Unpacking categorical and numerical data:
Teacher notes

http://topdrawer.aamt.edu.au/Statistics/Good-teaching/Data-collection/Types-of-data/Categorical-and-numerical-data

Random data from CensusAtSchool have been used to illustrate possible classroom responses

1. Answer the following two questions.

   (a) How do you usually get to school? Choose one of Car, Bus or Walk.

   (b) Estimate how many minutes it usually takes you to get to school.

2. Now enter the data from the responses of **all** the members of your class. You might want to use the headings below.

Name	Type of travel	Time (minutes)
	Walk	15
	Car	4
	Walk	5
	Bus	15
	Car	20
	Car	5
	Car	3
	Bus	20
	Car	5
	Bus	60
	Bus	15
	Bus	49
	Car	4
	Car	5
	Car	2
	Walk	2
	Car	2
	Car	5
	Walk	5
	Car	5
	Walk	3
	Walk	10
	Bus	25
	Bus	10
	Bus	15
	Car	5
	Bus	23
	Car	10
	Bus	15
	Car	5

(a) Looking at the data in the two columns, how could you go about ordering them?

Students might suggest 'alphabetical' for the type of travel and 'smallest to largest' for the time.

(b) Is there a difference in the way you would handle the two sets of data?

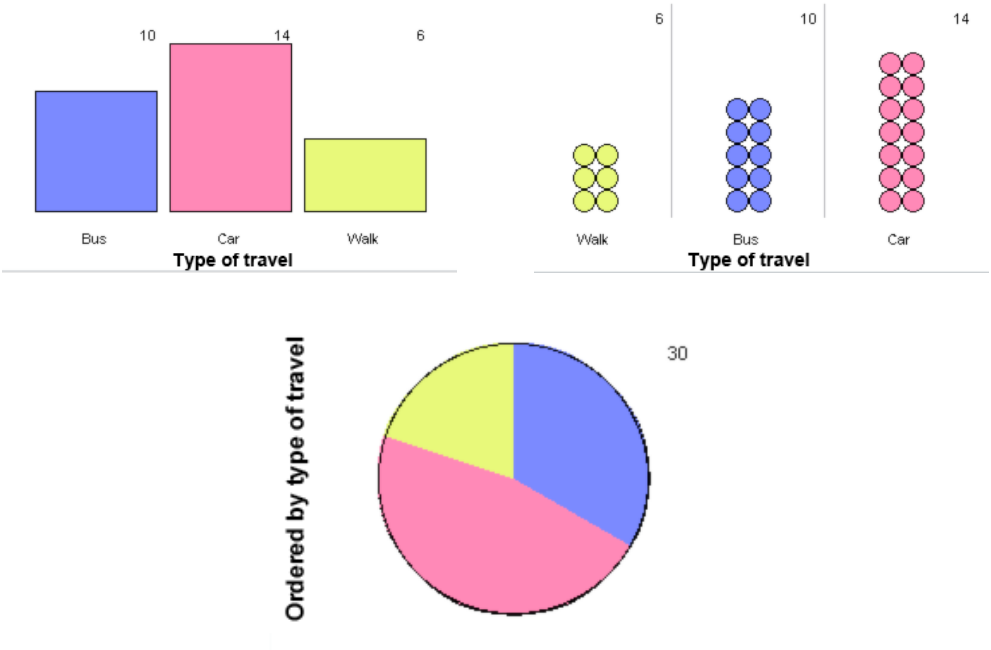
(c) What do you notice in the table about the **type** of travel students in your class use to get to school? Which is the most frequently used type of travel? Which is the least frequently used type?

For these data, the most frequently used is Car and the least frequent is Walk.

3. How could you summarise this information in a table?

Type of travel	Number
Bus	10
Car	14
Walk	6

4. How could you display this summary in a graph?



Although many students may want to draw a pie chart it could be difficult to estimate the size of the sections. Using a strip of paper with 1 cm for each Bus (lined up), each Car, and each Walk, students can then make the strip into a circle, connect the dividing lines to the centre, and draw the segments of the pie chart.
See http://topdrawer.aamt.edu.au/Statistics/Good-teaching/Data-representation/Creating-pie-graphs

5. (a) What do you notice in the original table about how **long** it takes students to get to school? What is the longest time it takes a student? What is the shortest time it takes a student?

