

A close-up photograph of a hand holding the top handle of a black backpack. The hand is positioned in the upper left, with fingers wrapped around the strap. The backpack is black with a visible zipper and textured fabric. The background is a plain, light color.

CensusAtSchool

NEW ZEALAND

Exploring Our Class Data

2021

LEVEL 3

This activity is designed to support students exploring the class data that the teacher can access following their students undertaking the 2021 CensusAtSchool survey.

Datasets

The teacher gets back several datasets for their class in comma separated variable files (.csv). The order of the data in these individual datasets has been randomised, which means you cannot match it to any individual student. Privacy concerns are the reason for the provision of individual datasets.

One dataset includes questions 8-19, which gives some of the survey questions about you, measurements, and survey questions about school and games.

Q.	Variable	Outcomes	Units	Type
8	eyecolour	blue, brown, grey, green, hazel		Categorical
9	handed	right, left, ambi		Categorical
10	height		cm	Numerical
11	footleft footright		cm	Numerical
12	wrist		cm	Numerical
13	thumb		cm	Numerical
14	travel	walk, motor, bus, train, bike, boat, scooter, skateboard, other		Categorical
15	timetravel		minutes	Numerical
16	bagweight		kg	Numerical
17	memory		seconds	Numerical
18	reaction		seconds	Numerical
19	standleftleg		seconds	Numerical

Table 1: Variable, outcomes, units, and type of variable for questions 8-19 from 2021 questionnaire

Individual datasets are provided for each of questions 20-29, survey questions about activities and opinions.

Q.	Variable	Outcomes	Units	Type
20	screentime		time (hh:mm:00)	Numerical
21	techphone techfacebook techinstagram techtwitter techtiktok techsnapchat techreddit techdiscord techroblox techminecraft techfortnite techyoutubewatch techyoutubeupload technone	yes no		Categorical
22	screentimeopinionphone screentimeopinionsocialmedia screentimeopinionvideogames	too much, about right, too little		Categorical
23	newstv newsradio newsprint newssites newssocialmedia newssearch newsbroadcasts	often, sometimes, rarely, never		Categorical
24	bedtime waketime sleeptime		24-hour time (hh:mm:00) hours	Numerical
25	favanimatedmovie			Categorical
26	learningathome	much better, better, about the same, worse, much worse		Categorical
27	climatechange	urgent, future, not a problem, don't know		Categorical
28	future	very positive, positive, neutral, negative, very negative		Categorical
29	included	always, very often, sometimes, rarely, never		Categorical

Table 2: Variable, outcomes, units, and type of variable for questions 20-29 from 2021 questionnaire

Statistical Enquiry Cycle

When students are provided with data to explore the statistical enquiry cycle, the PPDAC cycle, is rearranged to accommodate a different way of thinking about the statistical investigation (See Figure 1).

- We start with the **dataset**.
- We then interrogate the data to work out who the data was collected from, what was collected, when it was collected, and where it was collected. In other words, we understand the **plan** the original investigator had. In the plan phase we are identifying what variables we have and who the group of interest is for the dataset we have.
- Knowing what variables are available we can **pose investigative questions** that we can answer with the data we have.
- We then **analyse** the data; we display and then describe the data.
- Finally, we draw **conclusions**, answering the investigative question.

