

# A Modest Proposal for Developing Better Experience with TTC Subway\*

Terry Tu

January 22, 2024

This study analyzes the 2023 TTC subway delay data from OpenDataToronto to uncover delay patterns across different days and lines. Key findings reveal significant variances in delay durations and frequencies, offering practical insights for commuters and transit authorities. The analysis highlights the need for targeted strategies to improve subway efficiency, particularly on the most affected lines.

## Table of contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Data</b>	<b>2</b>
2.1	Initial Observation from TTC Subway Delay Data . . . . .	2
<b>3</b>	<b>Results</b>	<b>3</b>
3.1	Analysis on subway delay based on day of the week . . . . .	3
3.2	Analysis on subway delay based subway line . . . . .	5
<b>4</b>	<b>Conclusion</b>	<b>6</b>
<b>5</b>	<b>Appendix</b>	<b>6</b>
5.1	Data Cleaning Process . . . . .	6
	Reference . . . . .	6

---

\*Code and data supporting this analysis is available at: <https://github.com/TEJMaster/TTC-Subway-Delay-Analysis-2023.git>

# 1 Introduction

This project delves into the “subway\_delay\_data\_2023.csv” dataset from OpenDataToronto, focusing on the Toronto Transit Commission’s (TTC) subway delays. Through this analysis, we aim to discern patterns and implications of these delays, providing valuable information for Toronto’s commuters and transit planners.

## 2 Data

Data used in this paper are retrieved from Open Data Toronto Portal through the library `opendatatoronto` ([Gelfand 2022](#)).

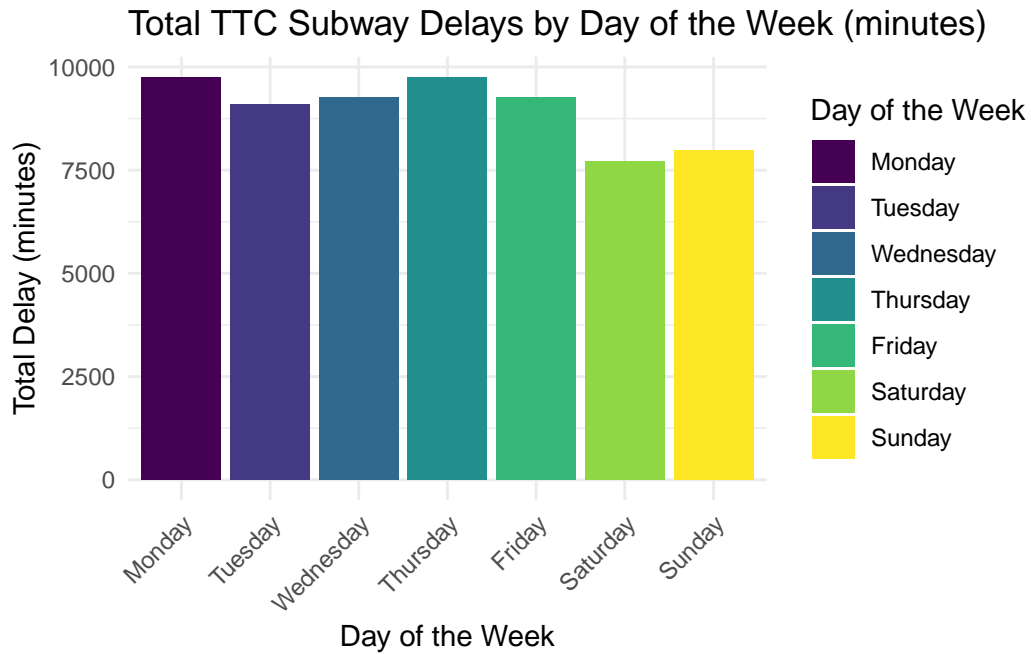
### 2.1 Initial Observation from TTC Subway Delay Data

The table below contains the first four rows from the “subway\_delay\_data\_2023.csv” dataset, retrieved from “opendatatoronto”. The table contains useful attributes including the date, time, day of the week, station name, specific delay codes, the recorded duration of delays (minimum delay), the shortest interval between consecutive trains (minimum gap), the designated direction (bound), the subway line affected, and the vehicle number. In our next step we will use this data to analyze the pattern of the current delays and help Toronto resident to plan for potential delays.

