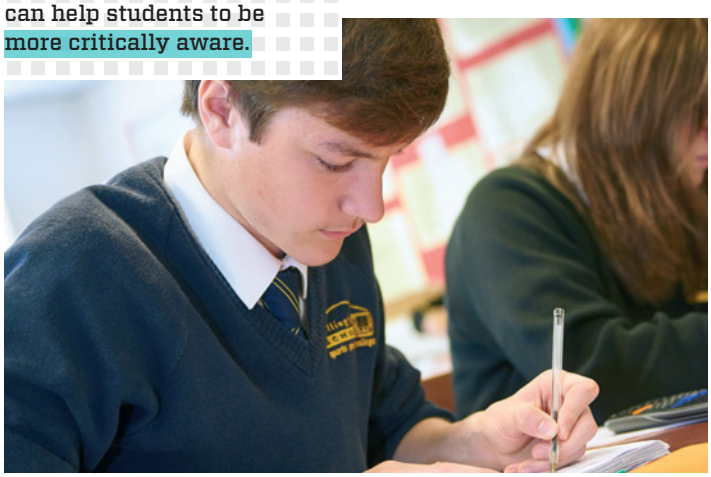
WHY TEACH THIS?

Representations of data in everyday life are often inaccurate or misleading. Knowing what to look for can help students to be more critically aware.

SPORTS CENTRE

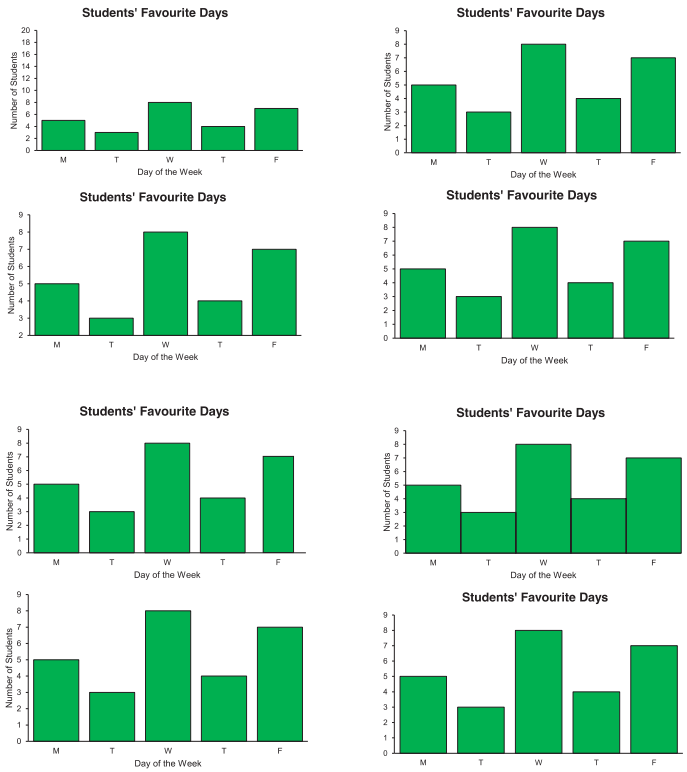
DRAWING BAR CHARTS SOUNDS SIMPLE, BUT WHEN THE CATEGORIES HAVE UNEQUAL WIDTHS HISTOGRAMS ARE NEEDED, SAYS COLIN FOSTER



Students learn about pictograms and bar charts at primary school, but when they are much older and are introduced to histograms and the idea of frequency density they don't always see the point. Drawing a bar chart in which some of the categories have different widths leads to misleading representations, which histograms correct for by dividing the frequency by the class width. To help students appreciate the importance of this, in this lesson students will do something unusual – they will deliberately draw *misleading* bar charts! By using the same data but splitting it into different categories, they will see that the bar charts look wildly different. Then by drawing histograms for each case they will appreciate the benefit of using frequency density if we want to get an accurate sense of the real distribution. This gives students a stronger understanding of why we have histograms, rather than seeing them merely as an alternative kind of chart with strange rules.

STARTER ACTIVITY

Q. What is wrong with these bar charts?
Find as many things for each one as you can.



The bar charts above are available at **tinyurl.com/tssportscentre**

+KEY RESOURCES

ATM's Exploring Area and Fractions with Square Geoboards is an excellent resource to develop understanding of fractions while promoting reasoning to develop fluency and problem solving skills. Go to www.atm.org.uk/shop/act095pk and Look Inside

Students may notice things such as:

- Bars not of equal widths
- Bars touching, which they should not for discrete data
- Missing title and/or axes labels
- Poor choice of scale, with a vertical axis extending unnecessarily far beyond the bars
- Repetition of "T" on the horizontal axis, although in this case it is probably clear that these indicate Tuesday and Thursday
- Vertical axis not beginning at zero, giving a misleading impression of the amounts represented by the bars.

