Behind the Delays: Exploring Variations in Toronto's Transit Disruptions*

A Comparative Analysis of Delay Duration Across TTC Subway, Streetcar, and Bus Systems

Claire Chang

November 28, 2024

This paper uses a Bayesian Linear Regression Model to examine which Toronto Transit Commission (TTC) transit mode—subway, streetcar, or bus—experiences the highest duration of delays. We found that buses experience significantly longer delays compared to subways and streetcars. This finding is important because it highlights critical areas where transit improvements could alleviate the impact of delays on daily commuters. Overall, our study provides valuable insights into TTC transit reliability, informing efforts to enhance urban transportation efficiency and passenger experience.

Table of contents

*Code and data are available at: [https://github.com/ClaireUoft/Toronto_TTC_Transportation).

1 Introduction

Urban transit systems play a crucial role in the daily lives of millions, influencing mobility, accessibility, and economic productivity in cities worldwide. However, transit delays can significantly impact passengers' schedules, urban congestion, and overall satisfaction with public transportation. In Toronto, the Toronto Transit Commission (TTC) operates a complex network of buses, streetcars, and subways, all of which are prone to delays caused by a variety of factors, including mechanical failures, traffic congestion, and operational disruptions. Despite the centrality of public transit to urban life, there is limited systematic analysis of how delay patterns vary across transit modes and the factors contributing to these delays.

This paper focuses on addressing this gap by analyzing the duration of delays across TTC transit modes—buses, streetcars, and subways—using a comprehensive dataset from Open Data Toronto. Specifically, it examines the effects of transit mode, time of day, and day of the week on delay duration. While previous studies primarily focused on historical data and specific transit types. Their work examined delays across the TTC system over eight years but lacked the inclusion of all three transit modes (subway, streetcar, and bus) in a single comprehensive study for the most recent years.

To fill this gap, we employed a Bayesian linear regression model to estimate the relationship between delay duration and its predictors. The analysis incorporates over 80,000 delay events, harmonized across transit modes and cleaned for consistency. The results reveal that buses experience the longest average delays, followed by streetcars and subways. Additionally, delays are longer during peak hours and vary by day of the week, with weekends exhibiting shorter delays on average.

These findings are significant for urban planners, transit operators, and policymakers, as they highlight critical areas for intervention to improve transit reliability. Additionally, daily TTC users can leverage these insights to make informed decisions about when to leave their homes and which mode of transportation to take. Moreover, addressing these delays could enhance passenger satisfaction, reduce economic losses caused by tardiness, and increase the overall efficiency and reliability of Toronto's transit network. This study contributes to the broader field of urban transit research by offering detailed insights into delay variability and providing a data-driven foundation for targeted improvements in public transportation systems.

The remainder of this paper is structured as follows. Section 2 discusses the data used for this analysis, including key variables and sources, with particular attention to the quality metrics that affect polling accuracy. Section 3 outlines our modeling approach…, incorporating lessons learned from recent electoral cycles. Our predictions are under section of each model. Section 5 discusses the implications of our findings and suggests directions for future research. Finally, Section A evaluates methodology and survey copy.

2 Data

2.1 Overview

In this project, we used data from the opendatatoronto dataset created by (Gelfand 2022). This dataset provided bus, subway, streetcar records essential for our analysis. In this project, we used R(R Core Team 2023) and several R packages for data processing, analysis, and visualization. Specifically, tidyverse (Wickham et al. 2019), arrow(Richardson et al. 2024), here(Müller 2020), ggplot2(Wickham 2016), while dplyr(Wickham et al. 2023) was key for data manipulation tasks. For dynamic report generation, knitr(Xie 2023) and kableExtra (Zhu 2024) used, providing enhanced formatting for outputs. Together, these packages enabled efficient data cleaning, analysis, and visualization throughout the study.

2.2 Measurement

Table 1: Preview of Cleaned Dataset for Bus, Subway, and Streetcar (One Example Each)

Date	Time	Day	Transit_mode	Line	Location	Incident	Min Delay
2024-01-01	02:08:00	Monday	Bus	89	KEELE AND GLENLAKE	Vision	10
2024-01-01	02:45:00	Monday	Streetcar	505	DUNDAS AND MCCAUL	Security	10
2024-01-01	02:00:00	Monday	Subway	1	SHEPPARD STATION	N/A	0

The measurement process refers to how real-world transit delays—such as a bus stuck in traffic, a subway experiencing a mechanical failure, or a streetcar held by emergency services—are captured and represented as entries in a dataset. Each entry corresponds to a distinct delay event recorded by the Toronto Transit Commission (TTC).