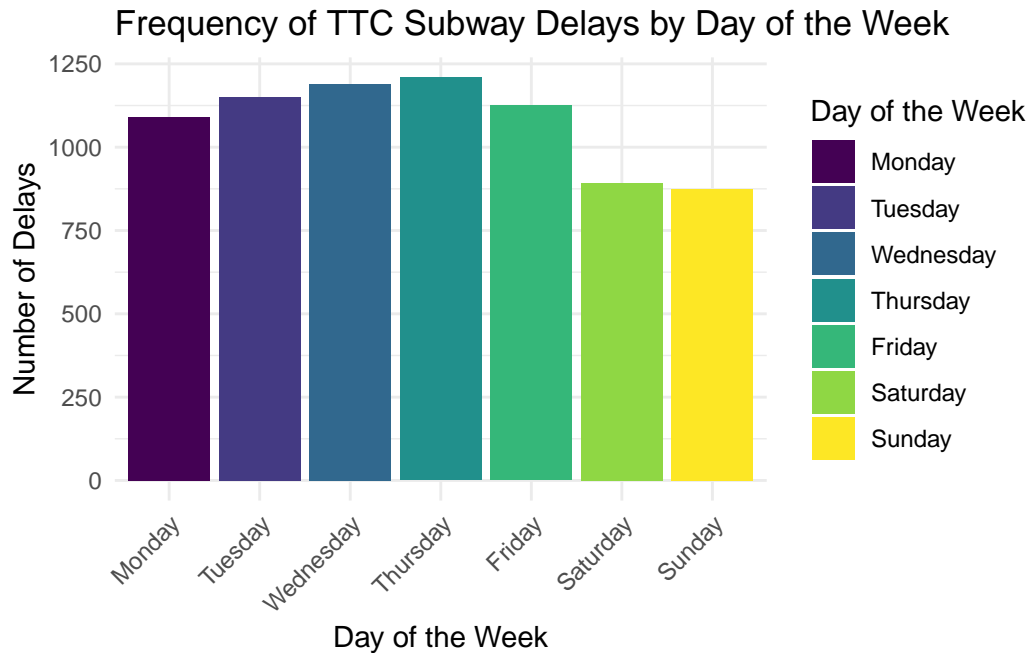
Date	Time	Day	Station	Code	Min Delay	Min Gap	Bound	Line	Vehicle
2023-01-01	02:22:00	Sunday	MUSEUM STATION	MUPAA	3	9	S	YU	5931
2023-01-01	02:30:00	Sunday	KIPLING STATION	MUIS	0	0	E	BD	5341
2023-01-01	02:33:00	Sunday	WARDEN STATION	SUO	0	0	W	BD	0
2023-01-01	03:17:00	Sunday	KEELE STATION	MUIS	0	0	NA	BD	0