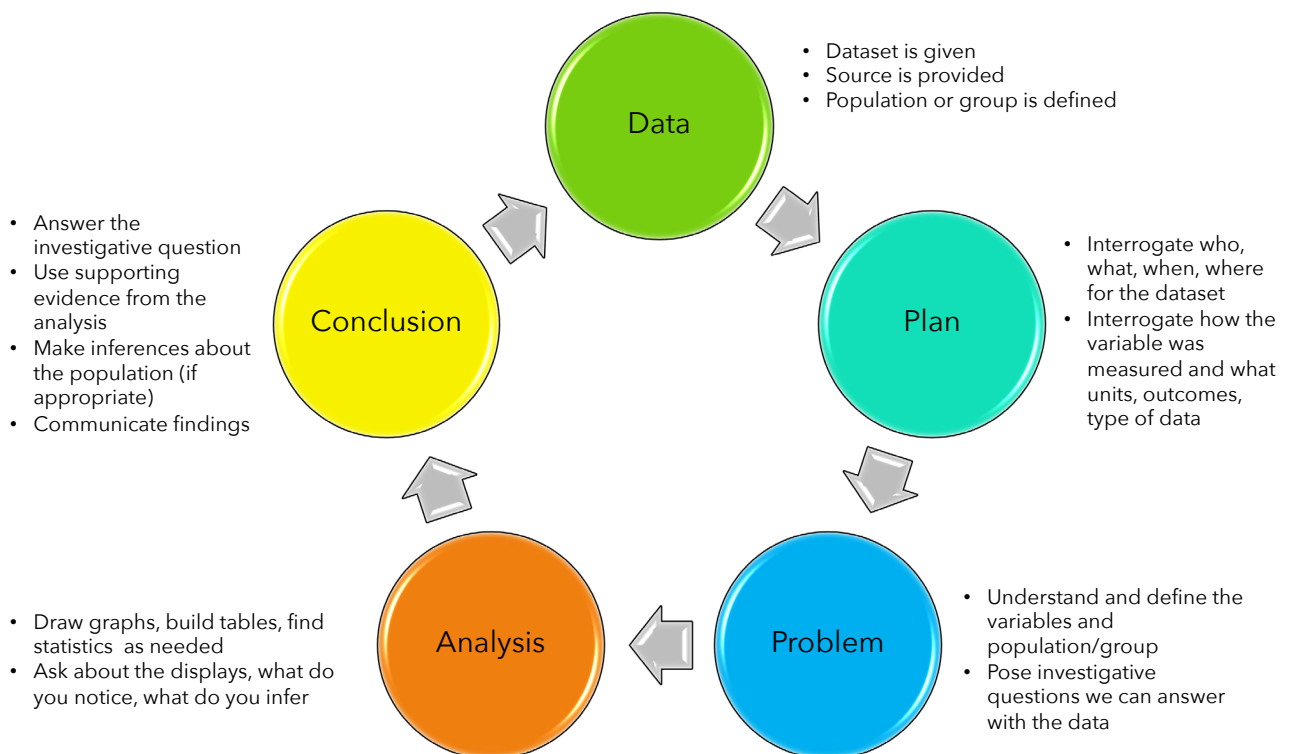


Figure 1: The statistical enquiry cycle for provided datasets (adapted from Arnold, P. (2013). Statistical Investigative Questions – An Enquiry into Posing and Answering Investigative Questions from Existing Data, (Doctoral thesis), Retrieved from <https://researchspace.auckland.ac.nz/handle/2292/21305>)

Statistical Software

It is assumed that students have access to technology, e.g., a Chromebook, desktop computer or laptop computer. Tablets can be used as well but the aforementioned options are preferable. Statistical software is recommended to create data visualisations. For example, the Common Online Data Analysis Platform (CODAP) <https://codap.concord.org/> is a free online tool for students (and teachers) to use. Students in year 5 have used this software successfully following an introduction to the tool.

Pip Arnold created a series of short videos to introduce students to using CODAP during lockdown in 2020. Teachers can make a copy of the Google doc and share it with their students. To access the Google doc see her blog post:

<https://karekareeducation.co.nz/codap/getting-your-students-started-with-codap/>.

The teacher can import their .csv files from CensusAtSchool that they received into one or more CODAP documents, for example, Q8-19 into one CODAP document and the remaining .csv files for the individual questions (Table 2) can all be imported into one CODAP document if you want. There will just be many separate tables in the CODAP document (see Figure 2). See this video <https://youtu.be/wdAjZEi7vtk> on how to import data into CODAP.

Once that is done, teachers need to save the CODAP document and share with their students (see <https://youtu.be/8LhcFWVs5k> on how to share CODAP documents with your students).

Figure 2 is a screenshot of a CODAP document set up with the data for questions 20-29 showing all of them in the one document. To do this, each .csv file is imported into the same CODAP document. In the screenshot all the tables have been changed to case card view and then minimised. Students can open the case card for the question or questions they want to explore. Q22 has been dragged out of the list and then opened to case card view as an example.

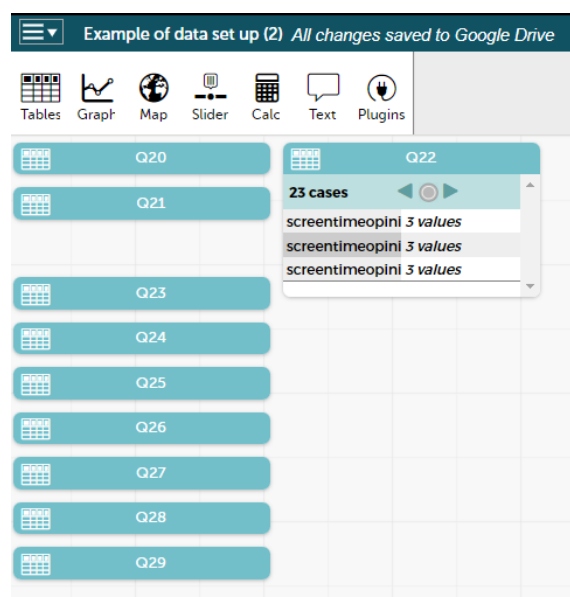


Figure 2. Showing multiple data tables in CODAP, all minimised except for Q22

Data

Making the dataset available to your students

There are several options for providing the data to your students to use:

1. Make up multivariate data cards with the class data
2. Provide a printed table of the raw data straight from the .csv files provided
3. Upload the dataset into statistical software and work with the data there
4. A combination of all the above

Multivariate data cards

Two Excel templates are provided to make your own multivariate data cards if you wish to use them with your class.

