

Working with data

Objectives

By the end of these sessions teachers will:

- have considered strategies for teaching students to use data in statistics;
- have discussed how data is analysed, represented and interpreted;
- have done some calculations with data.

Resources

For the trainer

- Computer with data projector, *Autograph*, Microsoft PowerPoint and Presentation 12.ppt
- Interlocking cubes and a number line
- *Curriculum Standards for mathematics: Grades K to 12*
- Copies of the evaluation form for Day 4

For each school group

- Computer with *Autograph*
- Graphics calculator

For each teacher

- *Teacher's pack*
Handouts 12.1–12.2
Handout 12.3 (to be given out at the relevant point of the session)
Evaluation form for Day 4
- Sheet of graph paper
- *Curriculum Standards for mathematics: Grades K to 12*

Session outline

The data handling cycle Slides 12.1–12.2	Whole group presentation	10 minutes
Pedagogical issues Slides 12.3–12.4 Handout 12.1	Whole group presentation and discussion Task 1: Answering questions	50 minutes
Teaching by doing 1 Handout 12.2 Handout 12.3	Whole group discussion Paired or individual work Task 2: Analysing data	30 minutes
Software demonstration	Software demonstration	20 minutes
Teaching by doing 2 Handout 12.3	Task 3: Using <i>Autograph</i>	40 minutes
Summary Slide 12.5 Evaluation form for Day 4	Whole group presentation	10 minutes

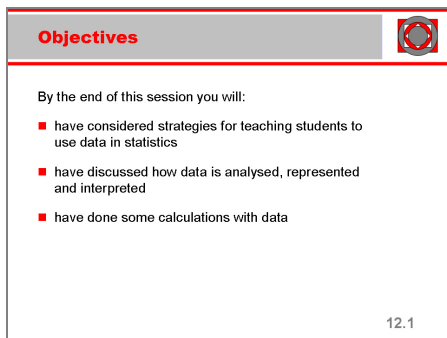
The data handling cycle

10 minutes

Tell everyone that this session will be about analysing data. Its main purpose is to consider how students might analyse data in statistical investigations. The focus is on Grades 7 to 10, so that the statistics will be relatively uncomplicated.

Explain that this is a double session and that there will be a break for lunch in the middle of it.

Show the objectives for the session, as on **slide 12.1**.



Objectives

By the end of this session you will:

- have considered strategies for teaching students to use data in statistics
- have discussed how data is analysed, represented and interpreted
- have done some calculations with data

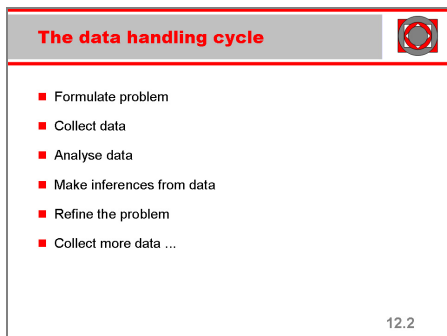
12.1

Before the session starts, brief any interpreter about the key points of the session.

Organise the IT equipment so that each school group can sit round a computer equipped with *Autograph*. Make sure that each school group has a graphics calculator.

Load **Presentation 12.ppt** and *Autograph* on your own computer.

Remind everyone about the importance of the data handling cycle, which was introduced in the previous session. Show **slide 12.2**.



The data handling cycle

- Formulate problem
- Collect data
- Analyse data
- Make inferences from data
- Refine the problem
- Collect more data ...

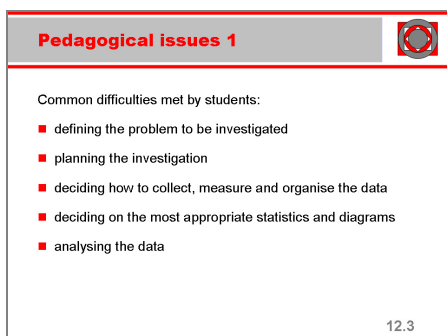
12.2

Point out that we now are in the middle of the cycle.

Pedagogical issues

50 minutes

Ask teachers to describe problems they have met when students have been set work to analyse statistical data. Hear all their comments, before comparing the sticking points with those on **slides 12.3 and 12.4**.

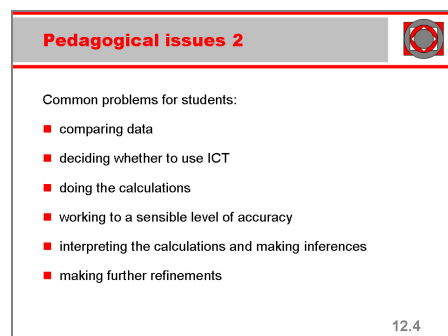


Pedagogical issues 1

Common difficulties met by students:

- defining the problem to be investigated
- planning the investigation
- deciding how to collect, measure and organise the data
- deciding on the most appropriate statistics and diagrams
- analysing the data

12.3



Pedagogical issues 2

Common problems for students:

- comparing data
- deciding whether to use ICT
- doing the calculations
- working to a sensible level of accuracy
- interpreting the calculations and making inferences
- making further refinements

12.4

Stress that teachers need to spend time helping students with their planning of a statistical investigation. Refer back to the writing frame mentioned in the previous session in **Handout 11.2**. Say that any enquiry needs to be focused. Students need to frame specific questions or hypotheses. They must consider what the population is and what samples to take. When they have collected the data, they need to define the type of data and decide what statistical techniques to apply. Students often embark on an enquiry without doing this level of thinking and waste a lot of time and effort applying techniques that are not appropriate or which duplicate each other.

Discuss with the whole group the pedagogy surrounding analysing data. Ask teachers how they have tried to resolve their difficulties. Record the key points of the discussion on a flipchart.

