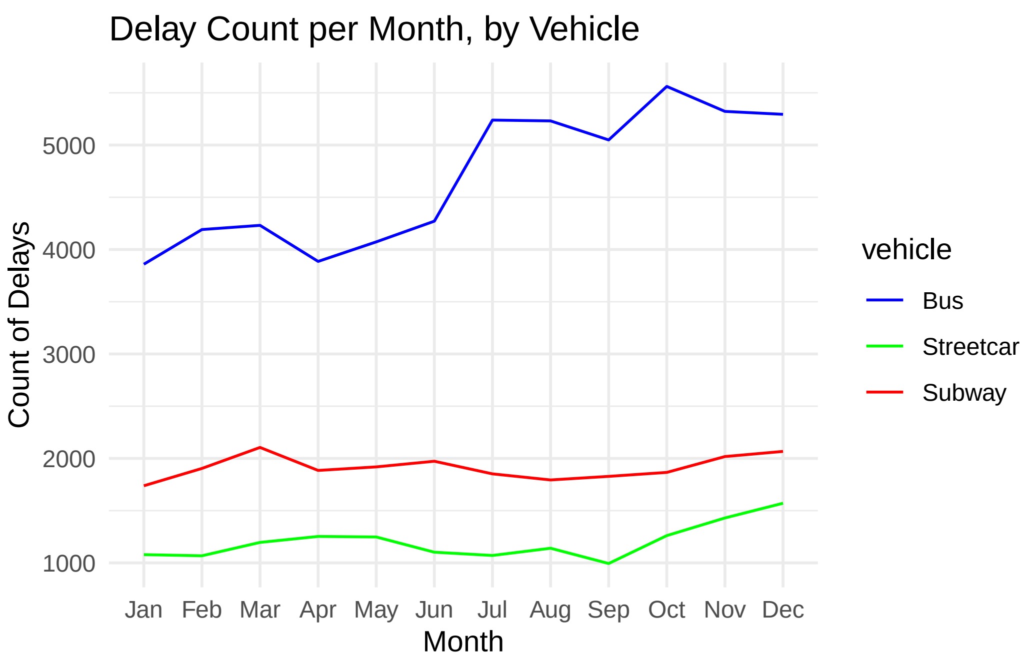


Figure 4: Delay Incidence by Month

In Figure 4 we see how the count of delays changes throughout 2023 for each mode of transportation. Buses have the highest incidence of delays throughout year with the count generally trending upward as the year progresses, peaking in November with 5322 incidents. Incidence of subway and street car delays are significantly lower, but also trend upwards as the year progresses. Overall, this figure presents similar findings to Figure 3 in section 2.3 - buses tend to have more incidences of delay than both streetcars and subways overall.

When looking at changes in average delay times throughout 2023, there are several observations we can make. With buses, we observe a slight downward trend in delay time as the year

