Analysis of Influencing Factors of Beijing Subway Passenger Flow Based on Data Analysis Method

Zuhao Liu *

The Affiliated High School of Peking University, Beijing, 100089, China

* Corresponding author: liuzuhao2024@i.pkuschool.edu.cn

Abstract. Subway passenger flow volume is a hot issue of concern to society. Some researchers develop predictive models to analyze short-time flow volume based on previous data. However, predictive ways to estimate the number of passenger flows before the subway construction still need to be completed. Therefore, by focusing on Beijing Subway Line 10, the article analyzes the factors influencing the passenger flow volume based on data analysis methods. The research finds that passenger flow volume on weekdays is greater than on weekends; as the number of stations and transfer stations increases, the volume increases; monofunctional stations have higher passenger flow volume. Before the construction of the subway, the design can be more appropriate according to relative factors. As a result, it is beneficial to release the stress of management and issues happening while the subway is running, decreasing operating costs and risks. In future studies, more detailed analyses in particular cases can be done.

Keywords: Beijing subway, passenger flow volume, influence factor, data analysis method.

1. Introduction

Subway passenger flow volume is a social problem that families and the government pay attention to. As the subway system connects more extensive places and becomes complete, the subway has become the first choice for most people to transfer. However, as the number of passengers significantly increases, issues like congestion and stampede accidents happen. Having a prediction of passenger flow volume will help prevent such issues. To make a prediction, the factors that influence the flow value are of significance.

The traits of passenger flow volume result from the integration and interaction of spatial and temporal passenger flow distributions at stations [1]. People choose to take the subway to work when the workplace is near the station [2]. According to previous research, there are morning and evening peaks on weekdays, and the flow direction differs between morning and evening [3]. According to passenger flow volume patterns, many algorithm models have been developed to predict short-time flow [4]. To create a better passenger' transfer experience, researchers like Zhou Yang, Zhang Huan, and Qu Shuling studied the transfer station [5, 6]. However, these previous researches are focused on prevention measures based on predicting passenger flow volume. Suppose stations and subway line designs can have suitable scales according to earlier predictions according to the factors of passenger flow volume. In that case, many issues like congestion can be solved without frequent measures. As a result, finding the factors is quite essential.

The article is focused on the passenger flow volume of the Beijing subway. This article first analyzes the attributes of the subway, dividing attributes into two dimensions: time and space. The paper uses data visualization to find the seasonal patterns on weekdays and weekends in the time dimension. The dimension of space has several traits, including the number of stations, the number of transfer stations, and different types of stations. In the space dimension, the article uses correlation analysis to explore how these traits influence the passenger flow volume. The paper aims to determine the factors that influence the passenger flow volume of the Beijing subway, thus helping the government better forecast the passenger flow volume and take relative actions to ensure passengers' comfort level while taking the subway.

2. Basic Information of Line 10

Line 10 in the Beijing subway is the most extensive line, with the most passenger flow volume almost daily. It has 45 stations and 16 transfer stations, connecting other lines like line 4. The essential elements in Line 10 are shown in Fig. 1.

Figure 1. Basic elements in the Beijing subway (Photo/Picture credit: Original).

3. Factor Analysis

3.1. Time

The research data from the official Weibo of Beijing Subway covers the passenger flow volume in Line 10 from May to November 2023 [7]. The data from this period is time-efficient in that it can fully present the recent seasonal situation.

The dataset is divided into two groups. The first group, presenting the situation on weekdays, contains most data from each Monday to Friday. In contrast, the second group, explaining the situation on weekends, includes data from each Saturday to Sunday. By visualizing these two groups of data, the case on weekdays and weekends can be shown. However, some particular dates can't fully present the normal function because of holidays like National Day. Therefore, every specific date is removed from the dataset.

Figs. 2 and 3 show the seasonal pattern.

