

Analysis of Transport Bus Initiative Usage*

Hoang Viet Nguyen

September 26, 2024

This paper presents an analysis of the transport bus initiative usage, with a specific focus on three key variables: clients, clients transported, and clients stationary. The dataset, consisting of daily records, is explored using various statistical distributions to understand the behavior of these variables and assess the performance of the bus initiative. By evaluating the distribution relationship between the three key variables, we aim to understand the efficiency of the transport system and identify patterns in client movement. The analysis reveals critical insights into how well the initiative meets transportation demands, contribute to more efficient public transport planning.

1 Introduction

Mass transportation is one of the essential elements of urban and rural mobility of people as it provides them with access to vital services and jobs and also aids in the enhancement of sustainable development. Of the various options of mass transport systems, bus initiatives are very critical in that they provide low-cost and easy to use options for the various populations. In recent years, many cities have implemented transport bus initiatives in a bid to improve the usage of public transport systems and the extent of the population covered by these transport systems. However, it is important to note that, although these initiatives are expected to bring in more targets in the formulation and execution of programs, determining how these initiatives affect the actual behavior of clients remains an area of interest.

This paper aims to analyze the usage of a particular transport bus initiative, based on the number of clients, clients transported and clients-stationary. In doing so, we seek to examine the usage of the services provided by the initiative and observe patterns in the use of the data which help to improve the services offered and show how the initiative fits within the wider aims of enhancing mobility within the area. The variables determining factors related to the

*Code and data are available at: [LINK](#).

clients' level of engagement with the service include time, the degree of service, and economic aspects.

The gap in existing literature on the effectiveness of transport bus initiatives necessitates a detailed analysis of client interactions. Understanding who uses the service and why is vital for transportation planners and policymakers aiming to create inclusive transit systems. Our analysis provides insights into the effectiveness of the bus initiative and its ability to meet the mobility needs of the community.

In the subsequent sections, we will detail the methodology used to collect and analyze the data, present key findings, and discuss the implications for public transportation strategies. By focusing on clients, clients transported, and clients stationary, this study contributes valuable knowledge that can guide future enhancements to transport bus initiatives and improve overall public transit experiences.

2 Data

Open Data Toronto provided the data used in this investigation, which covered the years 2023–2024, titled “Transport Bus Initiative Usage”. It offers comprehensive details on a transport bus program meant to enhance urban mobility. The number of clients using the transport service overall, the number of clients transported successfully, and the number of clients who stayed stationary make up the three primary variables in the dataset. Every entry is time- and date-stamped, thereby tracking variations in service efficiency and demand over time.

The information provides a comprehensive view of the initiative's performance, allowing for more in-depth examination of transportation patterns. During the data preparation stage, no significant anomalies or missing values were found in any of the variables, which were all organized and tidy. The dataset is a useful tool for assessing Toronto's public transportation system's performance and pinpointing areas that can use improvement.

The transport bus initiative usage analysis is based on three essential variables: clients, clients stationary, and clients transported. The clients variable gives a full picture of overall involvement by reflecting the entire number of people interacting with the transport system. Clients stationary in this sense refers particularly to people who stay at a certain place, emphasizing the need for transportation services and denoting times of high or low activity. Clients transported, on the other hand, measures the precise number of people who were moved, providing information on how well the transport system functions. Transport planners can increase the overall effectiveness of the bus initiative and optimize resource allocation by analyzing the interactions between these elements.

#Results ## Distribution of Clients Stationary The Distribution of clients who are stationary is represented by a histogram that offers significant insights into the behavior of users of the transport bus. The y-axis shows the frequency of those counts, and the x-axis shows the number of clients who stayed still at different locations.

