Scientific Research Publishing

Journal of Transportation Technologies, 2024, 14, 607-653
https://www.scirp.org/journal/jtts
ISSN Online: 2160-0481
ISSN Print: 2160-0473

Comparative Analysis of the Factors Influencing Metro Passenger Arrival Volumes in Wuhan, China, and Lagos, Nigeria: An Application of Association Rule Mining and Neural Network Models

Bello Muhammad Lawan[1]*, Jabir Abubakar[2], Shuyang Zhang[1]*

[1]School of Logistics and Transportation Engineering, Wuhan University of Technology, Wuhan, China
[2]Department of Surveying and Geoinformatics, Modibbo Adama University, Yola, Nigeria
Email: *bellomuhdlawanwhut@gmail.com, Nigeria Jabir@mau.edu.ng, *zhangshuyang@gmail.com

How to cite this paper: Lawan, B.M., Abubakar, J. and Zhang, S.Y. (2024) Comparative Analysis of the Factors Influencing Metro Passenger Arrival Volumes in Wuhan, China, and Lagos, Nigeria: An Application of Association Rule Mining and Neural Network Models. *Journal of Transportation Technologies*, **14**, 607-653.
https://doi.org/10.4236/jtts.2024.144033

Received: September 13, 2024
Accepted: October 15, 2024
Published: October 18, 2024

Abstract

This study explores the factors influencing metro passengers' arrival volume in Wuhan, China, and Lagos, Nigeria, by examining weather, time of day, waiting time, travel behavior, arrival patterns, and metro satisfaction. It addresses a significant research gap in understanding metro passengers' dynamics across cultural and geographical contexts. It employs questionnaires, field observations, and advanced data analysis techniques like association rule mining and neural network modeling. Key findings include a correlation between rainy weather, shorter waiting times, and higher arrival volumes. Neural network models showed high predictive accuracy, with waiting time, metro satisfaction, and weather being significant factors in Lagos Light Rail Blue Line Metro. In contrast, arrival patterns, weather, and time of day were more influential in Wuhan Metro Line 5. Results suggest that improving metro satisfaction and reducing waiting times could increase arrival volumes in Lagos Metro while adjusting schedules for weather and peak times could optimize flow in Wuhan Metro. These insights are valuable for transportation planning, passenger arrival volume management, and enhancing user experiences, potentially benefiting urban transportation sustainability and development goals.

Keywords

Metro Passenger Arrival volume, Influencing Factor Analysis, Wuhan and Lagos Metro, Neural Network Modeling, Association Rule Mining Technique

1. Introduction

To create more sustainable and livable urban environments worldwide, policy-makers must prioritize the transition of transportation development towards a more environmentally friendly future. This necessitates a comprehensive and coordinated approach to policy development and decision-making, to improve affordable, economically viable, people-centered, and environmentally sustainable transportation systems. Metro systems are critical in urban transportation, providing significant social and economic advantages. Enhancing service quality to meet passenger needs and ensure customer retention is crucial for the sustainable growth of metros. Currently, metro station evaluation systems mainly focus on planning and do not fully consider operational realities. An effective evaluation system should prioritize user experience, accurately assess operational quality and guide improvements accordingly. Thus, it is essential to identify the key factors that influence metro passenger arrival volumes to enhance service delivery [1]. Since the advent of metro rail transit, remarkable strides have been made in the realms of transportation and communication [2]. Electric trains, originally pioneered by London in 1890, have played a pivotal role in propelling technological, commercial, and socioeconomic progress over the years [3]. Furthermore, numerous developed nations, including Italy, France, Germany, Poland, the Netherlands, Spain, and Switzerland, have made substantial investments in metro and high-speed rail systems, yielding extensive benefits across diverse domains.

Urban metro systems have been implemented by numerous cities to tackle environmental and traffic challenges resulting from high-density urbanization [4] In China, metros are crucial underground public transit systems characterized by high capacity and dedicated rights-of-way, predominantly operating through underground tunnels [5]. In the past two decades, urban metro networks have expanded rapidly in China, with a total metro system length of 5180.6 km across 37 cities on the mainland by the end of 2019 [6]. In fact, four metro systems in China were among the top ten longest metro systems worldwide by the end of 2017 [7]. The swift development of metro systems in Chinese cities in recent years has resulted in substantial impacts on society, the economy, and the environment.

The Nigerian railway industry has been in a state of decline for many years, primarily due to inadequate funding and neglect. Since gaining independence in 1960, the railway system has seen minimal restructuring. This long-term neglect has led to a significant deterioration in both freight and passenger services, as well as rolling stock, drastically reducing the system's capacity and functionality. With Nigeria being Africa's most populous nation, with a current population of around 190 million and an expected increase to 260 million by 2030, the absence of reliable rail transport and freight services has contributed to a continual decline in socioeconomic development, decreased exports, elevated transportation costs, and increased strain on the road network, resulting in traffic congestion, accidents, and pollution [8].

As per the current global trends, major urban centers such as Lagos are developing efficient modern rail mass transit systems. The Lagos Metropolitan Area Transport Authority (LAMATA) has outlined plans for a comprehensive seven-line rail network spanning approximately 246 km to address the city's long-term transportation needs. The completion of this network is projected by 2025. The initial phase encompasses two operational lines, namely the Red Line (Agbado to Marina) and the Blue Line (Okokomaiko to Marina). The Red Line, covering a distance of 31 kilometers, is designed to include a six-kilometer spur leading to the Murtala Muhammed International Airport. On the other hand, the Blue Line extends over 27 km. These lines converge at Iddo and traverse the lagoon to reach the Marina via a specially constructed suspension bridge [9]. With the successful establishment of the Red and Blue lines, long-term plans are taking shape for an additional five lines to complete the 246 km network by 2025 [10]. The Green Line is set to run eastward from Marina to Lekki airport, running parallel to the coastline. Conversely, the Yellow Line diverges from the Blue Line at the National Theatre near Iddo and proceeds northwest to Otta in Ogun State. A short branch from the Red Line at Oshodi will cater to the international and domestic terminals at Murtala Mohammed International Airport. The Brown and Orange lines will cater to the northeast, sharing the Red Line's tracks from Marina to Jibowu before heading to another junction at Ojota. The Brown Line is slated to terminate at Mile 12, while the Orange Line will continue its route north across the Long Bridge to Redeem Camp in the satellite township of Mowe/Ibafo. Lastly, the Purple Line will provide an orbital route from Ojo in the west to the Lagos-Ibadan Expressway Toll gate in the northeast, where it will link up with the Orange Line tracks to reach Redeem. Interchanges have also been indicated by Yellow and Red lines in the northern suburbs. Additionally, a monorail encircling Lagos Island will serve as the city center [10]. Understanding how weather, time of day, waiting time, travel behavior, arrival pattern and metro satisfaction impact passenger usage is crucial, given the ambitious plans of the Lagos Metropolitan Area Transport Authority (LAMATA) to develop a comprehensive rail network covering about 246 km. The completion of the Red and Blue lines, as well as the proposed expansion to include the Green, Yellow, Brown, Orange, and Purple lines by 2025, highlights the importance of studying these factors [11] [12]. However, providing high-quality service in public transportation is essential to attract more passengers [13]. This study thoroughly examines the factors that affect passenger arrival volumes at metro stations, focusing on Yujiatou Station along Wuhan Metro Line 5 in China and the Lagos Light Rail Blue Line in Nigeria. It emphasizes the significant impact of weather conditions, time of day, waiting time, travel behavior, arrival patterns, and metro satisfaction. The analysis is based on a comprehensive review of existing literature that addresses the various influences on rail transit passenger volumes, highlighting the need for a thorough and detailed approach to urban transit planning and improving passenger satisfaction. Scholars' dedication to thoroughly assessing the multitude of factors influencing rail transit passenger volumes emphasizes the necessity for a comprehensive strategy in transit planning.

By exploring the interaction between temporal variations, weather conditions, waiting times, travel patterns, arrival patterns, and satisfaction levels, this study aims to improve metro services. Understanding and addressing these influencing factors is crucial, given that efficient metro transit systems are characterized by high service frequency. The literature review provided here sets the stage for an in-depth analysis of passenger arrival volumes, their determinants, and the critical assessment of passenger waiting times, which directly correlates with overall arrival volumes. This research will provide insights and enhance metro planning methods, leading to a more efficient and passenger-focused transit service.

2. Literature Review

2.1. Temporal and Meteorological Impacts on Metro Passenger Arrival Volume

According to recent studies, metro systems are impacted by various weather conditions, resulting in both positive and negative effects. For example, warmer weather typically leads to increased ridership, while cold and windy conditions tend to decrease transit use [14]. The impact of weather on public transportation varies widely across different systems. A study analyzing daily metro ridership in Nanjing from 2011 to 2014 found that certain transportation modes are more resilient to adverse weather than others. It was also noted that weekend travelers tend to be more affected by weather conditions compared to weekday passengers [15]. Additionally, the influence of weather on travel behavior depends on the mode of transportation and its perceived comfort. Extreme weather affects different modes differently, with private cars often seen as more comfortable and reliable during warmer conditions, potentially reducing subway usage. However, subways generally maintain steady ridership due to commuters' adaptability in switching between transportation modes [16]. Weather fluctuations can also impact metro operations, leading to reliability issues and increased operational costs [17]. The exacerbation of these challenges by climate change is noted, as severe weather events impact leisure travel more than daily commuting. For instance, in New York City, rain and snowfall typically decrease public transport usage, while lower-than-normal temperatures may have a positive effect [18]. Notably, based on hourly and station-level data, weather was observed to have the most significant impact on passenger volume in the afternoon, followed by midday and morning [19]. Moreover, the time of day significantly influences passenger arrival volumes, categorizing travelers into schedule-dependent and independent groups. This distinction is evident across morning, evening, and off-peak periods, affecting how passengers choose their travel times based on metro schedules [20]. The study also revealed that passenger behavior and arrival patterns vary accordingly, with peak hours typically experiencing higher volumes compared to off-peak times [21]. Therefore, to comprehensively understand and predict passenger arrival trends throughout the day, it's essential to analyze these variations across different timeframes.

2.2. Metro Passenger Travel Behavior and Arrival Patterns

As per the findings of Jolliffe *et al*. [22]-[24], the frequency of train services is a crucial factor in shaping passenger arrival patterns. The availability of gaps in public transportation significantly affects these patterns. The impact of train frequency on passenger arrivals is manifested in various ways. Higher frequencies result in less predictable arrival patterns, leading to an average waiting time of approximately half the headway. Conversely, lower frequencies prompt passengers to strategically time their arrivals to minimize waits [25]. Research has indicated that the shift from random to non-random arrival patterns typically occurs within headways ranging from 5 to 11 minutes, regardless of frequency levels. Passenger behavior also plays a significant role in shaping arrival patterns, with individuals categorized into those adhering to schedules and arriving on time, and those arriving randomly [26].

The unpredictability and fluctuation of passenger arrivals over time can significantly impact operational efficiency, particularly during peak hours when arrival flows vary dynamically. Addressing this uncertainty often involves assuming probability distributions for arrival volumes, although accurately defining such distributions can be challenging in practical applications. Therefore, considering a range of variability in arrival volumes is often deemed more practical than relying solely on precise probability distributions [27].

2.3. Metro Passengers Waiting Time

Recent studies highlight the challenges faced by commuters in major urban areas due to extended wait times and limited train capacity in crowded train stations [28]. The average waiting times for passengers vary based on factors such as station congestion, time of day, and trip purpose. The average waiting times for passengers vary based on factors such as station congestion, time of day, and trip purpose. During peak hours, passengers typically experience shorter wait times while random arrivals are more common during off-peak periods [20] [27]. Overcrowding can lead to longer waits at the original station as passengers vie for space on the train [29]. Furthermore, waiting time is a critical aspect of the passenger experience, significantly influencing behavior and potentially causing feelings of anxiety and dissatisfaction [30]. To address this issue, recent technological advancements have revolutionized the estimation of passenger waiting times. Smart card data has proven highly effective in collecting trip information at entry and exit points within the metro system, minimizing the need for manual data collection [31].

Conversely, passengers who refer to schedules with varying arrival frequencies aim to minimize wait times by arriving close to departure times. Previous studies modeling passenger waiting times often assumed random arrivals and calculated average waiting times by multiplying the average bus headway by twice the ratio of the average headway to headway variance [32]. Initial investigations into estimating rail transportation wait times were influenced by research in bus transportation.

Researchers explored models assuming uniformly distributed wait times using random passenger arrival models, advancing their studies in this area [33] [34]. A study conducted in Zurich, Switzerland, focused on measuring wait times for buses, trams, and trains at stations, recommending the use of a mixed uniform and Johnson SB distribution for modeling purposes [20]. Another study examining waiting times for both regional and metro lines in Copenhagen considered headway ranges from 2 to 60 minutes [35].

2.4. Metro Schedule and Headway Optimization

Generally, the interval between train arrivals plays a crucial role in identifying potential conflicts during train headway and timetable optimization. Optimizing the metro system timetable is widely acknowledged as a traditional decision-making challenge that requires balancing the needs of passengers and the metro operating company [36]. Numerous research initiatives have aimed to enhance the functionality and productivity of public transportation systems. The metro timetable, closely linked to the train schedule, details arrival and departure times along with stop durations at each station. To enhance public transportation efficiency during peak hours, Beijing Metro Lines 1 and 2 have significantly reduced headway times to just two minutes. Consequently, it is imperative for operated trains to maintain consistent service frequency and adhere strictly to predetermine arrival and departure times specified in the schedule [37]. Moreover, fixed headway train timetables are inadequate for handling variable demand due to significant fluctuations in transportation needs [38]. To tackle this issue, a mixed-integer linear programming approach has been proposed [39]. Furthermore, a stochastic programming model has been proposed and refined for metro train rescheduling as a decision-making method to address railway routing and scheduling challenges [40]. Additionally, most train timetable optimization models consider the minimum headway between consecutive trains as a fixed value, although they also allow this minimum headway to be influenced by the current track assignment conditions [41]. Moreover, a mixed-integer linear programming model has been developed to optimize train schedules and reduce passenger waiting time disparities [42].

2.5. Passenger Satisfaction in Metro Systems

The satisfaction and emotional responses of travelers are shaped by their perceptions and expectations of their journey, which are influenced by subjective feelings and views of various aspects of the travel experience [43]. Currently, there is no systematic method in place for evaluating passenger satisfaction at metro stations. Previous research has explored several factors contributing to passenger satisfaction, such as accessibility, information availability, time efficiency, customer service, comfort, safety, convenience, reliability, cost-effectiveness, and capacity considerations [44]. During the evaluation process, researchers typically choose indicators based on existing literature or personal experiences. However, these indicators can be subjective and may not accurately capture the evolving dynamics of metro services. Furthermore, opinions regarding the use of these indicators can vary depending on location, objectives, and time periods. Therefore, it is essential to carefully review and adjust evaluation criteria to align with specific research goals. Assessing passenger satisfaction with public transportation services is crucial for both transportation research and practical applications. To improve infrastructure, amenities, services, and increase public transport usage, transit agencies must understand how well they meet passenger expectations. Conducting customer surveys is critical as they provide valuable insights to transit agencies about aspects significant to passengers and specific areas of satisfaction or dissatisfaction [45]. In a study focused on Metro Rail Transit 3 (MRT3) stations in Metro Manila, Philippines, [46] identified reasons for low ridership from accessibility and intermodality perspectives. The main sources of passenger dissatisfaction include station congestion, relatively high fares, and inconveniences in connecting transport facilities to other modes of transit. The remainder of this paper is organized as follows: Section 2 details the materials and method used. In Section 3, we present the main results of the study. Section 4 discusses these results and future research directions. Finally, Section 5 provide the conclusion.

3. Materials and Methods

This study employs a comprehensive methodological framework to compare subway passenger arrival volumes between the Wuhan Metro and the Nigeria Blue Line. This comparison accounts for the disparities in time granularities and the scopes of the respective studies. Data was collected from hourly station-specific entries in Wuhan and daily line-level volumes in Lagos. The Lagos data was normalized by estimating hourly arrival volumes through consistent daily traffic assumptions to enable a meaningful comparison. Meanwhile, proportional analysis focused on relative changes in passenger arrival volume patterns acquired from station-specific hourly entries within Wuhan and daily volume metrics on a line-level basis in Lagos. To facilitate an analytical comparison, the data about Lagos were subjected to normalization by deducing hourly arrival volumes via assumptions of consistent daily traffic flows. This approach enabled a detailed examination of relative alterations in passenger arrival volume patterns, employing proportional analysis methodologies. Given the different scopes of station-specific data for Wuhan and line-level data for Lagos, contextual comparisons were made by focusing on crucial transit hubs in Wuhan and treating the entire line as a single unit of analysis in Lagos. The study compared these contexts by identifying peak hours in Wuhan and contrasting them with the busiest periods in Lagos. It also analyzed the proportional distribution of passengers across different times of the day in Wuhan and compared it with Lagos's daily flow. To summarize the total passenger arrival volumes for the Lagos Light Rail Blue Line, we analyzed daily data covering weekdays and weekends. From the daily records of passenger counts, the total weekday morning peak was 4713 passengers, while the weekend morning peak was 378. Weekdays had an average of 2485 passengers for the off-peak periods, and weekends had 216.5 passengers. During the evening peak, weekdays saw 3770 passengers, and weekends had 303 passengers.

