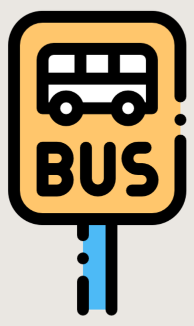
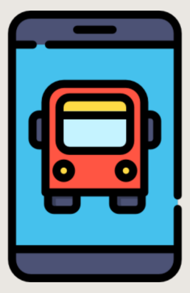


How reliable is my bus app?

Abstract

Our research problem is to investigate the accuracy of the bus apps in Singapore, specifically SGBuses, MyTransport.SG and Google Maps.

The motivation for this research was due to the observation on how most Singaporeans use bus apps to check for bus arrival timings but there are several occasions where the apps' estimated timing can be inaccurate, causing people to miss their bus and have to wait for a prolonged period, thereby wasting time. However, using an accurate bus app can significantly improve this problem as it allows one to have a better gauge on when the bus is arriving and from there plan his or her journey beforehand. Previous research done did not have data sets to prove the accuracy of the apps. Therefore, we would like to find out the accuracy of some of the most widely used bus apps in Singapore so that citizens can use the best app to assist them in their daily lives. To conduct our research, we collected primary data from almost all the bus interchanges in Singapore. We measured and calculated the difference between the estimated time of arrival of the bus on different apps and the actual time of arrival.



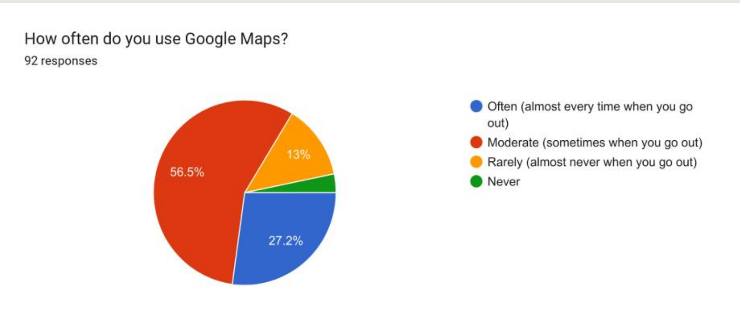


Figure 1: Pie chart showing how often people use Google Maps

Introduction

Did you know? In 2021, an average of approximately 3 million citizens use public transport per day in Singapore (Statista, 2022). Over the years, many Singaporeans have provided feedback that bus apps are becoming more and more inaccurate and unreliable (Mrbrown, 2015). This caused us to wonder if we could find the most accurate bus apps to satisfy Singaporeans' needs.

Purpose and Research Problem

We hope to find the most reliable bus app in Singapore. We are interested in this as bus apps are commonly used by people to assist them in their daily lives. Many Singaporeans use bus apps to check for the arrival timing of the bus they will be taking. This helps them plan their journey better, as they can roughly gauge the timing that they should leave their house, making bus trips and travelling more efficient.

At the same time, more people will be encouraged to take public transport instead of private transport, such as cars, as they can get to their destination with the same efficiency and without the need to drive. Taking public transport lowers our carbon footprint which in turn mitigates the impacts of global climate change (*Reducing your transportation footprint,* 2017).

Our project compares between MyTransport.SG, SGBuses, as well as Google Maps, and we hypothesise that MyTransport.SG and SGBuses are more accurate than Google Maps. We will be using SGBuses and MyTransport.SG as these apps are commonly used in Singapore to check for bus timings (Explore Singapore: A Handy Guide to Public Transport in the City, From Buses to Trains to Taxis, 2021). We will also be using Google Maps as it is highly used on a global scale. It is also used often locally, from a survey that we conducted using convenience sampling. From Figure 1, 83.7% from 92 people surveyed use Google Maps on a regular basis, and only 3.3% have never used it before, suggesting that it is quite a popular choice among Singaporeans.

Methodology

The approach to test our hypothesis is to collect raw data from the bus interchanges across Singapore. Afterwards, we will conduct hypothesis testing to draw a conclusion to our research question.

The three bus apps, SGBuses, MyTransport.SG, and Google Maps, will be our independent variables. The dependent variable will be the variation in timing of bus arrival, which will be the difference of timing from 6 minutes. We used 6 minutes as the benchmark, as it is the average time taken to walk to a bus stop. To derive the average time taken by Singaporeans to walk to public transport, we used the average walking distance to public transport in 2020, 576 m (Statista, 2022), divided by the average walking speed of Singaporeans, approximately 104 meters per minute (Berg, 2007). This resulted in an average time of 576/104 = 6 min (nearest whole number).