### 3 Results

#### 3.1 Analysis on subway delay based on day of the week



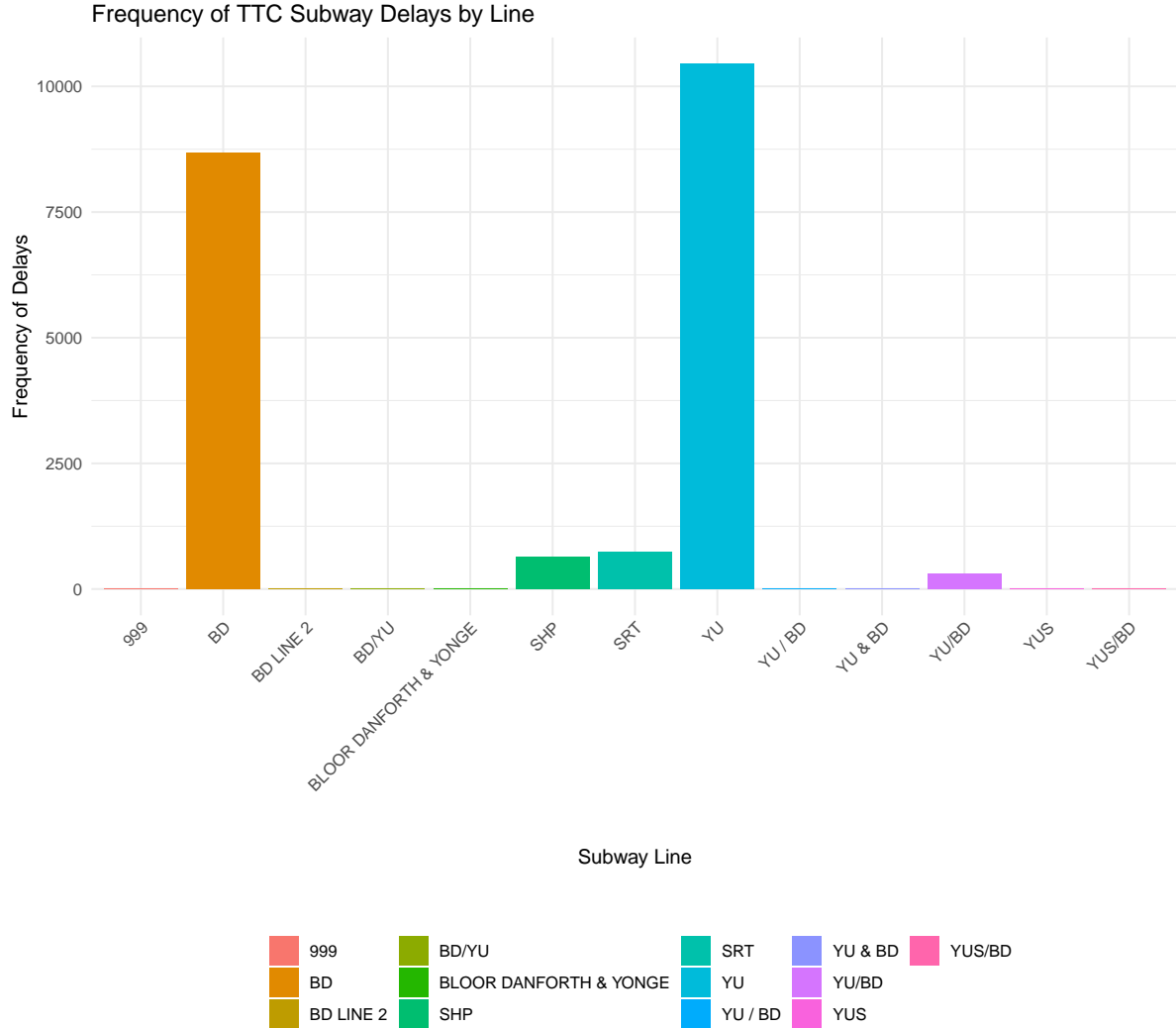
From the plot for total minutes delay we can clearly see that weekdays experience significantly longer cumulative delay duration as compared to weekends. Notably, Monday stands out with the highest aggregate minutes of delay, suggesting a peak in delay occurrences at the start of the workweek. Conversely, Saturday is the day with the least total delay time, indicating a smoother operational flow during weekend services.



Furthermore, the frequency analysis of delays corroborates the trend observed in total delay duration. Weekdays generally have a higher incidence of delays, with Friday leading in the number of reported delays. On the other hand, during Sunday the plot shows the least frequency of delays, aligning with the expected reduced demand for travel on the last day of the weekend.

These insights could serve as a cornerstone for targeted strategies aimed at improving service efficiency, such as enhanced resource allocation during identified peak times. Commuters might also benefit from planning their travel schedules around these insights to avoid potential delays.

### 3.2 Analysis on subway delay based subway line



The visualized data indicates a huge contrast in delay frequencies across different subway lines for the year 2023. Notably, certain lines such as '999' and 'BD Line 2' registered no delays, suggesting efficient operation or a lack of reporting. In contrast, other lines like 'SHP' and 'SRT' experienced a minimal number of delays, which could indicate isolated incidents or effective management of potential disruptions.

However, a significant concern is observed with the 'BD' (Bloor-Danforth) and 'YU' (Yonge-University) lines, which exhibited a substantially higher frequency of delays. This pattern highlights a critical area for operational focus. As a TTC subway manager, this insight necessitates an in-depth analysis of the causes behind the delays on these lines. Moreover, people taking these two lines are highly recommended to plan with extra time allowance in anticipation of the higher frequency of delays encountered on these routes.

## 4 Conclusion

Based on the data from “ttc-subway-delay-data-2023”, I have provided 2 suggestions. For commuters, especially those frequenting the more affected lines, and on weekdays it is prudent to anticipate potential delays and allocate additional time for travel. On the management side, the insights from this report should catalyze a thorough review of operational procedures and the implementation of strategic improvements to address the root causes of frequent delays.

Ultimately, the goal is to achieve a level of service that ensures efficiency, satisfaction, and the continued trust of the commuting public in the TTC subway system. This analysis lays the groundwork for the continuous improvement of Toronto’s subway operations and the enhancement of the city’s public transportation ecosystem.

## 5 Appendix

### 5.1 Data Cleaning Process

In preparing the dataset for analysis, the data is been cleaned by removing all lines with ‘na’ for their ‘Line’ column.

### Reference

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*.  
<https://CRAN.R-project.org/package=opendatatoronto>.