1. Example data cards for example year 5-6 dataset.xlsx

- a. This is an example dataset that was downloaded from CensusAtSchool to model what you can do with your class data.
- b. In this example dataset there are 32 year 5 and 6 students from the 2021 CensusAtSchool database. There are three sets of data cards available:
 - i. Height et al includes height, eye colour, handedness, left foot length, right foot length (Figure 3)

150	149	160	156	138
brown	left brown	left brown	right blue	right blue
23 / 23	28 / 28	25 / 24	24 / 24	22 / 20

Figure 3. Example data cards from height et al data card set

- ii. Travel et al includes travel method to school, travel time to school, bag weight, left wrist circumference and left thumb circumference
 - iii. Reaction et al includes reaction time, handedness, eye colour, memory time, left leg stand time
- c. You could use these data cards instead of creating class data cards

- d. Data cards are printed and cut up into bags or envelopes.



Figure 4. The three different sets of data cards from the example data cards, printed, and cut up into plastic bags. Pink is height et al set, green is the travel et al set, and blue is the reaction et al set.

- e. This dataset can be used to make up your own data cards for your class if the class is 32 students or less.
- Using the tabs read and adjust the *instructions* sheet. Note the *height et al* sheet is the same as the *copy your data here* sheet and the *height et al dataset* is the same as *example dataset 240* in the Excel file *Make your own data cards.xlsx*.
 - Paste your class dataset into the *year 5-6 example dataset*.

2. Make your own data cards.xlsx

- This is a template for making up to 240 data cards for five variables.
- See instruction page for how to do this.

Printed data table

Another way to give your students the data is to provide them with a printed copy of the spreadsheet (.csv file). Figure 5 shows the data table for the *year 5-6 example dataset* that is used throughout this lesson plan.

Eye Colour	Handed-ness	Height	Left foot length	Right foot length	Left wrist circumference	Left thumb circumference	Travel method to school	Travel time to school	Bag weight	Memory time	Reaction time	Time standing on left leg
green	right	148	21	21	14	14	6 motor	4	2	54	0.606	32
brown	ambi	140	21	22	15	15	6 other	3	1.2	179	2.06	2
blue	ambi	141	21	22	14	14	5 other	15	1.6	46	0.724	22
brown	right	159	24	25	16	16	7 walk	2	0	40	0.35	50
brown	right	144	22	22	15	15	6 motor	17	2	58	1.089	54
brown	right	140	23	23	16	16	7 bus	6	2.3	57	0.529	6
brown	right	147	22	22	10	10	4 motor	11	3	58	0.499	69
brown	right	134	21	21	14	14	5 motor	22	5.2	60	0.47	3
blue	right	146	22	23	16	16	6 motor	35	2	47	0.538	3
grey	left	136	21	22	14	14	4 motor	5	2	128	2.906	8
blue	right	174	31	30	17	17	7 motor	5	1.3	55	0.673	12
blue	right	134	20	21	14	14	5 walk	8	0.6	75	0.601	15
brown	right	146	21	20	14	14	6 motor	6	2	49	0.439	2
brown	right	140	24	24	17	17	5 bus	18	1	65	1.215	61
green	right	135	23	23	14	14	7 motor	2	2.8	49	0.469	2
blue	right	146	24	23	13	13	6 bus	30	5	63	0.726	26
blue	right	137	23	24	18	18	5 bus	15	0.5	33	3.656	180
blue	left	156	24	23	17	17	6 motor	15	2.5	65	3.053	13
blue	right	152	19	18	15	15	5 walk	10	6.3	59	0.66	3
brown	right	139	24	22	15	15	6 motor	5	1.2	53	0.419	12
brown	right	144	23	23	19	19	7 motor	2	2	79	0.455	70
brown	left	150	23	23	16	16	6 motor	2	0.2	52	0.613	104
green	right	150	23	23	17	17	7 motor	1	5	41	0.44	62
brown	left	149	28	28	15	15	8 motor	10	3	40	0.525	20
brown	right	153	23	23	14	14	7 walk	15	6.2	143	4.617	3
brown	right	160	25	24	16	16	6 motor	3	1.4	68	1.127	7
green	right	138	20	20	14	14	5 motor	15	3	47	0.414	20
blue	right	156	24	24	16	16	6 motor	5	3	63	0.546	120
green	right	176	24	24	14	14	5 motor	2	3	60	0.516	113
blue	right	138	22	20	17	17	6 scooter	6	1.8	36	0.547	31
brown	right	152	23	24	15	15	6 motor	2	0.5	45	2.206	120
hazel	right	139	20	21	13	13	6 motor	1	3	46	0.396	2

Figure 5. Data table showing the variables for questions 8-19 for the example year 5-6 dataset

Share the dataset(s)

Share the dataset(s) with your students using the format you have decided on, e.g., data cards, data table, statistical software. If you are using CODAP or other statistical software share the link to the data.

Explain that the dataset(s) contains the class data for the CensusAtSchool survey they participated in on [date]. Get students to share what they remember about doing the survey and collecting data for the survey.

