An Investigation of Subway Delays in Toronto: Insights from 2023 TTC Data*

Diving deep into subway delays

Charlie Zhang

September 23, 2024

This report presents an investigation into subway delays in Toronto using data from the Toronto Transit Commission (TTC) for the year 2023. The analysis examines patterns of delays across various subway stations, lines, and days of the week, with the goal of identifying critical factors that contribute to subway delays. Key findings show that Monday experienced the highest number of delays, with the Bloor-Danforth (BD) line being the most affected. These insights provide a data-driven foundation for improving operational efficiency and minimizing service interruptions in the future.

Table of contents

*Code and data are available at: https://github.com/Zqyyk11/Investigation-on-Toronto-Subway-Delay.

1 Introduction

Public transit plays a vital role in the daily lives of millions of people in urban centers around the world. In Toronto, the subway system, operated by the Toronto Transit Commission (TTC), serves as a backbone for many commuters. However, delays in subway services can cause significant inconvenience and disrupt the city's flow of traffic. In 2023, Toronto's subway system faced numerous delays, which are the focus of this report.

This analysis uses a comprehensive dataset of subway delays from the TTC to investigate the nature and extent of service disruptions. Specifically, the report aims to answer the following questions:

- Which days of the week are most prone to delays?
- What are the most affected subway lines and stations?
- What times of day experience the most delays?

By answering these questions, the report seeks to provide actionable insights for improving the reliability of Toronto's subway system, reducing the frequency and duration of delays, and enhancing the overall commuting experience.

The findings in this report are based on data-driven analyses and visualizations, which highlight delay trends throughout the year. This investigation serves as a foundation for further research into the causes of delays and offers guidance for policy makers and transit authorities in addressing key operational challenges.

2 Data

The data utilized in this report is sourced from Open Data Toronto and was imported using the opendatatoronto library (Gelfand 2022). The specific dataset used for analyzing TTC subway delays in Toronto is the "TTC Subway & SRT Train Service Delay Data" dataset [Toronto (2024). All data analysis was conducted using R (R Core Team 2023), with the assistance of several key packages, including tidyverse (Wickham et al. 2019), dplyr (Wickham 2023),here (Müller and Wickham 2020) janitor (Firke 2023), ggplot2 (Wickham 2016), knitr (Xie 2014), and lubridate (Grolemund and Wickham 2011).

To transform the raw data into a clean, analysis-ready format, several preprocessing steps were carried out using R. First, unnecessary columns that did not contribute to the analysis, were removed to clean the dataset. Using the dplyr package, specific columns were selected based on relevance to subway delays, such as date, time, station, line, and delay duration. The data was further cleaned by handling any missing or inconsistent values with the janitor package to ensure data integrity. Any outliers or erroneous entries were also identified and either corrected or removed as appropriate. These steps allowed for the creation of a clean dataset, which provided a solid foundation for the analysis below. The entire process was conducted

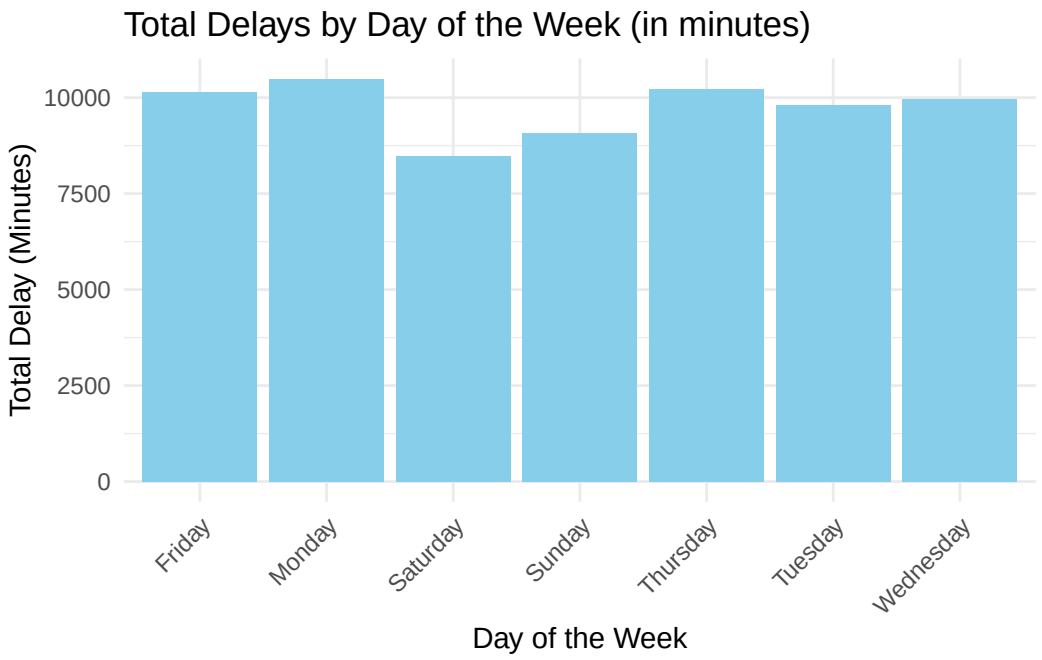
through efficient R code, leveraging the functionality of packages like dplyr for filtering and selection, and janitor for data cleaning.

