Choosing TTC Subway Service Provides the Most Time Stability during Weekday Rush-Hours*

Analysis on the Delay Magnitude of different TTC Services

Sandy Yu

September 23, 2024

This paper analyzes the magnitude of TTC delays in 2023 based on service types. TTC delay data for buses, streetcars, and subways are obtained through Open Data Toronto data portal and data is used to visualize delay magnitude during different days and hours throughout the year. We find that, although TTC streetcars offered the least number of delays in 2023, riding on TTC subways provides the least average time of delay during weekday rush-hour traffic. Our results suggest that, if the time required to arrive at one's office is the same for all TTC service options, choosing subway over other services could provide the most time stability.

Table of contents

*Code and data are available at: https://github.com/Jingying-yu/ttc_delay.git

1 Introduction

Public transportation serves as the transportation backbone for city residents, facilitating the daily commute of millions of people. For office workers, students, and the elderly, selecting the appropriate TTC service is crucial in maintaining a reliable and efficient schedule. Despite the importance of this choice, there has been limited effort to evaluate the reliability of different TTC services to guide commuters'choices.

Publicly accessible TTC delay data can be tracked back to 2014 on the Sharla Gelfand (2022) portal, yet a comprehensive visualization of how each type of service – bus, streetcar, and subway -— performs in terms of delay time remains narrowly explored. This paper addresses this gap by using data from the TTC delay data to analyze and visualize delay patterns across different services.

This paper specifically focuses on TTC bus, streetcar, and subway delay data from 2023, extracting and comparing them to identify which service type experienced the least amount of delay during weekday rush hours. Our findings indicate that, in 2023, the TTC subway system consistently demonstrated the shortest average delays.

Understanding these patterns is crucial for helping commuters make informed choices about which TTC service to use, especially when they have equal access to different types of stations. By providing this analysis, the paper aim to assist readers in optimizing their travel choices and enhancing their overall commuting experience.

Analyses and findings in this paper are structured into several sections: Section 2 – Data, Section 3 – Results, and Section 4 – Discussion. The Data section introduces the context of the datasets and examines all datasets and variables kept for analysis. The Result section focuses on visualizing and presenting the data results in Data section. The Discussion section further evaluate the implications behind the data results presented in the previous section, expands the topic beyond pure data, and touches on any weaknesses and next steps.

2 Data

Where the data comes from - All data used for and throughout the paper are extracted from Open Data Toronto Portal (Sharla Gelfand 2022). TTC Bus Delay data (Commission 2024), TTC Streetcar Delay (Transit Commission 2024), TTC Subway Delay (Toronto 2024) and TTC Subway Delay Codes (Toronto 2024) are retrieved to compare the delay magnitudes of different services. Data is cleaned and analyzed using the open source statistical programming language R (R Core Team 2022) and supporting packages tidyverse (Wickham et al. 2019), janitor (Firke 2023), dplyr (Wickham, François, et al. 2023), hms (Müller and Wickham 2021), ggplot2 (Wickham, Chang, et al. 2023), patchwork (Pedersen 2023), kableExtra (Zhu 2023), and knitr (Xie 2023).

For all delay dataset, only data in 2023 were chosen. There are three main reasons for this choice:

1. 2023 data is the most recent full-year data and offers the best reference for current (2024) delay patterns.

2. COVID-19 impacted TTC operating capacity and styles in other years close to 2024.

3. data from earlier years (before 2020) were structured very differently.

Therefore only 2023 delay data were chosen in this paper.

2.1 TTC Bus Delay

source of data (brief) structure of data & important attributes

why are these attributes kept for analysis (measurement)

Table 1: Sample of Cleaned TTC Bus Delay Dataset

Date	Line	Time	Day	Station	Cause of Delay	Min Delay
2023-01-01	91	02:30	Sunday	WOODBINE AND MORTIMER	Diversion	81
2023-01-01	69	02:34	Sunday	WARDEN STATION	Security	22
2023-01-01	35	03:06	Sunday	JANE STATION	Cleaning - Unsanitary	30

2.2 TTC Streetcar Delay

source of data (brief) structure of data & important attributes

why are these attributes kept for analysis (measurement)

Table 2: Sample of Cleaned TTC Streetcar Delay Dataset

Date	Line	Time	Day	Station	Cause of Delay	Min Delay
2023-01-01	509	02:37	Sunday	QUEENS QUAY AND SPADIN	Operations	0
2023-01-01	505	02:40	Sunday	BROADVIEW AND GERRARD	Held By	15
2023-01-01	504	02:52	Sunday	KING AND BATHURST	Cleaning - Unsanitary	10

2.3 TTC Subway Delay

source of data (brief)

2.3.1 Delay Code

purpose of extra dataset how it is related to subway_delay

Table 3: Sample of TTC Subway Delay Code

Code	Code Description
EUAC	Air Conditioning
EUAL	Alternating Current
EUATC	ATC RC&S Equipment

