Analysis of TTC Subway Delays in 2023: Uncovering Trends and Insights*

Zijun Meng

January 26, 2024

Subway plays an important role in many Toronto citizens' life, each delay may disturb them a lot. By analyzing the data of TTC subway delays, I found that most delays are because of customers' misbehavior or vehicles' problems. This may help TTC find why their subways delay a lot and consequently improve their service. It can also contribute to the safety of passengers.

1 Introduction

The subway system, as a part of the Toronto Transit Commission (TTC) service, plays an important role in the daily lives of commuters in Toronto . According to TTC (2024b), on average, the number of customer boardings on weekdays in 2022 is 746,200, its efficient functioning may benefit hundreds of thousands of people's life and contribute to the economy of the city. Subway delays can lead to substantial economic costs in terms of lost productivity and time, while also impacting the daily well-being and reliability experienced by thousands of commuters. Given its critical importance, understanding why subway delays is essential to improve the system.

To analyze the data uploaded by TTC (2024a), I worked on R (R Core Team 2022), using the tidyverse (Wickham et al. 2019), opendatatoronto (Gelfand 2022), and gitcreds (Csárdi 2022) packages. Tables in the report were created with kableExtra (Zhu 2021) and figures were created with ggplot2 which was built in tidyverse (Wickham et al. 2019). This paper first presents an analysis of the patterns of subway delays, followed by an exploration of the most frequent causes. We then discuss the implications of these findings for TTC operations and passenger safety. By studying on the causes, we learned that there are several factors not only make the subway late but also threats the safety of passengers, it is necessary for TTC to take actions.

---

*Code and data are available at: https://github.com/FrankMengZJ/TTC-Subway-Delay

Table 1: The first five rows from the cleaned TTC metro delay data

Date	Time	Day	Station	Code	Min_Delay	Line
2023-01-01	02:22:00	Sunday	MUSEUM STATION	MUPAA	3	YU
2023-01-01	08:02:00	Sunday	VAUGHAN MC STATION	TUNOA	6	YU
2023-01-01	08:11:00	Sunday	GREENWOOD STATION	MUNOA	5	BD
2023-01-01	08:15:00	Sunday	KEELE STATION	TUNOA	5	BD
2023-01-01	08:16:00	Sunday	GREENWOOD STATION	TUNOA	5	BD

2 Data

The TTC Subway Delay Data TTC (2024a) records the information of each delay they faced including time, location and reason which can be used to analyze and consequently find some insights to reduce the problems. Table 1 shows a sample of the cleaned data set. Variables "Date", "Time" and "Day" state when the delay happens, variable "Station" shows where the delay heppens, variable "Code" indicates what caused the delay, variable "Min_Delay" records how many minutes the subway delayed, for example, 'MUPAA' stands for Passenger Assistance Alarm Activations, and finally "Line" shows which subway line delayed, "YU" represents Line 1, BD stands for Line 2. Although TTC (2024a) provided the data from 2018 to 2023, the reason why I only work on the 2023 is the subway service was significantly influenced by COVID-19 from 2020 to 2022, while prepandemic data was too old, the problems may be different from what we have now. In this case, I will only focus on the post pandemic era, 2023.

Firstly, I analyzed the number of delays by line. As Figure 1 shows, majority of delays happened in YU (Line 1) and BD (Line 2), as they are the busiest subway lines in Toronto, this is not surprising.

There are 161 kinds of cases that caused the delay in 2023, among them, the top 16 kinds of incidents led to 6015 out of 8234 of all the delays, as they are the most important ones. Table 2 showed the number of cases happened and the mean of how long they delayed.

Table 2 showed that Disorderly Patrons causes the most delays, while Unauthorized at Track Level, which causes the longest delays among the most frequent caues in average, and Assault / Patron Involved are important factors as well. These kinds of delays usually come from people's misbehavior, and potentially may harm other customers as well as themselves. For the subways' efficiency and customers' safety, I suggest TTC focus more on the order in stations, maybe considering hiring more staff to decrease the number of such incidents.