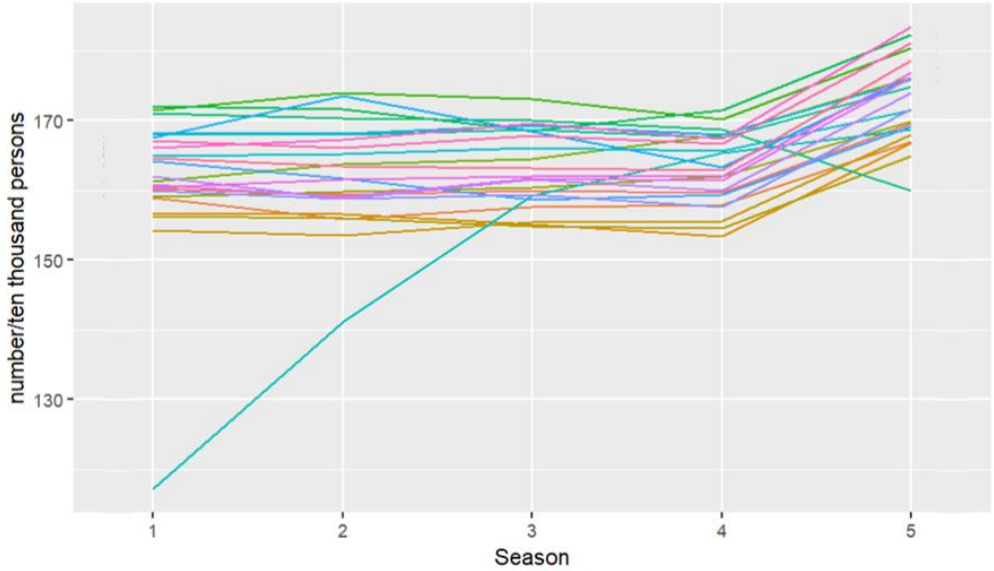

Figure 2. Passenger flow volume in Line 10 from Monday to Friday (Photo/Picture credit: Original).

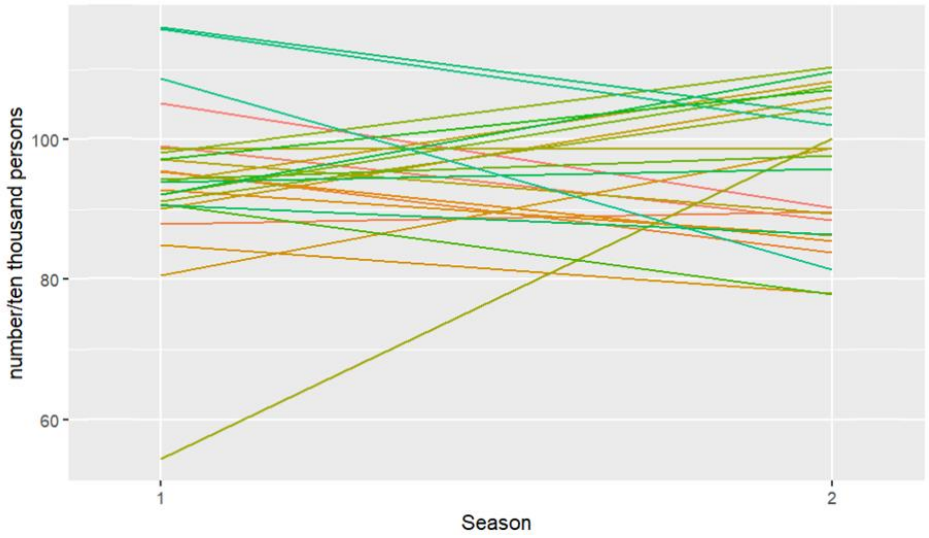

Figure 3. Passenger flow volume in Line 10 from Saturday to Sunday (Photo/Picture credit: Original).

In Figs. 2 and 3, the number aside lines indicate different weeks, while the x-axis shows the particular day of the week. In Fig. 2, every week, the pattern remains similar. The number is roughly 150-175 thousand persons from Monday to Thursday. The number increases on Friday, being between 160-185 thousand persons.

In Fig. 3, the pattern varies more than in Fig. 2, but the total pattern remains focused on 80-110 thousand persons on Saturday and Sunday.