Confirm that the data is the class data for the CensusAtSchool 2021 survey, or if you are using the example dataset provided confirm that the data is similar to the class data they collected for the CensusAtSchool survey.

For the examples given in this lesson plan, the focus is on questions 8-19 but you could use any of the other questions provided.

Plan

With provided data, such as this class data, take the time for students to familiarise themselves with the variables that are available to use. As they will have been involved in the data collection process, they should already be familiar with how the different variables were measured or the possible outcomes. However, it is worthwhile having both the questionnaire available <https://new.censusatschool.org.nz/2021-questionnaire> and the variable list (Tables 1 and 2), or a printout of the data table with the variable headings (e.g., Figure 5).

1. In pairs get the students to pick one categorical variable and one numerical variable.
2. For each variable ask them to find the following information:
 - a. State the variable
 - b. What was the survey question asked to collect the data?
 - c. How was the variable measured? E.g., were there specific instructions?
 - d. What are the units, if any, for the variable?
 - e. What are the possible outcomes for the variable?
 - f. What type of data is it? Categorical or numerical?

At this level students are mostly working with categorical and whole number data. Some of the data collected, e.g., reaction time, has up to three decimal places, bag weight is to one decimal place.

If working with data cards, students need to decide which variables are on the data cards. It is useful for students to try to work out what the variables are from the questionnaire and either Tables 1 and 2 or the data table (like Figure 5). Taking the time for students to work out the variables themselves allows them to become more familiar with the data, seeing the possible outcomes for the variable and connecting the outcomes to what they remember about collecting the data.

For example, if working with the data cards given in Figure 3 (repeated below), students might notice that there are three numbers (numerical variables) and two words (categorical variables). Ask them to think what the words might stand for; they might remember that they are eye colour and handedness.

150	149	160	156	138
brown	left brown	left brown	right blue	right blue
23 / 23	28 / 28	25 / 24	24 / 24	22 / 20

Figure 3 repeated. Example data cards from height et al data card set

Then ask them what they notice about the numbers. They might notice that the top number is three digits, to which you can respond, “what measurements did we take that were three digits?” (Height.) They might notice that the bottom two numbers are either the same or very similar, to which you can respond, “what did we measure that was the same or very similar?” (Left foot length, left number on card, right foot length, right number.)

Problem

At level 3 students are exploring multivariate categorical and whole number data to detect patterns, variations, and trends. They are working with data from the whole group – the class.

Students can explore summary investigative questions with the data that is provided to you, the teacher, from the 2021 CensusAtSchool questionnaire. You might start with doing one or two investigative questions together as a class and then allow the students to select their own investigative questions to explore.

Example teacher investigative questions

Summary investigative question (categorical data):

How do the children in our class get to school?

Summary investigative question (numerical data):

How heavy are the school bags of the children in our class?

Predicting class results

Before you start to answer the investigative question get students to predict what they think the class results will be. You could do this by giving students sticky notes to write their guesses on and they can stick their guesses to a chart with the investigative question on it (Figure 6).

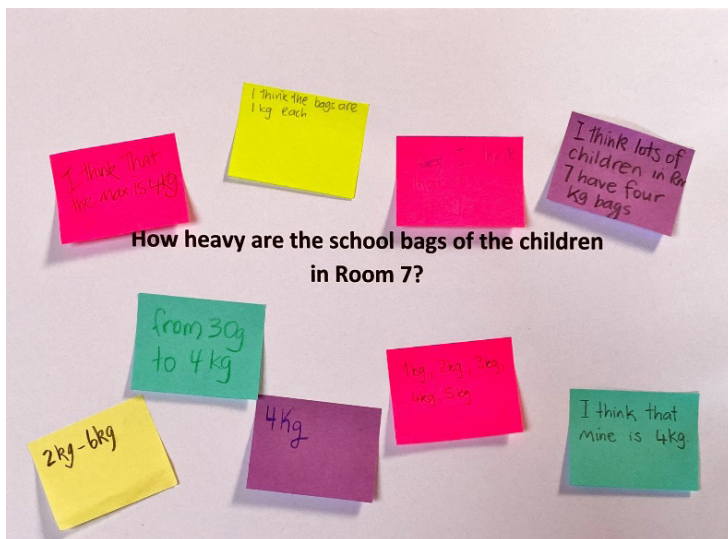


Figure 6. Example of student predictions for an investigative question.

In questions 8-19 there are three categorical variables (eye colour, handedness, and travel method to school) and the remainder are numerical variables. All the numerical variables are measurements, most of which are whole number measurements, with just bag weight and reaction time having decimal places.

Use your discretion as to where you think your students are in terms of working with the bag weight and reaction time variables, as numerical variables with decimal places can provide an opportunity to differentiate the activity, especially when working with the data cards or data tables. Using statistical software does not pose the same issues, as the software automatically places the data according to a suitable scale. When using data cards or data tables students create their own scale.