Delay Duration (`Min Delay`):
This variable represents the length of a transit delay, measured in minutes. It is derived from the difference between scheduled and actual event resolution times, as recorded by the TTC's operational systems. These durations are logged for each incident and standardized in the dataset as numerical values.

Transit Mode (`Transit_mode`):
This categorical variable identifies the mode of transit where the delay occurred: subway, streetcar, or bus. The transit mode is directly logged as part of the operational data.

Time (`Time`):
The timestamp of the delay event, recorded in the dataset in the format `HH:MM:SS`, specifies when the delay occurred. This variable is used to analyze temporal patterns in delays, such as peak hours.

Day (`Day`):
The day of the week when the delay occurred (e.g., Monday, Tuesday). This variable is derived from the event date and provides insights into potential weekly trends in delays.

Line (`Line`):
This variable represents the transit line or route affected by the delay (e.g., Line 1 for subway or specific route numbers for buses and streetcars). It is directly recorded in the TTC's data logs.

Location (`Location`):
The specific location or station where the delay occurred is captured in this variable, allowing for spatial analysis of delay patterns.

Incident (`Incident`):
This categorical variable captures the cause of the delay, such as "Mechanical," "Security," or "Collision." Raw incident descriptions are grouped into broader categories to ensure consistency across all transit modes.

2.3 Data Cleaning

The transit delay data was prepared for analysis by processing raw datasets for buses, streetcars, and subways. Key steps included:

1. **Adding a Transit Mode Column**:
   A `Transit_mode` column was added to differentiate between buses, streetcars, and subways.

2. **Removing Unnecessary Columns**:
   Irrelevant columns were dropped, retaining only essential variables like `Date`, `Time`, `Day`, `Transit_mode`, `Line`, `Location`, `Incident`, and `Min Delay`.

3. **Standardizing Subway Line Identifiers**:
   Subway lines were recoded into consistent and meaningful labels (e.g., `"YU"` to `"1"`).

4. **Harmonizing Column Names and Categories**:
   Column names were standardized across datasets, and incidents were grouped under broader categories (e.g., `"Collision - TTC Involved"` and `"Collision - TTC"` were both labeled `"Collision"`).

5. **Handling Missing and Mismatched Values**:
   Invalid or missing data (e.g., incidents labeled `"999"`) were reassigned as `"N/A"`.

6. **Combining Datasets**:
   All three datasets were merged into a unified dataset with consistent structure and formatting.

Subsampling the Dataset:
To address computational challenges with the original dataset of over 80,000 observations, a random subsample of 0.1% was selected. This ensured the subset remained representative of delay events across transit modes while enabling efficient model fitting without compromising analytical integrity.

Histogram of Delays:
Figure 1 Delays were capped at 45 minutes to focus on typical delay durations and exclude extreme values that could distort the distribution.

Boxplot of Delays:
Figure 2 The interquartile range (IQR) method was used to remove extreme outliers. The upper bound was calculated as `1.5 * IQR` above the third quartile, ensuring accurate representation of central tendencies and variability in delay durations.

2.4 Outcome variables

2.4.1 Delay Duration (Min Delay)

The main outcome variable in this study is the duration of delays, measured in minutes (Min Delay). This variable quantifies the length of each delay experienced by the TTC's different transit modes (subway, streetcar, bus). The primary goal of the analysis is to determine which transit mode exhibits the highest average delay duration.

To visualize the distribution of delay durations, a histogram Figure 1 is presented to provide an overview of the range of delay durations for each transit mode. This figure highlights the spread and common ranges of delay times, allowing us to observe whether there are particularly long delays for any given mode.