Table 1. Data summary

day of week	Monday-Thursday	Friday	Saturday-Sunday
number/thousand persons	150-175	160-185	80-110

Compared to Table 1, passenger flow volume on weekdays highly exceeds weekends. On weekdays, Friday slightly exceeds the volume on Monday to Thursday.

Here comes the conclusion. Time does play an essential role in affecting the passenger flow volume.

3.2. Space-stations

To find the relationship between the number of stations and transfer stations, this article uses the daily passenger volume of all weekday lines in the past year from the Beijing Transport Development Annual Report 2023 [8]. This data presents the usual working status, having generality. At the same time, according to the Beijing Tube Map, the number of stations and transfer stations can be acquired [9].

To find whether there is a correlation between the number of stations and passenger flow value, correlation analysis is used. In R studio, Fig. 4 shows the correlation.

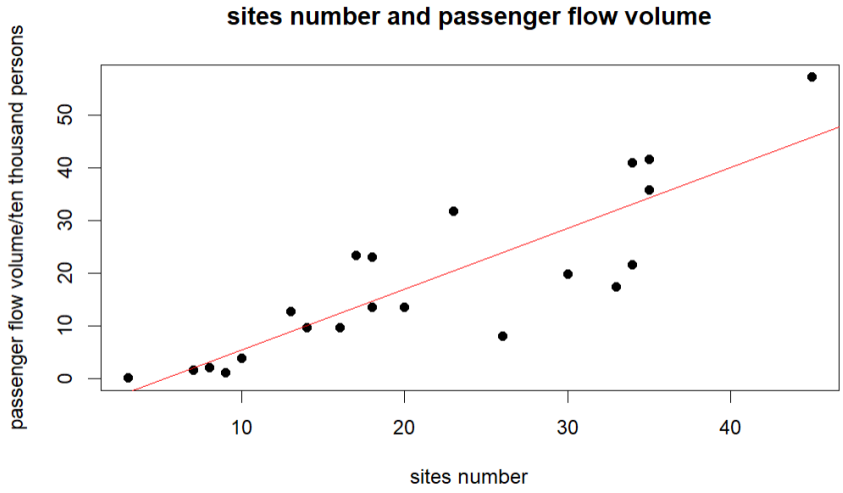

Figure 4. Stations number and passenger flow volume (Photo/Picture credit: Original).

In Figure 4, the points show one line's station number and passenger flow volume. As the station number increases, the volume is on the rise. The red line is fit. The points are roughly distributed around the line, which can show a strong correlation. Pearson's method of correlation test is used to better examine this correlation.

Table 2. Pearson's product-moment correlation test

p-value	6.452e-07
95 percent confidence interval	0.6777926—0.9412565
correlation coefficient	0. 8583355

Table 2 shows that the p-value is 6.452e-07, which is firmly smaller than 0.05. At the same time, the correlation coefficient is 0.8583355, whose absolute value is near 1. The 95 percent confidence interval is from 0.6777926 to 0.9412565. All prove the strong correlation between station number and passenger flow volume.

3.3. Space-transfer stations

Similarly, to find whether there is a correlation between the number of transfer stations and passenger flow value, correlation analysis is used. In R studio, Fig. 5 shows the correlation.