To collect the raw data, we went to 25 bus interchanges across Singapore to ensure that the data was not location-biased. At each interchange, the constant was the buses that we took the timings from. We used 2 buses that we obtained from a random number generator to ensure that there was no favouritism of buses. We then took the time taken for these 2 buses to arrive, for each app – SGBuses, MyTransport.SG, and Google Maps.

For our project, we made several assumptions. Firstly, we assumed that the arrival of the bus is when the bus doors open, as we observed that the doors of buses will be opened at interchanges even when there are no passengers boarding the bus. We also assumed that there is no difference between the bus apps between different brands of mobile phone (e.g., iOS and android). Additionally, we collected data from bus interchanges instead of going to bus stops as buses going to bus stops may be affected by unpredictable events such as traffic accidents or traffic jams, which will affect the accuracy of our results.

The process of data collection is as follows:

1. Start the timer when the estimated time of arrival for each app – SGBuses, MyTransport.SG, and Google Maps – reaches 6 minutes for the bus.
2. Stop the timer when the bus opens its doors after reaching the interchange.
3. Record down the timing in seconds.
4. Repeat steps 1 to 3 for the second bus.
5. Repeat steps 1 to 4 for the other buses at all 25 interchanges.

In total, 50 data sets were collected.

Data Analysis

To find out whether our hypothesis is accurate, we analysed the data using several methods. Firstly, we listed out all the values of the data collected and plotted them out on a graph, as shown in Figure 2.
A positive value means the bus is early, whereas a negative value means that the bus is late.