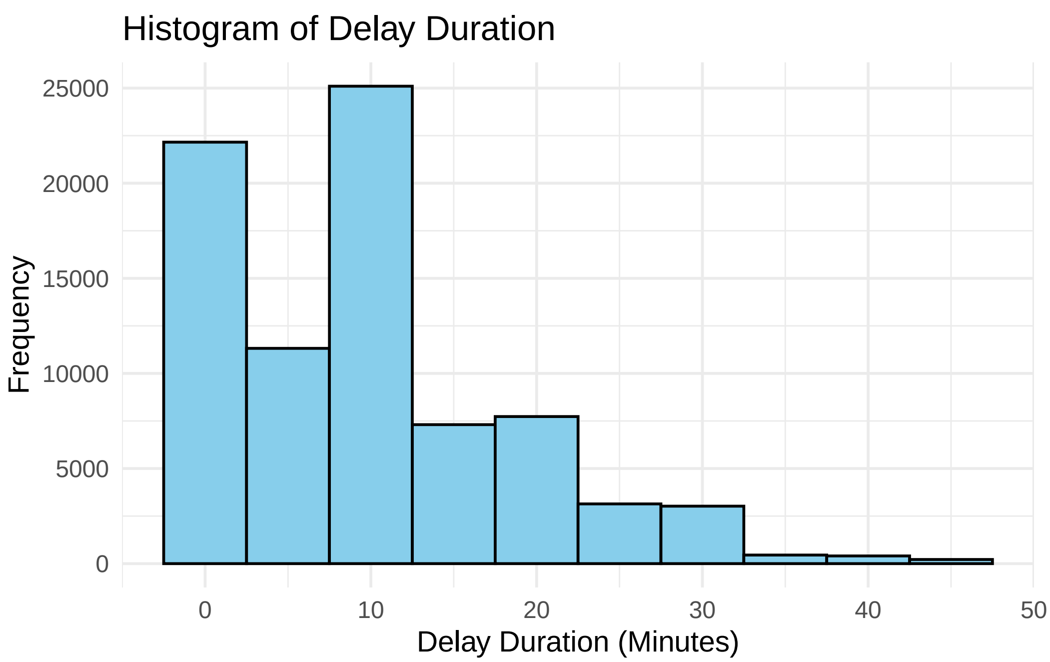


Figure 1: Histogram of Delay Duration (Filtered at 45 Minutes)

Lastly, Figure 2 displays a boxplot comparing delay durations across subway, streetcar, and bus transit modes. The boxplot offers insights into the median delays and potential outliers that may suggest unusual or extreme events, while showing variability in delay times across the different transit modes.

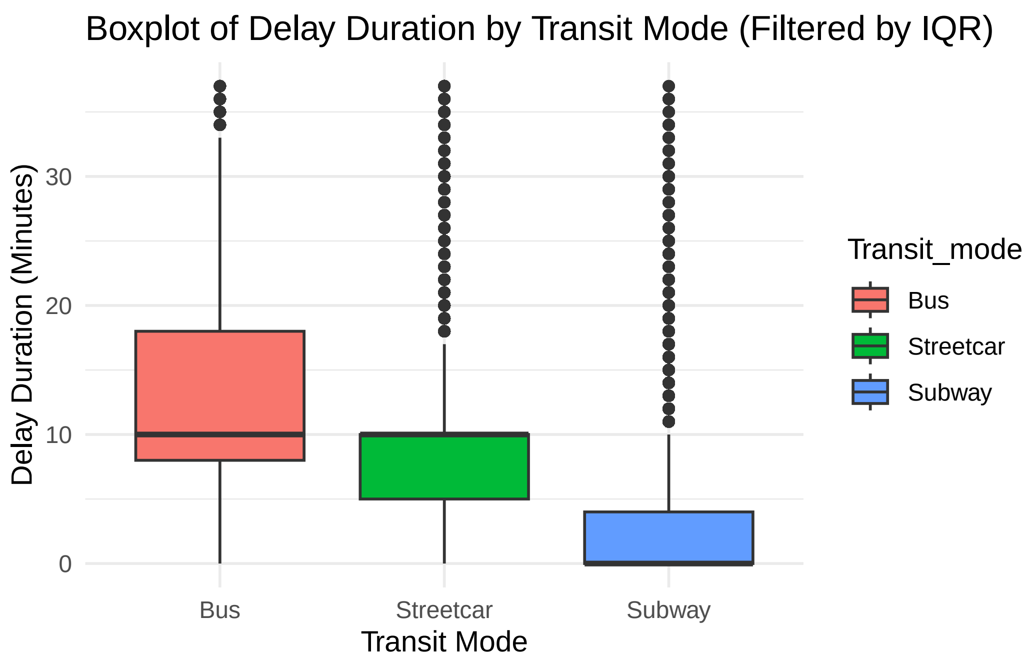


Figure 2: Boxplot of Delay Duration by Transit Mode (Filtered by IQR)

2.5 Predictor variables

2.5.1 Transit Mode

The main predictor variable is Transit Mode, which includes three categories: subway, streetcar, and bus. Each transit mode has distinct characteristics, service areas, and operational challenges that may contribute to variations in delay duration.

To examine this relationship, Figure 3 plots the average delay duration for each transit mode. This chart helps us visualize which transit mode experiences the longest delays on average and illustrates the extent to which each mode contributes to the overall delay picture.

2.5.2 Time of Day

Another key predictor is Time of Day shown in Figure 4. This variable plays a crucial role in understanding how congestion and operational challenges fluctuate throughout the day. For example, delays might be more frequent and prolonged during peak hours (morning and evening rush periods) due to increased ridership.

