The connection between pedestrians, vehicles and Toronto - The impact on urban transportation in Toronto based on pedestrian-vehicle correlations*

Xutao Chen

January 25, 2024

Pedestrian and vehicle-related data have an important impact on decisions such as transportation planning in Toronto, and their research plays an important role in everything from improving the efficiency of city operations to promoting resident well-being and environmental sustainability. This report is based on research on trends and correlations in pedestrian and vehicle counts over time to explore the current status of Toronto's transportation planning and discuss future transportation decisions.

Table of contents

*All data and scripts available at https://github.com/ZCZCZCNB/Toronto_Traffic_repo

1 Introduction

In Canada, more than 80% of the population lives in urban areas, with more than 57% of the population living in Canada's five largest urban areas, with Toronto being the most prosperous of the five (Bista, Hollander, and Situ 2021). The city's complexity can be seen in Toronto's busy streets and diverse transportation system, so a standardized transportation network not only provides opportunities for socio-economic activities within Toronto, but also has important implications for sustainable urban development and residents. Quality of life is also crucial (Xu et al. 2022).

In recent years, with population growth and urban expansion, Toronto's traffic flow has gradually increased, and the number of pedestrians and cars is the focus (Xu et al. 2022). Therefore, accurate and detailed pedestrian and vehicle flow data will inevitably affect the construction of urban transportation planning models. These data can not only show the traffic trends in recent years, but also provide guarantee for predicting future traffic needs and planning strategies. In addition, the data can help evaluate the effectiveness of traffic rules and policies and guide new transportation infrastructure projects.

This paper examines the relationship between pedestrian and vehicle flow data in Toronto by analyzing data collected between 2015 and 2019, and discusses how this data affects transportation planning and management decisions in Toronto. Because traditional transportation planning model systems estimate and calibrate in an unstructured way (Najmi et al. 2020), the research in this paper will focus on the correlation between data, and use ChatGPT4, a new AI technology, as an aid to discuss how to help Urban transportation planners can improve traditional model construction while promoting the sustainable development of cities and the quality of life of residents.

2 Data

2.1 Data Source

This report examines pedestrian and vehicle traffic data in Toronto from January 2015 to December 2019. These travel data are published by the Government of Toronto, and the dataset was last updated in March 2023. All relevant data was obtained using the "R" package "opendatatoronto" (Gelfand 2020).

2.2 Data collection methods

`R` (R Core Team 2021) was the language and environment used to collect and analyze this data, also the `tidyverse` (Wickham et al. 2019), `janitor` (Firke et al. 2021), `dplyr` (Wickham et al. 2021) and `kableExtra` (Zhu et al. 2021) packages.

Please note that these data include more than just pedestrian and car traffic, and some data are missing. So my workflow starts by extracting the year and month, and then for each year and month, calculate the sum of the number of various traffic participants such as pedestrians and cars. The result was saved to `clean_traffic_data.csv`.

Finally, I processed the organized data and created a new data frame, which contains the total number of pedestrians, cars and other traffic participants for each year and month combination. I save it in `toronto_traffic_10.csv`

2.3 Data Characteristics

The cleaned and calculated dataset showing the total number of pedestrians and vehicles per year can be seen in the table below Table 1. The data set shows the total number of pedestrians and vehicles respectively for each year from 2015 to 2019.

Table 1: Annual Total Number of Pedestrians and Vehicles (2015-2019)

year	total_cars	total_peds
2015	8581723	1582970
2016	15245012	3574675
2017	7030779	1303860
2018	9945781	1376034
2019	8729757	2240424

In addition, the average number of pedestrians and vehicles per month in the past few years can also be calculated through this data set, as shown in the table Table 2 below .