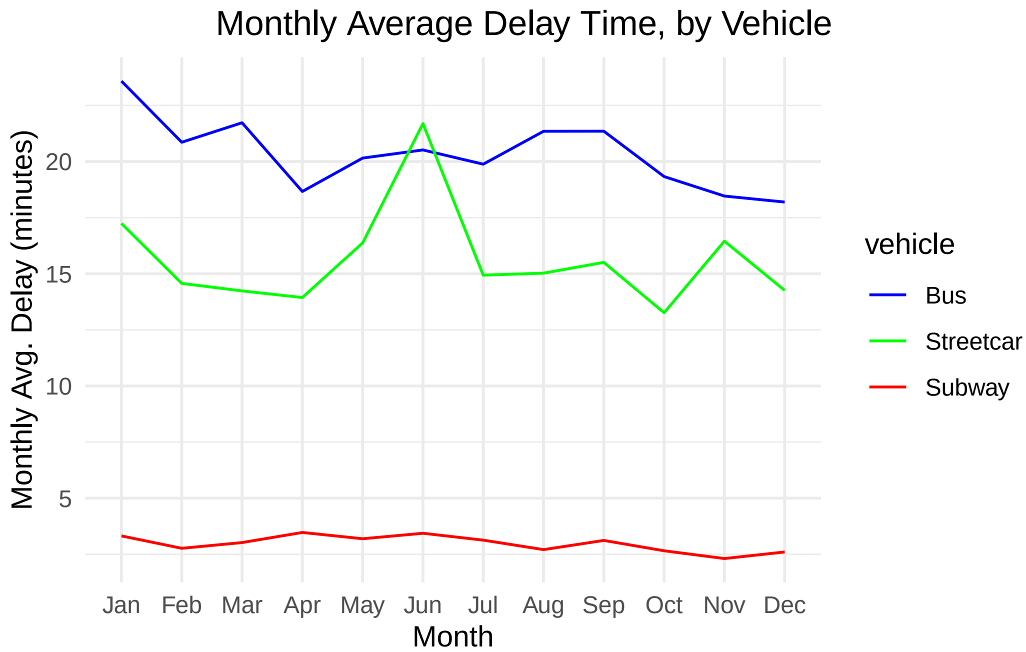


Figure 5: Average Delay Time, by Month

progresses, with delays going from an monthly average of 23.6 minutes to 18.2 minutes by the end of the year. Subway delay times stay more or less constant, on average, at around 3.2 minutes, the lowest average delay time of all three. With streetcars, however, although delay times tend to trend downward throughout the year, there is a noticeable change in delay times in the summer, with a spike in delay time in the spring and summer months that eclipses even buses in the month of June.

2.5 Line/Route

Finally, the delays on particular subway lines and bus/streetcar routes were examined. Given the large number of different transit routes offered by the TTC, this analysis focuses on the 10 routes with the highest average delay times and highest incidence of delay. These findings are illustrated in Figure 6 and Figure 7. A short guide to the route number nomenclature of the TTC is provided in the Appendix for the reader's convenience.

We observe a mix of bus, streetcar, and subway routes accounting for the highest incidences of delay. Interestingly, subway Lines 1 (Yonge-University) and 2 (Bloor-Danforth) occupy the top two positions in terms of delay frequency, with 11628 and 9511 recorded incidents, respectively. Following closely are Streetcar Lines 501 (Queen St.) and 504 (King St.), ranking third and fourth. Rounding out the top five is Bus Route 32 (Eglinton West), which holds the fifth position in delay incidence, with 1920 recorded incidents.