Figure 4 shows a line plot representing average delay duration across different hours of the day. This visualization allows us to determine the times at which each transit mode is most

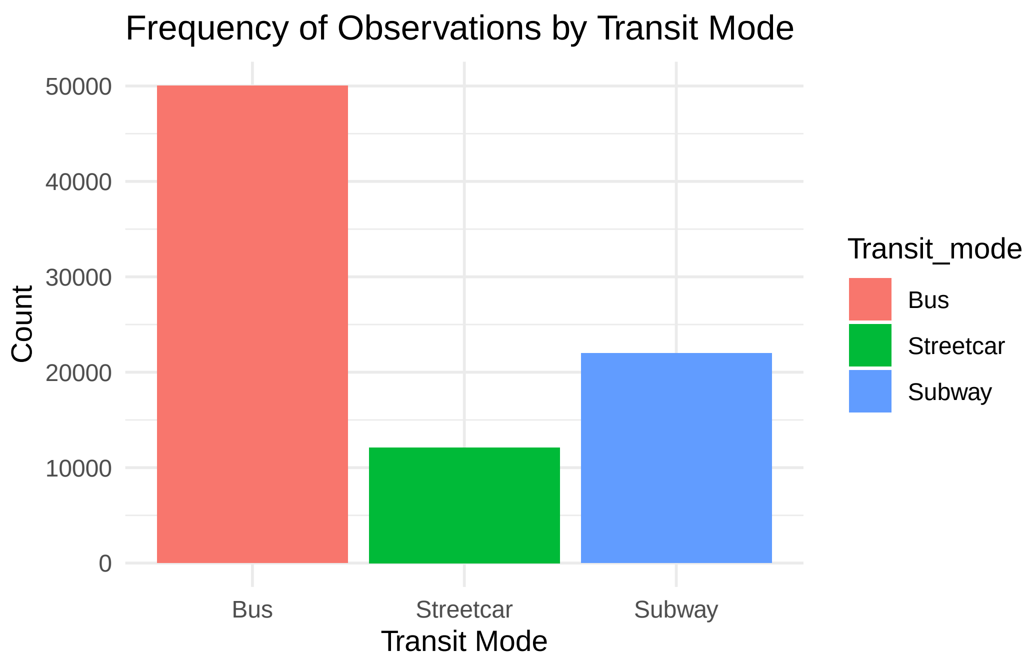


Figure 3: Average Delay Duration by Transit Mode

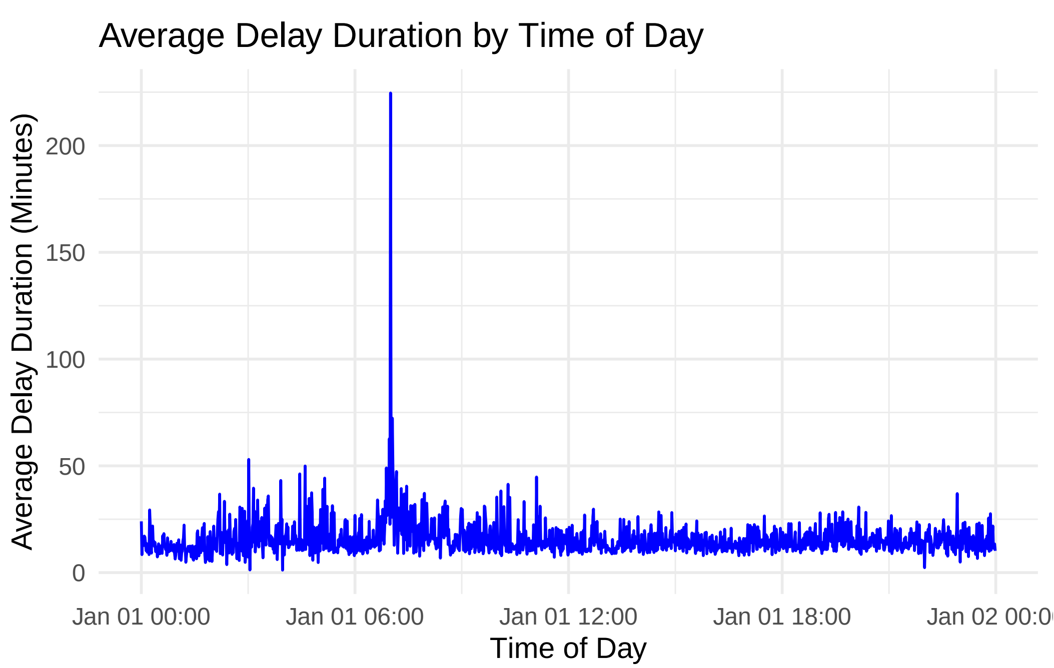


Figure 4: Average Delay by Time of Day

vulnerable to delays, and whether certain transit modes are disproportionately affected during rush hours.