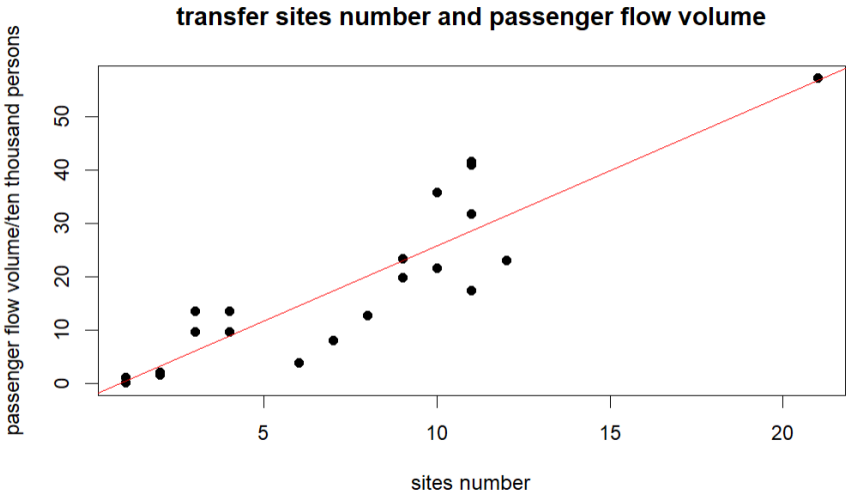

Figure 5. Transfer station number and passenger flow volume (Photo/Picture credit: Original).

In Fig. 5, the points show one line's transfer station number and passenger flow volume. As the transfer stations number increases, the volume is on the rise. The red line is fit. The points are roughly distributed around the line, which can show a strong correlation. Pearson's method of correlation test is used to examine this correlation better.

Table 3. Pearson's product-moment correlation test

p-value	7.063e-08
95 percent confidence interval	0.7426204—0.9545128
correlation coefficient	0.8892363

Table 3 shows that the p-value is 7.063e-08, which is firmly smaller than 0.05. At the same time, the correlation coefficient is 0.8892363, whose absolute value is near 1. The 95 percent confidence interval is from 0.7426204 to 0.9545128. All prove the strong correlation between transfer station numbers and passenger flow volume.

The results show that both correlation tests show a strong correlation that the more stations and transfer stations a line has, the more passenger flow volume it will have.

3.4. Space-types of stations

The types of stations can be recognized according to the surrounding buildings. In research, the professor uses K-medoids to divide the stations into six types: residential, job, mixed partial residential, mixed partial job, mixed residential and job, recreational, or pivotal [10]. The type indicates one station's general function. Fig. 6 shows the types of Line 10's stations.

Figure 6. Line 10's station's types graph (Photo/Picture credit: Original).

As shown in Fig. 7, according to the Beijing Transport Development Annual Report 2023, there are 12 transfer stations in Line 10 in the top 20 passenger flow volume transfer stations [8]. What influence the type of station has on the volume can be shown.

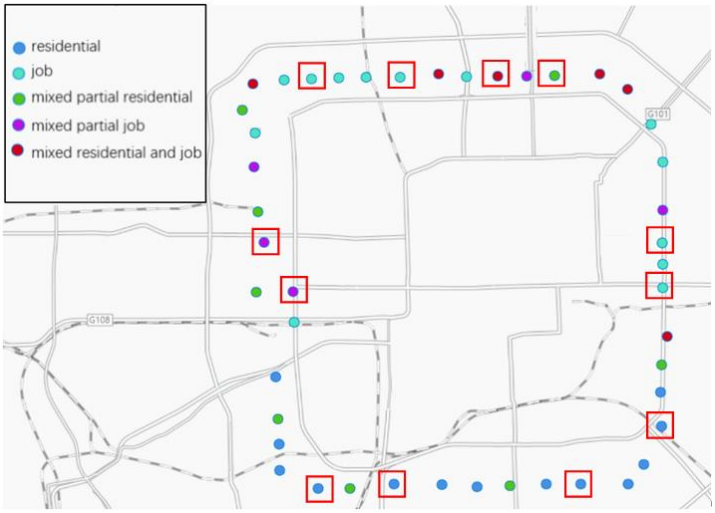

Figure 7. Line 10's transfer station's types graph (Photo/Picture credit: Original).

Table 4. 12 stations in the top marked

types	amount
residential	4
job	4
mixed partial residential mixed partial job mixed residential and job	1 2 1

By summarizing the amounts of different types of stations, it is clear that the residential stations and job stations have the most passenger flow volume. The types of stations influence the volume. Monofunctional stations, including residential and job, are more likely to have large volumes (Table 4).

