# Unpacking categorical and numerical data:
Student worksheet

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

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) |
|---|---|---|
| | | |
| | | |

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

   (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?

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

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

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?

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

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

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

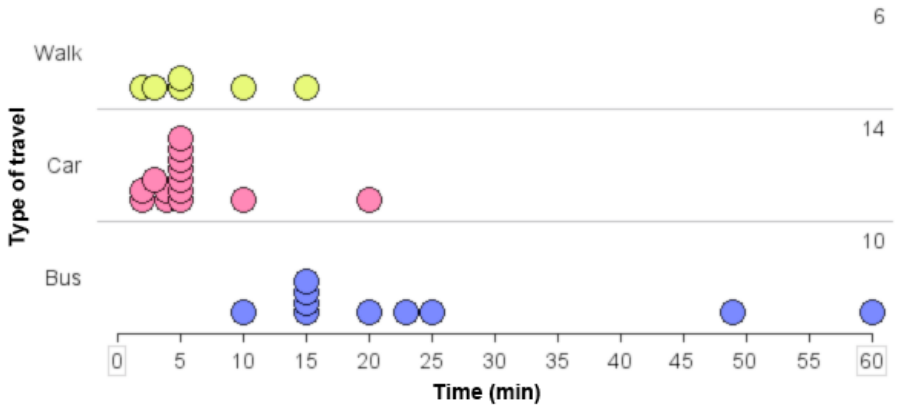
   (c) Explain what the graph tells you about the number of minutes it takes students in your class to get to school. Discuss clusters and gaps in your graph.

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

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

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?