Q. Which of these things do you think are more serious problems? Which are less serious?

Depending on the context, missing labels may make the chart completely incomprehensible. The vertical axis not beginning at zero can be very misleading, as can unequal-width bars.

MAIN ACTIVITY

Give out the Task Sheets A, B and C (available at **tinyurl.com/tssportscentre**). Give about a third of the students sheet A, a third sheet B and a third sheet C.

Q. *Now we are going to look at some other problems that you can have with bar charts. Read through the sheet. Do you understand what to do? Sharifah had a problem with her bar charts – there was something that she was doing that was wrong. You're going to repeat what she did, so you're going to be doing something wrong, and you have to work out what it is!*

> Last Tuesday, 20 people visited the local Sports Centre.
> Sharifah wrote down the age in years of each person:
>
> 2, 2, 4, 7, 14, 16, 16, 17, 19, 21, 28, 34, 36, 38, 45, 52, 63, 79, 82, 93.
>
> She decided to draw a bar chart for this data.
> She wasn't sure what categories to use, so she tried it three different ways:
>
> **A:** $0 < age \leq 20$; $20 < age \leq 40$; $40 < age \leq 60$; $60 < age \leq 80$; $80 < age \leq 100$
> **B:** $0 < age \leq 12$; $12 < age \leq 18$; $18 < age \leq 35$; $35 < age \leq 60$; $60 < age \leq 100$
> **C:** $0 < age \leq 3$; $3 < age \leq 15$; $15 < age \leq 20$; $20 < age \leq 40$; $40 < age \leq 100$
>
> Fill in the frequencies in the table.
> Draw the bar chart.

See if students can make sense of the situation and the instructions on their sheet – each sheet has one of the three sets of categories A, B and C. Students may need reminding that "frequency" means the number of people. (There is sometimes confusion because this is different from "frequency" in science, which tells you how often something happens per unit time interval.)

Give students time to complete the table and draw the bar chart. They should check that their frequencies add up to 20. If anyone finishes early, give them the "Sports Centre" sheet that doesn't have A, B or C at the top, and they can choose their own categories and see what happens.

If you have students sitting next to each other who are working on different sheets, they are likely to notice the different appearances of their bar charts. If not, then at this point mix up the students so that they go to look at each other's charts.

Q. *Why do the charts look so different?*

Students will realise that the different categories that have been chosen are dramatically affecting the look of the bar charts. They may prefer version A, where the bars are equal width, but can they say why the others come out wrongly?

Frequency density corrects for the problem that, other things being equal, a category with larger width is more likely to contain more people. It does this by dividing by the width of the category. This gives a fair comparison between the different categories.
The formula to use is:

$$\text{frequency density} = \frac{\text{frequency}}{\text{class width}}$$

A chart with frequency density plotted vertically is called a **histogram**. This time the bars **should** touch.

Students can now complete the "Histograms" sheet (also available at **tinyurl.com/tssportscentre**). They should use the same categories as those on whichever "Sports Centre" sheet they did, but this time work out **frequency density** for each category and plot this.

Q. *In what ways does the histogram look different from the bar chart? Why?*

Encourage students to compare the shape of the histogram with their bar charts, and the bar charts drawn by other students. Those who used the categories on sheet A should find that their histogram has exactly the same shape, just with a different vertical scale. Other students should find more differences.

Q. *What does the chart tell you about the people who went to the sports centre? Were they mostly old people or mostly young people or what?*

DISCUSSION

This is a very important part of this lesson, so it is worth making sure that you have a decent amount of time left for this.