2.5.3 Day of the Week

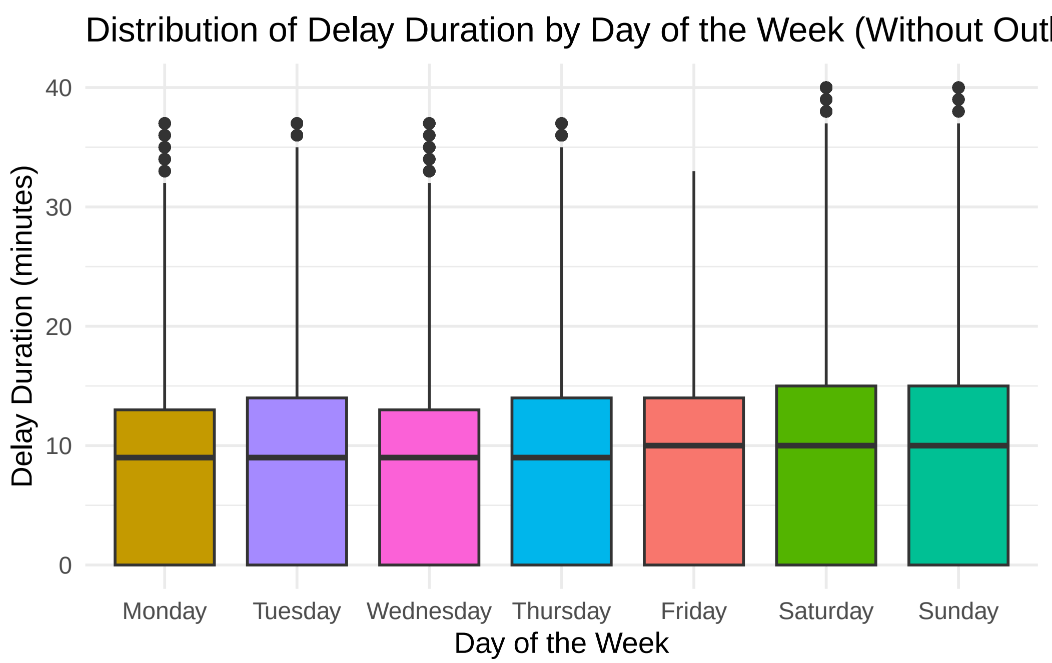


Figure 5: Delay Duration by Day of the Week

The Day of the Week is an important predictor that may influence delays due to varying passenger demand and maintenance schedules. Weekdays typically see different patterns compared to weekends, with increased rider demand and therefore greater potential for delays.

Figure 5 is a bar chart that shows the average delay duration for each day of the week, grouped by transit mode. It allows us to observe any weekly trends in delay duration, such as consistent weekday delays or weekend-specific issues, giving more insight into the operational aspects of the different modes. The predictor variables, including transit mode, time of day, and day of the week, help in assessing the specific conditions that affect the delay duration.

3 Model

The Bayesian linear regression model is designed to predict the duration of delays experienced by different TTC transit modes (subway, streetcar, or bus) based on various predictors such as transit mode, time of day, and day of the week. Background details and diagnostics are included in Appendix C.

3.1 Model Set-Up

Define y_i as the duration of the delay (in minutes) for a given transit event. Let β_i represent the transit mode (subway, streetcar, or bus), and γ_i and δ_i represent the time and day of the week, respectively. The generative model is specified as follows:

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) & (1)\\
\mu_i &= \alpha + \beta_i + \gamma_i + \delta_i & (2)\\
\alpha &\sim \text{Normal}(0, 2.5) & (3)\\
\beta &\sim \text{Normal}(0, 2.5) & (4)\\
\gamma &\sim \text{Normal}(0, 2.5) & (5)\\
\delta &\sim \text{Normal}(0, 2.5) & (6)\\
\sigma &\sim \text{Exponential}(1) & (7)
\end{aligned}
$$

Here:

- y_i is the observed delay duration for a specific transit event.
- α is the baseline average delay duration.
- β_i captures the effect of transit mode (subway, streetcar, or bus).
- γ_i captures the effect of the time of day (e.g., peak or off-peak).
- δ_i captures the effect of the day of the week.
- σ is the standard deviation, representing unexplained variability in y_i.

This Bayesian model allows for estimating the contributions of transit mode, time, and day to delay duration while accounting for uncertainty in the data.