2.1 Overview of the data

Date	Time	Day	Station	Min Delay	Min Gap	Line
2023-01-01	02:22:00	Sunday	MUSEUM STATION	3	9	YU
2023-01-01	02:30:00	Sunday	KIPLING STATION	0	0	BD
2023-01-01	02:33:00	Sunday	WARDEN STATION	0	0	BD
2023-01-01	03:17:00	Sunday	KEELE STATION	0	0	BD
2023-01-01	07:16:00	Sunday	BATHURST STATION	0	0	BD
2023-01-01	07:44:00	Sunday	JANE STATION	0	0	BD

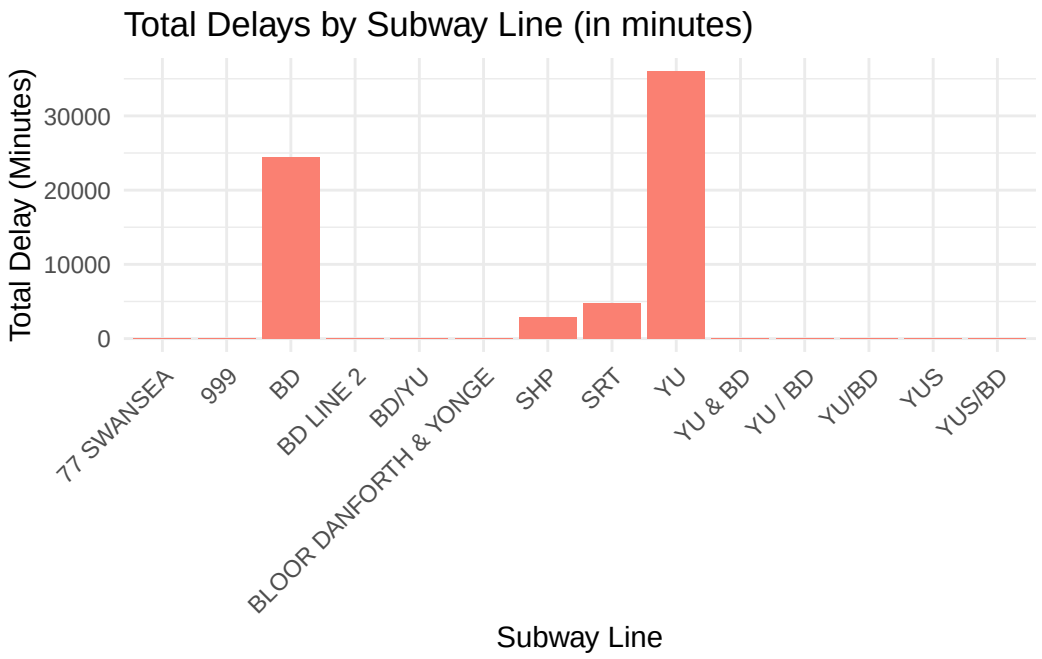
Sample of the Toronto subway's delay timer