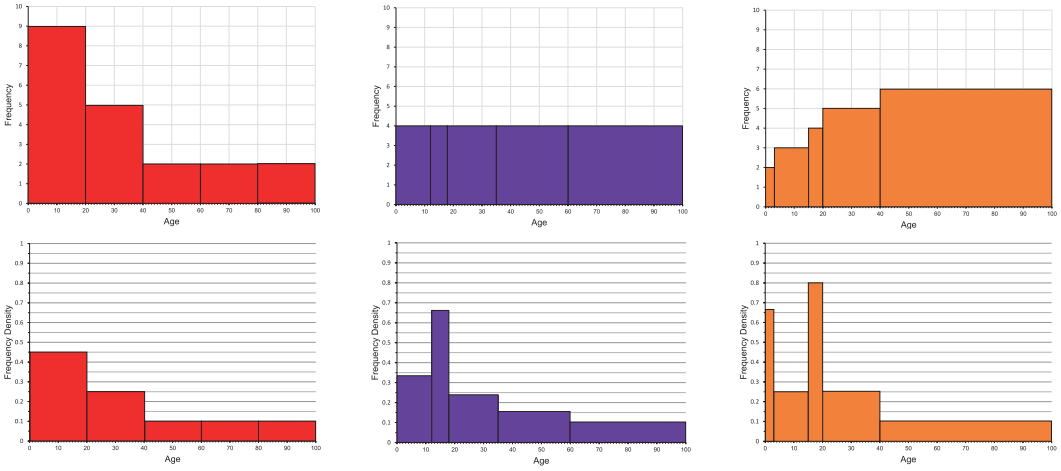
Q. What did your charts look like? Why do you think they looked different from each other?

The correct frequencies (f) and frequency densities (fd) are given in the table below:

A			B			C		
Age (years)	f	fd	Age (years)	f	fd	Age (years)	f	fd
0 < age ≤ 20	9	0.45	0 < age ≤ 12	4	0.33	0 < age ≤ 3	2	0.67
20 < age ≤ 40	5	0.25	12 < age ≤ 18	4	0.67	3 < age ≤ 15	3	0.25
40 < age ≤ 60	2	0.1	18 < age ≤ 35	4	0.24	15 < age ≤ 20	4	0.8
60 < age ≤ 80	2	0.1	35 < age ≤ 60	4	0.16	20 < age ≤ 40	5	0.25
80 < age ≤ 100	2	0.1	60 < age ≤ 100	4	0.1	40 < age ≤ 100	6	0.1

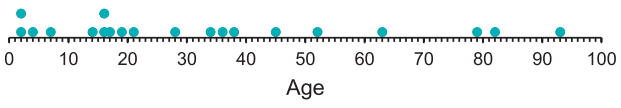
(Frequency densities are rounded to 2 decimal places where necessary.)

The bar charts and histograms for A, B and C are shown below (without titles, to save space):



Bar chart A is not misleading, because all of the class widths are equal, and this means that its corresponding histogram is identical in shape, with just a different scale on the vertical axis. Bar charts B and C, however, are both very misleading, and don't accurately show us the distribution of ages of people who used the sports centre. B appears to be a uniform distribution, and C gives the impression that older people used the sports centre more, which is not true. The histograms for B and C are much more faithful to the data and show us that younger people used the sports centre more. They have some spikes in the younger ages, but correctly show the general trend of smaller numbers of older people.

The scenario in this lesson is a little artificial, as with just 20 pieces of data it might have been more revealing to use a stem-and-leaf plot, or a dot plot, as shown below:



This avoids splitting the data into categories and shows clearly that the majority of the users (14 out of 20) were actually under 40. (In fact, the median age is 24.5 years.) In this lesson the splitting into categories was purely to help students understand the differences between bar charts and histograms.

Q. What is the difference between a histogram and a bar chart? Why are histograms necessary?
Pupils could write down a summary answer to these questions.

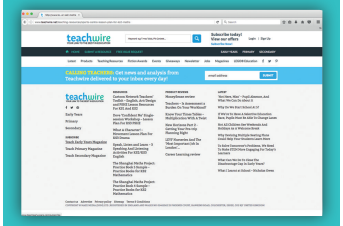
INFORMATION CORNER

ABOUT OUR EXPERT

Colin Foster is an assistant professor in mathematics education in the School of Education at the University of Nottingham. He has written many books and articles for mathematics teachers (see www.foster77.co.uk).

ADDITIONAL RESOURCES

TASK SHEETS CONTAINING ALL OF THESE RESOURCES ARE AVAILABLE AT TINYURL.COM/TSSPORTSCENTRE



STRETCH THEM FURTHER

CONFIDENT STUDENTS COULD BE ASKED TO CONSIDER OTHER WAYS OF REPRESENTING THIS DATA, SUCH AS PICTOGRAMS, PIE CHARTS, STEM-AND-LEAF PLOTS, BOX-PLOTS AND CUMULATIVE FREQUENCY GRAPHS – AND THE PROS AND CONS OF THESE DIFFERENT REPRESENTATIONS