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.2 Model justification

The specified Bayesian generative model is designed to estimate how transit mode, time of day, and day of the week contribute to delay duration. The choice of a Bayesian framework allows for explicit modeling of uncertainty in the data and provides posterior distributions for all parameters of interest, offering more nuanced insights than simple point estimates.

We expect that the transit mode (β_i) will have a significant impact on the delay duration (y_i), as different modes (subway, streetcar, and bus) are subject to varying operational challenges and infrastructure constraints. For instance, subways might experience shorter average delays due to dedicated tracks, while buses could encounter higher delays due to traffic congestion.

The time of day (γ_i) is hypothesized to influence delays, with peak hours (e.g., rush hour) likely contributing to longer delays due to higher demand and operational stress. Similarly, the day of the week (δ_i) may play a role, as weekends often have reduced schedules and lower demand, potentially resulting in shorter delays.

The prior distributions for the parameters are chosen to be weakly informative, specifically $\text{Normal}(0, 2.5)$ for α, β, γ, and δ, reflecting prior beliefs that the effects are likely centered around zero but allowing for moderate deviations. The standard deviation (σ) is modeled with an Exponential(1) prior, which is appropriate for ensuring positive values while discouraging overly large variances in unexplained delay durations.

3.3 Assumptions and Limitations

The analysis assumes a linear relationship between delay duration and predictors (transit mode, time of day, and day of the week), with delays following a normal distribution. Independence between delay events is assumed, though real-world factors like cascading delays might violate this. While the IQR method removes extreme outliers, it may exclude rare but valid events, and the model explains only a small fraction of the variation ($R^2 = 0.033$), leaving many influencing factors unaccounted for. Simplistic representations of time and day fail to capture interactions or seasonal trends, and external influences like weather or events are not considered. Additionally, data inconsistencies, such as missing values and aggregated incident types, may impact accuracy, limiting the model's ability to fully explain delay dynamics.

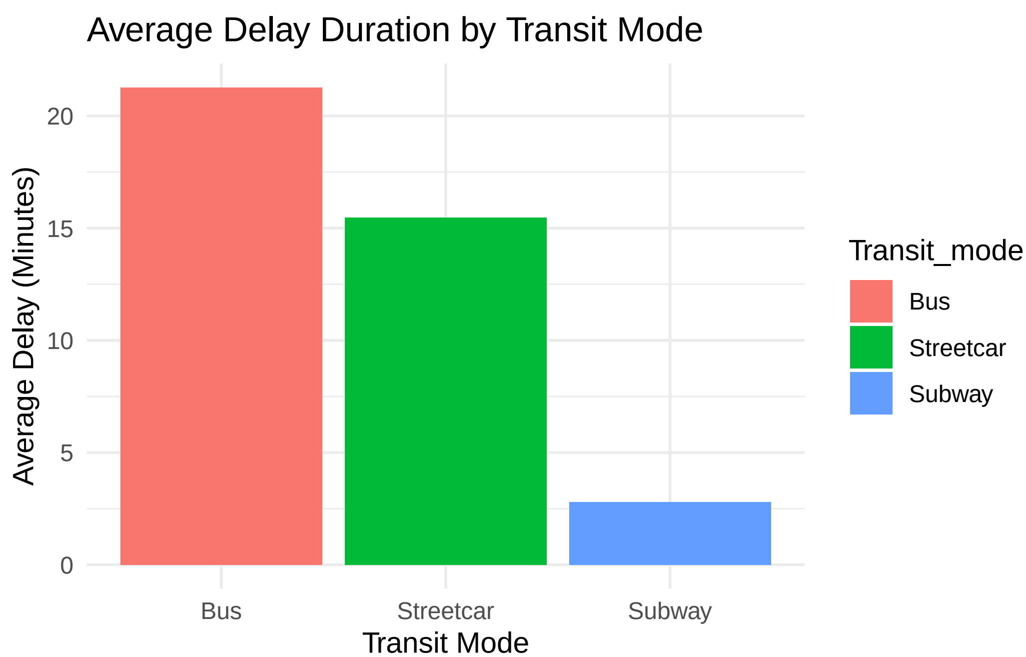
4 Results

These statistics indicate that, on average, TTC streetcars experienced a delay duration of approximately 15.7 minutes, while buses faced an average delay of 12.5 minutes, and subways experienced the shortest average delay of 9.3 minutes. The total number of delay events recorded was 50,039 for buses, 21,979 for subways, and 12,107 for streetcars.

Table 2: Summary statistics of delay events for TTC transit modes

Transit_mode	Average Delay Duration (Minutes)	Count of Events
Bus	21.273167	50039
Subway	2.806497	21979
Streetcar	15.476666	12107

The bar plot visually confirms that streetcars have consistently higher average delay durations compared to buses and subways, aligning with the numerical analysis."