2.2 Subway delays by day of the week



The study of subway delays on days of the week uncovers a noticeable trend linked to how weekdays and weekends unfold in terms of commuter patterns and travel demand shifts within the city's transit system. Observations show that Mondays encounter the most substantial delays and also other weekdays like Thursdays and Fridays due to heightened rider activity during the typical workweek when individuals travel to their workplaces. The uptick in passenger traffic during peak hours exerts pressure on subway operations by causing trains and busier platforms along, with prolonged boarding processes that collectively contribute to amplified service disruptions. During peak hours when there is demand for subway services issues like mechanical problems or signal interruptions tend to cause bigger disruptions compared to weekends. A reason weekends experience fewer delays possibly because fewer people use the subway due to less frequent service and reduced ridership. With a number of commuters relying on the subway during weekends the system runs more smoothly with less strain resulting in better train operations. Moreover weekend schedules often incorporate planned maintenance activities that have impact since there are fewer passengers traveling. The difference, between weekdays and weekends shows how commuter traffic affects the reliability of subway services. Weekdays experience frequent and severe delays during rush hour demands compared to weekends when operations are lighter and services run more smoothly.

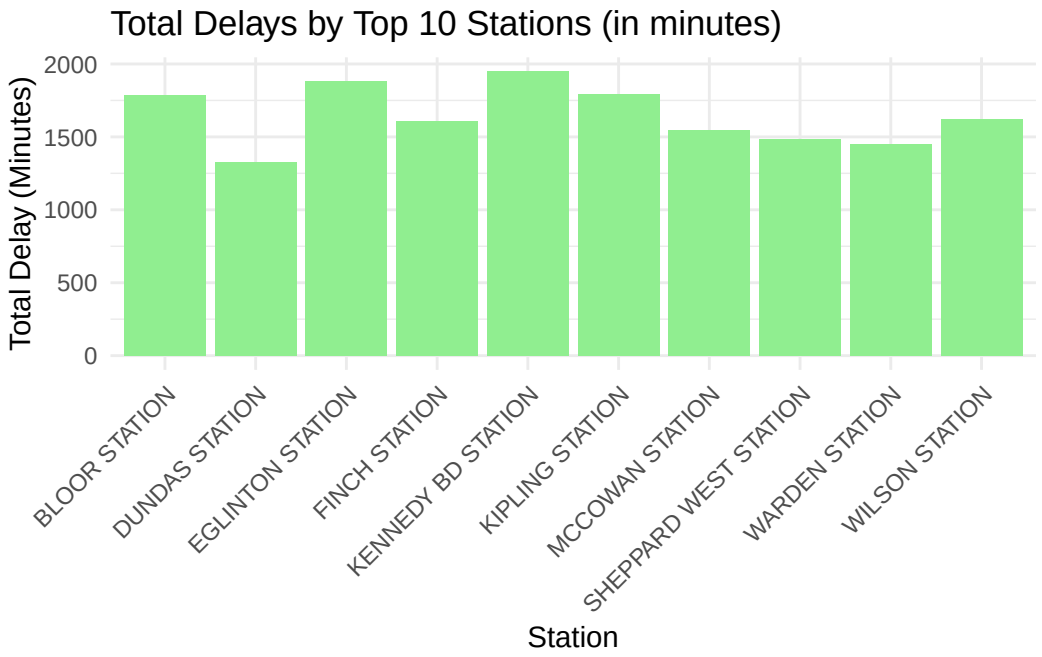
2.3 Subway delays by subway line



The Yonge University (Y U) and Bloor Danforth (B D) subway routes encounter delays mainly because they are the most heavily used and vital transportation links in Toronto city center and surrounding regions. These subway lines pass through populated neighborhoods and

major transit junctions like downtown Toronto serving as crucial lifelines for the millions of people who rely on them daily for their commutes. Specifically the Y U Line provides access to locations such, as Union Station, financial centers and popular spots leading to congestion during busy periods. The BD Line runs through the city connecting neighborhoods from east to west. Is essential for many commuters getting to work or school and carrying out their daily routines efficiently. The large number of passengers using these lines presents difficulties like longer boarding times and crowded platforms along with frequent train delays. Furthermore the substantial ridership on these routes amplifies the effects of disruptions as any delays have a cascading effect throughout the system given the heavy dependence, on these crucial lines. On the hand the Sheppard line (SRT) has fewer delays compared to others mainly because it caters to smaller communities and operates with less frequent service thus easing the burden, on infrastructure and operations.