Table 2: Monthly Average Number of Pedestrians and Vehicles (2015-2019)

month	avg_cars	avg_peds
1	418.3732	113.89993
2	323.1051	63.98008
3	417.4212	56.49034
4	447.2600	77.47597
5	419.4937	60.09342
6	345.3558	63.27023
7	315.8936	72.73317
8	296.2951	154.78440
9	379.7417	252.68467
10	357.4684	44.47930
11	408.8571	101.19362
12	469.1415	26.97046

A more intuitive data representation is the annual average of the five-year total, as shown in the table Table 3 below.

Table 3: Average Annual Number of Pedestrians and Vehicles Over Five Years (2015-2019)

avg_cars_per_year	avg_peds_per_year
9906610	2015593

3 Discussion

3.1 Mean comparison

Based on the above data, the histogram shown in the figure Figure 1 below can be obtained. It can be clearly seen that the average monthly total number of vehicles is higher than the total number of pedestrians. The monthly average of the total number of vehicles is relatively stable without excessive fluctuations. On the other hand, the average number of pedestrians changes significantly, reaching a peak in August and September and a trough in December. It is generally believed that this is due to Toronto's special weather conditions (too cold winter, etc.) (Hewer 2020).

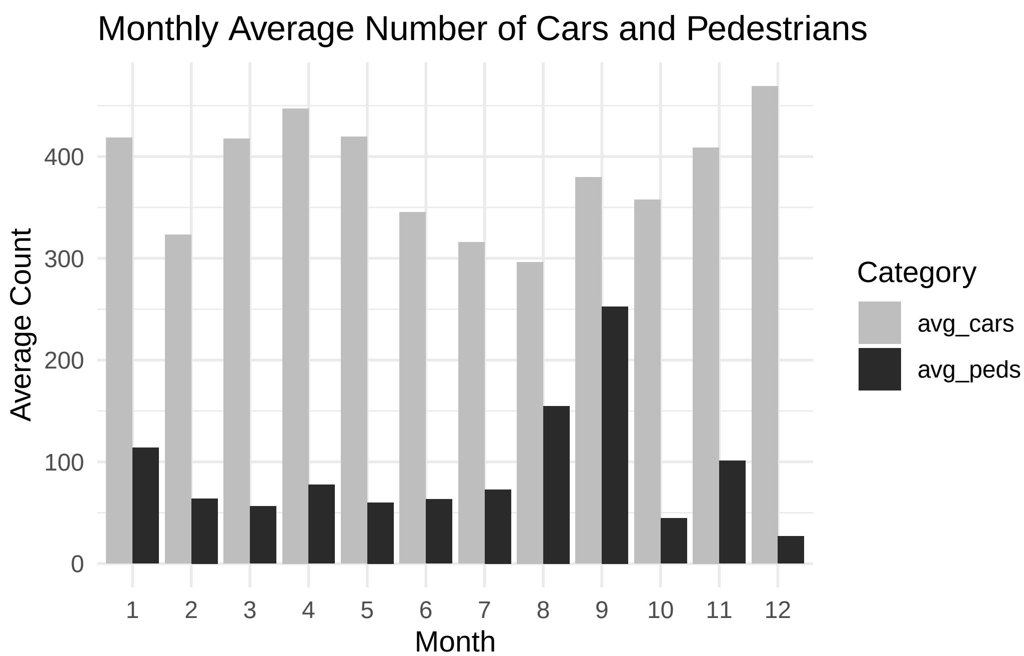

Figure 1: **?(caption)**

3.2 Total comparison

Although the average number of vehicles is usually much greater than the average number of pedestrians, a magical phenomenon is that the total number of pedestrians can even catch up with the total number of vehicles during some periods of time. As shown in the figure Figure 2, the number of pedestrians exceeded the number of vehicles twice in 2016 and 2019. In addition, there were several similar situations in early 2015 and mid-2017. It can be seen from this that the relationship between the two data is not simple, and the correlation between them is worth thinking about.