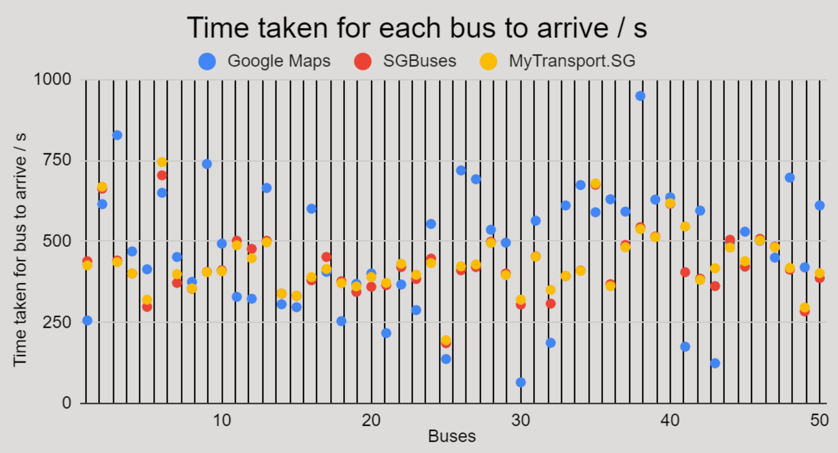


Figure 2: Scatter graph showing the time taken for each bus to arrive / s

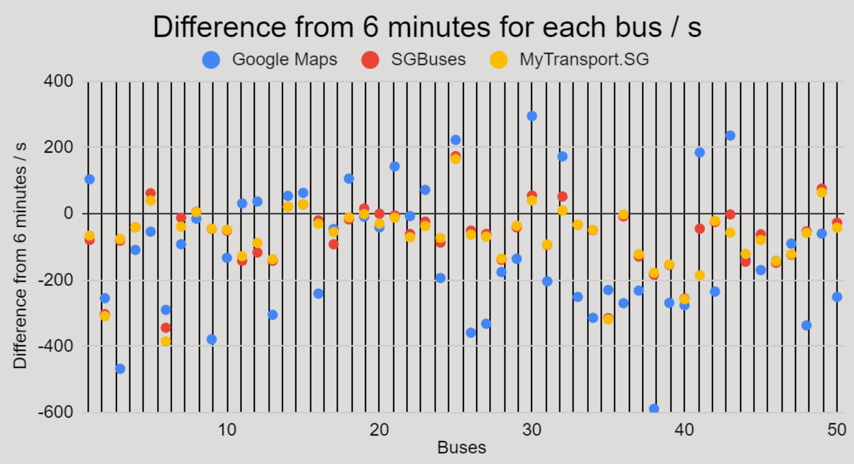


Figure 3: Scatter graph showing the difference from 6 minutes for each bus / s

From Figure 2, we can observe that the timings for SGBuses and MyTransport.SG are less widespread and more concentrated around 360 seconds, whereas the timings for Google Maps are more varied, with its highest at 949 seconds, suggesting that Google Maps is less reliable. Generally, all 3 apps have identical trends and buses take longer to come than the expected timing. With these data sets, we calculated the difference of the timings from the buses from 6 minutes by subtracting the time taken for the buses (in seconds) from 360 seconds i.e. 6 minutes. The calculated values were then plotted on another graph, as shown in Figure 3.

As seen in Figure 3, some of the datasets have great differences from each point to another. Hence, we wanted to determine whether our data is significant enough to conclude that SGBuses and MyTransport.SG are more accurate than Google Maps.

For our calculations, we decided to use the absolute value of the timings instead of taking into account the negative values as we will not be taking into consideration whether the bus will arrive earlier or later than the estimated time, since either way it will cause one to miss the bus or waste their time.

Firstly, we calculated the mean of the absolute difference. SGBuses had the lowest mean absolute difference of 86.02 seconds, which is then followed closely by MyTransport.SG at 87.88 seconds, and lastly Google Maps at 188.42 seconds. This suggests that the local bus apps are more accurate estimates of the arrival timing of the buses in Singapore than global ones. Hence, this suggests that our hypothesis may be true based on our collected data.

Apart from the quantitative analysis done above, we also conducted hypothesis testing to determine whether the local bus apps are actually more accurate than the global one. As we wanted to determine if 2 paired groups are significantly different from each other, and with our sample size greater than 30, population variance unknown, we used paired-sample t-test.

Let G, B and M represent Google Maps, SGBuses and MyTransport.SG respectively.

Let n be the sample size, which is 50.
Let x_i be the absolute difference from 6 minutes for each data point.
Let $\bar{x}$ be the mean of the absolute difference from 6 minutes for the buses.

$\bar{x}_G = 188.42$
$\bar{x}_B = 86.02$
$\bar{x}_M = 87.88$

With significance level $\alpha = 0.05$, t-test critical values = ± 2.010 and +1.677 for two-tailed and one-tailed test respectively,

Hypothesis	Testing	Conclusion
Google Maps vs MyTransport.SG $H_0: \bar{x}_G - \bar{x}_M = 0$ $H_1: \bar{x}_G - \bar{x}_M > 0$	$t = \frac{\bar{x}_G - \bar{x}_M}{\left(\sqrt{\frac{\sum_{i=1}^{n}[(x_{iG} - x_{iM}) - (\bar{x}_G - \bar{x}_M)]^2}{n-1}}\right)/\sqrt{n}}$ $= \frac{188.42-87.88}{\sqrt{\frac{798638}{49}}/\sqrt{50}}$ $= 5.57\ (3s.f.)$ 5.57 > 1.677	Reject H_0. Difference is significant
Google Maps vs SGBuses $H_0: \bar{x}_G - \bar{x}_B = 0$ $H_1: \bar{x}_G - \bar{x}_B > 0$	$t = \frac{\bar{x}_G - \bar{x}_B}{\left(\sqrt{\frac{\sum_{i=1}^{n}[(x_{iG} - x_{iB}) - (\bar{x}_G - \bar{x}_B)]^2}{n-1}}\right)/\sqrt{n}}$ $= \frac{188.42-86.02}{\sqrt{\frac{813042}{49}}/\sqrt{50}}$ $= 5.62\ (3s.f.)$ 5.62 > 1.677	Reject H_0. Difference is significant
MyTransport.SG vs SGBuses $H_0: \bar{x}_M - \bar{x}_B = 0$ $H_1: \bar{x}_M - \bar{x}_B \neq 0$	$t = \frac{\bar{x}_M - \bar{x}_B}{\left(\sqrt{\frac{\sum_{i=1}^{n}[(x_{iM} - x_{iB}) - (\bar{x}_M - \bar{x}_B)]^2}{n-1}}\right)/\sqrt{n}}$ $= \frac{87.88-86.02}{\sqrt{\frac{33844.0}{49}}/\sqrt{50}}$ $= 0.500\ (3s.f.)$ -2.010 < 0.500 < 2.010	Accept H_0. Difference is insignificant