However, the dataset utilized in this research was obtained from a questionnaire distributed to respondents in Wuhan, China, and Lagos, Nigeria. Additionally, it included manual data from Yujiatou station and two other stations, as well as online data, specifically the Lagos Light Rail Blue Line Passenger arrival volume Statistics. The Wuhan questionnaire was initially drafted in English, translated into Chinese, and uploaded as a Wenjuanxing Form. Conversely, the Lagos questionnaire was created and uploaded in English as a Google form, with hyperlinks and QR codes generated for both. These questionnaires were then emailed to randomly selected respondents in both cities via social media and emails. The data were collected between May 2024 and July 2024, yielding 365 valid responses for the Lagos questionnaire and 403 for the Wuhan questionnaire.

The questionnaire gathered information on various variables, including weather, time of day, waiting time, arrival pattern, travel behavior, and metro satisfaction. The manual collection was conducted at Yujiatou Station and two other stations over six weeks to obtain the necessary data. Data were collected during both peak and off-peak hours to provide a comprehensive understanding of passenger behavior. A total of 24 questions were designed to cover all the essential variables of the study, such as passenger arrival volume, time of day, weather conditions, waiting time, travel behavior, arrival pattern, and metro satisfaction. The questionnaire aimed to gather comprehensive and meaningful data by incorporating questions targeting these variables. The questions included:

1) Demographic, gender, age, and level of education (Questions 1 - 3)

2) Metro passenger arrival volume (Questions 4 - 6)

3) Weather-related aspects (Questions 7 - 9)

4) Time of day-related aspects (Questions 10 - 12)

5) Waiting time-related aspects (Questions 13 - 15)

6) Metro satisfaction (Questions 16 - 18), which asked:

Whether satisfactory service leads to higher passenger arrival volumes

Whether satisfaction with the metro influences the decision to use it

Whether overall satisfaction with the metro positively affects the number of passengers

7) Travel behavior (Questions 19 - 21), which sought to determine:

How frequently do participants use the metro for commuting to work, social events, or school/university?

To what extent do participants agree that their travel behavior, such as taking alternative modes of transportation or adjusting travel time, influences the rate of passenger arrivals at the metro station?

8) Arrival patterns (Questions 22 - 24), which asked:

To what extent do participants agree that the pattern of passenger arrivals influences the overall passenger volume at the metro station?

How important is understanding the passenger arrival pattern when planning a trip to the metro station?

Whether participants observe a consistent pattern in the arrival of passengers at metro stations throughout the day?

As presented in the **Appendix C**.

Additionally, the following steps were taken to ensure the accuracy of the subsequent analyses: the collected data were cleaned to remove inconsistencies or missing values. Additionally, data preprocessing involved normalizing the data and encoding categorical variables as needed. However, to evaluate the internal consistency of the questionnaire, we calculated Cronbach's alpha, which is a widely used measure of reliability. Table 1 below shows the internal consistency of the dependent variable passenger arrival volume is 0.817, and independent variables such as weather is 0.834, time of day is 0.826, waiting time as 0.811, metro satisfaction as 0.808, travel behaviour as 0.837 and arrival pattern as 0.842 which indicates high internal consistency, exceeding the recommended threshold of 0.7.

Table 1. Reliability analysis results.

Variable	Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Passenger Arrival volume	PAR1	50.48	92.192	0.735	0.808
	PAR2	50.64	97.169	0.392	0.825
	PAR3	50.66	94.276	0.511	0.818
Weather	WTH1	50.66	98.862	0.343	0.828
	WTH2	50.78	97.719	0.331	0.829
	WTH3	52.48	110.132	-0.170	0.846
Time of day	TOD1	52.56	108.038	-0.044	0.841
	TOD2	52.33	107.579	-0.015	0.840
	TOD3	50.75	87.197	0.809	0.799
Waiting time	WAT1	50.92	93.279	0.527	0.817
	WAT2	50.94	90.321	0.644	0.810
	WAT3	50.83	89.523	0.674	0.808
Metro satisfaction	MS1	50.99	89.980	0.638	0.810
	MS2	50.88	88.142	0.719	0.804
	MS3	50.91	90.470	0.640	0.810
Travel behaviour	TB	51.26	100.956	0.203	0.837
Arrival pattern	AP	51.73	106.380	0.021	0.842

3.1. Model Selection

3.1.1. Association Rule Mining

The Apriori algorithm is a data mining technique used to discover patterns and associations within a dataset. It identifies frequent item sets and generates association rules based on these item sets. The algorithm evaluates the strength and relevance of the generated association rules using support, confidence, and lift measurements. The algorithm operates by incrementally enlarging the item sets until no additional frequent item sets can be discovered. It uses the following measures of significance and interest:

- Support ($S(x)$): The proportion of responses in the dataset containing the item set.
- Confidence: The likelihood of the rule being true.
- Lift: The ratio of the observed support to the expected support if the two items were independent. This technique has been effectively used in studies by [47] [48]. Association rule mining is a well-established method employed to uncover relationships among variables within a large dataset, offering flexibility and not requiring a dependent variable. In this study, the Apriori algorithm was selected for its flexibility [49]. The arules package in R was used for analysis.

The specifics of the algorithm are as follows: Let $I=\{i_1,i_2,\cdots,i_n\}$ represent the set of factors influencing metro passenger arrival volume referred to as the item set, and $D=\{t_1,t_2,\cdots,t_n\}$ denote the set of responses from individual respondents, known as the data set; every response in D possesses a unique identifier and includes a subset of the items in I. A rule of an item set is expressed as $\{X\} \Rightarrow \{Y\}$ where: $X,Y \subseteq I$ and $X \cap Y = \varphi$ (X and Y are disjoint items). The sets of items X are called antecedent (or Left-hand side LHS) and sets of items Y consequent (or Right-hand side RHS) of the rules. There are mainly three measures of significance and interest which are the support, confidence, and lift. The support $S(x)$ of an item set is the proportion of responses in the dataset which contains the item set given as:

$$S(x)=\frac{f(x)}{N} \tag{1}$$

$$S(x)=\frac{f(y)}{N} \tag{2}$$

$$S(x \Rightarrow y)=\frac{f(x \cap y)}{N} \tag{3}$$

where:

$f(x)$ = Number of Instances with x

$f(y)$ = Number of Instances with y

N = Total number of Instances

$f(x \cap y)$ = Number of Instances with both x and y

$S(x)$ = Support of x item set

$S(y)$ = Support of y item set

$S(x \Rightarrow y)$ = Support of the association

The confidence is an estimate of probability $P(xy)$ of finding Consequent (RHS) of the rule in instances under the condition that these instances also contain the antecedent (LHS). The confidence is given as:

$$\text{Confidence}(x \Rightarrow y) = \frac{S(x \Rightarrow y)}{S(x)} \tag{4}$$

The Lift is the deviation of the support of the whole rule from the support expected under independence given the support of LHS and RHS, with a greater value indicating a better association.

$$\text{Lift}(x \Rightarrow y) = \frac{S(x \Rightarrow y)}{S(x)S(y)} \tag{5}$$

The Apriori algorithm, implemented using the open-source R programming language and the "arules" and "arulesviz" packages, was employed to investigate the relationship between high, moderate and low metro passenger arrival, with other factors such as waiting time, weather, and time of day, arrival pattern, travel behaviour and metro satisfaction.

3.1.2. Neural Network Models

Neural network models were deployed to detect nonlinear patterns in the data. The neural network architecture featured multiple hidden layers, and the training process incorporated backpropagation and the Adam optimizer. The effectiveness of the neural network models was assessed using metrics such as accuracy, precision, recall, and F1-score. The input data is transmitted to the hidden layers for processing, and the final hidden layer forwards the processed information to the output layer and receives the outcomes. This investigation employs a fully connected neural network, where each neuron in one layer is connected sequentially to every neuron in the subsequent layers, encompassing the input, hidden, and output layers. This approach is consistent with studies by [50] [51]. The process of deriving output data is delineated by the following equation:

$$Y_k^{n+1} = f\left(\sum_{i=1}^{N} X_i^n w_{ki}^n + b_i^n\right) \tag{6}$$

where, Y_k^{n+1} X is output of unit k in the nth layer, f is the function of activation, X_i^n is the input vector, w_{ki}^n is a weight vector, b_i^n is the bias weight.

The initial weights are typically assigned randomly at the beginning of neural network training, which involves adjusting these weights through backpropagation comprises two main phases: feedforward and backward propagation. A training set, consisting of input vectors and corresponding target output vectors, is provided to the network for learning. The network's actual output is compared with the target output to calculate an error, which is then used to update the weights by propagating them. Iterative weight adjustments are performed for each training set until a stopping condition, such as a predefined number of epochs or a specified threshold, is met. The backpropagation algorithm consists of three key stages:

1) Feedforward Stage: The input layer computes the output by summing the weighted inputs and biases up to the output layer using a specified activation function.

2) Backpropagation stage: The error, obtained by comparing the network output with the target output, is calculated and propagated backward through the network starting from the output layer.

3) Weight and Bias Update Stage: In this final phase, the weights are adjusted to minimize errors based on the back propagated error signals.

3.1.3. Comparison of Models

Evaluation metrics were used to compare the performances of linear regression, neural network models, and association rule mining to determine the most effective approach for predicting passenger arrival. The complete methodology process is presented in Figure 1 below.

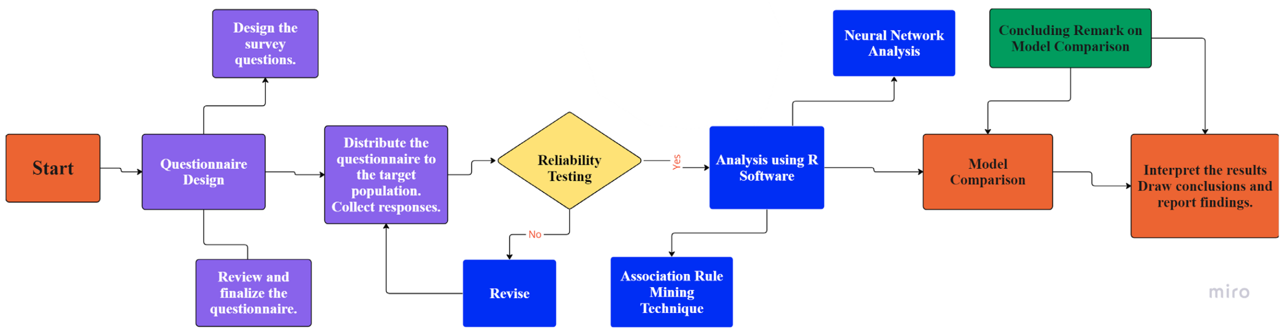

Figure 1. Research methodology framework.

3.1.4. Abbreviations and Acronyms

Considering your recent experiences how would you rate the overall volume of passengers arriving at metro stations?	PAR1
During the busiest times of the day, how would you rate the level of crowding on the metro?	PAR2
Comparing to a year ago, how would you describe the change in metro passenger arrival volume?	PAR3
Weather conditions (e.g. rain, snow, heat) impact metro passenger arrival volumes	W1
The metro passenger arrival volume significantly increases during peak summer/winter.	W2
What type of weather-related disruptions are most likely to stop you from coming to the metro station?	W3
How likely are you to use alternative transportation or adjust your travel schedule due to congestion or delays during peak hours?	TOD1
How would you rate the difference in metro passenger arrivals between peak and off-peak hours?	TOD2
I perceive a significant increase in metro usage during evening peak hours.	TOD3
Longer waiting times at metro stations result in more people using the metro.	WAT1
Shorter waiting times lead to higher passenger arrival volumes at metro stations.	WAT2
Perceived waiting time influences the number of passengers arriving at metro stations.	WAT3
Satisfactory metro service leads to higher passenger arrival volumes.	MSS1
I am more likely to use the metro when satisfied with its service.	MSS2
Overall satisfaction with the metro positively influences the number of passengers using it.	MSS3
How often do you change your travel plans in response to real-time information about metro passenger volume or congestion, which may impact passenger arrival volumes?	TB1

Continued

To what extent do you agree that your travel behaviour, such as taking alternative modes of transportation or adjusting your travel time influences the rate of passenger arrivals at metro station?	TB2
How frequently do you use the metro for commuting to work, social events, or school/university?	TB3
To what extent do you agree that the pattern of passenger arrivals influence the overall passenger volume at the metro station?	AP1
How important is understanding the passenger arrival pattern to you when planning your trip to the metro station?	AP2
Do you observe a consistent pattern in the arrival of passengers at metro stations throughout the day?	AP3

4. Results

4.1. Direct Field Observation of Metro Passengers Arrival Volume

During a period of six weeks, we collected passenger arrival volume data at Yujiatou, Qingnian Road, and Jiangshe 2nd Road stations, focusing on both weekday and weekend passenger arrival volume during morning peak hours, off-peak hours, and evening peak hours. The data reveals some significant patterns. You can find the summary of this data in the provided Table 2 and Figure 2, Figure 3 and Figure 4 below. For a comprehensive view of the entire dataset, which includes all stations and specific time intervals, please refer to the **Appendix A**.

Table 2. Summary of metro passenger arrival volume observations.

Station	Time Period	Highest Passenger Count	Lowest Passenger Count	Remarks
Yujiatou	Morning Peak (Week 1)	1754 (Monday)	407 (Sunday)	Highest on Monday, sharp drop on weekends.
	Evening Peak (Week 4)	1257 (Friday)	653 (Sunday)	Friday peak, with gradual decline toward Sunday.
Qingnian Road	Morning Peak (Week 5)	1670 (Monday, Friday)	490 (Sunday)	Consistent high on weekdays, lowest on Sunday.
	Evening Peak (Week 5)	2041 (Friday)	1134 (Saturday)	Highest passenger arrival volume on Friday evening.
Jiangshe 2nd Rd	Morning Peak (Week 6)	1131 (Monday)	613 (Sunday)	Higher weekday traffic, Sunday low.
	Off-Peak (Week 6)	624 (Saturday)	555 (Wednesday)	Off-peak hours show relatively consistent traffic.

The pie charts in Figure 2 show that weekday passenger volumes are consistently higher than weekends, reflecting typical commuter patterns. During summer holidays, more off-peak hour travel on weekends indicates increased leisure travel. This highlights the impact of summer holidays on metro travel patterns, with heightened off-peak and weekend travel volumes due to non-commuter activities.

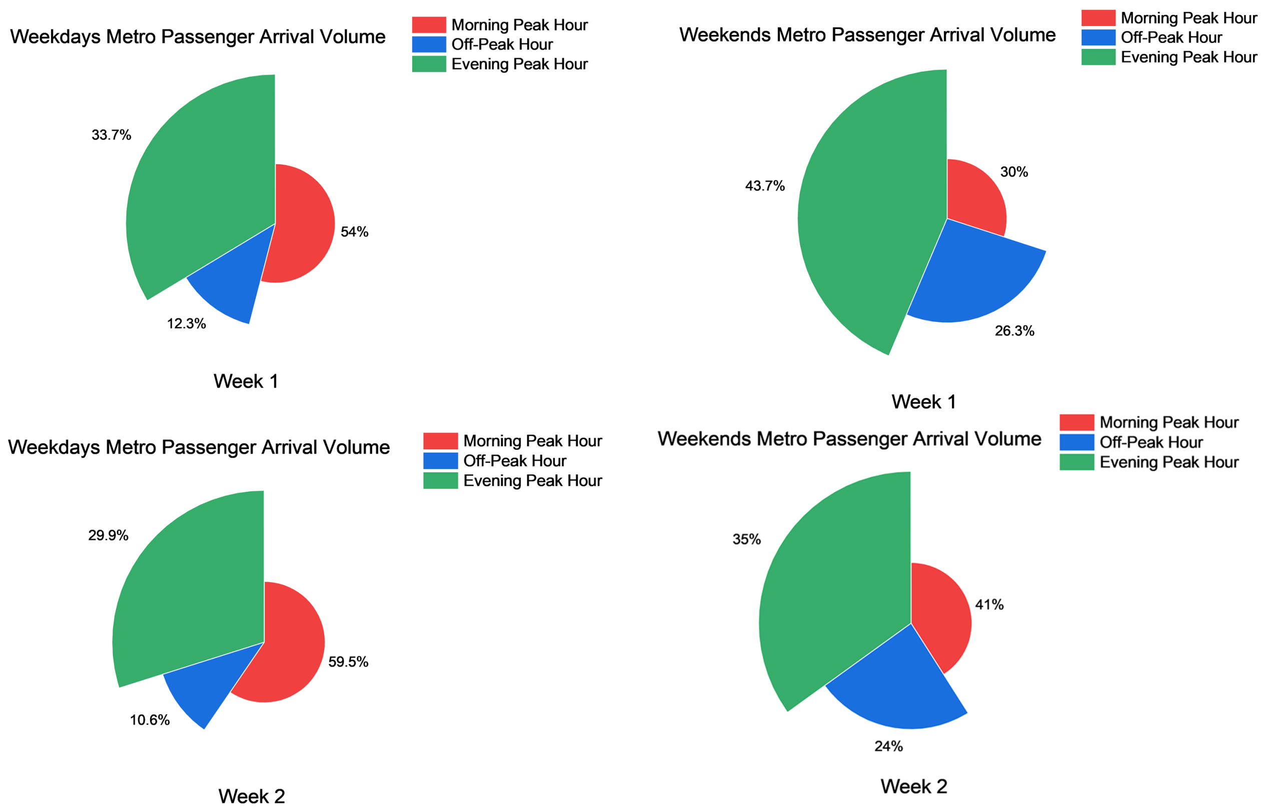

Figure 2. Weekdays & weekends metro passenger arrival volume for both week 1 & week 2.

