Achieved
Intended for teacher use only

For Achievement the student response includes:

I am going to answer the question, I wonder if the heights of year 11 students who completed the 2023 Census at School survey tend to be taller than the heights of year 9 students who completed the same survey"

For my answer I am using a sample that had 100 year 9 students and 100 year 11 students. I worked with a group to take the sample from Census at School 2023 data, our teacher helped us. Our whole class had done the survey so I was able to use that to help me with my sources of variation.

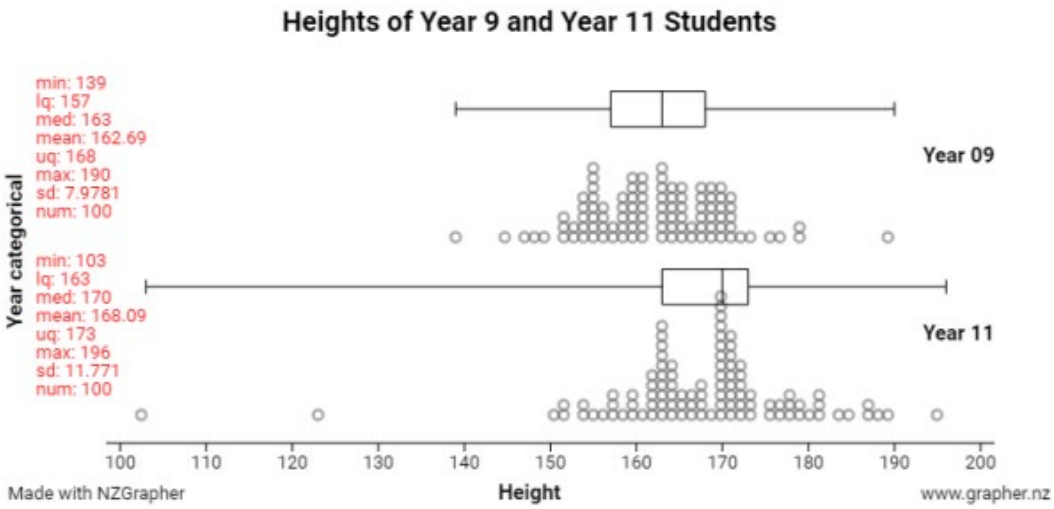
Sources of variation

1. Occasion-to-occasion variation. As part of our planning we found out that that your height changes during the day. Because of this we decided to measure our heights as early in the day as we could. This would give us our tallest height for the day. We collected our heights during our period 1 class on Monday. We think the survey could manage this type of variation better by including an instruction to measure your height as early in the day as possible.

2. Measurement variation. There are lots of ways the actual measurements can be done so we agreed to all measure by taking off our shoes, keeping our heels flat against the wall, and looking straight ahead. We got one person to stand on a chair and use a book to find the top of each persons head. We all used the same tape measure that as attached to the wall. We think the survey should have included better instructions or a diagram on how to measure your height to try and manage this type of variation.

3. Sampling variation. We used a sample of data that was shared by people looking at lots of different things. We agreed to use a sample that had 100 year 9s and 100 year 11s. The students came Auckland and Wellington but there weren't an even amount from each (138 from Auckland). I don't think this matters as we are comparing year 9s and 11s and don't think there will be any difference based on where they live in NZ. I think that 100 from each year group is a big enough sample that if I took another sample I would get a pretty similar answer to what I get with this sample.

My Graph – I have shown a dot plot with a high box plot. I used NZGrapher to make my graph.



Features in my graph.

In my graph I can see that the two middle 50% sections (the boxes) crossover but the box for year 11s is more to the right than the boxes for year 9s. For my sample this means that the middle 50% of year 11s were taller than the middle 50% of year 9s.

In my graph I can see that the middle value for Year 11 is higher than the middle value for year 9. This makes sense because year 11s are older than year 9s and have had more time to grow taller. In my graph I can see two values for year 11 that seem quite short – one near 110 cm and one a bit bigger than 120cm. I think that these seem quite short for a year 11 but they might be real values so I have left them in the graph. They make the whisker for year 11 really long.

Inference

For my sample the two box sections overlap but the median for year 11 is outside of the overlap and the median for year 9 is on the line for the lower quartile for year 11. I can make a call using the half three quarters rule and say that back in the population (everyone who did census at school in year 9 and 11 in 2023) the year 11 students are likely to be taller than the year 9 students.

91944 Achieved Exemplar Notes:

Whilst the student has not named the statistical enquiry process that they are following, the work links closely to the PPDAC cycle.

The student has a clearly worded question that they have used to write an informal inference. The student had attempted to write the question themselves and was supplied this question by their teacher. The teacher should have more carefully identified the population by adding “from Auckland and Wellington”. The student worked with a group of others to take a sample from the Census at School website with guidance from their teacher.

The student has explained three different sources of variation in the data collection process, noting that their own experience with the survey made it easier to help them write this part. The explanations list the type of variation, how the variation could occur in the collection

process when completed across the country, and for the first two, gives some recommendations for the survey instructions. Only two different sources of variation were required. This level of discussion was very well written and likely informed by participation in a planning session (as indicated by the use of “we” statements). Note that although the CensusAtSchool teacher guidance document refers to some of this variation, these ideas were independently suggested by the student. The guidelines were not used when the students participated in the activity.

The student presented two visualisations in one graph, a comparative dot plot and a comparative box and whisker plot. They have included the summary statistics for the data, which are not required for achieved.

The student has described three different features they can see in the visualisations, although only two required. The student has made connections to the context for each by using the words taller, time to grow, and quite short. This could have been improved by referring to the variable “height” throughout.

Whilst not required by Explanatory Note 1, it is good to see that the student has made an informal sample to population inference. The informal inference does not use the word “tend” but has suggested the probabilistic nature of the inference by saying “likely”. This would have been an acceptable inference if each group had a sample of size 30 (maybe at the lowest level). Given the student had sample sizes of 100, they should have used, visually, the distance between the medians as a proportion of overall visible spread to make a call for a completed investigation.

To reach Merit, the student would need to consider their enquiry process as a whole. A brief introduction to the exploration would include what data to collect, who to collect it from and why it is important. This student talks about how they chose their sample (what and who) but they have not linked to the usefulness or importance. The student also needs to justify the features using both at least one appropriate visualisation and one measure for two features. This could be done by referencing the lower and upper quartiles for both groups when looking at position of the middle 50% (shift and overlap of two groups) and stating the median for each group. The student needed to use DBM and OVS visually to make the call for their sample to population inference due to samples of size 100 per group.