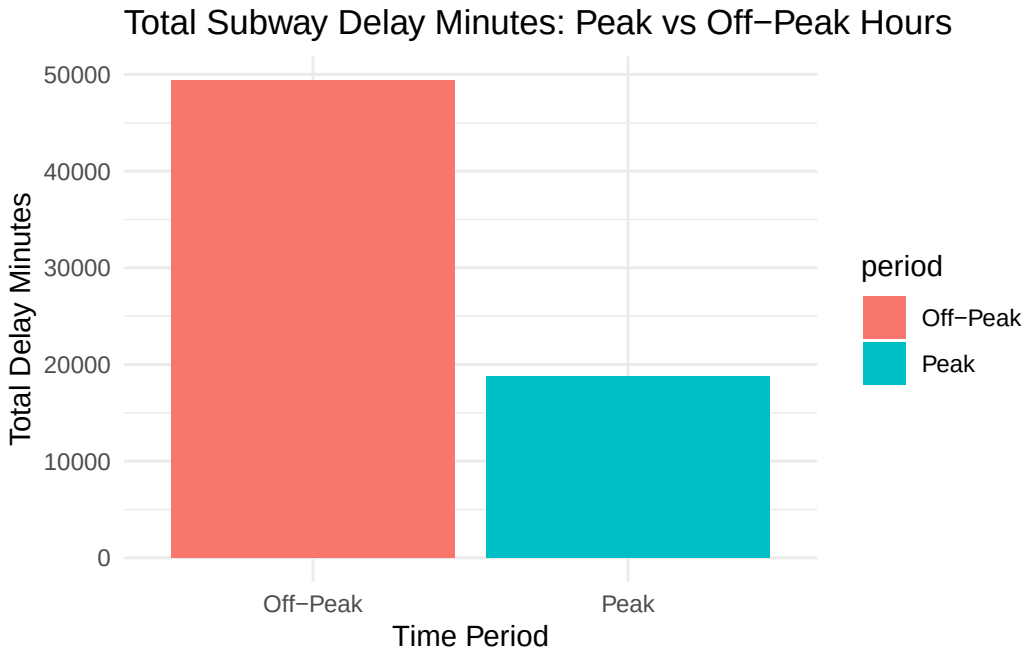
2.4 Subway delays by top 10 stations



The top ten stations experiencing the delays are mainly major terminals and transfer points in Torontos subway system – with Kennedy BD Station leading the list due to its operational intricacies and heavy passenger traffic flow through connections with other lines, like the Scarborough RT and the Bloor Danforth line. Currently at Eglinton Station construction is underway for the Eglinton Crosstown LRT project causing disruptions that affect passenger flow and train operations resulting in increased delays. Kipling and Bloor stations are also facing delays as they serve as transfer points for various subway and bus routes, with many commuters switching lines or modes of transportation. The clustering of travelers at

these locations results in crowded platforms and increases the likelihood of delays since even small interruptions can lead to a domino effect that impacts the reliability of the service provided there.This scenario occurs mainly in stations serving as central transit points due to their passenger flow and intricate operations along, with occasional infrastructure issues that collectively add up to prolonged delay times.

2.5 Subway delays Peak and off peak hours



The examination of subway disruptions in Toronto indicates a contrast between busy and quieter times of the day when it comes to delays. With less crowded periods encountering much longer total delay durations compared to peak hours despite the higher number of passengers and operational demands during those busier times when the transit system seems to focus on efficiency and manages to have fewer delays than, during the less busy times. Delays are more common during quieter hours due to frequent services and maintenance work being carried out then to avoid disruptions during prime commute times also fewer resources being available could lead to longer delays when problems occur. It seems that the system handles high demand situations effectively but could possibly do better at reducing interruptions during quieter times to maintain service all day long.

3 Results

The analysis of Toronto subway delays was conducted across four key aspects: delays by day of the week, subway lines, stations, and peak vs. off-peak hours. Each of these aspects offers unique insights into the operational challenges faced by the Toronto Transit Commission (TTC) in managing subway services.

Delays by Day of the Week: The analysis revealed that Mondays experience the highest total delay minutes, likely due to the surge of passengers at the start of the workweek. Both Thursday and Friday also show elevated delays, potentially reflecting increased traffic as commuters prepare for the weekend. Weekends, on the other hand, see fewer delays, which can be attributed to lower ridership and reduced service frequency.