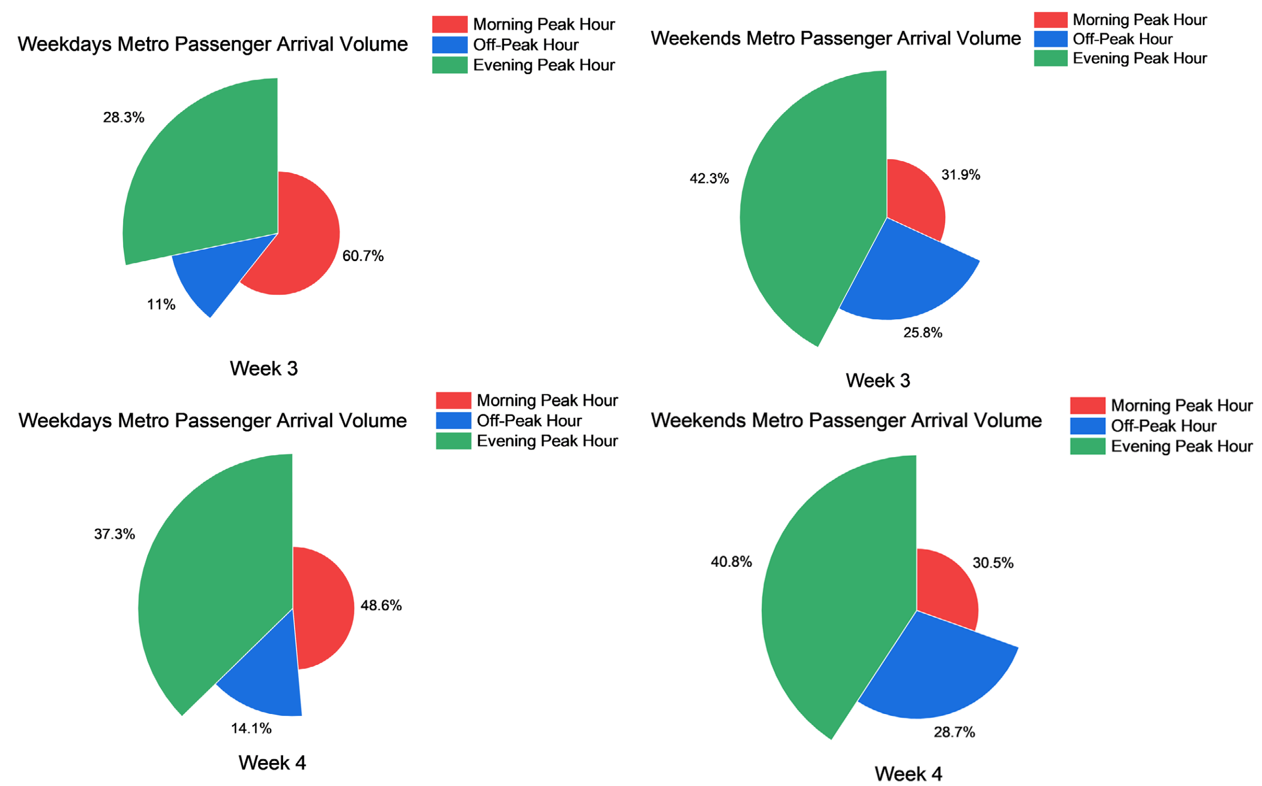

Figure 3. Weekdays & weekends metro passenger arrival volume for both week 3 & week 4.

Figure 3 Shows that In July 2023, there was increased off-peak hour and weekend travel, indicating more leisure travel due to school closures and vacations. In May 2024, there were dominant weekday morning peak hours and increased evening peak hour travel, suggesting more evening activities or later commutes.

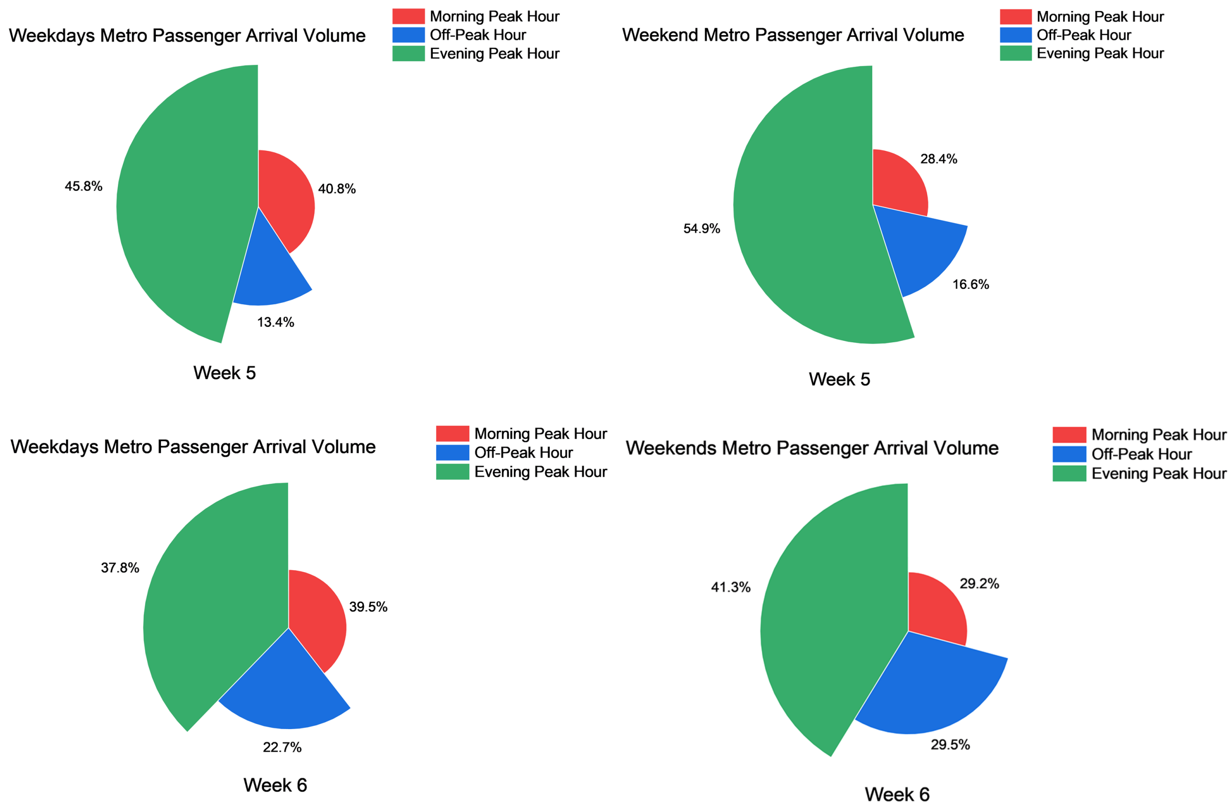

Figure 4. Weekdays & weekends metro passenger arrival volume for both week 5 & week 6.

Figure 4 above shows typical weekdays at the Qingnian Road Station. The higher percentage of evening peak hour dominance suggests that many commuters used this station to return home from work or school. In contrast, the lower off-peak percentage indicates that the station primarily serves as a hub for commuting, with less leisure or non-work-related travel occurring during the middle of the day. Compared to the Qingnian Road Station, the Jiangshe 2nd Road Station I had a noticeable increase in off-peak travel during weekends. A higher off-peak percentage on weekends indicates more leisure travel, with individuals using the station for activities such as shopping, dining, or visiting friends and family. However, evening peak hours remained dominant, suggesting significant evening activity.

4.2. Lagos Blue Line Passenger arrival volume

Table 3 and Figure 5, Figure 6 and Figure 7 below, shows a summarized information on significant fluctuations in average daily passenger arrival volume over several months, with peaks near 9000 passengers and troughs around 3000. Weekdays have a much higher average daily passenger arrival volume (approximately

5886 passengers) compared to weekends (around 1765 passengers), indicating the dominance of work and school commutes. Passengers flow is relatively balanced between mornings and evening on weekends, reflecting more flexible travel patterns. The detailed information is in the **Appendix B**.

Table 3. Summary of the Lagos Light Rail Blue Line Passenger arrival volume.

Metric	Value
Total Days Recorded	200
Total Passenger arrival volume	778,422
Average Daily Passenger arrival volume	3892.11
Max Daily Passenger arrival volume	10,901 (11/3/2023)
Min Daily Passenger arrival volume	0 (10/15/2023)
Max Daily Train Trips	54 (from 10/16/2023 onward)
Min Daily Train Trips	10 (9/4/2023)
Total Train Trips Recorded	8812
Average Daily Train Trips	44.06
Max Daily Revenue (Confidential)	Confidential
Min Daily Revenue (Confidential)	Confidential
Highest Passenger arrival volume on a Weekday	10,901 (11/3/2023)
Lowest Passenger arrival volume on a Weekday	343 (9/10/2023)
Highest Passenger arrival volume on a Weekend	6440 (11/4/2023)
Lowest Passenger arrival volume on a Weekend	159 (2/11/2024)

Source: LAMATA Lagos Nigeria.

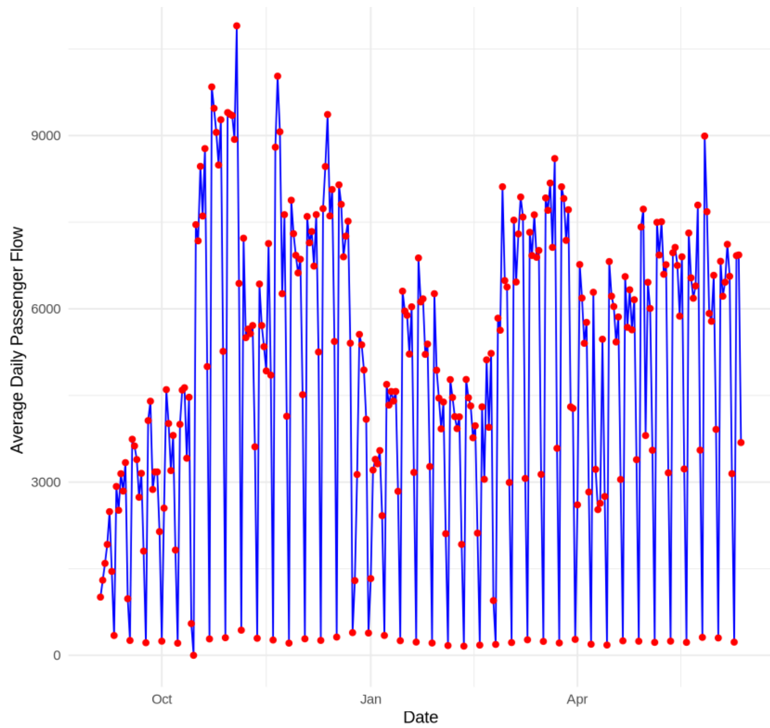

Figure 5. Average daily passenger arrival volume over time in lagos blue line.

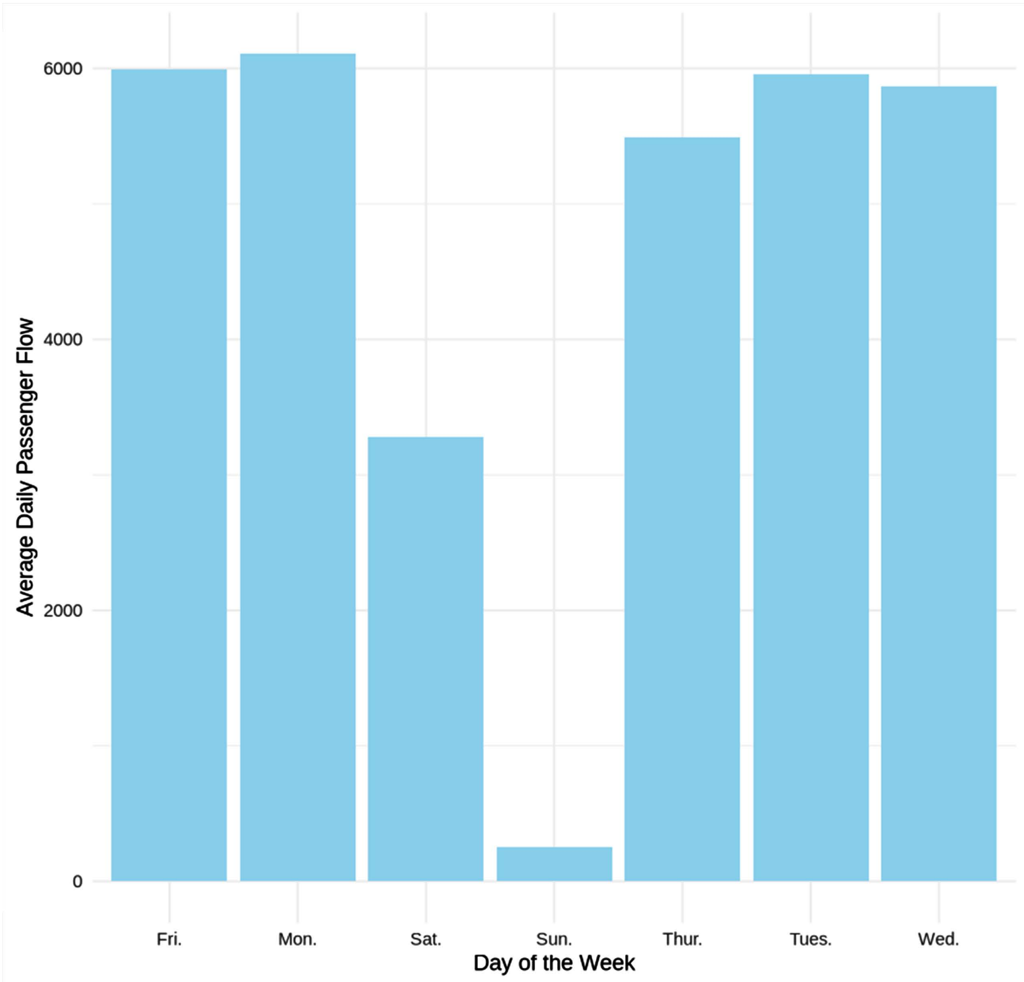

Figure 6. Average daily passenger arrival volume by day of the week in lagos blue line.

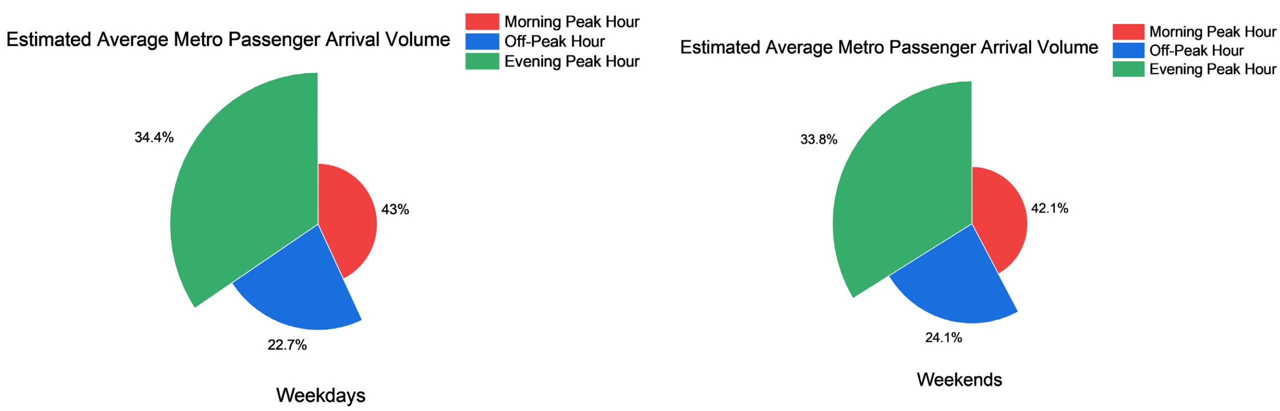

Figure 7. Estimated average metro passenger arrival volume in lagos blue line.