2.3.2 Delay Data

structure of data & important attributes

why are these attributes kept for analysis (measurement) mention “code” column

Table 4: Sample of Cleaned TTC Subway Delay Dataset

Date	Time	Day	Station	Delay Code	Min Delayed	Line
2023-01-01	02:22	Sunday	MUSEUM STATION	MUPAA	3	YU
2023-01-01	02:30	Sunday	KIPLING STATION	MUIS	0	BD
2023-01-01	02:33	Sunday	WARDEN STATION	SUO	0	BD

3 Results

3.1 Highest Number of Total Delayed

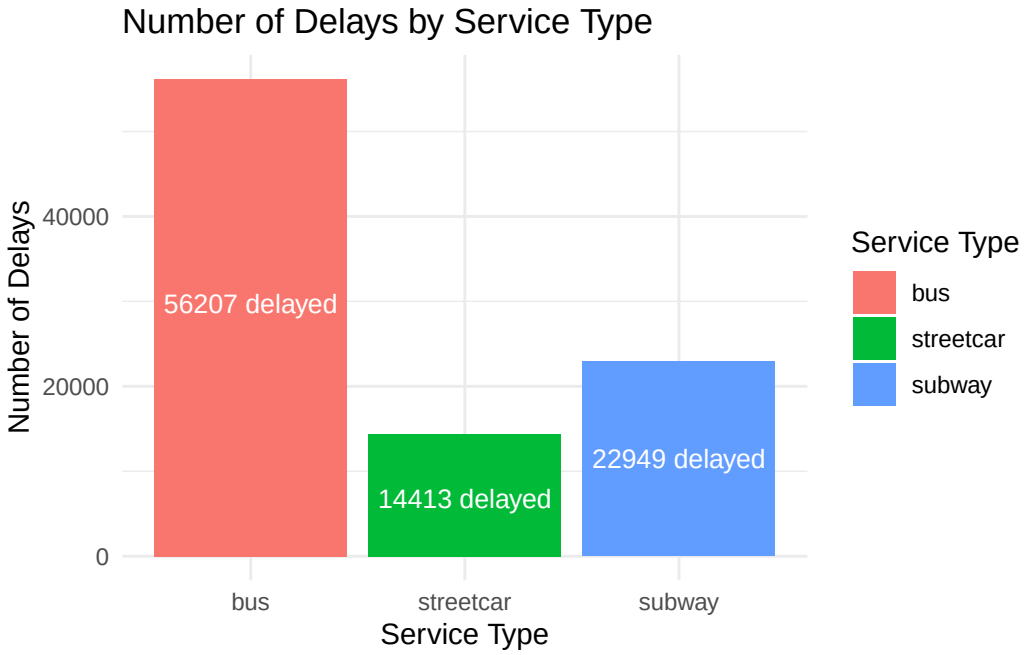

Figure 1: Visualization to show the TTC service that has the highest number of delays in year 2023

3.2 Most Delayed during Weekday

Define Rush Hours: Rush Hours: • Morning Rush-Hour: Approximately 7:00 AM to 9:00 AM • Evening Rush-Hour: Approximately 4:00 PM to 7:00 PM