Analysis

EIS-T pedagogical approach

The dataset can be used to introduce students to using software for data visualisations, moving from enactive mode (manipulating objects) to technology mode (using statistical software). The EIS-T pedagogical approach to introducing students to using software is a combination of Jerome Bruner's (Graham, 2006) education theory about thought processes and Laura Martignon's (2008) extension to include technology.

Simply put the EIS-T approach involves an enactive mode, iconic mode, symbolic mode, and technology mode.

- **Enactive** mode involves introducing students to data visualisations using physical materials, e.g., data cards. Here students can physically manipulate the data cards into data visualisations as they start to organise and sort the data.
- **Iconic** mode involves students starting to use a more structured form of data visualisation, where the structures are provided by the teacher, e.g., a graph outline with scales/labels included. The students add the data to the graph outline. In the iconic mode, data, especially categorical data, may be pre-sorted into a frequency table.
- **Symbolic** mode involves the students abstracting the process of data visualisation and creating their own graphs from scratch, including frequency tables if required.
- **Technology** mode involves using statistical software to create data visualisations, for students who already have knowledge about the inherent structure of the different data visualisations for different types of data, e.g., knowing that a bar graph is an appropriate data visualisation for categorical data.

The EIS-T pedagogical approach is shown in this lesson sequence using the same two investigative questions. In an actual teaching and learning activity it would be more interesting to students to use different investigative questions as they progress through enactive to iconic to symbolic to technology.

Enactive mode – using data cards

Data cards can be used to introduce students to data visualisations and is recommended as a starting point. Data cards allow students to flexibly sort and organise the data for a particular variable.

Another data cards activity: <https://nzmaths.co.nz/resource/data-cards-level-3>

Students may sort the data cards into piles, but by this level students should be sorting data cards so that frequencies and features of the data can be easily seen. This is likely to be a form of picture graph or dot plot.

Once the data visualisation is made, a record should be captured, e.g., a photo taken, and then the photo can be part of the recorded information for the investigative question.

For each data visualisation students should write statements about what they “notice” in the graph. Using the starter “I notice...” is a good writing prompt. You will need to model writing descriptions with your students, which includes “thinking out loud” when you are doing the description. Thinking out loud could be a series of analysis questions that you are asking yourself to support the description. For example:

- What is the most common way students in our class travel to school and how many students travel that way?
- What is the next most common travel method to school... and how many students chose that?
- What is the least common travel method to school...?

Including the context in the description is important. A good description occurs by including the variable, the group, and the values and units for numerical variables, or the outcomes and counts for categorical variables. For example:

- I notice that 21 [count] students in our class [group] travel to school [variable] by car [outcome].
- I notice that the bag weights [variable] for students in our class [group] vary from 0.2 [value] kg [units] to 6.3 [value] kg [units].

Examples of possible data visualisations and accompanying descriptions for the two investigative questions are given below. Note your class data will very likely look different. As the examples are repeated the description is included the first time. An additional note is included for bag weights after the first dot plot is created.

How do the children in our class get to school?

Figure 7 shows the data cards sorted and organised for travel method. From left to right the modes of transport are motor car, walk, bus, other, scooter.

Note that the bars are in order from most frequent to least frequent. This is a common way to organise categories when there is no order to the categories. Watch out for students sorting numerical data in this way, or sorting categories with an obvious order, e.g., year 1, year 2, year 3 when looking at class levels.



Figure 7. Travel method to school for our class (travel method is the variable at the top)

Description

I notice that... 21 students in our class travel to school by car.

I notice that... four students in our class travel to school by walking and four travel by bus.

I notice that... only one student in our class travels to school by scooter.

I notice that... two students in our class come by “other” method to school.

How heavy are the school bags of the children in our class?

Figure 8 shows the bag weights for the students in the class. The values for the variable bag weight have one decimal place. The student who made this display has sorted the bag weights on the whole kilogram value only.

Note that this means that all the bags that weigh 1.something are in the same column. This effectively has created a histogram with the bars being $[0,1)$, $[1,2)$, $[2,3)$, $[3,4)$, $[4,5)$ – no values, $[5,6)$ and $[6,7)$.



Figure 8. Bag weight for students in our class (bag weight is the variable on the right-hand side)

Description

I notice that... the most common interval of bag weights for students in our class is from 2 and up to 3 kg. Six of the nine students in this interval have a bag weight of exactly 2 kg.

I notice that... bag weights for students in our class vary from 0.2 kg to 6.3 kg.

I notice that... most students in our class have a bag that weighs between 1 kg and 4 kg, 22 students have bag weights in this interval.

I notice that... no one in the class has a bag that weighs between 4 and 5 kg.

Iconic mode – making data visualisations using a graph template

Students are introduced to making their own bar graphs and dot plots through a provided graph template. See, for example, the bar graph and dot plot templates from the Fridge pickers activity on nzmaths.