Figure 5 and Figure 6 show significant fluctuations in average daily passenger arrival volume over several months, with peaks nearing 9000 passengers and troughs around 3000. Weekdays have a much higher average daily passenger arrival volume (approximately 5886 passengers) than weekends (around 1765 passengers). Figure 7 reinforces this, showing a more significant portion of weekday passenger arrival volume. This emphasizes the need for optimized metro services to accommodate higher weekday demand while maintaining balanced weekend operations.

4.3. Association Rule Mining Technique

The Apriori algorithm analyzed metro passenger arrival volumes, generating 420 association rules grouped into high, moderate, and low classes. Strong correlations were found between high passenger volumes and longer waiting times, increased evening peak usage, high metro satisfaction, shorter waiting times, perceived waiting times, and weather conditions. Moderate volumes had significant associations with longer waiting times, disagreement on waiting times, neutral stance on perceived waiting times, and disagreement on weather conditions affecting arrival volumes. Low volumes strongly disagreed with perceived waiting times, shorter waiting times, overall metro satisfaction, increased evening peak usage, and weather conditions, all of which were strongly linked to low passenger volumes. These results highlight the impact of waiting times, peak periods, satisfaction levels, and weather conditions on metro passenger volumes. As presented in Table 4 and Figure 8, Figure 9 and Figure 10 below.

Table 4. Association rule mining result for PAR1.

	PAR1		Support	Confidence	Lift
Considering your recent experiences, how would you rate the overall volume of passengers arriving at metro stations?	High	{WAT1 = Strongly Disagree}	0.246	0.346	1.369
		{TOD3 = Strongly Agree}	0.315	0.442	1.363
		{MSS3 = Strongly Agree}	0.249	0.35	1.352
		{WAT2 = Strongly Agree}	0.239	0.336	1.350
		{WAT3 = Strongly Agree}	0.292	0.410	1.345
		{WTH2 = Strongly Agree}	0.295	0.415	1.332
	Moderate	{MSS1 = Neutral}	0.095	0.518	2.589
		{WAT1 = Disagree}	0.046	0.25	2.061
		{WAT3 = Neutral}	0.082	0.446	2.032
		{WTH1 = Disagree}	0.046	0.25	2.007
		{WAT2 = Neutral}	0.082	0.446	2.002
	Low	{WAT3 = Strongly Disagree}	0.052	0.5	5.259
		{WAT2 = Strongly Disagree}	0.056	0.531	4.629
		{MSS3 = Strongly Disagree}	0.049	0.469	4.205
		{TOD3 = Strongly Disagree}	0.033	0.313	4.144
		{WTH1 = Strongly Disagree}	0.016	0.156	3.971

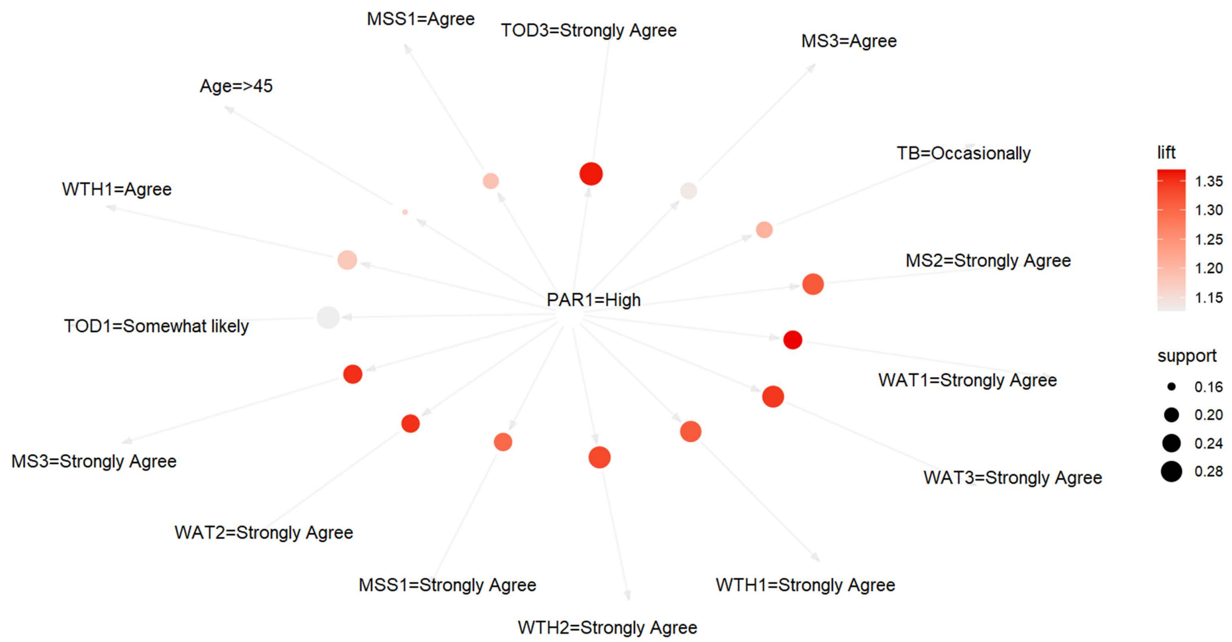

Figure 8. PAR 1 High.

Figure 8 above shows a strong correlation between passenger volumes and factors such as waiting times, weather impact, and peak hour usage at metro stations. It uses color intensity to represent the strength of these associations visually. Darker colors indicate stronger connections, while lighter colors indicate weaker ones. This visualization helps identify the factors most strongly correlated with high passenger volumes at metro stations.

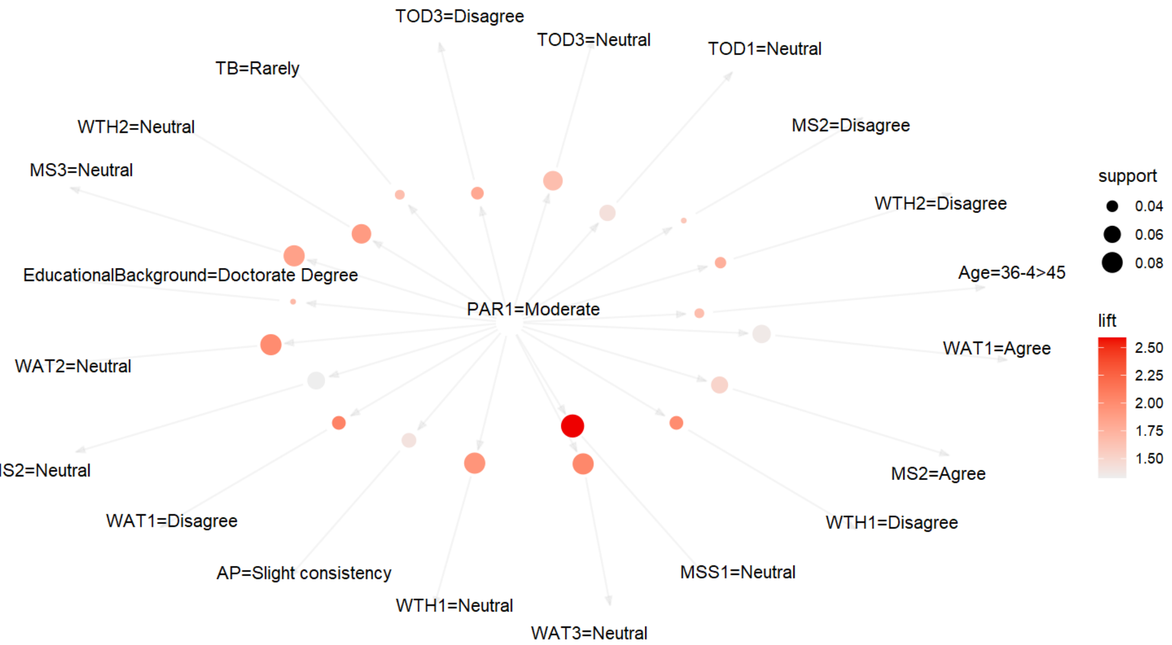

Figure 9. PAR 1 Moderate.

From **Figure 9** above, the visualization highlights that waiting times (both agreement and neutral stances) and weather conditions (neutral and disagreement stances) are significant factors that influence moderate passenger volumes at metro stations.

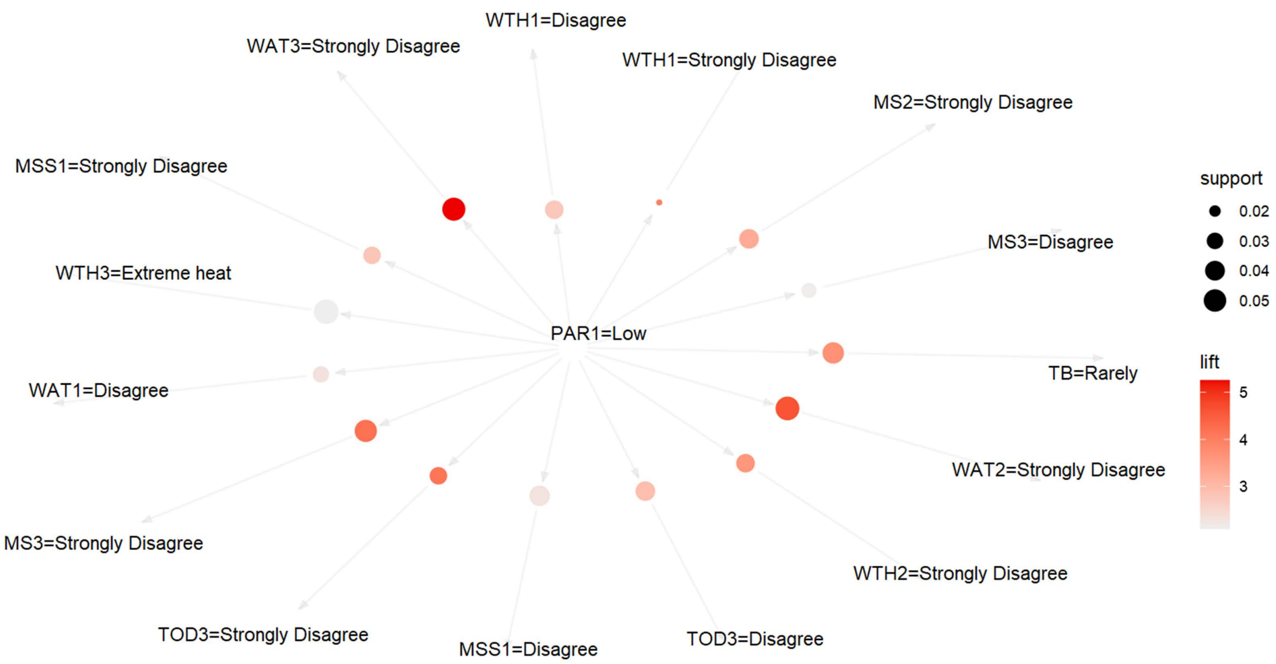

Figure 10. PAR 1 Low.

The figure highlights that the strongest correlations with low passenger volumes are between "WAT2 = Strongly Disagree" and "WAT3 = Strongly Disagree," indicated by the highest lift values. Moderate correlations are seen in relationships such as "WTH1 = Strongly Disagree" and "MSS2 = Strongly Disagree." Meanwhile, "WTH2 = Strongly Disagree" and "WTH3 = Extreme heat" show weaker, still significant, associations with lower lift values.

4.4. Descriptive Statistics

In the forthcoming analysis, **Table 5** provides a comprehensive summary of the socio-demographic data obtained for this study. Three hundred sixty-five valid responses were collected from participants in Lagos, Nigeria. Among the respondents, 48.2% were male and 51.8% were female. The age distribution of the participants is as follows: under 18 years (19.2%), 18 years old (28.2%), 18 - 25 years (18.9%), 26 - 35 years (24.7%), and 36 - 45 years (9.0%). Regarding educational background, 46.8% of the participants are high school students, 36.4% are bachelor's students, 11.5% are master's students, and the remaining 5.2% are pursuing doctoral degrees. Additionally, 403 valid responses were obtained from participants in Wuhan, China. Among them, 48.4% were male and 51.6% were female. The age distribution in this group is as follows: under 18 years (18.9%), 18 years old (27.3%), 18 - 25 years (19.6%), 26 - 35 years (25.6%), and 36 - 45 years (8.7%). Regarding educational status, 47.6% of the participants are high school students, 30.0% are pursuing bachelor's degrees, 11.4% are master's students, and 5.0% are working towards doctoral degrees.

Table 5. Demographic distribution.

Variable	Code	Description	N	Percentage %	N	Percentage %
			Lagos		Wuhan	
Demographic Distribution	Gender	MALE	176	48.2	195	48.4
		FEMALE	189	51.8	208	51.6
	Age	under18	70	19.2	76	18.9
		18	103	28.2	110	27.3
		18 - 25	69	18.9	79	19.6
		26 - 35	90	24.7	103	25.6
		36 - 45	33	9.0	35	8.7
	Educational level	high school or below	171	46.8	192	47.6
		bachelor	133	36.4	145	30.0
		master	42	11.5	46	11.4
		doctorate	19	5.2	20	5.0

4.5. Neural Network Model for Lagos Data Set

Neural networks are a type of machine learning model that closely resembles the human brain in structure and function. They excel at processing complex patterns and large amounts of data, making them particularly well-suited for tasks such as image recognition, natural language processing, and predictive analytics.

4.5.1. Regression Metrics

A lower MSE indicates superior model performance because it signifies that the predictions are closely aligned with the actual values. In this instance, an MSE of 0.1493 suggests that the model's predictions are fairly accurate, although there is still potential for enhancement. R-squared values range from 0 to 1, with higher values denoting better model performance. An R-squared value of 0.8666 implies that approximately 86.66% of the variation in the dependent variable (Passenger Arrival) can be accounted for by the model. This suggests a robust association between the predictors and response variables. As presented in the Table 6 below.

Table 6. Regression metric results.

Metric	Value
Mean Squared Error	0.1493
R-squared	0.8666

4.5.2. Feature Importance

In this instance, feature importance is calculated using a method that involves shuffling the values of each feature and then measuring the increase in the model's error. The greater the error increase, the more significant the feature. The results indicate that waiting time was the most crucial feature in this model, with an importance score of 7.0581, signifying its substantial contribution to prediction accuracy, followed by MetroSatisfaction, weather conditions, time of day, ArrivalPattern, and general travel behavior. As presented in Table 7 and Figure 11 below.

Table 7. Feature importance results.

Feature	Importance
Weather	1.0316
Timeofday	0.8904
WaitingTime	7.0581
ArrivalPattern	0.8571
TravelBehaviour	0.6643
MetroSatisfaction	6.1193

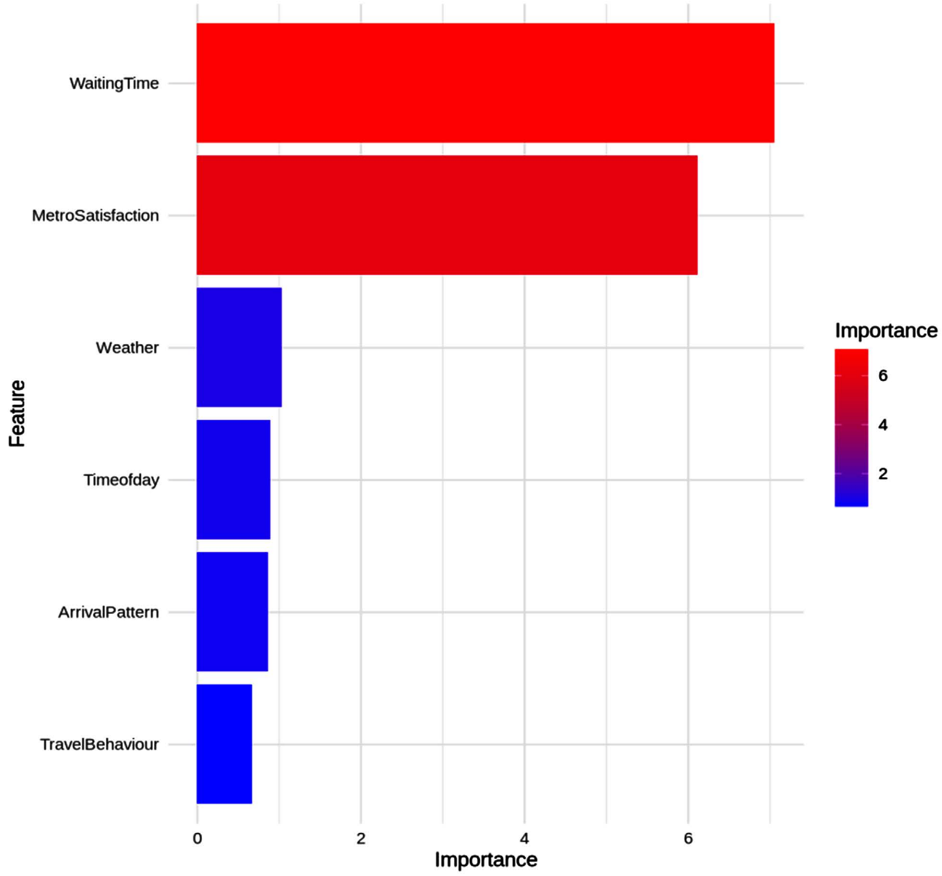

Figure 11. Feature importance in neural network.