Passengers mistriggering alarms is also a problem worth considering, apart from making the trains late, they may make other customers panic and may cause harm to them, TTC can consider improve how users use the alarm. What's more, the problems of trains themselves are also critical, there are more than 300 delays caused by door problems, which is possible to
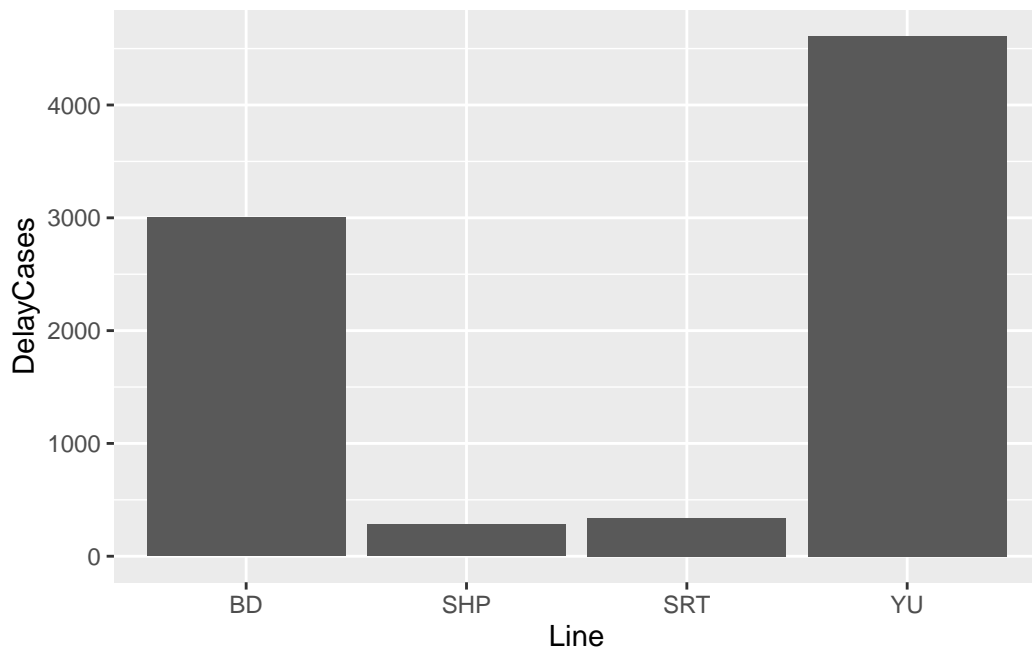
Figure 1: Delay by line

Table 2: Delay by reason

Code	Mean_Delay	n
Disorderly Patron	7.49	1229
Passenger Assistance Alarm Activated - No Trouble Found	4.18	736
OPTO (COMMS) Train Door Monitoring	5.19	664
Passenger Other	11.04	465
Injured or ill Customer (On Train) - Medical Aid Refused	7.03	420
Unauthorized at Track Level	14.03	415
Unsanitary Vehicle	4.65	294
Injured or ill Customer (On Train) - Transported	10.86	288
ATC Project	6.32	285
Transportation Department - Other	5.95	203
Door Problems - Faulty Equipment	6.19	189
Miscellaneous Other	7.17	189
Assault / Patron Involved	10.73	167
Emergency Alarm Station Activation	12.01	164
Door Problems - Debris Related	5.75	163
Misc. Transportation Other - Employee Non-Chargeable	4.30	144

hurt the passengers, TTC is responsible for keeping the vehicles safe to use, maybe they need more overhauls.

In conclusion, this study highlights key areas for TTC improvement, particularly in managing passengers' behavior and vehicle maintenance. Future research might explore the effectiveness of specific interventions in reducing these delays.

References

Csárdi, Gábor. 2022. *Gitcreds: Query 'Git' Credentials from 'r'.* https://gitcreds.r-lib.org/.

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

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 Subway Delay Data.* https://open.toronto.ca/dataset/ttc-subway-delay-data/.

———. 2024b. *TTC Transit Planning.* TTC. https://www.ttc.ca/transparency-and-accountability/transit-planning.

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.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax.* http://haozhu233.github.io/kableExtra/.