Dot plot template

<https://nzmaths.co.nz/sites/default/files/2020-10/fridge-pickers-2.pdf>

Bar graph template

<https://nzmaths.co.nz/sites/default/files/2020-10/fridge-pickers-3.pdf>

Students can use the data cards or the data table (Figure 5) as data sources to create their graphs in both the iconic and symbolic modes.

How do the children in our class get to school?

Figure 9B is a bar graph of the same data as shown in Figure 7. In this instance the students have created the bar graph using a provided template (Figure 9A).

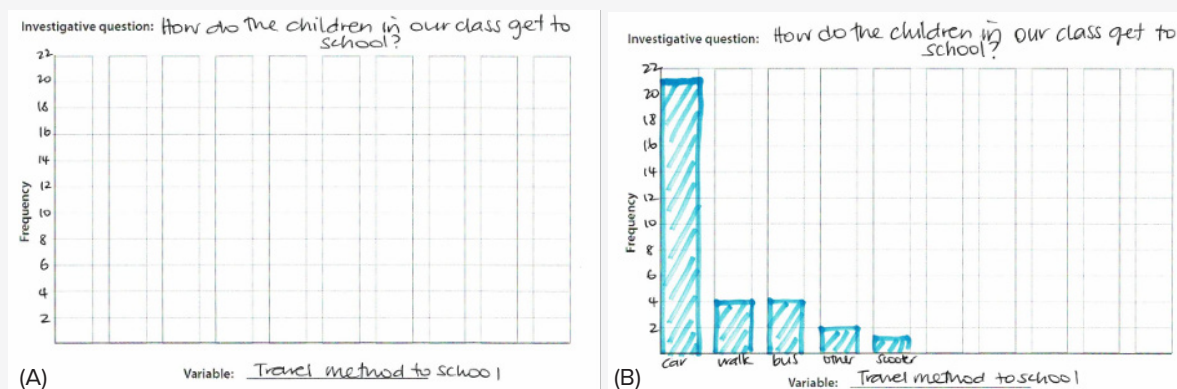


Figure 9. Travel method to school for our class.

(A) blank template provided with labels, (B) completed graph

How heavy are the school bags of the children in our class?

Figure 10B is a dot plot of the same data as shown in Figure 8. In this instance the students have created the dot plot using a provided template (Figure 10A).

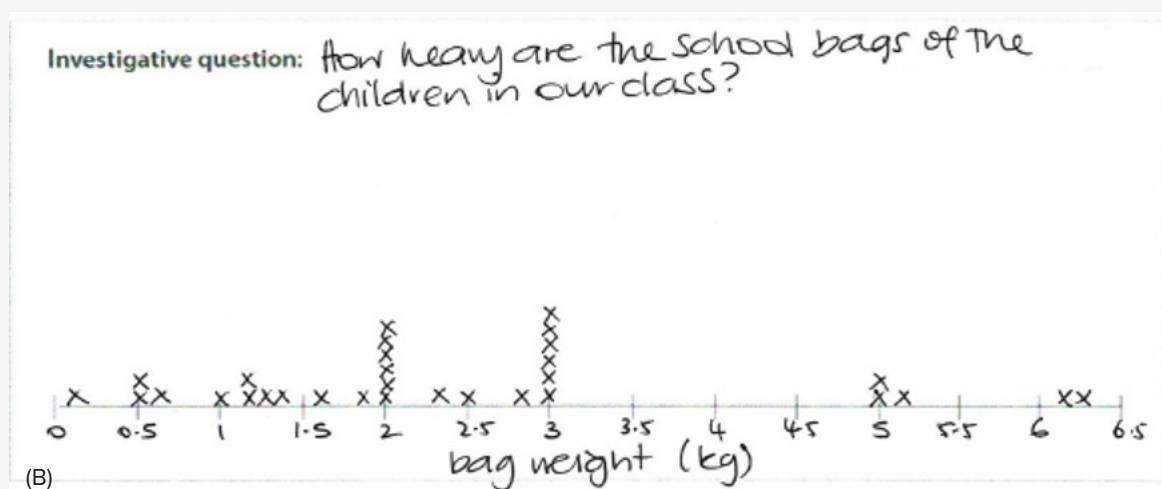
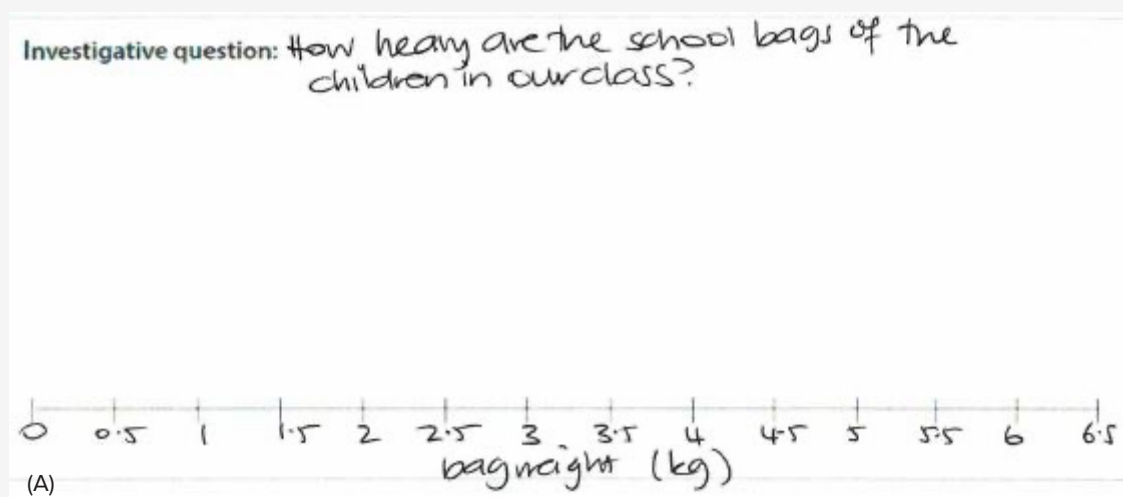