The longest time is took someone to get to school was 60 minutes. The shortest time was 2 minutes. It looks like many students arrive at school quite quickly.

(b) Does it appear there are some clumps of times close together? Are there any extreme values, either really small or really large?

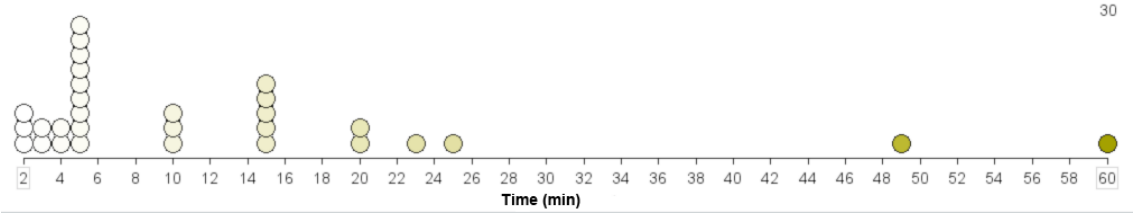
60 minutes might be considered an outlier but since this time is for the bus, the student may live a very long way from the school. There are a few times clumped around 15 minutes and around 20 minutes.

6. (a) Make a list of the values in order from smallest to largest.

2	2	2	3	3	4	4	5	5	5	5	5	5	5	5

5	10	10	10	15	15	15	15	15	20	20	23	25	49	60

(b) How could you display these values in a graph?



(c) Explain what the graph tells you about the number of minutes it takes students in your class to get to school. Discuss clusters and gaps in your graph.

The big gap makes one think that 49 minutes and 60 minutes could be considered outliers. Lots of children seem to get to school in 5 minutes or less.

7. (a) What are the differences in the way the data are presented in the graphs for type of travel and the time taken?

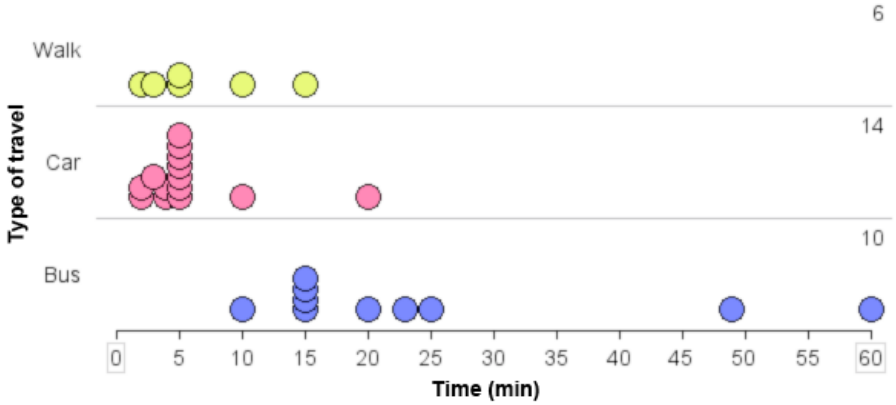
The type of travel graph is reporting categories of data and the time graph reports the numerical values of the times.

(b) Does order make a difference in the type of travel graph? What about in the time taken graph?

Order does not matter in the graphs for type of travel. The first type of travel graph (see answers for question 4) on the right is just as correct as the second one on the left. Order does matter for time, in order to understand the variation in time for different students.

8. The data related to the way students get to school are called *categorical* because they represent categories that cannot be ordered. The data related to the time it takes students to get to school are called *numerical* because we can order them and plot them in a graph on a scale from smallest to largest.

We sometimes use categorical data sets to compare sets of numerical data. The graph below shows the time it takes 30 randomly selected children from around Australia to get to school by the three types of travel: car, bus, and walk.



(a) Create a similar graph for your class.

(b) Write a summary comparing the type of travel and the number of minutes it takes to get to school for your class, to the data from around Australia. Are there differences? If so, why?