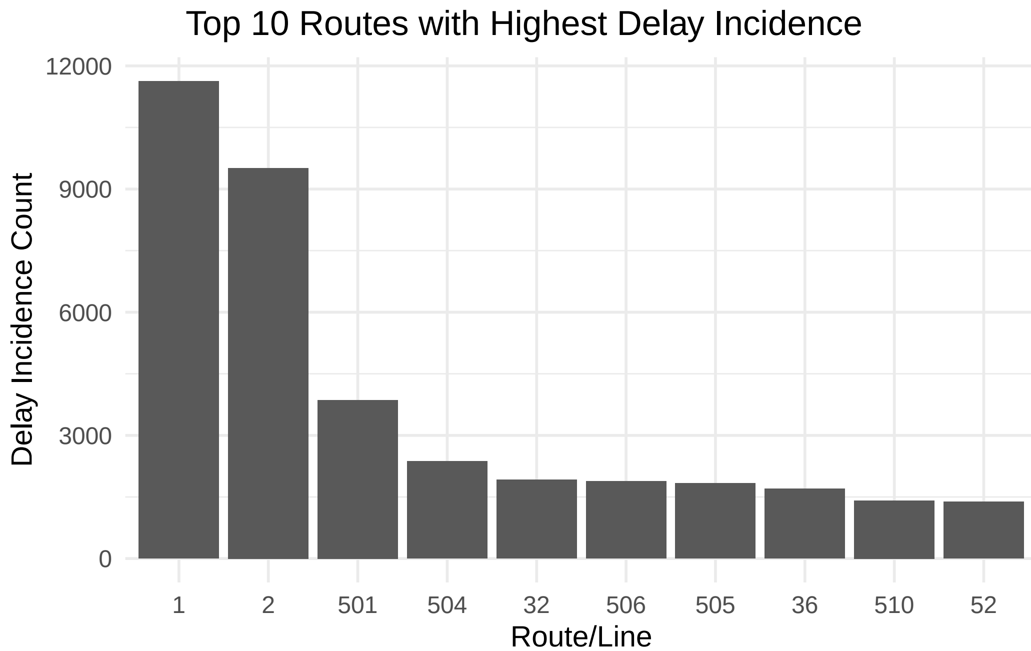


Figure 6: Routes/Lines with Highest Delay Incidence

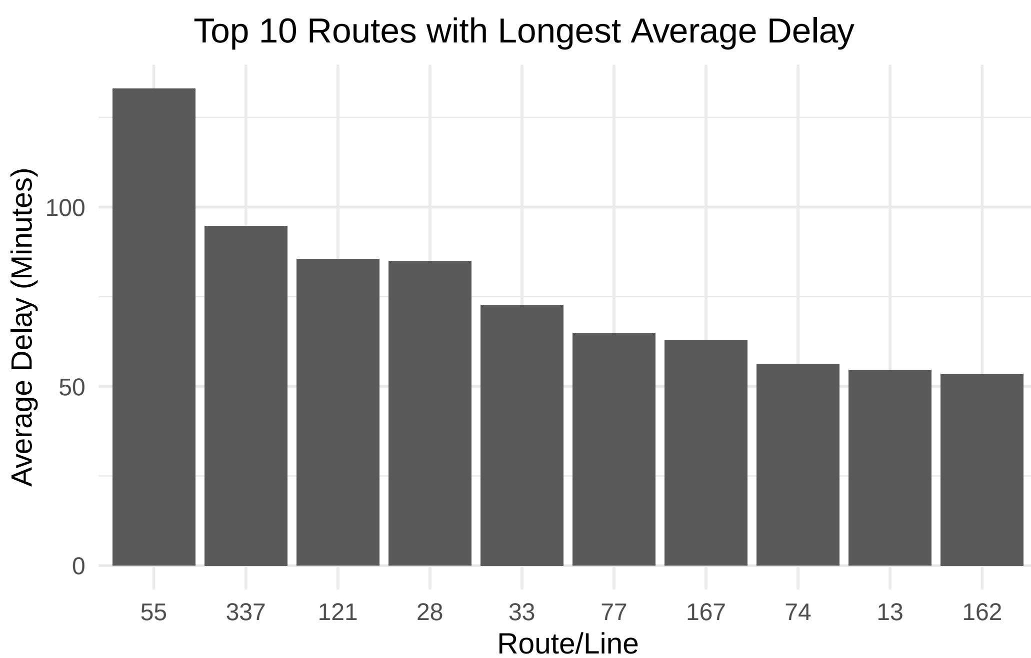


Figure 7: Routes/Lines with Highest Average Delay Time

However, the picture changes entirely when routes are looked at in terms of average delay time in Figure 7. Here, we see that bus routes entirely make up the 10 routes with the longest average wait times, ranging from 53.4 minutes for Route 162 (Lawrence-Donway) and up to 133.0 minutes for Route 55 (Warren Park). These routes were checked for accuracy in the TTC's route finder (TTC 2024a)).

3 Discussion

3.1 Results and Implications

The analysis of Toronto Transit Commission delays reveals several key findings and implications that could be useful in the public agency's operations. The analysis found that buses experience the longest delay times among all modes of transport, suggesting a dire need to reduce bus waiting times across the board, but particularly on the routes analyzed in section 2.5. This could be accomplished in several ways, such as deploying more buses on these routes, or using tactics normally reserved for rail service like bus bridging, where buses are dispatched from scheduled services to act as temporary shuttles (Itani et al. 2019). However, it is important to consider the relative passenger volume of these routes relative to the system as a whole. Perhaps there is such a high delay that the benefit of employing these tactics against the cost to other services provided by the TTC is simply not worth it to the agency.