Figure 10. Bag weight for students in our class.
(A) blank template provided with labels, (B) completed graph

Additional description from dot plot

Students may notice the more pronounced gap between the bags weighing 0.2-3 kg and 5-6.3 kg. They may also notice the clustering of bag weights at 2 kg and 3 kg.

Symbolic mode - creating data visualisations from scratch

Students are not given a template, rather they make their own bar graphs and dot plots, building on previous work in the iconic mode. Structures of data visualisations such as scales, labels, titles (e.g., investigative question) are confirmed with the students. If needed, they can make frequency tables of variables to support creating bar graphs.

How do the children in our class get to school?

Figure 11A is a bar graph, drawn from scratch, of the same data as in Figure 7 and Figure 9B. Figure 11B is its corresponding frequency table.

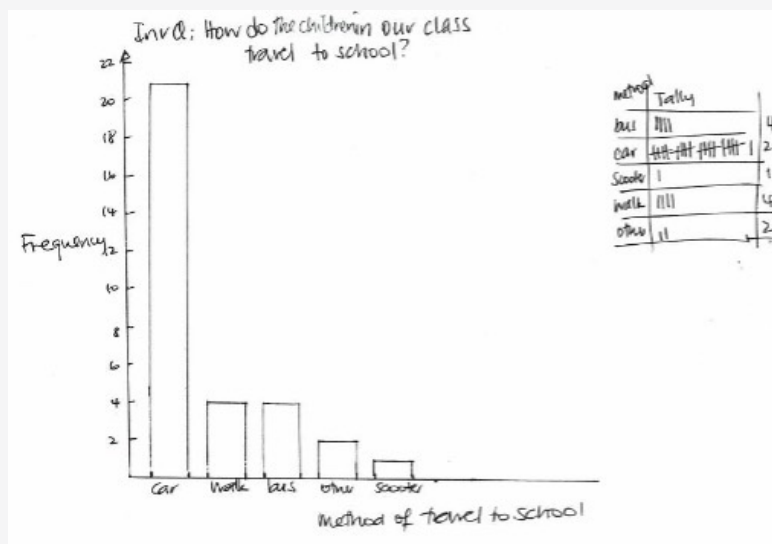


Figure 11. Travel method to school for our class

How heavy are the school bags of the children in our class?

Figure 12 is a dot plot, drawn from scratch, of the same data as Figure 8 and Figure 10B.

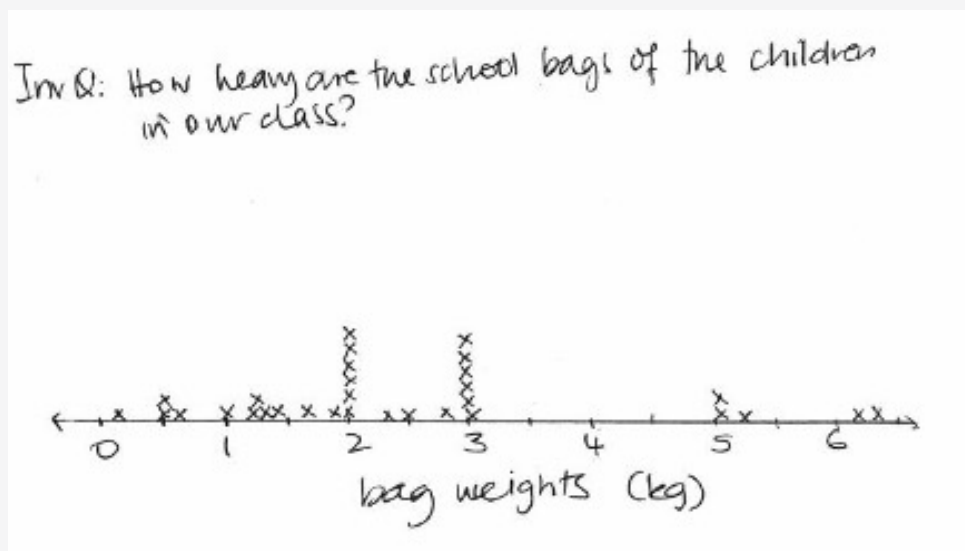


Figure 12. Bag weight for students in our class

Technology mode - using statistical software

Prior to moving to technology, the dataset (in .csv format) is imported into CODAP. The example dataset is available here <https://codap.concord.org/app/static/dg/en/cert/index.html#shared=https%3A%2F%2Fcfm-shared.concord.org%2FJBzTbjFqZNPfyvVcpOuK%2Ffile.json>. The link can be shared with students and the dataset will open as a new document for them to use and make their own data visualisations.