The aforementioned Figure 11 illustrates the relationship between various features and passenger arrivals. Among these, Waiting Time is the most influential factor, implying that reducing waiting time can substantially enhance passenger arrival. Similarly, Metro Satisfaction has a high level of influence, suggesting that improving overall satisfaction can have a positive effect on passenger arrival. The impact of the weather was considered moderate, highlighting the importance of implementing weather adaptation strategies. The time of day also had a moderate effect, emphasizing the significance of optimizing operations based on time-of-day data. Finally, arrival patterns and travel behavior have a moderate influence, indicating the need for further analysis and a deeper understanding of these factors.

4.5.3. Model Prediction

The model's predictions were impressively accurate, with minimal residuals for each prediction. The model consistently demonstrated a balanced mix of positive and negative residuals, indicating well-calibrated performance. Additionally, there were no discernible patterns in the residuals, suggesting that the model effectively captured the underlying data patterns without overfitting or under fitting. As presented in **Table 8** and **Figure 12** and **Figure 13** below.

Table 8. Predicted and residual results.

Predicted	Residual
3.9926	0.0074
4.9648	0.0352
2.9851	0.0149
1.3190	0.0143
4.0079	-0.0079
2.3303	0.0030

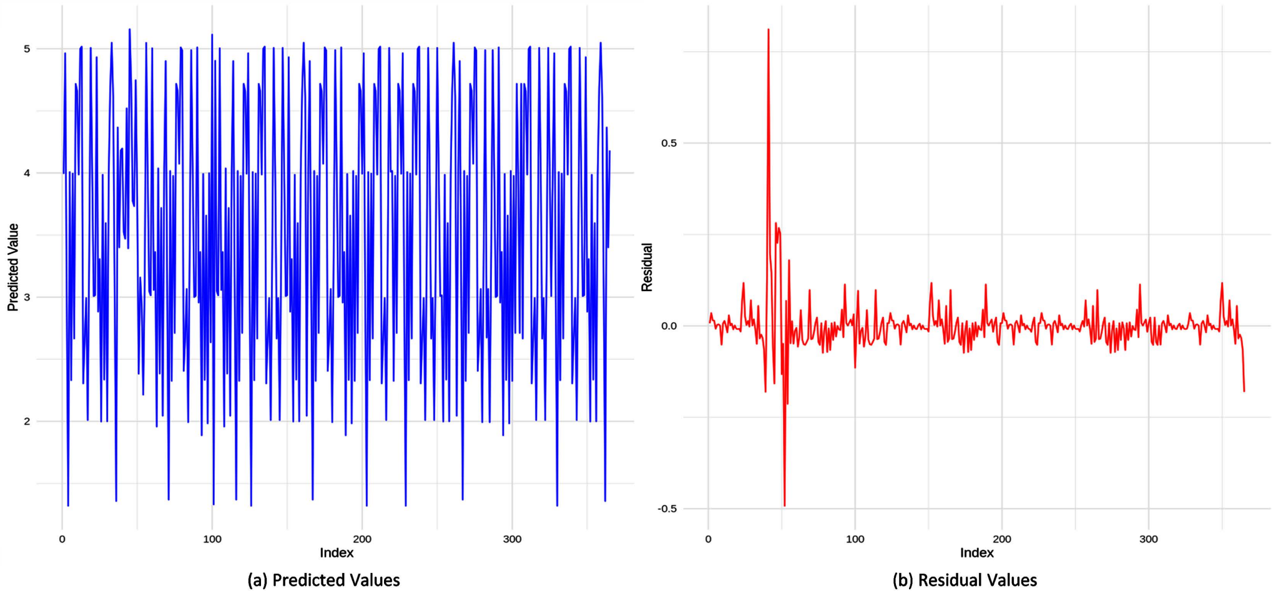

Figure 12. Predicted and residual values.

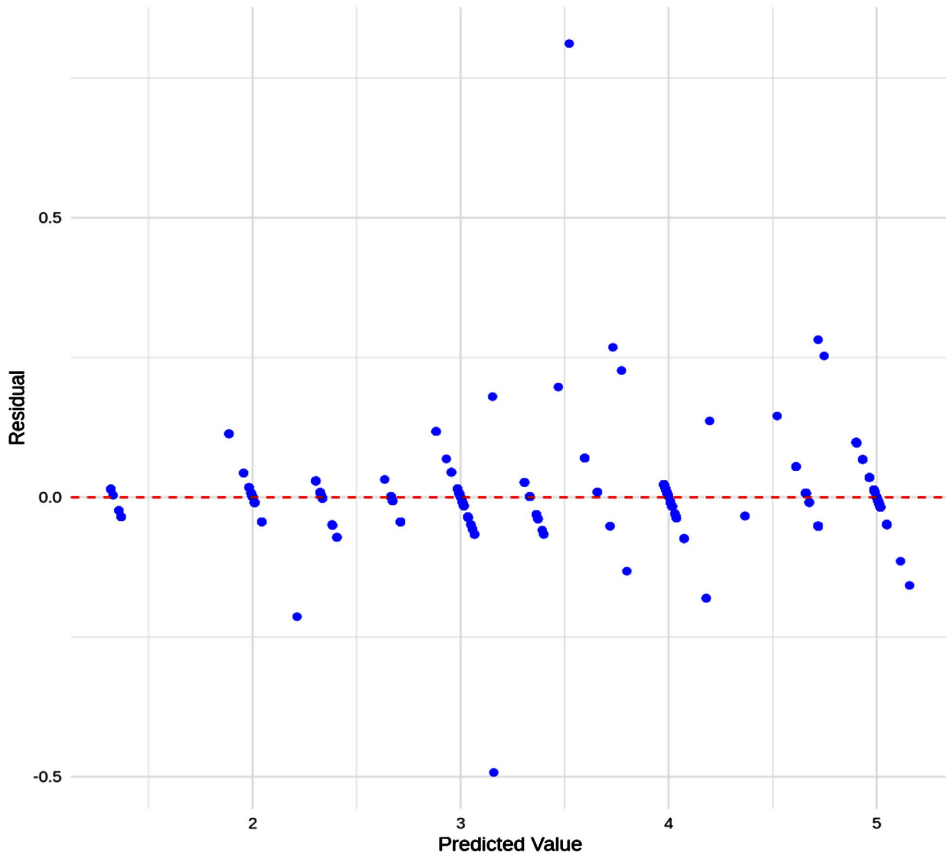

Figure 13. Predicted vs residual values.

The predicted value plot above shows the target variable's forecasted values and the residual plot displays variances between actual and predicted values. The scatter plot indicates that the residuals are randomly distributed without any systematic pattern, confirming the model's high accuracy and absence of significant biases.

4.6. Neural Network Model for Wuhan Data Set

Neural networks are a type of machine learning model that closely resembles the human brain in structure and function. They excel at processing complex patterns and large amounts of data, making them particularly well-suited for tasks such as image recognition, natural language processing, and predictive analytics.

4.6.1. Regression Metrics

The regression metric results provide an evaluation of the model's performance. The Root Mean Squared Error (RMSE) is 0.1970, indicating the average magnitude of the prediction errors, with lower values suggesting better model performance. The Mean Absolute Error (MAE) is 0.0921, representing the average absolute difference between the predicted and actual values, with smaller values indicating more accurate predictions. The R-squared value is 0.5615, which means that approximately 56.15% of the variance in the target variable is explained by the model. This suggests a moderate level of explanatory power, indicating that while the model captures some of the variability in the data, there is still room for improvement. As presented in **Table 9** below.

Table 9. Regression metric results.

Metric	Value
RMSE	0.1970
MAE	0.0921
R-squared	0.5615

4.6.2. Permutation Feature Importance

The importance of each feature is determined by its influence on the model's predictions, known as feature importance. A higher value indicates a more significant impact on the model's output. The feature "ArrivalPattern" has the highest importance score of 0.0659, "Weather" follows with a score of 0.039, "Time of day" has a score of 0.0331, "TravelBehaviour" with a score of 0.0322, and "MetroSatisfaction" has the lowest importance score of 0.0277. As presented in **Table 10** and **Figure 14** below.

Table 10. Permutation feature importance results.

Feature	Importance
Weather	0.039
Time of day	0.0331
Arrival Pattern	0.0659
Travel Behaviour	0.0322
Metro Satisfaction	0.0277

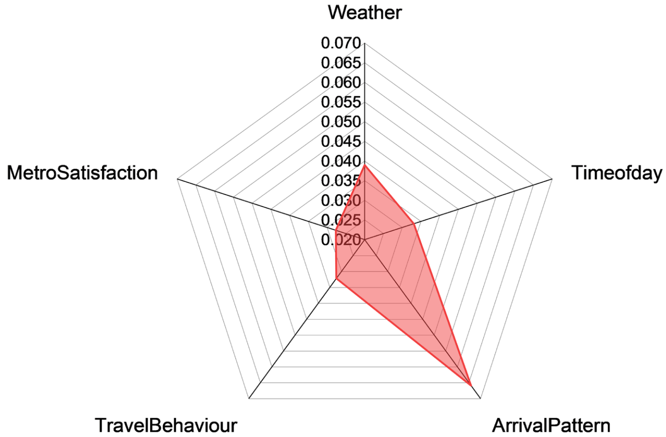

Figure 14. Permutation feature importance.

The radar chart above visually represents the relative importance of five features (Weather, Time of day, arrival pattern, travel behavior, and MetroSatisfaction) in predicting the target variable. Each feature is plotted on an axis radiating from the center, with importance scores ranging from 0 to 0.07. The polygon formed by connecting the data points shows that ArrivalPattern has the highest importance, followed by Weather and Timeofday, while TravelBehaviour and MetroSatisfaction have lower importance scores. The chart's shape and size provide a quick visual comparison, highlighting the most influential features clearly and intuitively.

4.6.3. Model Prediction

The Predicted and Residual Results table compares the predicted and actual values for passenger arrival, demonstrating that the neural network model's predictions are generally close to the exact values. For instance, the expected value of 2.8202 is near the actual value of 2.6667, and the predicted value of 3.0717 is close to 3.0000. The minor discrepancies between the predicted and actual values, known as residuals, indicate slight overestimations by the model, such as a residual of 0.1535 for the first row and 0.0717 for the second row. These small residuals suggest that the model's predictions are pretty accurate, but there is still room for improvement. The consistency of the residuals across different rows indicates that the model is systematically close to the actual values, but fine-tuning the model or incorporating additional features could further enhance its predictive accuracy. As presented in Table 11 and Figure 15 below.

Table 11. Predicted and residual results.

Predicted	Actual
2.8202	2.6667
3.0717	3.0000
2.2914	2.3333
2.6445	2.6667
2.5540	2.6667
2.6034	2.6667

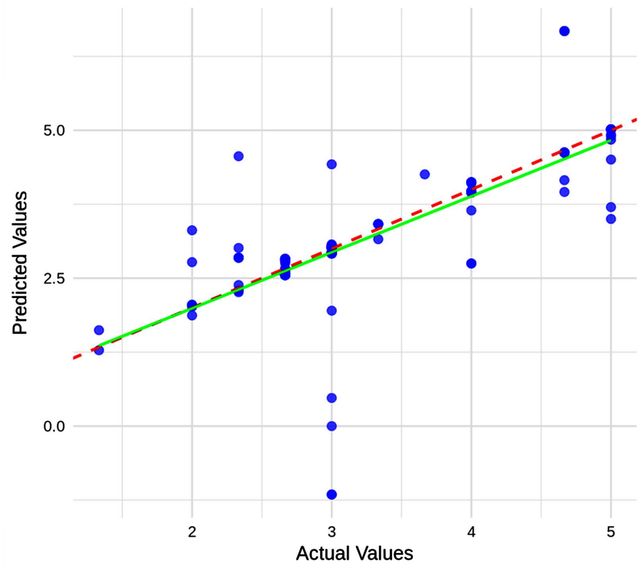

Figure 15. Model prediction vs actual values.

The scatter plot includes more prominent, semi-transparent blue points to represent the model's predictions, a red dashed line indicating the ideal scenario where predicted values match actual values, and a green solid line showing the linear regression fit. The plot also features a title, subtitle, and improved axis labels for better readability.

However, the feature Arrival Pattern shows positive and negative impacts on the model's predictions, with contributions ranging from −0.0565 to 0.7223 and a variable value of 0.03585, indicating its variability in influencing passenger arrival. Metro Satisfaction consistently has a negative impact, with contributions ranging from −0.7273 to −0.7105 and a variable value of −0.9683, suggesting that lower satisfaction decreases passenger arrival predictions. Time of Day consistently has a positive impact, with contributions ranging from 0.0057 to 0.4969 and a variable value of 1.081, indicating that certain times of the day increase passenger arrival predictions. Travel Behaviour consistently has a negative impact, with contributions ranging from −0.2914 to −1.4838 and a variable value of 1.306, suggesting that certain travel behaviors decrease passenger arrival predictions. Lastly, Weather consistently has a positive impact, with contributions ranging from 0.1149 to 0.2093 and a variable value of 0.6323, indicating that certain weather conditions increase passenger arrival predictions. As presented in **Table 12** and **Figure 16** below.

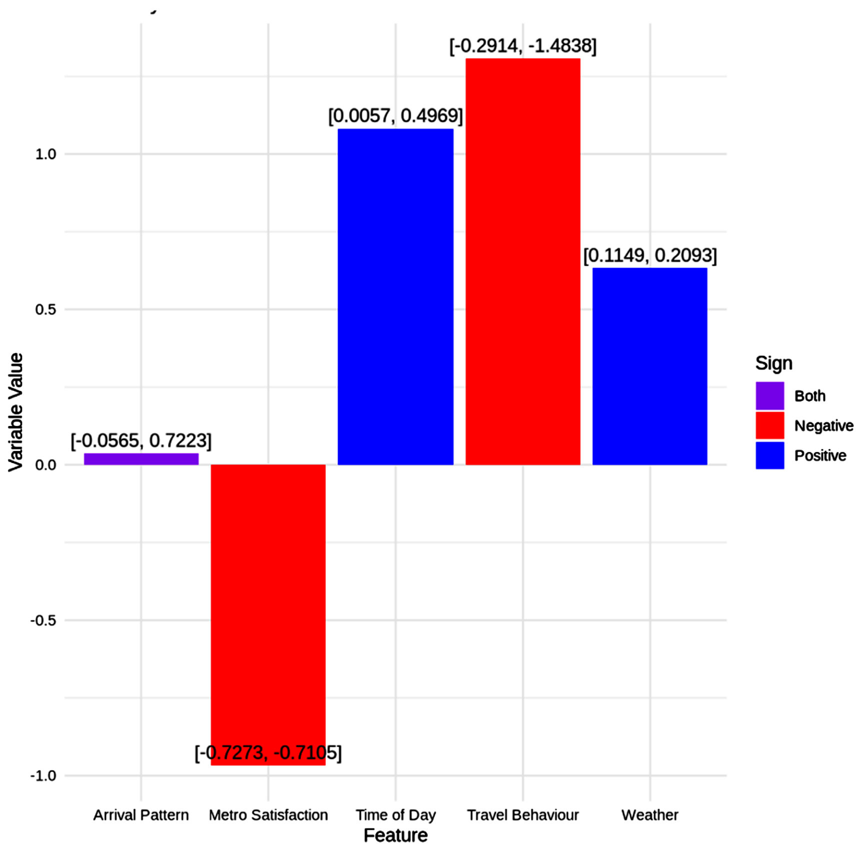

Figure 16. Absolute SHAP values for the most influential features.

Table 12. Absolute SHAP values for each features.

