

Merit
Intended for teacher use only

For Achievement with Merit the student response includes:

Problem:

I am going to investigate the total number of passengers hopping on buses each month in Auckland between July 2005 and September 2023. This is a time series investigation and I am going to use the PPDAC cycle. I am going to investigate if the total number of people hopping on buses each month in Auckland has gone up between 2005 and 2023.

Data to collect: I found the data on Figure.NZ. It was part of data set that had all of the public transport data (including ferry and trains) but I am only going to be looking at the total number of people hopping on the buses each month. [Data from Figure.NZ.](#)

Importance: I think the results might be useful a couple of groups:

People who use cars might like to see if buses are becoming more popular. It could lead to changes in bus services. Maybe the same for people who go on the bus as well.

The bus companies would like to know if they need to buy more buses or hire more staff, or less if the trend goes down.

People who care about the environment might like to use the results to help them with their actions.

Plan:

This is about how the data will be collected.

I found this data on Figure.NZ. I downloaded the data and kept the bus total column. I changed the months from 2025-07 to 2025M07 so that I could use NZGrapher to make my graphs.

This is supposed to be all of the information and so it's not really a sample. I think there might be missing data from people jumping on the bus without paying or tagging or a bus company not giving its information properly.

Natural (or real) variation: Things that have happened at different times could play a big part in the data collection. A good example is covid when everyone was locked up. Because people couldn't go to work or shopping or other stuff, the bus numbers would have really big variation. I think this will make a big difference from 2020 with what I see in the data. Other things that I've noticed in Auckland that could influence the data include road works, public holidays, storms and other big weather interruptions – we've had lots of flooding recently as an example. These could have an impact on the total number of people hopping on the bus each month, usually less passengers.

Measurement variation: This data is based on counting so how the person or machine doing the counting needs to be kept the same over time. I saw that before 2014 the data was reported by the bus company – there is no info on how they did that. I think that each bus driver would have had to keep a record of how many passengers came onboard the bus during the day. They probably did this by looking at the ticket numbers each morning and



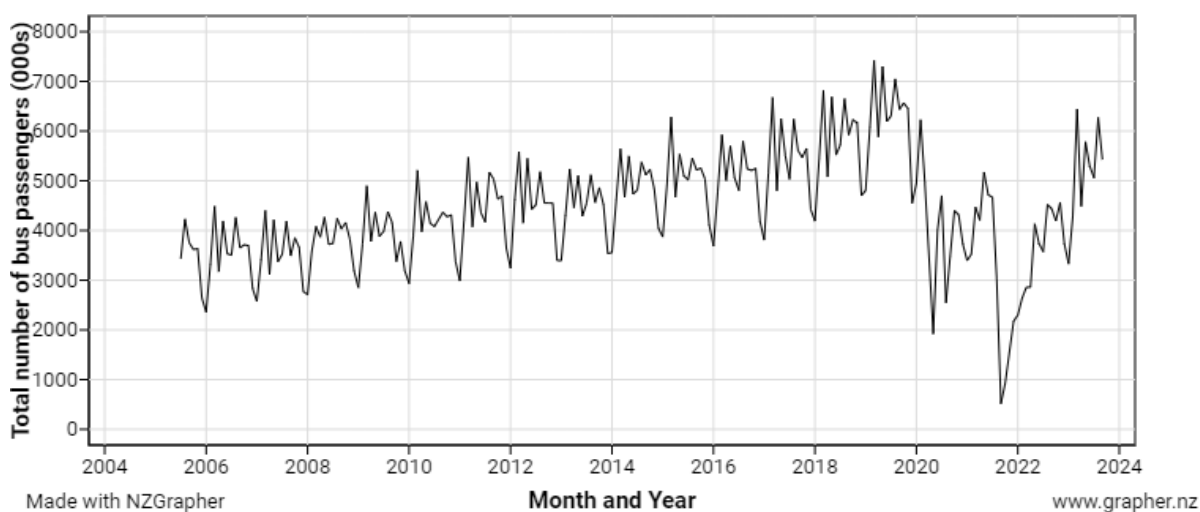
Expected Student Response for Mathematics and Statistics Achievement Standard 1.1

night and seeing how many had been used during the day. It would make sense for each driver to report this data when they signed off at the end of the shift. Because this is done by humans there is a lot of space for mistakes. Eg not reporting, doing bad maths, people being let on for free, people sneaking on. From 2014 the data comes from AT HOP (the tag on/off cards). This data will be more accurate because it is probably just a digital report each day. It would have variations if the driver didn't watch everyone tag on properly.