Interestingly, delay incidents tend to occur more frequently on weekends, indicating a need for better trained and/or more response staff on call during those time to address these issues. Given the sensitive nature of some incidents, enhanced personnel training could potentially prevent certain situations.

Subways, despite being a popular mode of transport, have the highest incidence of delay. This could point to underlying issues in subway service that warrant further investigation. As said in Section 2, more data is need to investigate this further, and a good first would be to record the reason for a particular delay when a delay is recorded in the dataset.

The most common reasons for delay are mechanical issues, operations, and route diversions. This suggests that subway trains might need to be replaced, and staff could benefit from improved training to reduce delays and flexibility when dealing with route diversions and traffic.

In Toronto's warmer months, April-June particularly, streetcar delay times tend to spike significantly , possibly due to increased pedestrian and cyclist traffic in the downtown core, where streetcars are predominantly located. Although delays during that time are unavoidable, it could be worth exploring if training programmes and/or other strategies are being effectively deployed so that streetcars are running efficiently.

3.2 Further Areas of Exploration

Future analyses could be more comprehensive with better data on the reasons for subway delays. Additionally, given more time and resources to explore additional tools and R packages, location could be another factor of analysis within the scope of TTC delays. This would allow us to identify which neighbourhoods experience the most delays, thereby helping to prioritize maintenance crew dispatch and policy decision with respect to adding more vehicles and service to a particular area.

Using data-driven methods to analyze the the TTC's performance of can enable the agency to make decisions that target key routes that need the most assistance, allowing public funds to be used in the most efficient way possible. This approach not only improves the efficiency of the TTC but will enhance the commuting experience for the public that depends on it every day to conduct their business in the ci

4 Appendix

Route nomenclature has gone through many changes throughout the history of the TTC. As James Bow, a TTC historian, said to BlogTO, "[The current system] is a kludge of several numbering schemes that were added to or replaced elements of the original scheme, with many elements grandfathered in". Hence, he says "there's very little rhyme or reason" to the transit numbering, making it confusing to almost everyone. (Flack 2023)

Luckily, in the same interview, conducted in late 2023, he provides a general overview of TTC nomenclature that can provide helpful context when understanding some of the graphs in this paper. The numbering is as follows:

- 1-4 - Subway and LRT routes
- 5-139 - Regular bus routes, assigned willy-nilly
- 140-159 - Premium express bus services
- 160-189 - Regular bus routes, assigned willy-nilly
- 300-399 - Night Bus and Streetcar services
- 400-499 - Accessible Community Bus services
- 500-599 - General Streetcar routes.
- 900-999 - Express bus routes

References

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.

Diab, E., M. Badami, and A. El-Geneidy. 2015. "Bus Transit Service Reliability and Improvement Strategies: Integrating the Perspectives of Passengers and Transit Agencies in North America." *Transport Reviews* 23 (3): 292–328.

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

Flack, Derek. 2023. "Here's the Secret Meaning Behind TTC Route Numbers." *blogTO*. blogTO. https://www.blogto.com/city/2015/11/whats_the_meaning_behind_ttc_route_numbers/.

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/.

Itani, Alaa, Aya Aboudina, Ehab Diab, Siva Srikukenthiran, and Amer Shalaby. 2019. "Managing Unplanned Rail Disruptions: Policy Implications and Guidelines Towards an Effective Bus Bridging Strategy." *Transportation Research Record* 2673 (4): 473–89. https://doi.org/10.1177/0361198119838838.

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.ca*. Toronto Transit Commission. https://www.ttc.ca/routes-and-schedules/.

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

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

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

Wickham, Hadley. 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.

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