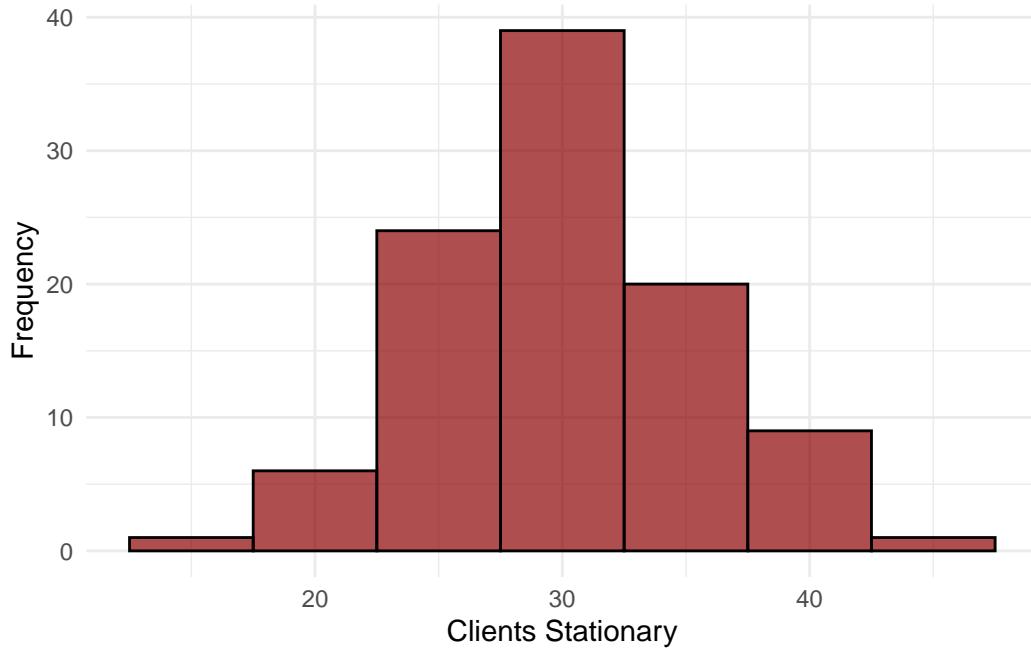


Figure 1: Distribution of Clients Stationary

The histogram shows that the majority of entries are between 0 and 40 clients, with 30 clients stationary showing the largest frequency. The data seems to be no skew, indicating a perfectly symmetrical distribution. The increase of clients in the beginning stopped until it reaches a peak of 30, then decrease beyond that point. This suggests that while there are times when many clients remain in one location, such occurrences are less frequent. By identifying important areas and times when the demand for transport services may be higher due to stationary clients, these insights might assist transport planners in formulating better service availability strategies during these times.

2.1 Distribution of Clients Transported

The second histogram displays the distribution of clients transported. Similar to the previous visualization, the x-axis shows the number of clients who were successfully transported, while the y-axis indicates the frequency.

The histogram reveals that the majority of clients transported lies heavily in the negative interval, it indicates that the amount of none clients transported are higher in comparison with the amount of clients that are transported. There is a noticeable drop in frequency as the number of clients transported increases beyond this point. This indicates that while smaller groups are frequently transported, larger groups are transported less often. Understanding



Figure 2: Distribution of Clients Transported

this distribution is critical for optimizing bus capacity and scheduling, as it points to typical transport loads and potential under utilization of buses during peak transport times.

2.2 Number of Clients

This scatterplot illustrates the relationship between date and the number of clients in the dataset, providing a visual representation of client engagement over time. The x-axis represents the date, showing the timeline of observation, while the y-axis indicates the number of clients.

This visualization demonstrated the amount of clients over the observed period. Peaks are observed in the range of February and March when the number of clients are surged to 60. Conversely, several dates show a drop to below 50 clients, indicating lower usage. This variation implies that a few occasions or elements, such the state of the weather or neighborhood events, have a big impact on how many customers use the transportation service. Comprehending these trends is essential for allocating supplementary resources during periods of high demand and improving service accessibility during periods of low demand.

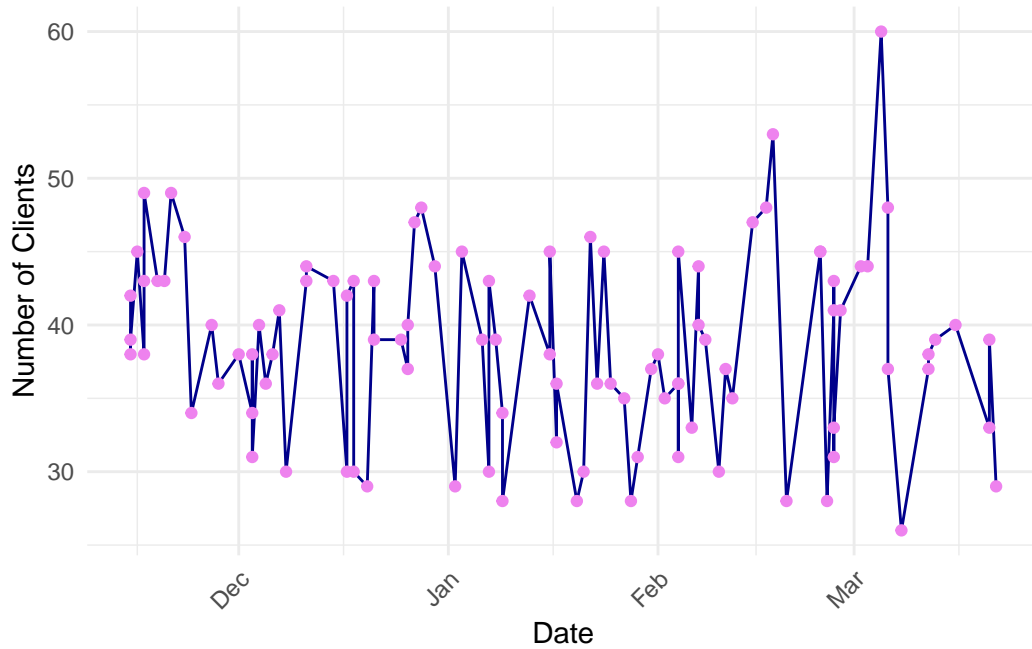


Figure 3: Number of Clients over time

2.3 Relationship between Clients Stationary and Clients Transported

The bar chart examines the relationship between clients stationary and clients transported, with the x-axis represents clients stationary while the y-axis shows clients transported.