Variable	Contribution	Variable value	Sign	Label	B
ArrivalPattern = 0.03585	−0.05651111	0.03585	−1	Neural Network	0
MetroSatisfaction = −0.9683	−0.72726619	−0.9683	−1	Neural Network	0
Timeofday = 1.081	0.00566966	1.081	1	Neural Network	0
TravelBehaviour = 1.306	−0.29139648	1.306	−1	Neural Network	0
weather = 0.6323	0.2093084	0.6323	1	Neural Network	0
MetroSatisfaction = −0.9683	−0.71054409	−0.9683	−1	Neural Network	0

The bar plot above, shows the variable values for each feature, with colors indicating the sign of the contribution (positive, negative, or both). The contribution range is annotated above each bar, providing additional context on each feature's impact variability. This visualization helps summarize and compare each feature's influence on the model's predictions. Please let me know if you have any further questions or need additional analysis.

4.7. Model Comparison

4.7.1. Association Rule Mining

High passenger volumes are commonly associated with extended waiting times, weather-related disruptions, and peak-hour usage. On the other hand, moderate passenger volumes tend to have neutral stances on waiting times and weather impact. In contrast, low passenger volumes are strongly correlated with strong disagreement on waiting times, weather impact, and metro satisfaction. The use of color intensity and point size in visualizing association rules offers a clear and concise understanding of the strength of the relationships between these various factors and passenger traffic volumes.

4.7.2. Neural Network Model

A Mean Squared Error (MSE) measure of 0.149 indicates that the model's predictions were fairly precise. The R-squared value of 0.866 suggests that the model can account for approximately 86.6% of the variation in passenger arrival volume. Among the features, waiting time had the highest importance score of 7.058, followed by metro satisfaction, with a score of 6.119 and weather with a score of 1.031. Although, time of day, arrival pattern, and travel behavior also played a significant role, their importance scores were lower. The model predictions were remarkably close to the actual values, with residuals ranging between -0.007 and 0.035 units for the Lagos Blue Light Rail Line.

However, for the Wuhan data set, the R-squared value of 0.561 suggests that the model can account for approximately 56.1% of the variation in passenger arrival volume. Among the features Arrival pattern had the highest importance score of 0.065, weather had 0.039 followed by time of day, with a score of 0.033. The expected value of 2.820 is near the actual value of 2.666, and the predicted value of 3.071 is close to 3.000, suggesting that the neural network model's predictions are generally close to the exact values.

4.7.3. Concluding Remark

After a comprehensive analysis of the two models, it is clear that the neural network model surpasses in terms of predictive capabilities due to its superior accuracy and reduced error rates. Additionally, association rule mining provides additional insights by discovering intricate relationships that might not be apparent through conventional regression analysis. The neural network model's superior predictive performance, as demonstrated by its increased accuracy and decreased error rates, makes it the preferred choice among the evaluated models. With its more dependable and precise approach to forecasting passenger arrival volumes, this model can be invaluable for transportation planning and operations.

5. Discussion

During the field observation at the Yujiatou, Jiangshe 2nd Road line 5, and Qingnian Road line 2 in Wuhan, there were distinct patterns in passenger arrival volumes across different times of the day and weather conditions. On weekdays, morning and evening peak hours consistently showed higher passenger volumes than off-peak hours, reflecting typical commuter patterns. Visual representations in **Figure 1**, **Figure 2**, and **Figure 3** indicate the proportion of passenger volume at different times of the day, categorized by weekdays and weekends for each week. The data consistently showed that passenger volumes on weekdays were more significant than on weekends, reflecting typical commuter patterns. In July 2023, during the summer vacation, there was an increase in off-peak hour travel on weekends, indicating more leisure travel. In May 2024, not coinciding with the peak holiday season, there was a more balanced distribution of travel times, with a noticeable increase in evening peak hour travel compared to July 2023. This is in line with the findings of previous studies [15].

The association rule mining analysis revealed significant correlations between various factors and passenger arrival volumes. One rule showed a strong association between rainy weather in the evening and higher passenger arrival volumes, with a confidence level of 41.5%. Another rule indicated that extreme weather conditions, such as rain and heat, have a minor impact on passenger arrivals despite a low support value but significantly and negatively impact passenger arrivals when they occur. These findings are consistent with previous studies, as supported by the manual data **Table 2**. However, the neural network model, which offers accuracy with an MSE of 0.149 and an R-squared value of 0.86, suggests that weather is not the strongest predictor of passenger volume as shown in **Table 6**. The study further notes that the impact of weather on travel behavior varies depending on the mode of transportation used and is contingent on it. The seasonal variations observed in the data, with increased off-peak and weekend travel during the summer holiday, suggest that seasonal factors can significantly influence passenger arrival patterns and the overall usage of the metro system. These findings are consistent with the of previous studies [16] [18] [19].

Similarly, association rule mining reveals that passengers are 1.258 times more likely to travel during peak hours than off-peak hours **Table 4**. This is consistent with typical commuter patterns, where passenger volumes are higher during peak hours owing to work schedules and daily routines. The neural network model also supports this, with importance value of 0.89. The manual data, as shown in **Table 2**, also supports this finding, with higher passenger volumes recorded during the morning and evening peak hours compared to off-peak hours. Moreover, the arrival patterns and travel behaviour possess a lift value of 1.125 and 1.150 as shown in **Figure 8**. This study also demonstrated that passengers exhibit both random and non-random arrival patterns, which is supported by the literature [22]-[24]. The results however, indicate a strong correlation between shorter waiting times and higher passenger arrival volumes. The neural network model shows waiting time as the most significant feature, with an importance rating of 7.058 as shown in **Figure 11**. Additionally, association rule mining revealed a positive correlation, with a 33.6% probability of observing higher passenger arrival volumes when waiting times are shorter and a lift of 1.350, indicating that passenger arrival volumes are 1.350 times more likely to be higher when waiting times are shorter as presented in **Table 4**.

Moreover, the neural network model developed for the Wuhan dataset demonstrates promising performance, with a Root Mean Squared Error (RMSE) of 0.1970, indicating an average prediction error magnitude of 0.1970. The Mean Absolute Error (MAE) is 0.0921, suggesting that the model's predictions are, on average, within 0.0921 of the actual values. The R-squared value of 0.5615 indicates that the model explains approximately 56.15% of the variance in the target variable, suggesting a moderate explanatory power **Table 9**. The permutation feature importance analysis reveals that the Arrival Pattern is the most influential feature, with an importance score of 0.0659, followed by Weather (0.039), Time of Day (0.0331), Travel Behaviour (0.0322), and Metro Satisfaction (0.0277) **Table 10** and **Figure 14**. While the SHAP analysis shows that Arrival Pattern has positive and negative impacts on the model's predictions, with contributions ranging from −0.0565 to 0.7223 and a variable value of 0.03585, indicating its variability in influencing passenger arrival. Metro Satisfaction consistently has a negative impact, with contributions ranging from −0.7273 to −0.7105 and a variable value of −0.9683, suggesting that lower satisfaction decreases passenger arrival predictions. Time of Day consistently has a positive impact, with contributions ranging from 0.0057 to 0.4969 and a variable value of 1.081, indicating that certain times of the day increase passenger arrival predictions. Travel Behaviour consistently has a negative impact, with contributions ranging from −0.2914 to −1.4838 and a variable value of 1.306, suggesting that certain travel behaviors decrease passenger arrival predictions. Lastly, Weather consistently has a positive impact, with contributions ranging from 0.1149 to 0.2093 and a variable value of 0.6323, indicating that certain weather conditions increase passenger arrival predictions **Table 12** and **Figure 16**. Based on the findings of this study, the following recommendations are proposed for future research:

Replicating the study in other metropolitan areas or transportation systems could help validate the findings and explore the potential influence of regional or cultural differences on passenger arrival patterns.

Exploring advanced data collection techniques, such as sensor-based systems or intelligent card data, could provide more comprehensive and reliable data, enabling a deeper analysis of the research problem.

Investigating the impact of socioeconomic, demographic, and other contextual factors on passenger arrival patterns could yield additional insights and enhance the understanding of the underlying dynamics.

Conducting a longitudinal study over an extended period could provide valuable insights into the long-term trends and the influence of seasonal or other temporal factors on passenger arrival patterns.

Expanding the research to include the interactions between different modes of transportation, such as buses, trains, and private vehicles, could offer a more comprehensive understanding of passenger travel behavior and its implications for the overall transportation system.

By addressing these recommendations, future research can build upon the foundations laid by this study and contribute to the ongoing efforts to optimize metro systems and enhance the overall passenger experience.

6. Conclusions

In conclusion, the impact of weather, time of day, waiting time, metro satisfaction, arrival pattern, and travel behavior differs significantly between Wuhan Metro and Lagos Light Rail Blue Line, mainly due to efficiency, reliability, and advancement variations. The advanced nature of Wuhan Metro minimizes the influence of these factors, whereas, in the developing and underdeveloped Lagos Metro, they play a significantly impactful role. Within this context, this research has chosen to adopt the Neural Network Model for analysis in both cities. However, it's crucial to acknowledge the study's limitations that could affect its findings and future applications:

In the absence of Automatic Fare Collection Data, the reliance on direct field observation, passenger arrival volume statistics, revenue, and other online sources from Wuhan and Lagos may limit the depth and accuracy of the analysis.

Despite neural networks' promising performance and other machine learning models' accuracy and generalizability could be enhanced by incorporating more diverse factors.

The failure to consider the potential influence of land use, socio-economic, and demographic factors has possibly omitted critical context that could offer a more comprehensive understanding of the observed phenomena.

Acknowledging these limitations is essential for interpreting the results and guiding future research directions, especially in enhancing the Lagos Metro's development and improvement, where the insights from this study are expected to be most beneficial.

Acknowledgements

I am profoundly thankful for the collective efforts of all those involved, making this research a truly rewarding and fulfilling endeavor. The successful completion of this study is a testament to the support, guidance, and contribution of many individuals. I am deeply grateful for their involvement with this study.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Li, L., Gao, T., Yu, L. and Zhang, Y. (2022) Applying an Integrated Approach to Metro Station Satisfaction Evaluation: A Case Study in Shanghai, China. *International Journal of Transportation Science and Technology*, **11**, 780-789. https://doi.org/10.1016/j.ijtst.2021.10.004

[2] Kurniawati, W. (2023) The Impact Analysis of First Phase Construction Project of MRT (Mass Rapid Transit) on Economic Growth in Jakarta Using Input-Output Analysis.

[3] Duffy, M.C. (2003) Electric Railways.

[4] Lin, D., Nelson, J.D. and Cui, J. (2021) Exploring Influencing Factors on Metro Development in China from Urban and Economic Perspectives. *Tunnelling and Underground Space Technology*, **112**, Article ID: 103877. https://doi.org/10.1016/j.tust.2021.103877

[5] Krüsmann, V. (2019) Mobility in 21st Century China: Snapshots, Dynamics & Future Perspectives. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

[6] China Association of Metros (CAMET) (2020) Statistical Analysis Report on Urban Rail Transit in China 2019. China Association of Metros. (In Chinese)

[7] UITP-2018-Travel-for-All-2. https://cms.uitp.org/wp/wp-content/uploads/2020/08/UITP-2018-Travel-for-all-2.pdf

[8] Ayilaran, H. (2019) Structural Analysis of Nigeria's Rail Industry: A Switch from Nationalisation to Privatisation. https://www.researchgate.net/publication/331894621

[9] Alao, T. (2008) Lagos in $2b Rail Lines Project Drive. http://www.ngrguardiannews.com/business/article01//indexn2_html?pdate=101108&ptitle=Lagos%20in%20$2b%20rail%20lines%20project%20drive

[10] Mobereola, D. (2008) Africa's Megacity Needs an Urban Rail Backbone. https://www.railwaygazette.com/africas-megacity-needs-an-urban-rail-backbone/33447.article

[11] Mannino, C. and Mascis, A. (2009) Optimal Real-Time Traffic Control in Metro Stations. *Operations Research*, **57**, 1026-1039.

https://doi.org/10.1287/opre.1080.0642

[12] Wang, Y., Tang, T., Ning, B., van den Boom, T.J.J. and De Schutter, B. (2015) Passenger-Demands-Oriented Train Scheduling for an Urban Rail Transit Network. *Transportation Research Part C*: *Emerging Technologies*, **60**, 1-23. https://doi.org/10.1016/j.trc.2015.07.012

[13] Watkins, K.E., Ferris, B., Borning, A., Rutherford, G.S. and Layton, D. (2011) Where Is My Bus? Impact of Mobile Real-Time Information on the Perceived and Actual Wait Time of Transit Riders. *Transportation Research Part A*: *Policy and Practice*, **45**, 839-848. https://doi.org/10.1016/j.tra.2011.06.010

[14] Arana, P., Cabezudo, S. and Peñalba, M. (2014) Influence of Weather Conditions on Transit Ridership: A Statistical Study Using Data from Smartcards. *Transportation Research Part A*: *Policy and Practice*, **59**, 1-12. https://doi.org/10.1016/j.tra.2013.10.019

[15] Li, J., Li, X., Chen, D. and Godding, L. (2018) Assessment of Metro Ridership Fluctuation Caused by Weather Conditions in Asian Context: Using Archived Weather and Ridership Data in Nanjing. *Journal of Transport Geography*, **66**, 356-368. https://doi.org/10.1016/j.jtrangeo.2017.10.023

[16] Wu, J. and Liao, H. (2020) Weather, Travel Mode Choice, and Impacts on Subway Ridership in Beijing. *Transportation Research Part A*: *Policy and Practice*, **135**, 264-279. https://doi.org/10.1016/j.tra.2020.03.020

[17] Guo, Z., Wilson, N.H.M. and Rahbee, A. (2007) Impact of Weather on Transit Ridership in Chicago, Illinois. *Transportation Research Record*: *Journal of the Transportation Research Board*, **2034**, 3-10. https://doi.org/10.3141/2034-01

[18] Cravo, V.S., Cohen, J.E. and William, A. (2009) The Impact of Weather on Transit Revenue in New York City. *Proceedings of the* 88*th Annual Meeting of the Transportation Research Board*, Vol. 3036, 14 p. https://trid.trb.org/View/882059

[19] Singhal, A., Kamga, C. and Yazici, A. (2014) Impact of Weather on Urban Transit Ridership. *Transportation Research Part A*: *Policy and Practice*, **69**, 379-391. https://doi.org/10.1016/j.tra.2014.09.008

[20] Marco, L., Ulrich, W. and Andrew, N. (2007) Passenger Arrival Rate at Public Transport Stations. New York City. *Proceedings of the* 86*th Annual Meeting Compendium of Papers of the Transportation Research Board*, Washington, 01-05 January 2007.

[21] Li, S., Yang, L. and Gao, Z. (2018) Optimal Switched Control Design for Automatic Train Regulation of Metro Lines with Time-Varying Passengers Arrival Flow. *Transportation Research Part C*: *Emerging Technologies*, **86**, 425-440. https://doi.org/10.1016/j.trc.2017.11.025

[22] Jolliffe, J.K. and Hutchinson, T.P. (1975) A Behavioural Explanation of the Association between Bus and Passenger Arrivals at a Bus Stop. *Transportation Science*, **9**, 248-282. https://doi.org/10.1287/trsc.9.3.248

[23] O'Flaherty, C.A. and Mangan, D.O. (1990) Bus Passenger Waiting Time in Central Areas. *Traffic Engineering and Control*, **11**, 419-421.

[24] Seddon, P.A. and Day, M.P. (1974) Bus Passenger Waiting Time in Great Manchester. *Traffic Engineering and Control*, **15**, 422-445.

[25] Singh, R., Graham, D.J., Hörcher, D. and Anderson, R.J. (2021) The Boundary between Random and Non-Random Passenger Arrivals: Robust Empirical Evidence and Economic Implications. *Transportation Research Part C*: *Emerging Technologies*, **130**, Article ID: 103267. https://doi.org/10.1016/j.trc.2021.103267

[26] Frumin, M. and Zhao, J. (2012) Analyzing Passenger Incidence Behavior in Heterogeneous Transit Services Using Smartcard Data and Schedule-Based Assignment. *Transportation Research Record*: *Journal of the Transportation Research Board*, **2274**, 52-60. https://doi.org/10.3141/2274-05

[27] Yin, J., Tang, T., Yang, L., Gao, Z. and Ran, B. (2016) Energy-Efficient Metro Train Rescheduling with Uncertain Time-Variant Passenger Demands: An Approximate Dynamic Programming Approach. *Transportation Research Part B*: *Methodological*, **91**, 178-210. https://doi.org/10.1016/j.trb.2016.05.009

[28] Li, W., Yan, X., Li, X. and Yang, J. (2020) Estimate Passengers' Walking and Waiting Time in Metro Station Using Smart Card Data (SCD). *IEEE Access*, **8**, 11074-11083. https://doi.org/10.1109/access.2020.2965155

[29] Lüthi, M., Weidmann, U. and Nash, A. (2007) ETH Library. Passenger Arrival Rates at Public Transport Stations. Conference Paper.