3.5. Discussion

The results show that passenger flow volume differs on weekdays and weekends. The need for commuting is a significant cause. On weekdays, many people need to go to their workplace or school; on weekends, most people don't need to commute, and the purpose of taking the subway is various, like traveling. As a result, passenger flow volume on weekdays is more significant and stable, while the volume on weekends is low and changeable. As for the number of stations, more stations mean the Line covers more places. More people would take the subway. Line 10 is annular, covering between the Beijing Forth and Third Ring Road, and most people and workplaces in Beijing concentrate on this area. Line 10 significantly connects with other lines and this area, performing substantial transportation value. As the number of transfer stations increases, Line 10 connects more places, so the passenger flow volume rises. When it comes to the types of stations. Monofunctional stations are more likely to have large volumes. Because these places only mainly offer workplaces or residential areas, people working must go through a long distance to return to their homes. As a result, the need increases.

Recently, the development of China strongly motivated infrastructure demand. More subway lines and stations are being built in China. Designing a suitable station size is fundamental to ensuring the station can accommodate a large passenger volume. Previous research often focused on predicting daily passenger flow volume and making relative actions. If the number and types of stations indicate the volume before building the station, the stress of preventing passenger congestion while running will be significantly released. It is beneficial for subway management and passengers' well-being. As a result, a previous search for building one station based on the factors analyzed in this article is needed.

It should be noted that the data based on the Beijing Transport Development Annual Report is macroscopical. First of all, many holidays are expelled in the dimension of time. The seasonal plot on weekdays and weekends could be discussed more specifically, including Friday's passenger flow volume. What's more, in the space dimension, the daily passenger flow volume can be talked about at a smaller scale, like focusing on the morning and evening peaks. More specific, more reference value. In addition, more factors remain to be excavated.

4. Conclusion

The research finds that passenger flow volume on weekdays is more significant than on weekends; the more stations and transfer stations there are, the larger the number of passengers will be; monofunctional stations will lead to more passenger flow volume than other types of stations. The main contribution of this article is that while planning the construction of a new subway, the result could help better estimate the number of passengers, thus designing more suitable and appropriate stations and lines. This will release the stress of management and issues while the subway runs, decreasing operating costs and risks. Future studies can be focused on detailed patterns of stations.

Monofunctional stations have different patterns and structures. While delving deeper into the particular case, the results and predictions can be more accurate.

References

[1] Wang Jing, Liu Jianfeng, Ma Yilin, et al. Temporal and Spatial Passenger Flow Distribution Characteristics at Rail Transit Stations in Beijing. Urban Transport of China, 2013, 11 (06): 18 - 27.

[2] Tian Jingjing. Analysis of Influencing Factors of Subway Travel Flow in Beijing. Capital University of Economics and Business, 2016.

[3] Chen Yaoning, Ji Min, Ren Jing. Spatial and temporal Distribution of Passenger Flow in Beijing Rail Transit. Geospatial Information, 2021, 19 (03): 105 - 108+8.

[4] Zhou Qingmei, He Xiping. Review of Methods for Short-Term Prediction of Subway Passenger Flow. Journal of Chongqing Technology and Business University (Natural Science Edition), 2020, 37 (01): 25 - 32.

[5] Zhou Yang. Research on Transfer Mode and Transfer Facilities between New Subway Station and Existing Station. Shanxi, Architecture, 2020, 46 (22): 112 - 113.

[6] Zhang Huan, Qu Shuling. Optimization Analysis of Subway Transfer Function by Static Functional Evaluation System. Journal of Municipal Technology, 2014, 32 (01): 97 - 102.

[7] Beijing Subway. Short term distributed load forecasting method based on big data. November 11, 2023. Retrieved on November 11, 2023. Retrieved from Ma Kunlong. Changsha: Hunan University, 2014.

[8] Beijing Transport Insititute. 2023 Beijing Transport Development Annual Reprot. August 1, 2023. Retrieved on November 11, 2023. Retrieved from https://www.bjtrc.org.cn/List/index/cid/7.html.

[9] Beijing Subway. Route Map Query. November 19, 2023. Retrieved on November 19, 2023. Retrieved from https://map.bjsubway.com/.

[10] Lu Dongliang. Research on the relationship between the passenger flow of subway station and the surrounding built environment. Kunming University of Science and Technology, 2021.