Data:

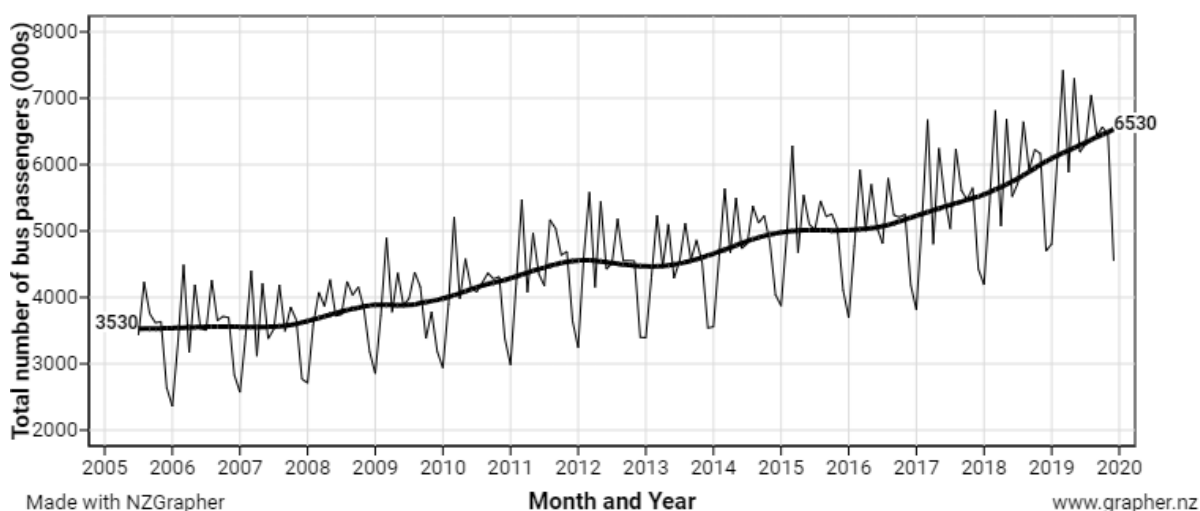
This is my first graph with all of the data included:

Total number of bus passenger boardings in Auckland (000s)



From looking at this graph I have decided to make another graph, which ends at the end of 2019 because it's a bit of mess from 2020 on (probably covid). I decided to take the 2020 stuff off when adding a trend line.

Total number of bus passenger boardings in Auckland (000s)



Expected Student Response for Mathematics and Statistics Achievement Standard 1.1

This graph is for July 2005 until December 2019. I have added a long term trend line and the start and stop points are shown.

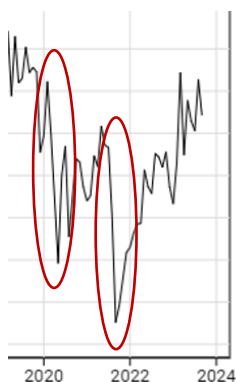
Analysis:

In my first graph I can see that the trend of passengers on the buses in Auckland increased over the months from July 2005 until the end of December 2019. I think that data looks like it slightly curving upwards. This is shown on the second graph where the trend line has been added. The trend starts at 3530000 passengers in July 2005 and ends at 6530000 in December 2019.

So the long term trend for bus passenger boardings in Auckland from 2005 until 2019 is increasing.

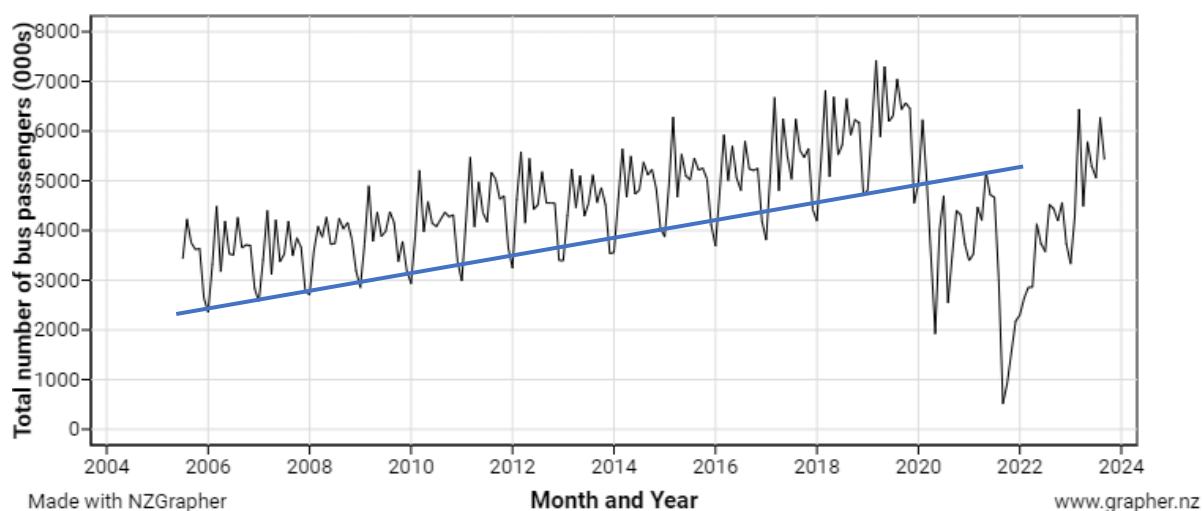
Patterns; Each year there is a pattern in the data, the lowest values happen in January each year and the highest numbers are in March. January makes sense because lots of people will be on holiday and not going to work or school.