Background notes for trainers

It is almost certain that students' difficulties will centre on planning the investigation, including making decisions about the appropriate statistics to use and which diagrams and charts are appropriate to support the chosen statistics. Another difficulty is how to make meaningful comparisons using different data displays for similar data from different sources, or for data from different samples. It is therefore important to hear from teachers what they do to resolve these difficulties. The importance of using correct vocabulary should be stressed here, so that teachers can engage in meaningful dialogue on technical issues.

Some school groups might volunteer some case histories but some teachers may have no experience of students doing whole-scale statistical enquiries. It is possible that students will have been taught specific techniques in isolation and merely completed exercises from textbooks.

Explain that statistical techniques are used mainly for the purpose of comparison. Show teachers two sets of interlocking cubes representing heights: set A {2, 5, 6, 7, 10} and set B {2, 2, 6, 10, 10}. Describe each set slowly and deliberately. For example, say:

Each set has five members. We can see that the sets are different in appearance. However, when we begin to work out some summary statistics we find that the values are the same. Each has a mean of 6 [demonstrate this by moving cubes from the tallest towers to the shortest so that all the towers are the same height], a median of 6 and a range of 8. One set has a mode of 10 while the other has no mode. This is a good example of how the modal value is not really of relevance in most situations. How, then, can we summarise each set so that we can compare them?

Place the towers of cubes on a number line. This will demonstrate that one set is more spread out than the other.

Point to the tower of cubes that represents the mean. Discuss how the other towers deviate from the mean. Slowly work through the calculations, putting the results in a table, for deviation, square of deviation, sum of squares, mean of sum of squares, and finally the square root of the mean of the sum of square of deviations. Work through these again and build up in reverse the formula for standard deviation.

Task 1: Answering questions

Refer to **Handout 12.1**. Ask teachers to discuss the questions in small groups. Allow 10 minutes. Take feedback, asking them to present their answers with explanations.

Teaching by doing 1

30 minutes

Suggest that is essential not to do statistics in a vacuum: the more that teaching and learning is linked in with doing statistics centred around real data, the more meaningful it will become.

Tell teachers that they are now to do some statistics for themselves. They will need **Handout 12.2**.

First, discuss data set 1 on the handout.

Ask teachers:

- What sort of data is this? (Quantitative, discrete and measured on a ratio scale)
- What can you do with this data? (There are many possibilities here. Press teachers to give as many examples as possible.)

Task 2: Analysing data

Ask teachers to work in pairs. Give out **Handout 12.3** and a sheet of graph paper to each pair. Remind them that each school group has a graphics calculator. Ask the pairs to work on Task 2. Help any teacher having difficulties with the questions. (The questions are straightforward, but question 10 should not be attempted without a calculator with statistical functions.)

Allow enough time for the pairs to answer all the questions, with the exception of question 10. Discuss their solutions. Ask them for any comments or reactions.

Say that there is much more to the data than meets the eye. Teachers could, for example, draw a frequency polygon, a dot plot, a box-and-whisker plot (and answer questions on all that relate to quartiles and the range of the data), a cumulative frequency distribution, and they could estimate the mean and variance of the sentence lengths in the whole of Darwin's book.

Software demonstration

20 minutes

Ask teachers to move into school groups around a computer. Demonstrate the ideas using *Autograph*. Ask each school group to follow the demonstration on their computer.

Tell them they have just explored the possibility of working with what is really a comparatively simple data set, but it has great potential to illustrate many statistical ideas.

Teaching by doing 2

40 minutes

Task 3: Using Autograph

Ask teachers to work in school groups to do Task 3 on **Handout 12.3**. Help them as necessary. Explain that when they have completed the task you will ask one group to do a presentation.

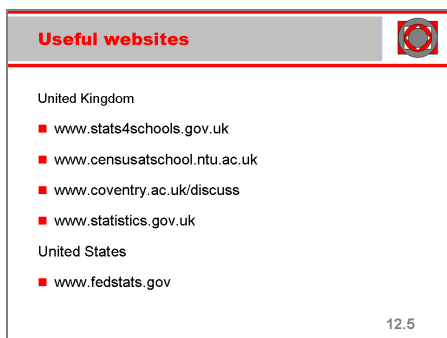
Take feedback, choosing one group to present their findings. Tell the other groups to ask questions and provide assistance so that everyone gets involved in a discussion about the results.

Summary

10 minutes

Ask teachers to recall some of the common difficulties experienced by their students, the potential of data sets to develop the statistical understanding of their students, and what sorts of inputs they themselves will have to make to ensure positive learning takes place

End by showing **slide 12.5**, which shows some useful websites in the UK and the USA, which can be used as sources of large data sets.



Explain that the first three UK ones are specifically school related, whereas the last two bullets are the government statistic sites for the UK and the USA.

Background notes for trainers

The site www.stats4schools.gov.uk is a large UK database developed for schools by the Department of Education. Anyone can use it to draw samples from, and the project is centred around data of interest to students.

The site www.censusatschool.ntu.ac.uk may be of even more interest to schools in Qatar. This is an international project run by the Royal Statistical Society Centre for Statistical Education, and has comparative data from schools in the UK (from the UK alone there are more than 60 000 entries in the database), in Canada, in Australia and in South Africa. Any school, in any country, is encouraged to join this project to use and to pool data. Random samples from the database can be downloaded for analysis.

Related to CensusAtSchool is the website www.coventry.ac.uk/discuss, which informs teachers how to use Excel spreadsheets to handle data collected by the CensusAtSchool project.

Ask teachers to complete the evaluation sheet for Day 4. Thank them for their participation during the day.