

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

Zijun Meng

January 24, 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. Given its critical importance, understanding why subway delays is essential to improve the system.

To analyze the data, 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). 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 happens, variable “Code” indicates what caused the delay, variable “Min\_Delay” records how many minutes the subway delayed, and finally “Line” shows which subway line delayed. 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 causes 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 mistriggerring 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 hurt the passengers, TTC is responsible for keeping the vehicles safe to use, maybe they need more overhauls.

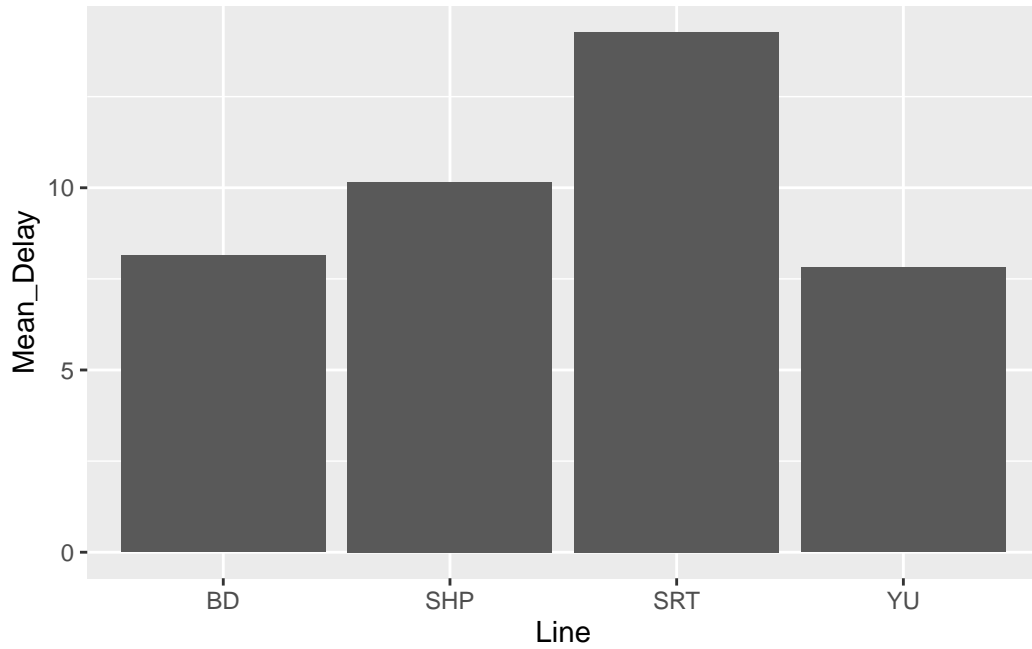


Figure 1: Delay by line

Table 2: Delay by reason

| Code   | Mean Delay | n    |
|--|------------|------|
| Disorderly Patron  | 7.488202   | 1229 |
| Passenger Assistance Alarm Activated - No Trouble Found  | 4.184783   | 736  |
| OPTO (COMMS) Train Door Monitoring                       | 5.192771   | 664  |
| Passenger Other  | 11.038710  | 465  |
| Injured or ill Customer (On Train) - Medical Aid Refused | 7.026191   | 420  |
| Unauthorized at Track Level                              | 14.026506  | 415  |
| Unsanitary Vehicle                                       | 4.646259   | 294  |
| Injured or ill Customer (On Train) - Transported         | 10.857639  | 288  |
| ATC Project  | 6.315790   | 285  |
| Transportation Department - Other                        | 5.950739   | 203  |
| Door Problems - Faulty Equipment                         | 6.185185   | 189  |
| Miscellaneous Other                                      | 7.169312   | 189  |
| Assault / Patron Involved                                | 10.730539  | 167  |
| Emergency Alarm Station Activation                       | 12.006098  | 164  |
| Door Problems - Debris Related                           | 5.748466   | 163  |
| Misc. Transportation Other - Employee Non-Chargeable     | 4.298611   | 144  |

In conclusion, by solving these problems, TTC can not only decrease the number of delays but also improve the safety of all the passengers.

## 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 Golemund, 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/>.