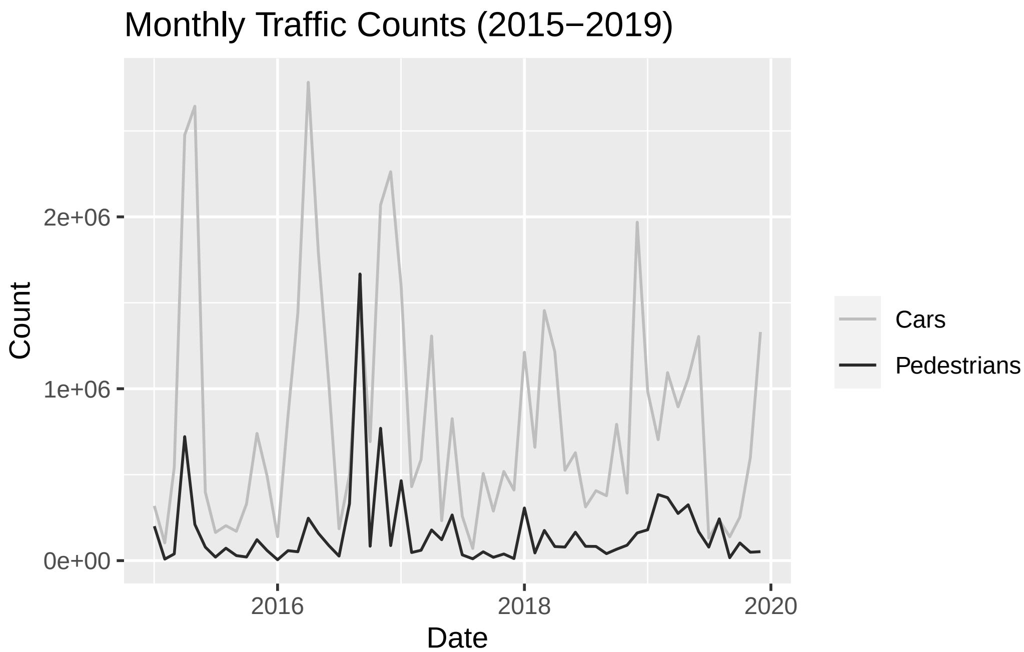

Figure 2: **?(caption)**

3.3 Linear fitting

When the relationship between two data needs to be considered, it is most intuitive to use scatter plots and linear fitting lines. As shown in the figure Figure 3 below, it can be seen from the figure that there seems to be a certain degree of positive correlation between the number of pedestrians and the number of cars. This means that, in some cases, when the number of cars increases in a given month, the number of pedestrians also increases. However, it is worth noting that this relationship is not very obvious because the data points are relatively widely distributed in the graph, which also proves that the relationship between the two data is not simple.

Generally speaking, the number of pedestrians and vehicles may be related to a variety of factors, which may affect the number of pedestrians and vehicles at the same time (Najmi et al. 2020). But from linear analysis, we can see that even if there are multiple factors, the relationship between the two still tends to be positive. Although such results are beneficial to transportation planning decisions, various influences still need to be paid attention to, because fluctuations in these two data are not uncommon, and the positive linear relationship is not direct.