[30] Csikos, D. and Currie, G. (2007) Impacts of Transit Reliability on Wait Time: Insights from AFC Data. *Transportation Research Board* 86*th Annual Meeting*, Washington DC, 21-25 January 2007, Paper No. 07-0544. https://trid.trb.org/view/1154348

[31] Qu, H., Xu, X. and Chien, S. (2020) Estimating Wait Time and Passenger Load in a Saturated Metro Network: A Data-Driven Approach. *Journal of Advanced Transportation*, **2020**, Article ID: 4271871. https://doi.org/10.1155/2020/4271871

[32] Ingvardson, J.B., Nielsen, O.A., Raveau, S. and Nielsen, B.F. (2018) Passenger Arrival and Waiting Time Distributions Dependent on Train Service Frequency and Station Characteristics: A Smart Card Data Analysis. *Transportation Research Part C*: *Emerging Technologies*, **90**, 292-306. https://doi.org/10.1016/j.trc.2018.03.006

[33] Welding, P.I. (1957) The Instability of a Close-Interval Service. *Journal of the Operational Research Society*, **8**, 133-142. https://doi.org/10.1057/jors.1957.21

[34] Osuna, E.E. and Newell, G.F. (1972) Control Strategies for an Idealized Public Transportation System. *Transportation Science*, **6**, 52-72. https://doi.org/10.1287/trsc.6.1.52

[35] Abkowitz, M.D. (1981) An Analysis of the Commuter Departure Time Decision. *Transportation*, **10**, 283-297. https://doi.org/10.1007/bf00148464

[36] Hu, Y., Li, S., Wang, Y., Zhang, H., Wei, Y. and Yang, L. (2023) Robust Metro Train Scheduling Integrated with Skip-Stop Pattern and Passenger Flow Control Strategy under Uncertain Passenger Demands. *Computers & Operations Research*, **151**, Article ID: 106116. https://doi.org/10.1016/j.cor.2022.106116

[37] Le, Z., Li, K., Ye, J. and Xu, X. (2014) Optimizing the Train Timetable for a Subway System. *Proceedings of the Institution of Mechanical Engineers*, *Part F*: *Journal of Rail and Rapid Transit*, **229**, 852-862. https://doi.org/10.1177/0954409714524377

[38] Li, X. and Lo, H.K. (2014) Energy Minimization in Dynamic Train Scheduling and Control for Metro Rail Operations. *Transportation Research Part B*: *Methodological*, **70**, 269-284. https://doi.org/10.1016/j.trb.2014.09.009

[39] Pellegrini, P., Marlière, G. and Rodriguez, J. (2014) Optimal Train Routing and Scheduling for Managing Traffic Perturbations in Complex Junctions. *Transportation Research Part B*: *Methodological*, **59**, 58-80. https://doi.org/10.1016/j.trb.2013.10.013

[40] Yang, X., Chen, A., Ning, B. and Tang, T. (2016) A Stochastic Model for the Integrated Optimization on Metro Timetable and Speed Profile with Uncertain Train Mass. *Transportation Research Part B*: *Methodological*, **91**, 424-445. https://doi.org/10.1016/j.trb.2016.06.006

[41] Lee, Y. and Chen, C. (2009) A Heuristic for the Train Pathing and Timetabling Problem. *Transportation Research Part B*: *Methodological*, **43**, 837-851. https://doi.org/10.1016/j.trb.2009.01.009

[42] Wu, Y., Yang, H., Zhao, S. and Shang, P. (2021) Mitigating Unfairness in Urban Rail Transit Operation: A Mixed-Integer Linear Programming Approach. *Transportation Research Part B*: *Methodological*, **149**, 418-442. https://doi.org/10.1016/j.trb.2021.04.014

[43] Eboli, L. and Mazzulla, G. (2009) A New Customer Satisfaction Index for Evaluating Transit Service Quality. https://creativecommons.org/licenses/by-nc/4.0

[44] del Castillo, J.M. and Benitez, F.G. (2013) Determining a Public Transport Satisfaction Index from User Surveys. *Transportmetrica A*: *Transport Science*, **9**, 713-741. https://doi.org/10.1080/18128602.2011.654139

[45] Le-Klähn, D.-T., Hall, C.M. and Gerike, R. (2014) Analysis of Visitors' Satisfaction with Public Transport in Munich, Germany. *Journal of Public Transportation*, **17**, 68-85. https://www.researchgate.net/publication/269109228

[46] Doi, K. and Aoki, T. (2003) Quantification of User's Preference for the Improvement of Railway Stations Considering Human Latent Traits: A Case Study in Metro Manila. *Doboku Gakkai Ronbunshu*, **625**, 15-27.

[47] Agrawal, R., Imieliński, T. and Swami, A. (1993). Mining Association Rules between Sets of Items in Large Databases. *Proceedings of the* 1993 *ACM SIGMOD International Conference on Management of Data*, Washington DC, 26-28 May 1993, 207-216. https://doi.org/10.1145/170035.170072

[48] Furth, P.G. and Muller, T.H.J. (2006) Service Reliability and Hidden Waiting Time: Insights from Automatic Vehicle Location Data. *Transportation Research Record*: *Journal of the Transportation Research Board*, **1955**, 79-87. https://doi.org/10.1177/0361198106195500110

[49] Agrawal, R. and Srikant, R. (1994) Fast Algorithms for Mining Association Rules.

[50] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 27-30 June 2016, 770-778. https://doi.org/10.1109/cvpr.2016.90

[51] Goodfellow, I., Bengio, Y. and Courville, A. (2016) Deep Learning. MIT Press. http://www.deeplearningbook.org/

Appendices

Appendix A

Table A1. Direct field observation of metro passenger arrival volume.

Date: 15th-21st May 2023 (Week 1)

Yujiatou Station

	Weekdays					Weekends	
	Morning peak hour						
Weather	23˚C/Fair	24˚C/cloudy	22˚C/Heavy rainfall	22˚C/Fair	24˚C/Fair	24˚C/Fog	24˚C/Fair
Time interval	Mon.	Tue.	Wed.	Thurs.	Fri.	Sat.	Sun.
7:00am - 7:20am	420	422	200	394	450	195	103
7:20am - 7:40am	573	547	350	552	539	247	149
7:40 am - 8:00 am	761	624	360	582	457	300	155
Total	**1754**	**1593**	**910**	**1528**	**1446**	**742**	**407**
	Off-peak hour						
	30˚C/cloudy	24˚C/cloudy	24˚C/cloudy	29˚C/Fair	29˚C/Sunny	27˚C/Sunny	26˚C/Light rain
12:00pm - 12:20pm	129	77	80	120	110	240	125
12:20pm - 12:40pm	87	147	110	93	123	158	90
12:40pm - 01:00pm	104	128	106	128	108	203	192
Total	**320**	**352**	**296**	**341**	**341**	**601**	**407**
	Evening peak hour						
	23˚C/cloudy	28˚C/rainfal	24˚C/Fair	31˚C/Sunny	30˚C/Sunny	29˚C/Sunny	24˚C/Cloudy
05:00pm - 05:20pm	287	284	275	311	421	330	203
05:20pm - 05:40pm	229	286	272	319	433	297	208
05:40pm - 06:00pm	336	167	220	267	403	324	309
Total	**852**	**737**	**767**	**897**	**1257**	**951**	**720**

Date:10th - 16thJuly 2023 Week 2

	Weekdays					Weekends	
	Morning peak hour						
Weather	30˚C/Sunny	33˚C/sunny	33˚C/sunny	33˚C/sunny	28˚C/sunny	30˚C/sunny	28˚C/fair
Time interval	Mon.	Tue.	Wed.	Thurs.	Fri.	Sat.	Sun.
8:00 am - 8:20am	615	526	495	503	562	289	162
	Off-peak hour						
	30˚C/sunny	36˚C/sunny	36˚C/sunny	36˚C/sunny	28˚C/Sunny	32˚C/fair	31˚C/fair
12:00pm - 12:20pm	105	81	87	95	113	145	119
	Evening peak hour						
	30˚C/sunny	36˚C/sunny	37˚C/sunny	35˚C/sunny	26˚C/Fair	32˚C/fair	31˚C/fair
06:00pm - 06:20pm	287	272	247	244	307	225	160

Date: 17th - 24thJuly 2023 Week 3

Continued

	Weekdays					Weekends	
Morning peak hour							
Temperature	27˚C/Partly cloud	28˚C/Partly cloud	28˚C/fair	28˚C/light rain	29˚C/Partly cloud	28˚C/cloudy	28˚C/sunny
Time interval	Mon.	Tue.	Wed.	Thurs.	Fri.	Sat.	Sun.
8:00 am - 8:20am	527	501	478	548	525	197	159
Off-peak hour							
Temperature	30˚C/Partly cloud	31˚C/Partly cloud	30˚C/light rain	30˚C/Partly cloud	33˚C/sunny	29˚C/Partly cloud	33˚C/Partly cloud
12:00pm - 12:20pm	94	110	95	89	80	159	132
Evening peak hour							
Temperature	30˚C/fair	29˚C/Heavy rain	27˚C/Heavy rain	28˚C/Heavy rain	29˚C/cloud	32˚C/cloudy	34˚C/Partly cloud
06:00pm - 06:20pm	245	234	223	195	307	246	230

Date: 20th -26th May 2024 (Week 4)

	Weekdays					Weekends	
Morning peak hour							
Weather	24˚C/Fair	24˚C/cloudy	22˚C/Heavy rainfall	24˚C/Fair	26˚C/Fair	24˚C/Fog	27˚C/Fair
Time interval	Mon.	Tue.	Wed.	Thurs.	Fri.	Sat.	Sun.
7:00am - 7:20am	339	295	274	312	289	186	126
7:20am - 7:40am	454	439	338	437	419	226	159
7:40 am - 8:00 am	638	550	483	580	467	245	175
Total	**1431**	**1238**	**1095**	**1329**	**1446**	**657**	**460**
Off-peak hour							
	27˚C/fair	24˚C/cloudy	24˚C/cloudy	27˚C/Fair	27˚C/Sunny	27˚C/Sunny	33˚C/partly cloud
12:00pm - 12:20pm	148	201	130	79	116	185	162
12:20pm - 12:40pm	132	133	127	90	128	167	160
12:40pm - 01:00pm	127	133	138	100	113	193	185
Total	**407**	**467**	**395**	**269**	**357**	**545**	**507**
Evening peak hour							
	27˚C/cloudy	28˚C/cloudy	27˚C/Fair	31˚C/Sunny	32˚C/Sunny	29˚C/Sunny	33˚C/Cloudy
05:00pm - 05:20pm	291	297	245	326	287	265	230
05:20pm - 05:40pm	241	296	237	466	277	289	208
05:40pm - 06:00pm	283	326	321	437	354	287	215
Total	**815**	**919**	**803**	**1229**	**1257**	**841**	**653**

Date: 27th May-02 June 2024 (Week 5)
Qingnian Road Station

	Weekdays					Weekends	
Morning peak hour							
Weather	23˚C/cloudy	21˚C/cloudy	23˚C/sunny	19˚C/light rain	21˚C/Fog	23˚C/clear	24˚C/sunny
Time interval	Mon.	Tue.	Wed.	Thurs.	Fri.	Sat.	Sun.
7:00am - 7:20am	380	340	320	215	450	220	130
7:20am - 7:40am	500	470	380	246	590	280	170
7:40 am - 8:00 am	790	850	520	480	630	340	190
Total	**1670**	**1660**	**1230**	**941**	**1670**	**840**	**490**
Off-peak hour							
	28˚C/sunny	28˚C/cloudy	31˚C/sunny	19˚C/light rain	24˚C/fog	29˚C/passing cloud	28˚C/passing cloud
12:00pm - 12:20pm	150	140	160	110	130	120	110
12:20pm - 12:40pm	170	160	180	130	150	140	102
12:40pm - 01:00pm	190	170	200	150	160	170	136
Total	**510**	**470**	**540**	**390**	**440**	**430**	**348**
Evening peak hour							
	28˚C/cloudy	30˚C/cloudy	29˚C/cloudy	21˚C/light rain	26˚C/haze	29˚C/Sunny	26˚C/sunny
05:00pm - 05:20pm	491	400	396	302	561	330	322
05:20pm - 05:40pm	361	403	500	433	600	389	415
05:40pm - 06:00pm	726	668	898	445	880	415	700
Total	**1578**	**1468**	**1794**	**1180**	**2041**	**1134**	**1435**

Date: 03rd - 09th June 2024 (Week 6)
Jiangshe 2nd Road

	Weekdays					Weekends	
Morning peak hour							
Weather	21˚C/cloudy	22˚C/cloudy	19˚C/cloudy	20˚C/fog	24˚C/clear	25˚C/sunny	27˚C/Fair
Time interval	Mon.	Tue.	Wed.	Thurs.	Fri.	Sat.	Sun.
7:00am - 7:20am	339	350	304	294	350	200	206
7:20am - 7:40am	354	300	328	352	339	199	220
7:40 am - 8:00 am	438	380	303	382	357	215	187
Total	**1131**	**1030**	**935**	**1028**	**1046**	**614**	**613**
Off-peak hour							
	26˚C/cloud	26˚C/haze	20˚C/light rain	25˚C/cloudy	26˚C/cloudy	31˚C/cloudy	33˚C/cloudy
12:00pm - 12:20pm	200	190	180	199	185	215	200
12:20pm - 12:40pm	210	200	190	210	200	205	198
12:40pm - 01:00pm	205	195	185	206	220	204	220
Total	**615**	**585**	**555**	**615**	**605**	**624**	**618**

Continued

Evening peak hour							
	25˚C/cloudy	22˚C/sunny	21˚C/light rain	27˚C/cloudy	27˚C/sunny	31˚C/sunny	34˚C/sunny
05:00pm - 05:20pm	300	299	276	330	350	289	210
05:20pm - 05:40pm	315	288	289	300	390	297	230
05:40pm - 06:00pm	305	400	387	308	410	324	387
Total	**920**	**987**	**952**	**938**	**1150**	**910**	**827**

Appendix B

Table B1. Lagos light rail blue line passenger arrival volume statistics and revenue.

Safety Production Day	Date	Days	average daily passenger arrival volume	daily train trips	daily revenue (confidential)
1	9/4/2023	Mon.	1008	10	353,825
2	9/5/2023	Tues.	1301	12	467,525
3	9/6/2023	Wed.	1592	12	583,225
4	9/7/2023	Thur.	1920	12	700,375
5	9/8/2023	Fri.	2488	12	911,350
6	9/9/2023	Sat.	1453	12	531,500
7	9/10/2023	Sun.	343	12	123,600
8	9/11/2023	Mon.	2924	12	1,074,975
9	9/12/2023	Tues.	2511	12	917,900
10	9/13/2023	Wed.	3145	12	1,156,225
11	9/14/2023	Thur.	2844	12	1,048,050
12	9/15/2023	Fri.	3337	12	1,201,125
13	9/16/2023	Sat.	980	12	361,425
14	9/17/2023	Sun.	257	12	89,700
15	9/18/2023	Mon.	3742	12	1,383,850
16	9/19/2023	Tues.	3626	12	1,335,900
17	9/20/2023	Wed.	3390	12	1,242,650
18	9/21/2023	Thur.	2737	12	997,575
19	9/22/2023	Fri.	3151	12	1,161,950
20	9/23/2023	Sat.	1805	12	657,925
21	9/24/2023	Sun.	218	12	79,525
22	9/25/2023	Mon.	4065	12	1,462,425
23	9/26/2023	Tues.	4400	12	1,625,175
24	9/27/2023	Wed.	2872	12	1,046,925
25	9/28/2023	Thur.	3176	12	1,175,525
26	9/29/2023	Fri.	3176	12	1,718,600
27	9/30/2023	Sat.	2142	12	790,550
28	10/1/2023	Sun.	245	12	87,975