How do the children in our class get to school?

Figure 13 is a bar graph, created in CODAP, of the same data as in Figure 7, Figure 9B and Figure 11

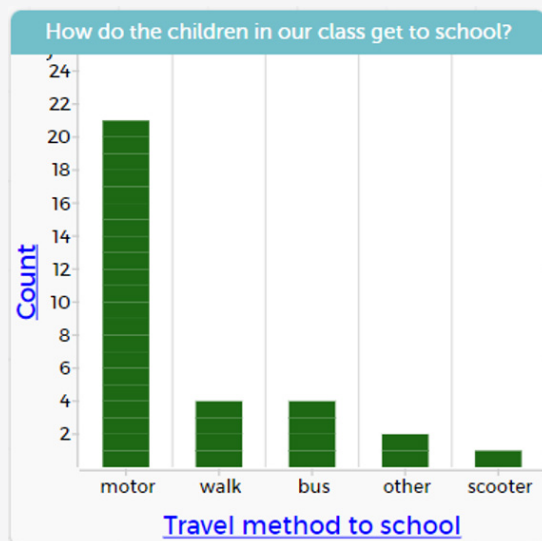


Figure 13. Travel method to school for our class

How heavy are the school bags of the children in our class?

Figure 14 is a dot plot, created in CODAP, of the same data as Figure 8, Figure 10B and Figure 12.

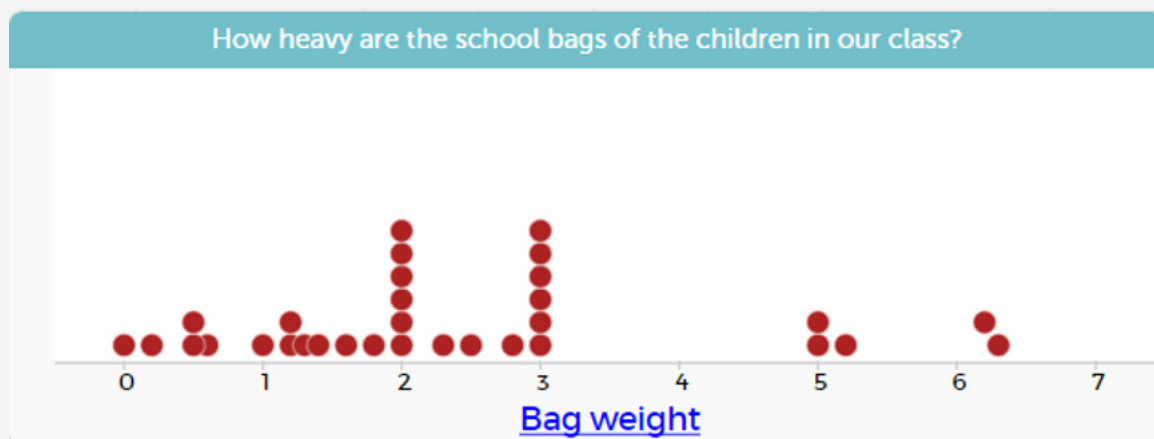


Figure 14. Bag weight for students in our class

Further Support on Using CODAP and Describing Data Visualisations: Curriculum Level 3

Fridge Pickers

In the unit, using our fridge as the context, students collect data and present these as dot plots and bar graphs. They start to learn about using the computer to display their data. They learn to describe data visualisations.

<https://nzmaths.co.nz/resource/fridge-pickers>

Planning a Statistical Investigation

In the unit students will identify how to plan and carry out a statistical investigation about a topic of interest.

<https://nzmaths.co.nz/resource/planning-statistical-investigation-level-3>

Conclusion

The last stage of the statistical enquiry cycle is the conclusion, where we answer our investigative question(s). We use evidence from our analysis to answer the investigative question. Examples for each of the investigative questions are given below. Student responses will differ. There is no **one** right answer, but the answers must align with their evidence in the analysis.

How do the children in our class get to school?

Most of the children in our class get to school by car, 21 out of 32 children come by car. Other methods of transport to school include walking (4 children), bus (4 children), scooter (1 child) and other (2 children).

How heavy are the school bags of the children in our class?

Bag weights of the children in our class vary from 0.2 kg to 6.3 kg. Most of the bag weights are between 1.5 kg and 3 kg, as 17 out of 32 children have bags in this interval. Five children have bags that are heavier than the rest of the class, their bags vary from 5 kg to 6.3 kg.