Present Results During Rush Hours & at Off-Peak Hours

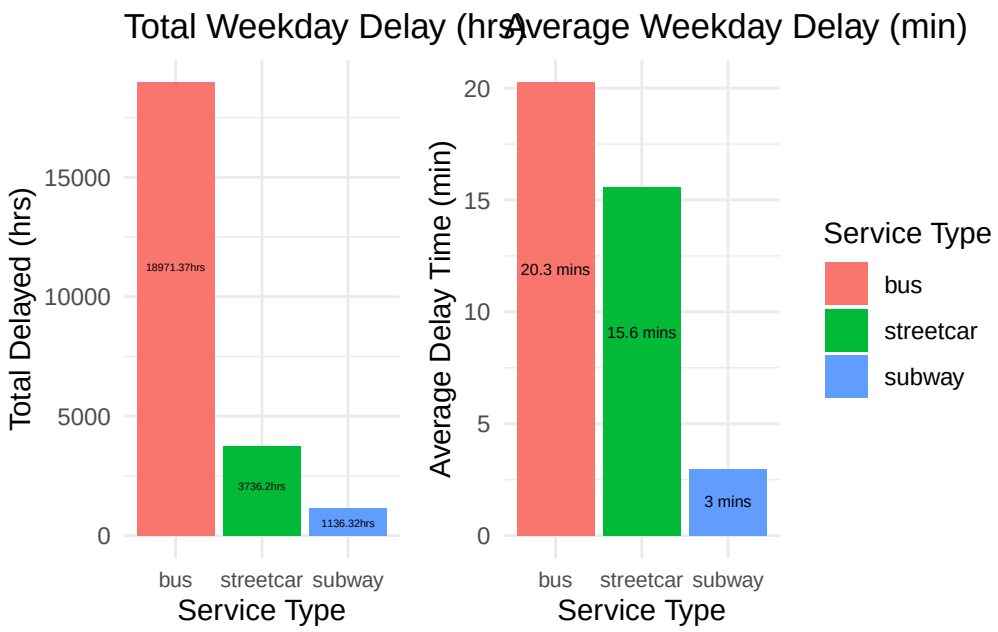

Figure 2: Visualization to show the total delayed time during weekdays in 2023

3.3 Most Delayed during Rush-Hours

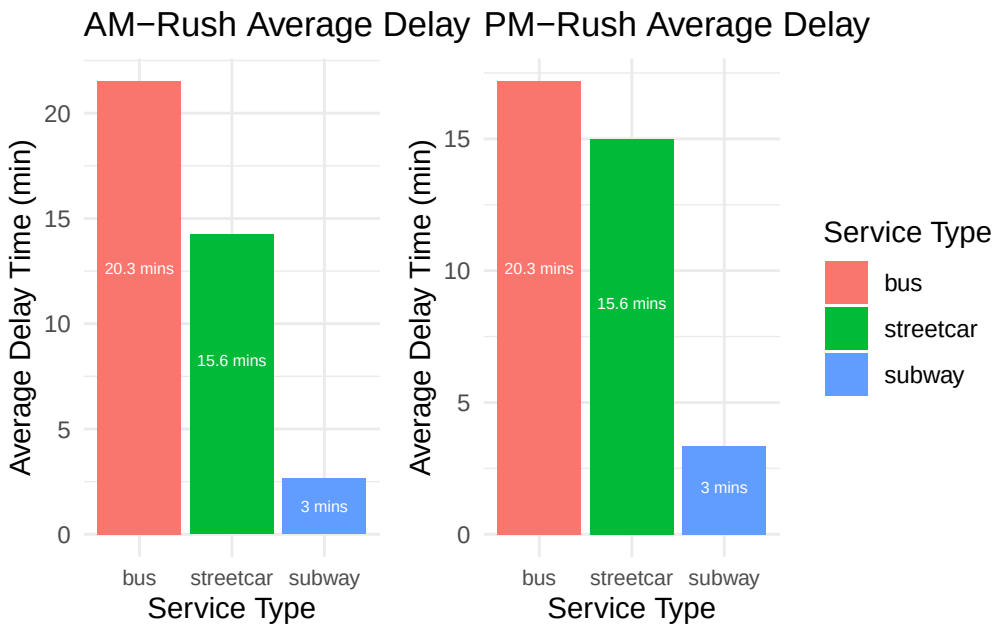

Figure 3: Barplot used to visualize the longest average delay time for different TTC services during Rush-Hours and Off-Peak Hours

3.4 Most Popular Delay Reasons

- 1 table for all three services
- separate into rush & off-peak hours

Table 5: Top 5 Most Popular Reason for Delay

Bus	Streetcar	Subway
Mechanical	Operations	Disorderly Patron
Operations - Operator	General Delay	Injured or ill Customer (In Station) - Transported
Security	Security	Passenger Other
Collision - TTC	Mechanical	Passenger Assistance Alarm Activated - No Trouble Found
Diversion	Diversion	Injured or ill Customer (In Station) - Medical Aid Refused

3.5 Most Popular Delayed Lines

- 1 table for all three services
- separate into rush & off-peak hours

Table 6: Most delayed Lines for each TTC Service

bus	streetcar	subway
121	506	Line 3

4 Discussion

4.1 TTC Service Prices

- one-fare

4.2 Importance of TTC Services

- capacity of TTC daily TTC service capacities
- rush-hour traffic

4.3 TTC Service Amenities

- The PATH (for financial district)
- AC in Summer & Winter

4.4 Errors in analysis, Limitations and Next Steps

- only included 2023 data, should include more years for fairer inspection
- unable to distinguish between on-time rate for different district
- next step: explore GO services (GO trains)

5 Conclusion

References

Commission, Toronto Transit. 2024. "TTC Bus Delay Data." https://open.toronto.ca/dataset/ttc-bus-delay-data/.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Müller, Kirill, and Hadley Wickham. 2021. *Hms: Pretty Time of Day*. https://hms.tidyverse.org.

Pedersen, Thomas Lin. 2023. *Patchwork: The Composer of Plots*. https://patchwork.data-imaginist.com.

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

Sharla Gelfand, City of Toronto. 2022. "Opendatatoronto: Access the City of Toronto Open Data Portal." https://open.toronto.ca.

Toronto, City of. 2024. "Toronto Subway Delay Data." https://open.toronto.ca/dataset/ttc-subway-delay-data/.

Transit Commission, oronto. 2024. "TTC Streetcar Delay Data." https://open.toronto.ca/dataset/ttc-streetcar-delay-data/.

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, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, and Hiroaki Yutani. 2023. *Ggplot2: Elegant Graphics for Data Analysis*. https://ggplot2.tidyverse.org.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org.

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

Zhu, Hao. 2023. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://cran.r-project.org/package=kableExtra.