The analysis shows that the number of clients transported and the number of clients stationary appear to have little to no relationship. Throughout the observed range, both variables typically exhibit independent and consistent trends. For example, even when ten clients are stationary, the number of clients carried can fluctuate a lot yet still stay mostly steady with no discernible rising or decreasing trend. This lack of correlation shows that the number of clients transferred at any particular time is not directly influenced by the number of clients who remain immobile.

Transport planners must understand this independence because it suggests that strategies to increase service efficiency—instead of depending exclusively on stationary client data—may need to take into account factors other than client stationary, such as route optimization and service frequency adjustments.

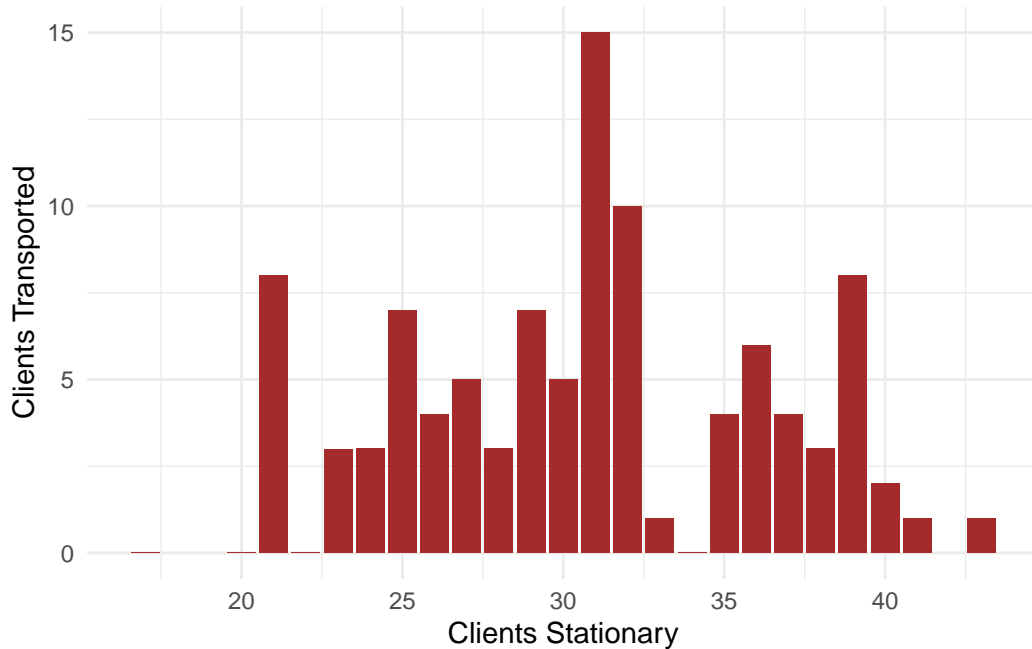


Figure 4: Analysis of Clients Stationary vs Clients Transported

3 Discussions

An examination of the transport bus initiative’s usage provides important new information about customer behavior and service dynamics. The absence of a substantial correlation between the number of clients moved and the number of clients sedentary is one of the most important findings. Although both variables show oscillations, they don’t seem to have a big impact on one another. This independence calls into question theories on the possible correlation between service demand and client behavior by indicating that immobile clients do not directly affect the number of clients being carried at any given time. As a result, transportation planners and regulators must take into account the possibility that factors other than the sheer presence of clients in stationary places may also impact transportation consumption.

On the other hand, a historical study of clientele shows distinct patterns and trends. Variations in customer involvement can be seen throughout a range of time periods, suggesting that outside variables—like neighborhood gatherings or seasonal changes—have a significant impact on transportation consumption. Significant increases in the number of clients on specific dates, for example, indicate that strategically timed interventions, like raising bus frequency or capacity, could be put into place during peak hours to improve service efficiency. To effectively build public transport policies that adapt to actual demand rather than assumptions based purely on stationary counts, it is imperative to comprehend these temporal trends.

4 Limitations

Although the results add significant insights to the subject of transportation planning, it is essential to recognize some obvious limits within the dataset. The raw data is significant since it shows cases in which no clients were transported, which raises questions regarding the data's reliability. These gaps in the transport data could be caused by a number of events, such as possible reporting errors or neglect to gather data during periods of low activity. These restrictions could compromise the analysis's robustness by masking actual usage patterns and resulting in incorrect interpretations of client behavior. To increase the accuracy and dependability of the results, future research should make an effort to guarantee more thorough data gathering procedures and include extra factors that take into account outside impacts on transportation usage.

In conclusion, this analysis highlights how complex customer behavior is within the transit bus system. The absence of a relationship between clients who are transported and clients who remain stationary implies that when organizing transport services, a comprehensive strategy is required. Simultaneously, knowing the seasonal trends in client provides useful information for improving service delivery. In order to ensure that public transportation programs successfully serve community needs and to improve the credibility of future research in this field, it will be necessary to address the constraints of the dataset.