Conclusion

For Google Maps and MyTransport.SG, the calculated t-value rejects the null hypothesis, thus the difference between their timings are significant. For Google Maps and SGBuses, the calculated t-value also rejects the null hypothesis, thus the difference between their timings are significant. As SGBuses and MyTransport.SG both have smaller means of the difference at 86.02 seconds and 87.88 seconds respectively, they are more accurate than Google Maps at 188.42 seconds. Both hypothesis tests suggest that the mean of the absolute difference from 6 minutes for Google Maps is higher than that of the other app, thus Google Map's absolute difference in timing is larger, meaning that it is less reliable. For MyTransport.SG and SGBuses, the calculated t-value accepts the null hypothesis, thus the difference between their timings is insignificant.

This proves that our hypothesis is correct. MyTransport.SG and SGBuses are indeed more accurate than Google Maps. This may be due to several reasons. Local bus apps such as MyTransport.SG and SGBuses focus only on Singapore, whereas Google Maps is an international app, hence its focus may not be on Singapore. Additionally, the main function of Google Maps is not for the estimation of bus arrival timings but navigation instead, which may have resulted in lower accuracy for estimation of bus arrival timing.

Limitations

Though the chances of unexpected events affecting the data were already reduced by collecting the data at interchanges, unexpected events may still occur. For example, a bus driver could have chosen to start the drive earlier than the timing given, resulting in the bus being early. Next, although the 3 apps investigated are the popular choices, they may not be fully representative of all the other apps. In addition, the data is only representative of the arrival time for bus interchanges. Due to other factors such as road accidents, traffic jams and bad weather conditions, the arrival time of the buses for bus stops may be delayed even if the bus was on time at the interchange. Lastly, consider a case where the app states at 0800 that the bus will arrive in 8 minutes. If the app states again at 0802 that the bus will arrive in 8 minutes and the bus indeed does arrive at 0810, the app having shown a timing of 8 minutes at 0800 is actually inaccurate. As we only took the reading from 6 minutes, any inaccuracy before the 6 minutes is not taken into account, meaning that the app timings at different times may be inconsistent, which may not be accounted for in our data.

Future Work

Data can be collected from bus stops to observe how the timings vary at bus stops instead of limiting it to only be at bus interchanges. At the bus stops, we will be able to observe whether real-time traffic conditions are taken into account by the apps, showing how accurate the app is even when there are unforeseen circumstances.

Similarly, more apps can be surveyed and analysed. This way, we will be able to find the most accurate bus app locally, and potentially see some trends in these apps. Through analysing the timings from more bus apps, we can also see if the apps are using the same data and algorithm to estimate the bus arrival times.

References

Average walking distance to public transport on a commute trip in Singapore from 2019 to 2020. (2022, July 22). Statista. Retrieved September 18, 2022, from https://www.statista.com/statistics/1232845/singapore-average-walking-distance-on-commute-trip/

Berg, N. (2007, May 3). *Which Cities Have The Fastest Walkers?* Planetizen. Retrieved September 18, 2022, from https://www.planetizen.com/node/24247

Mrbrown. (2015, June 29). *Why has the singapore bus arrival info become unreliable?* mrbrown.com. Retrieved September 18, 2022, from https://www.mrbrown.com/blog/2015/06/why-has-the-singapore-bus-arrival-info-become-unreliable.html

Statista. (2022, April 13). *Daily ridership of public buses Singapore 2015-2021*. Retrieved September 18, 2022, from https://www.statista.com/statistics/1006198/singapore-daily-public-bus-ridership/#:%7E:text=In%202021%2C%20an%20average%20of,COVID%2D19%20and%20its%20measures

Explore Singapore: A handy guide to public transport in the city, from buses to trains to taxis. (2021, May 6). Honeycombers. Retrieved September 18, 2022, from https://thehoneycombers.com/singapore/public-transport-singapore/

Statista. (2022, July 22). *Average walking distance on commute trip Singapore 2019-2020*. Retrieved September 18, 2022, from https://www.statista.com/statistics/1232845/singapore-average-walking-distance-on-commute-trip/

Reducing your transportation footprint. Center for Climate and Energy Solutions. (2017, October 27). Retrieved September 18, 2022, from https://www.c2es.org/content/reducing-your-transportation-footprint/