Delays by Subway Line:The Yonge-University (YU) and Bloor-Danforth (BD) lines experienced the most delays. As two of the busiest lines serving major commuter hubs in downtown Toronto, they face greater operational challenges and are more susceptible to disruptions due to overcrowding and high service frequency. In contrast, smaller lines, like the Sheppard and Scarborough RT, showed fewer delays, likely due to lower ridership and fewer trains.

Delays by Subway Station: Kennedy Station on the Bloor-Danforth (BD) line had the highest recorded delays, followed by Eglinton and Kipling stations. As key transfer and terminal points, these stations handle large volumes of passengers, which can cause bottlenecks, particularly during rush hours. Eglinton Station, which is undergoing heavy construction due to transit expansion projects, is another hotspot for delays.

Peak vs. Off-Peak Hours:The final analysis compared peak and off-peak hours, revealing that off-peak periods experienced significantly higher total delay minutes. This might be due to reduced service frequency during off-peak hours, which makes each delay more impactful. Scheduled maintenance, often performed during off-peak hours, and fewer operational resources could also contribute to the higher delay times.

The overall analysis gives an overview of subway disruptions in Toronto for the year 2023.The frequency of delays is greatly impacted by the day of the week – particularly Mondays and the latter part of the workweek witness the highest delays.The Yonge-University and Bloor-Danforth lines face delays owing to their popularity and frequent train services. Specific stations like Kennedy and Eglinton serve as transition points which accumulates more delay minutes due to their vital positions in the transit system. During times when there are passengers traveling such as off hours or low traffic periods of the day there tends to be more delays in terms of total minutes being experienced due to possibly having fewer trains running and maintenance work taking place.The analysis points out areas for operational enhancements to concentrate on such as improving off hours service and handling delays at busy stations like Kennedy and Eglinton to enhance the overall efficiency of the system better. Likewise tackling the issues on subway lines like YU and BD could lead to fewer disruptions especially during peak hours, with heavy passenger traffic.

4 Discussion

The examination of delays, in the Toronto subway system reveals the challenges of managing a public transportation network efficiently and reliably. A notable discovery is the increase in delays during busy hours compared to peak periods – a surprising trend that goes against the assumption of more disruptions during high demand peak hours. The information indicates that when there are trains running and passengers must wait longer during off hours due to reduced service frequency which cause delays to have an impact. Moreover scheduled maintenance activities that are typically carried out during these quieter times may worsen the situation by causing delays to build up progressively over time. This situation underscores a chance for the TTC to reassess its off hour schedules and resource distribution in order to minimize delays during periods of passenger traffic.

During hours the transportation system faces stress but appears focused on maintaining punctuality for trains. This indicates a strategy in handling rush periods by addressing operational challenges to prevent minor issues from causing major delays. Nevertheless the clustering of delays on weekdays, like Mondays and towards the weeks end implies that commuter patterns greatly impact the systems effectiveness. On Mondays, in particular there may be delays because of the influx of commuters heading back to work after the weekend break; while on Thursdays and Fridays we tend to see increased delays as folks get ready, for their weekend plans ahead of time.It might be helpful for the TTC to take note of these trends and adjust their resources accordingly to manage operations smoothly.

In summary even though Toronto's subway system seems to handle peak hours despite the number of passengers. The examination reveals weaknesses during less busy times and, at specific stations and routes. To improve the situation the TTC could address these issues by improving services during off hours targeting stations with frequent delays and better managing resources on busy weekdays. These steps would greatly improve the system's effectiveness making the subway a more dependable and sturdy transportation option for the people of Toronto.

References

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

Gelfand, A. 2022. *Opendatatoronto: Access Toronto Open Data*. https://cran.r-project.org/package=opendatatoronto.

Grolemund, Garrett, and Hadley Wickham. 2011. *Lubridate: Make Dealing with Dates a Little Easier*. https://CRAN.R-project.org/package=lubridate.

Müller, Kirill, and Hadley Wickham. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Toronto. 2024. *Non Regulated Lead Sample*. https://open.toronto.ca/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Mara Averick, Jennifer Bryan, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.