The model summary Table 3 results indicate that, compared to buses, streetcars and subways have shorter delays, with average reductions of 6.8 and 18.8 minutes, respectively. Delays also vary by day of the week, with Tuesday and Wednesday showing slightly shorter delays. However, the model explains only about 3% of the variance in delay duration ($R^2 = 0.033$), and the RMSE of 44.88 suggests moderate prediction error. While transit mode and day of the week have some influence, much of the variation in delay duration remains unexplained.

The Figure 6 graph indicates that bus face the highest average delay durations, peaking at over 40 minutes in the early morning. Streecar have moderate delays throughout the day, averaging between 15-20 minutes, while subways consistently experience the lowest delays, averaging below 10 minutes. The significant early morning peak for streetcars suggests this as a critical period for intervention to reduce delays.

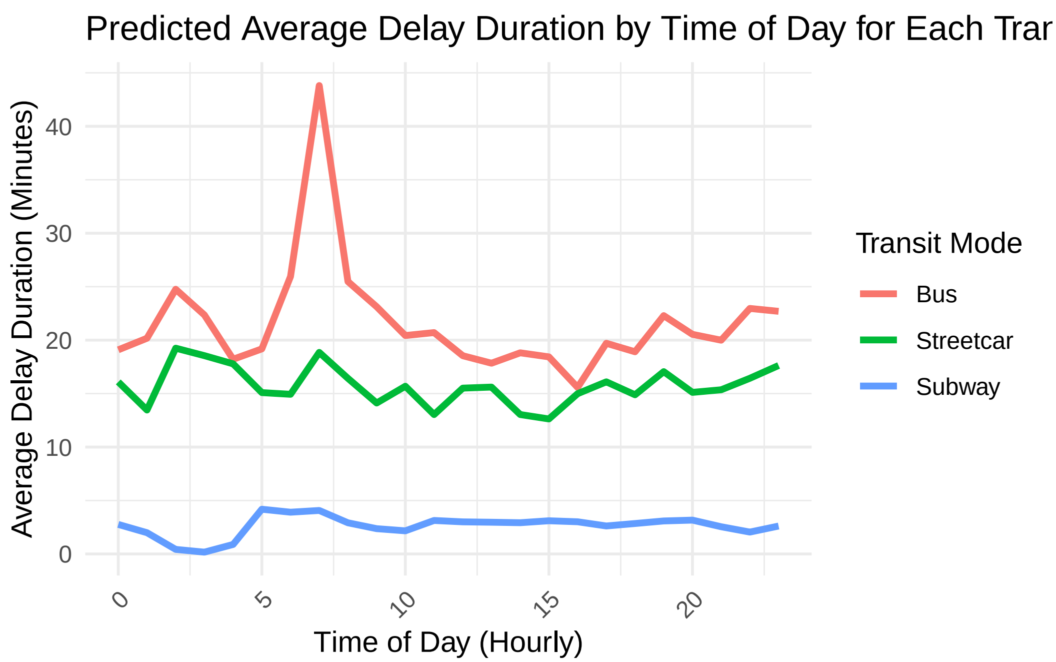


Figure 6: Predicting average delay duration by time of day for each transit mode

5 Discussion

5.1 Understanding Transit Delays Across Modes

This study aimed to analyze delay durations across Toronto Transit Commission (TTC) transit modes—subway, streetcar, and bus—using data from 2024. The findings reveal critical differences in delay patterns, with buses experiencing the longest delays on average, followed by streetcars and subways. These results align with expectations given the shared-road infrastructure of buses and streetcars, which are more exposed to traffic congestion and external disruptions. Subways, operating on dedicated tracks, exhibit fewer and shorter delays, highlighting the reliability of grade-separated transit systems.

Table 3: Summary of model results evaluating the relationship between transit mode and delay duration

	model1
(Intercept)	25.051
Transit_modeStreetcar	−6.799
Transit_modeSubway	−18.849
Time	0.000
DayMonday	−1.043
DaySaturday	−2.160
DaySunday	−1.572
DayThursday	−1.161
DayTuesday	−3.601
DayWednesday	−3.712
Num.Obs.	8412
R2	0.033
R2 Adj.	0.031
Log.Lik.	−43 938.441
ELPD	−43 986.3
ELPD s.e.	622.5
LOOIC	87 972.7
LOOIC s.e.	1245.0
WAIC	87 985.6
RMSE	44.88

5.2 Temporal and Weekly Patterns in Delays

The analysis also identified temporal trends, with peak hours contributing to longer delays across all modes. This reflects the increased strain on transit networks during rush hours, driven by higher ridership and demand. Additionally, delays were shorter on weekends, likely due to reduced traffic congestion and lower service frequency. These findings underscore the importance of tailoring operational strategies to address the unique challenges of peak periods.