Unusual: As I thought at the start, covid made a big impact on the graph. There is two really be dips/lows/troughs shown below. These were in May 2020 (1925300 passengers) and September 2021 (520500). Before covid, the lowest value was 2360300 in January 2006. I don't think anyone would have predicted this would happen.



Prediction: I think if covid hadn't happened, or we hadn't been forced to stay home or if people weren't worried, that the number of people on the bus in January 2022 would have been somewhere about 5250000. I made my prediction by drawing a straightline through the January points and continuing it on until 2022.

Total number of bus passenger boardings in Auckland (000s)



Conclusion:

At the start I said I was going to investigate the total number of passengers hopping on buses each month in Auckland between July 2005 and September 2023 to see if the numbers had gone up. I saw how bad the data was from covid lockups so chose not to use this to make my long term trend. It didn't make sense. If I had put a trend line over the entire data set, the more recent data would have a strong impact over where that line would go. So I have only looked until the end of December 2019. Overall, I can say from July 2005 until December 2019 there was an increase in the total number of bus passengers hopping on the Auckland bus system each month. This was impacted by covid lock ups from 2020 onwards. By the end of the data (September 2023) the total number of people on the bus each month had not got back to same level as before covid hit.

91944 Merit Exemplar Notes:

The student has explicitly used The Statistical Enquiry Cycle (PPDAC) and used each term as a heading for writing their report. They have given reasons for why the investigation is useful (who may benefit) and how the data was collected.

In explaining sources of variation, the student has given very good responses for two different sources of variation that they have named. They have drawn on their own experiences of COVID 19. In considering the counting of passengers the student has gone into good detail about the variation and what happens here. The student has shown good insight into both statistics and the context in this section, demonstrating a firm grasp on how data is collected over time and the need for consistency.

In the analysis section the student has used the original data set from Figure.NZ. They could have used the chart directly from the site but have chosen to reformat the data (for the use with the program they were using, not necessary in other programs) to produce their visualisation. The student has chosen to focus only on boardings for bus passengers from

Expected Student Response for Mathematics and Statistics Achievement Standard 1.1

within this data set, and does not look at the ferry or train data. They have used the automatic trend line feature but could have chosen to manually add the line and read the figures from the graph. These would be more ball-park than those they have given. The student has correctly used scaled values (passenger numbers in millions) for contextual communication. The student uses the phrase “hopping on” as an appropriate substitute for “boarding”.

The student has removed a section of data from the original data set and given a limited explanation. They refer back to the removed data when discussing features. Whilst there are some interesting/questionable statistical practices demonstrated, the student has holistically met the requirements for Merit. Further descriptions of their thinking would take this work to a higher level of achievement or would require a level of knowledge beyond the scope of the Learning Matrix for NCEA Level 1. It is noted that their forecast falls within the “messy” COVID19 period. The student could have equally acknowledged making a prediction based on the original data, including the “messy” data, is inappropriate. This is acceptable for Merit. In their conclusion, the student has acknowledged the difference between their original question and the answer they provided.

To reach Excellence, the student needed to show statistical and contextual knowledge in more areas than they did. They could have done this by reflecting on the process they have completed and the method of making their prediction such as using a manual model through the mid points/peaks/troughs or other, or reflecting the usefulness of the prediction. They could have given a better explanation for the removal of the data and reflected on what a reworked purpose for the investigation could look like as part of their conclusion.