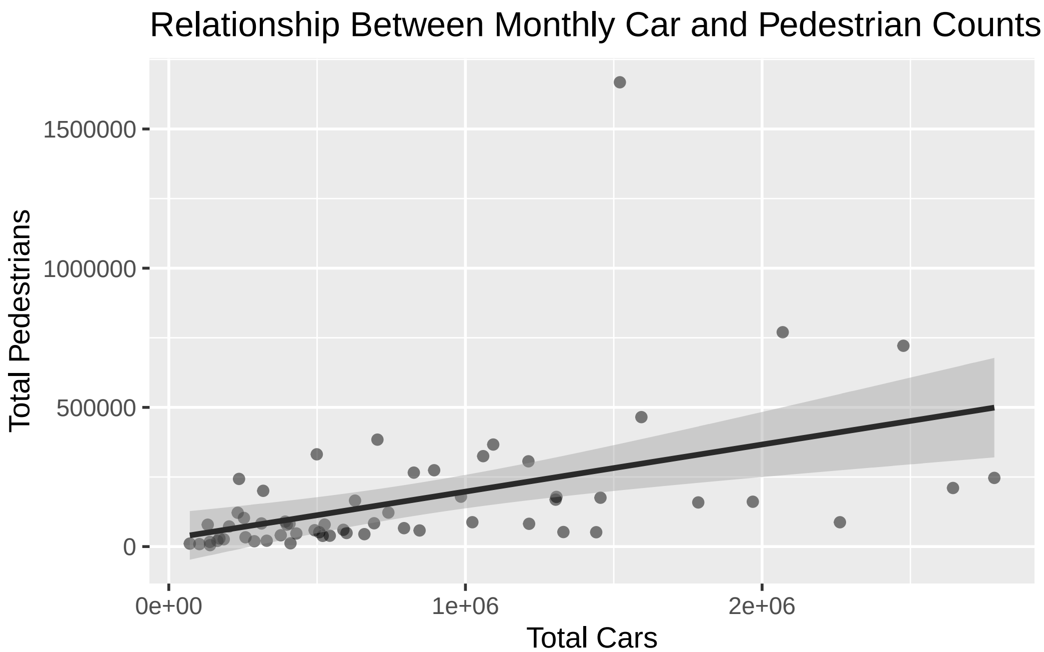

Figure 3: **?(caption)**

4 Conclusion

After this series of data analysis, we can conclude that the relationship between the total number and the mean of pedestrians and vehicles is not the same: the total number of pedestrians and vehicles will be affected by various factors, such as urban activities, holidays, weather changes, etc. And the fluctuations are huge. There are certain rules in the changes of the two averages. Not only that, there is also a certain linear positive correlation between the two. The results of the two different relationships are also different, but for Toronto's urban transportation planners, both conclusions are quite important. Based on these two conclusions, policymakers should first consider the overall situation when making urban planning decisions. Considering that in most cases the increase and decrease trends of pedestrians and vehicles are similar, they can be based on experience in advance at predictable time points (Such as winter, holidays, etc.) Make preparations in advance.

From another perspective, decision makers also need to be prepared to deal with emergencies. Based on multiple fluctuations in the total number and a weak linear relationship, the number of pedestrians and vehicles in Toronto can easily lead to different results than predicted. Therefore, urban transportation planners must have room to deal with unexpected situations, and even need to So that the corresponding police force can be dispatched at any time to deal with some accidents that may occur. In addition, based on the above data, some time points at which emergencies may occur can also be roughly derived. Before these special periods, decision-makers can prepare in advance to maintain traffic and protect the safety of citizens.

References

Bista, Shabnam, Justin B Hollander, and Minyu Situ. 2021. "A Content Analysis of Transportation Planning Documents in Toronto and Montreal." *Case Studies on Transport Policy* 9 (1): 1–11.

Firke, Sam, Bill Denney, Chris Haid, Ryan Knight, Malte Grosser, and Jonathan Zadra. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://cran.r-project.org/package=janitor.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://cran.r-project.org/package=opendatatoronto.

Hewer, Micah J. 2020. "Determining the Effect of Extreme Weather Events on Human Participation in Recreation and Tourism: A Case Study of the Toronto Zoo." *Atmosphere* 11 (1): 99.

Najmi, Ali, Taha H Rashidi, James Vaughan, and Eric J Miller. 2020. "Calibration of Large-Scale Transport Planning Models: A Structured Approach." *Transportation* 47: 1867–1905.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xu, Xiaocong, Dachuan Zhang, Xiaoping Liu, Jinpei Ou, and Xinxin Wu. 2022. "Simulating Multiple Urban Land Use Changes by Integrating Transportation Accessibility and a Vector-Based Cellular Automata: A Case Study on City of Toronto." *Geo-Spatial Information Science* 25 (3): 439–56.

Zhu, Hao, Thomas Travison, Timothy Tsai, Will Beasley, Yihui Xie, GuangChuang Yu, Stéphane Laurent, et al. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=dplyr.