Continued

29	10/2/2023	Mon.	2549	12	956,350
30	10/3/2023	Tues.	4601	12	1,684,350
31	10/4/2023	Wed.	4014	12	1,474,775
32	10/5/2023	Thur.	3199	12	1,184,100
33	10/6/2023	Fri.	3808	12	1,416,775
34	10/7/2023	Sat.	1823	12	675,150
35	10/8/2023	Sun.	210	12	74,000
36	10/9/2023	Mon.	4000	12	1,574,900
37	10/10/2023	Tues.	4593	12	1,672,950
38	10/11/2023	Wed.	4634	12	1,750,025
39	10/12/2023	Thur.	3413	12	1,280,725
40	10/13/2023	Fri.	4469	12	1,669,675
41	10/14/2023	Sat.	549	5	205,275
42	10/15/2023	Sun.	0		
43	10/16/2023	Mon.	7458	54	2,747,600
44	10/17/2023	Tues.	7177	54	2,642,625
45	10/18/2023	Wed.	8467	54	3,102,375
46	10/19/2023	Thur.	7607	54	2,813,075
47	10/20/2023	Fri.	8775	54	3,224,625
48	10/21/2023	Sat.	5000	54	1,845,175
49	10/22/2023	Sun.	284	22	99,025
50	10/23/2023	Mon.	9843	54	3,612,800
51	10/24/2023	Tues.	9472	54	3,502,225
52	10/25/2023	Wed.	9055	54	3,319,675
53	10/26/2023	Thur.	8490	54	3,102,475
54	10/27/2023	Fri.	9275	54	3,428,100
55	10/28/2023	Sat.	5264	54	1,941,750
56	10/29/2023	Sun.	305	22	109,325
57	10/30/2023	Mon.	9399	54	3,466,175
58	10/31/2023	Tues.	9369	54	3,448,800
59	11/1/2023	Wed.	9348	54	3,436,575
60	11/2/2023	Thur.	8934	54	3,294,175
61	11/3/2023	Fri.	10,901	54	3,997,575
62	11/4/2023	Sat.	6440	54	2,355,925
63	11/5/2023	Sun.	435	22	152,375
64	11/6/2023	Mon.	7223	54	5,267,600
65	11/7/2023	Tues.	5501	54	3,054,310
66	11/8/2023	Wed.	5649	54	3,090,620
67	11/9/2023	Thur.	5570	54	3,086,210
68	11/10/2023	Fri.	5711	54	3,118,040

Continued

69	11/11/2023	Sat.	3611	54	1,929,810
70	11/12/2023	Sun.	295	22	157,675
71	11/13/2023	Mon.	6430	54	3,549,745
72	11/14/2023	Tues.	5711	54	3,177,480
73	11/15/2023	Wed.	5345	54	2,958,755
74	11/16/2023	Thur.	4925	54	2,725,410
75	11/17/2023	Fri.	7131	54	3,978,220
76	11/18/2023	Sat.	4852	54	2,695,950
77	11/19/2023	Sun.	265	22	144,925
78	11/20/2023	Mon.	8799	54	4,911,265
79	11/21/2023	Tues.	10,029	54	5,588,870
80	11/22/2023	Wed.	9066	54	5,016,905
81	11/23/2023	Thur.	6262	54	3,479,570
82	11/24/2023	Fri.	7631	54	4,232,750
83	11/25/2023	Sat.	4140	54	2,300,740
84	11/26/2023	Sun.	211	22	112,860
85	11/27/2023	Mon.	7881	54	4,344,570
86	11/28/2023	Tues.	7301	54	4,049,930
87	11/29/2023	Wed.	6927	54	3,823,455
88	11/30/2023	Thur.	6622	54	3,601,145
89	12/1/2023	Fri.	6859	54	3,782,400
90	12/2/2023	Sat.	4513	54	2,485,380
91	12/3/2023	Sun.	286	22	154,825
92	12/4/2023	Mon.	7599	54	4,225,365
93	12/5/2023	Tues.	7144	54	3,984,410
94	12/6/2023	Wed.	7339	54	4,098,160
95	12/7/2023	Thur.	6740	54	3,743,230
96	12/8/2023	Fri.	7631	54	4,257,030
97	12/9/2023	Sat.	5252	54	2,920,020
98	12/10/2023	Sun.	258	22	138,360
99	12/11/2023	Mon.	7735	54	4,322,580
100	12/12/2023	Tues.	8464	54	4,706,765
101	12/13/2023	Wed.	9365	54	5,166,555
102	12/14/2023	Thur.	7609	54	4,228,940
103	12/15/2023	Fri.	8066	54	4,499,540
104	12/16/2023	Sat.	5435	54	3,038,295
105	12/17/2023	Sun.	316	22	170,280
106	12/18/2023	Mon.	8147	54	4,554,440
107	12/19/2023	Tues.	7809	54	4,374,705
108	12/20/2023	Wed.	6900	54	3,825,805

Continued

109	12/21/2023	Thur.	7259	54	4,035,275
110	12/22/2023	Fri.	7517	54	4,199,005
111	12/23/2023	Sat.	5405	54	3,010,850
112	12/24/2023	Sun.	393	22	213,165
113	12/25/2023	Mon.	1294	54	698,810
114	12/26/2023	Tues.	3131	54	1,735,280
115	12/27/2023	Wed.	5557	54	3,093,460
116	12/28/2023	Thur.	5377	54	2,981,455
117	12/29/2023	Fri.	4941	54	2,752,680
118	12/30/2023	Sat.	4087	54	2,270,525
119	12/31/2023	Sun.	386	22	211,545
120	1/1/2024	Mon.	1329	54	710,120
121	1/2/2024	Tues.	3208	54	1,773,730
122	1/3/2024	Wed.	3394	54	1,881,360
123	1/4/2024	Thur.	3317	54	1,827,120
124	1/5/2024	Fri.	3546	54	1,958,955
125	1/6/2024	Sat.	2418	54	1,339,490
126	1/7/2024	Sun.	344	22	184,140
127	1/8/2024	Mon.	4691	54	2,612,635
128	1/9/2024	Tues.	4333	54	2,412,815
129	1/10/2024	Wed.	4569	54	2,549,080
130	1/11/2024	Thur.	4403	54	2,450,030
131	1/12/2024	Fri.	4569	54	2,493,870
132	1/13/2024	Sat.	2840	54	1,580,835
133	1/14/2024	Sun.	253	22	132,755
134	1/15/2024	Mon.	6306	54	3,509,490
135	1/16/2024	Tues.	5961	54	3,317,670
136	1/17/2024	Wed.	5888	54	3,271,320
137	1/18/2024	Thur.	5216	54	2,855,745
138	1/19/2024	Fri.	6035	54	3,349,895
139	1/20/2024	Sat.	3167	54	1,751,205
140	1/21/2024	Sun.	229	22	121,055
141	1/22/2024	Mon.	6881	54	3,812,185
142	1/23/2024	Tues.	6120	54	3,400,770
143	1/24/2024	Wed.	6173	54	3,417,775
144	1/25/2024	Thur.	5211	54	2,881,130
145	1/26/2024	Fri.	5390	54	2,981,680
146	1/27/2024	Sat.	3269	54	1,795,685
147	1/28/2024	Sun.	213	22	115,580
148	1/29/2024	Mon.	6262	54	4,216,515

Continued

149	1/30/2024	Tues.	4937	54	3,606,865
150	1/31/2024	Wed.	4454	54	3,275,840
151	2/1/2024	Thur.	3922	54	2,860,280
152	2/2/2024	Fri.	4386	54	3,206,580
153	2/3/2024	Sat.	2106	54	1,545,725
154	2/4/2024	Sun.	169	22	127,785
155	2/5/2024	Mon.	4775	54	3,510,960
156	2/6/2024	Tues.	4465	54	3,242,685
157	2/7/2024	Wed.	4134	54	3,157,450
158	2/8/2024	Thur.	3926	54	2,861,575
159	2/9/2024	Fri.	4130	54	3,043,080
160	2/10/2024	Sat.	1921	54	1,412,415
161	2/11/2024	Sun.	159	22	114,160
162	2/12/2024	Mon.	4777	54	3,509,055
163	2/13/2024	Tues.	4461	54	3,255,380
164	2/14/2024	Wed.	4319	54	3,141,260
165	2/15/2024	Thur.	3767	54	2,762,510
166	2/16/2024	Fri.	3975	54	2,964,300
167	2/17/2024	Sat.	2116	54	1,565,750
168	2/18/2024	Sun.	177	22	125,550
169	2/19/2024	Mon.	4303	54	3,157,550
170	2/20/2024	Tues.	3048	54	2,192,540
171	2/21/2024	Wed.	5117	54	3,795,250
172	2/22/2024	Thur.	3949	54	2,841,350
173	2/23/2024	Fri.	5228	54	3,828,100
174	2/24/2024	Sat.	948	31	698,660
175	2/25/2024	Sun.	189	22	136,210
176	2/26/2024	Mon.	5838	54	3,258,430
177	2/27/2024	Tues.	5629	54	3,108,600
178	2/28/2024	Wed.	8114	54	4,452,305
179	2/29/2024	Thur.	6489	54	3,360,915
180	3/1/2024	Fri.	6378	54	3,512,825
181	3/2/2024	Sat.	2993	54	1,648,445
182	3/3/2024	Sun.	221	22	119,580
183	3/4/2024	Mon.	7536	54	4,110,765
184	3/5/2024	Tues.	6464	54	3,543,045
185	3/6/2024	Wed.	7296	54	3,837,440
186	3/7/2024	Thur.	7936	50	4,344,100
187	3/8/2024	Fri.	7589	54	4,191,565
188	3/9/2024	Sat.	3063	54	1,713,815

Continued

189	3/10/2024	Sun.	269	22	144,960
190	3/11/2024	Mon.	7325	54	4,056,305
191	3/12/2024	Tues.	6923	54	3,844,725
192	3/13/2024	Wed.	7628	54	4,236,665
193	3/14/2024	Thur.	6892	54	3,806,375
194	3/15/2024	Fri.	7011	54	3,880,100
195	3/16/2024	Sat.	3133	54	1,735,130
196	3/17/2024	Sun.	241	22	126,180
197	3/18/2024	Mon.	7920	54	4,384,950
198	3/19/2024	Tues.	7706	54	4,214,705
199	3/20/2024	Wed.	8178	54	4,499,450
200	3/21/2024	Thur.	7064	54	3,884,635
201	3/22/2024	Fri.	8603	54	4,763,905
202	3/23/2024	Sat.	3585	54	2,008,820
203	3/24/2024	Sun.	214	22	116,185
204	3/25/2024	Mon.	8113	54	4,517,645
205	3/26/2024	Tues.	7910	54	4,058,220
206	3/27/2024	Wed.	7185	54	3,982,150
207	3/28/2024	Thur.	7716	54	4,295,145
208	3/29/2024	Fri.	4303	54	2,403,850
209	3/30/2024	Sat.	4276	54	2,221,410
210	3/31/2024	Sun.	276	22	149,465
211	4/1/2024	Mon.	2606	54	1,441,730
212	4/2/2024	Tues.	6767	54	3,740,185
213	4/3/2024	Wed.	6187	54	3,433,135
214	4/4/2024	Thur.	5403	54	3,004,715
215	4/5/2024	Fri.	5765	54	3,206,170
216	4/6/2024	Sat.	2828	54	1,576,940
217	4/7/2024	Sun.	192	22	103,350
218	4/8/2024	Mon.	6287	54	3,485,580
219	4/9/2024	Tues.	3220	54	1,801,395
220	4/10/2024	Wed.	2523	54	1,389,880
221	4/11/2024	Thur.	2632	54	1,449,840
222	4/12/2024	Fri.	5474	54	3,025,605
223	4/13/2024	Sat.	2751	54	1,538,200
224	4/14/2024	Sun.	178	22	98,450
225	4/15/2024	Mon.	6819	54	3,776,655
226	4/16/2024	Tues.	6219	54	3,440,670
227	4/17/2024	Wed.	6040	54	3,329,650
228	4/18/2024	Thur.	5426	54	3,014,570

Continued

229	4/19/2024	Fri.	5860	54	3,242,060
230	4/20/2024	Sat.	3045	54	1,704,635
231	4/21/2024	Sun.	251	22	135,120
232	4/22/2024	Mon.	6558	54	3,632,680
233	4/23/2024	Tues.	5680	54	3,139,400
234	4/24/2024	Wed.	6330	54	3,495,865
235	4/25/2024	Thur.	5636	54	3,075,670
236	4/26/2024	Fri.	6158	54	3,354,505
237	4/27/2024	Sat.	3388	54	1,882,535
238	4/28/2024	Sun.	244	22	133,430
239	4/29/2024	Mon.	7417	54	4,080,905
240	4/30/2024	Tues.	7726	54	4,277,825
241	5/1/2024	Wed.	3806	54	2,115,215
242	5/2/2024	Thur.	6459	54	3,494,065
243	5/3/2024	Fri.	6007	54	3,246,635
244	5/4/2024	Sat.	3551	54	1,981,455
245	5/5/2024	Sun.	224	22	121,565
246	5/6/2024	Mon.	7498	54	4,142,005
247	5/7/2024	Tues.	6932	54	3,836,915
248	5/8/2024	Wed.	7508	54	4,165,400
249	5/9/2024	Thur.	6598	54	3,651,870
250	5/10/2024	Fri.	6764	54	3,743,140
251	5/11/2024	Sat.	3160	54	1,761,710
252	5/12/2024	Sun.	245	22	132,830
253	5/13/2024	Mon.	6971	54	3,846,025
254	5/14/2024	Tues.	7064	54	3,593,820
255	5/15/2024	Wed.	6751	54	3,574,985
256	5/16/2024	Thur.	5873	54	3,117,090
257	5/17/2024	Fri.	6901	54	3,604,185
258	5/18/2024	Sat.	3227	54	1,716,580
259	5/19/2024	Sun.	224	22	117,665
260	5/20/2024	Mon.	7313	54	3,834,029
261	5/21/2024	Tues.	6537	54	3,594,330
262	5/22/2024	Wed.	6183	54	3,394,930
263	5/23/2024	Thur.	6396	54	3,523,461
264	5/24/2024	Fri.	7796	54	4,313,655
265	5/25/2024	Sat.	3553	54	1,978,420
266	5/26/2024	Sun.	311	22	167,175
267	5/27/2024	Mon.	8992	54	4,929,160
268	5/28/2024	Tues.	7682	54	4,254,235

Continued

269	5/29/2024	Wed.	5920	54	3,096,745
270	5/30/2024	Thur.	5784	54	3,130,825
271	5/31/2024	Fri.	6580	54	3,611,070
272	6/1/2024	Sat.	3912	54	2,176,470
273	6/2/2024	Sun.	301	22	165,290
274	6/3/2024	Mon.	6823	54	4,938,305
275	6/4/2024	Tues.	6218	54	4,571,485
276	6/5/2024	Wed.	6461	54	4,748,825
277	6/6/2024	Thur.	7117	54	5,252,405
278	6/7/2024	Fri.	6563	54	4,841,745
279	6/8/2024	Sat.	3145	54	2,321,345
280	6/9/2024	Sun.	228	22	162,915
281	6/10/2024	Mon.	6921	54	5,101,270
282	6/11/2024	Tues.	6933	54	5,085,865
283	6/12/2024	Wed.	3684	54	2,709,445

Appendix C

Table C1. Questionnaire content.

No	Questions	Abbreviation of the question name
1	What is your age?	
2	What is your gender?	
3	What is your education level?	
4	Considering your recent experiences how would you rate the overall volume of passengers arriving at metro stations?	PAR1
5	During the busiest times of the day, how would you rate the level of crowding on the metro?	PAR2
6	Comparing to a year ago, how would you describe the change in metro passenger arrival volume?	PAR3
7	Weather conditions (e.g. rain, snow, heat) impact metro passenger arrival volumes	W1
8	The metro passenger arrival volume significantly increases during peak summer/winter.	W2
9	What type of weather-related disruptions are most likely to stop you from coming to the metro station?	W3
10	How likely are you to use alternative transportation or adjust your travel schedule due to congestion or delays during peak hours?	TOD1
11	How would you rate the difference in metro passenger arrivals between peak and off-peak hours?	TOD2
12	I perceive a significant increase in metro usage during evening peak hours.	TOD3
13	Longer waiting times at metro stations result in more people using the metro.	WAT1
14	Shorter waiting times lead to higher passenger arrival volumes at metro stations.	WAT2
15	Perceived waiting time influences the number of passengers arriving at metro stations.	WAT3
16	Satisfactory metro service leads to higher passenger arrival volumes.	MSS1

Continued

17	I am more likely to use the metro when satisfied with its service.	MSS2
18	Overall satisfaction with the metro positively influences the number of passengers using it.	MSS3
19	How often do you change your travel plans in response to real-time information about metro passenger volume or congestion, which may impact passenger arrival volumes?	TB1
20	To what extent do you agree that your travel behaviour, such as taking alternative modes of transportation or adjusting your travel time influences the rate of passenger arrivals at metro station?	TB2
21	How frequently do you use the metro for commuting to work, social events, or school/university?	TB3
22	To what extent do you agree that the pattern of passenger arrivals influence the overall passenger volume at the metro station?	AP1
23	How important is understanding the passenger arrival pattern to you when planning your trip to the metro station?	AP2
24	Do you observe a consistent pattern in the arrival of passengers at metro stations throughout the day?"	AP3

PAR = Passenger Arrival Volume, TOD = Time of Day, W = Weather, WAT = Waiting Time, MSS = Metro Satisfaction, TB = Travel Behaviour, AP = Arrival Pattern.