5.3 Computational Constraints and Data Subsampling

While the original dataset contained over 80,000 observations, computational challenges necessitated subsampling to 0.1% of the data. This approach ensured the analysis remained feasible without compromising the representativeness of the sample. However, this step may have excluded rare but significant delay events, potentially influencing the findings.

5.4 Implications for Policy and Practice

These findings have practical implications for transit planning and policy. Prioritizing interventions for buses and streetcars, such as dedicated lanes or signal prioritization, could reduce delays significantly. Understanding temporal patterns can guide resource allocation, with more staff and services deployed during peak hours. For daily TTC users, insights into delay variability can aid in better trip planning and mode selection.

5.5 Limitations and next steps

The model accounted for only 3% of the variation in delay durations ($R^2 = 0.033$), suggesting the need to include additional factors, such as weather, traffic incidents, and ridership levels, in future analyses. Temporal variables were modeled independently, without exploring interactions or seasonal effects, which could further refine the results. Missing or inconsistent entries (e.g., incidents labeled as "N/A") may have introduced biases, while the IQR method for outlier removal may have excluded rare but valid delay events. Addtionally, future research should expand on these findings by incorporating additional predictors, such as weather and traffic data, to enhance model accuracy. Analyzing interactions between variables, such as the combined effects of time of day and transit mode. Lastly, exploring the impact of infrastructure improvements, such as bus rapid transit (BRT) or streetcar signal prioritization, on delay reduction.

A Appendix

Surveys, sampling, and observational data appendix Please include an appendix where you focus on some aspect of surveys, sampling or observational data, related to your paper. This should be an in-depth exploration, akin to the "idealized methodology/survey/pollster methodology" sections of Paper 2. Some aspect of this is likely covered in the Measurement sub-section of your Data section, but this would be much more detailed, and might include aspects like simulation and linkages to the literature, among other aspects.

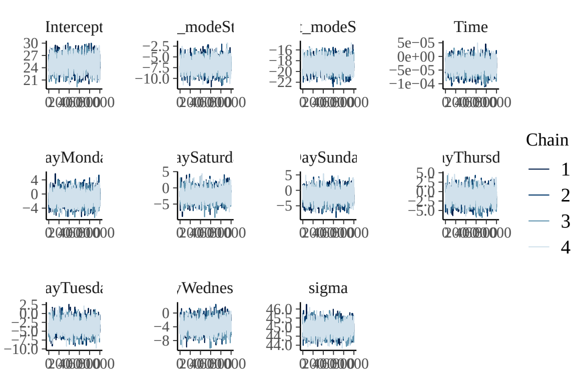
B Additional data details

C Model details

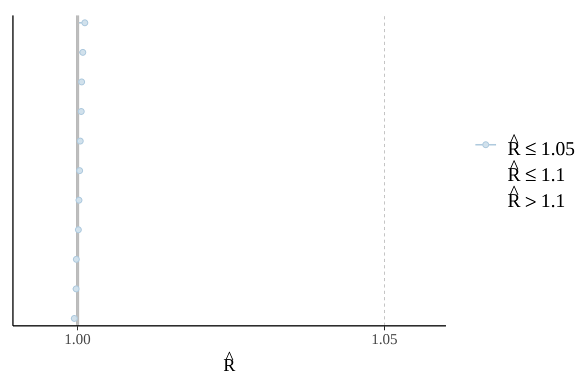
C.1 Diagnostics

Figure 7a is a trace Plot that shows the sampled values of each parameter from the MCMC chains over the iterations. Ideally, these chains should overlap well and display consistent movement across the same regions. From the trace plot, the chains for all parameters appear to have converged and are overlapping consistently, without any clear trends or divergences. This indicates good mixing and suggests the MCMC algorithm has likely converged.

Figure 7b is a Rhat Plot that shows the potential scale reduction factor, which assesses convergence of the MCMC chains. The Rhat statistic should be close to 1, typically below 1.1, to indicate good convergence. In this plot, the Rhat values for all parameters are very close to 1, suggesting the MCMC algorithm has converged, and there is no significant difference between the variances within and across chains.



(a) Trace plot



(b) Rhat plot

Figure 7: Checking the convergence of the MCMC algorithm

References

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

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

Müller, Kirill. 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/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

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

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, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://CRAN.R-project.org/package=knitr.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.