A Guide to Working with Data Handling

Teaching Approach

When teaching data handling skills, it's important to talk about the statistical process. The six interconnected stages of the process are: posing a question, collecting data, organizing data, summarizing data, representing data and interpreting (analyzing) data.

In each video lesson, we explain and demonstrate only one or two examples of each concept or method that we use. We leave it to you to reinforce the learning with many examples for learners to work on in class and for homework. You can use the activities given in the task video. Full answers are also provided for every worked example in each lesson and for every activity in the Task Answers. In addition there are online resources provided where you can find more examples.

To get the full benefit of the lessons, your learners need to engage actively with the concepts presented. So, when you preview the videos, think about how to introduce each lesson and what follow up activities will be useful. Also watch out for places in the video where you can pause to have a class discussion. We have used this pause icon to suggest some of these places to you.

This is an important section and it can be fun and enjoyable. Data Handling is a section where you as the teacher can incorporate many real life issues or topics from your learners' local communities. It is also a section where you as the teacher can expose the learners to statistics and how they play a prevalent role in the local communities, for example, the percentage of learners who drink while underage, or the number of people in the community voting for each political party. They can look at how it affects the everyday lives of the learners and members of the community. As a teacher, you can use Data Handling as a tool to educate your learners, and it could be especially effective if you could pair it up with their subject Life Orientation.

We have included additional activities to supplement the learning in the videos and the skills covered in the task video. Use these as class activities or as homework exercises.

Video Summaries

Some videos have a 'PAUSE' moment, at which point the teacher or learner can choose to pause the video and try to answer the question posed or calculate the answer to the problem under discussion. Once the video starts again, the answer to the question or the right answer to the calculation is given.

Mindset suggests a number of ways to use the video lessons. These include:

- Watch or show a lesson as an introduction to a lesson
- Watch of show a lesson after a lesson, as a summary or as a way of adding in some interesting real-life applications or practical aspects
- Design a worksheet or set of questions about one video lesson. Then ask learners to watch a video related to the lesson and to complete the worksheet or questions, either in groups or individually
- Worksheets and questions based on video lessons can be used as short assessments or exercises
- Ask learners to watch a particular video lesson for homework (in the school library or on the website, depending on how the material is available) as preparation for the next days lesson; if desired, learners can be given specific questions to answer in preparation for the next day's lesson

1. How to Collect Data

This lesson focuses on the different methods that can be used to collect data. Advantages and disadvantages, usefulness and appropriateness of each method will be discussed through the use of examples.

2. Developing Questions

Learners are introduced to how to develop suitable, reliable and unbiased questions that can be used to collect data. The lesson also introduces the concept of sample and population, as well as what constitutes a good questionnaire.

3. Organising Data

The focus of this first lesson lies in the organizing and sorting of the data. This is done through the use of frequency tables and tallies, after the data has already been collected. The use of intervals in frequency tables is explored.

4. Trends in Data

The focus of this lesson is working with summarizing data and finding averages, in order for various conclusions concerning the data to be made. We will also be looking into the effects that outliers might have on the mean, median, mode and range.

5. Types of Graphs

The focus of this lesson is on the different types of graphs one can use to represent data and how each type of bar graph offers a different overall picture of the data.

6. Manipulating Graphs

In this lesson we will focus on the different graphs that have been drawn incorrectly or that have been altered to misrepresent data.

Resource Material

Resource materials are a list of links available to teachers and learners to enhance their experience of the subject matter. They are not necessarily CAPS aligned and need to be used with discretion.

1. How to Collect Data	http://prezi.com/vijvu05p40pl/weydon-school-maths-10y1/	A Prezi presentation on different types of sampling.
	http://www.mindset.co.za/resources//0000062933/0000136510/0000139974/LN%20Calculating%20with%20Data_L1.pdf	A Mindset worksheet on collecting data, aligned to NCS Mathematics.
	http://www.mindset.co.za/resources//0000062935/0000136512/0000139976/LN%20Calculating%20with%20Data_L2.pdf	A Mindset note on how to collect data, aligned to NCS Mathematics.
2. Developing Questions	http://www.flickr.com/photos/83006968@N06/7635447286/	Examples of types of questionnaires and types of questions.
	http://www.flickr.com/photos/47194823@N02/4327782900/	
	http://www.bbc.co.uk/schools/gcsebitesize/dida/managing_projects/evaluatingrev6.shtml	
3. Organising Data	http://www.bbc.co.uk/schools/gcsebitesize/maths/statistics/representingdata3hirev1.shtml	Classifying data into different types.
	http://www.flickr.com/photos/61891052@N02/5678591925/	Organising Data using a tally table.
4. Trends in Data	http://www.bbc.co.uk/schools/gcsebitesize/maths/statistics/measuresofaveragerev3.shtml	Understanding the measures of central tendencies (mean, median and mode).
	http://www.bbc.co.uk/schools/gcsebitesize/maths/statistics/measuresofaveragerev6.shtml	Advantages and disadvantages of mean, median and mode.
5. Types of Graphs	http://www.basic-mathematics.com/bar-graphs.html	Understanding bar graphs
	http://www.basic-mathematics.com/bar-graphs.html	
	http://www.emathzone.com/tutorials/basic-statistics/component-bar-chart.html	Understanding compound bar graphs.
	http://www.mathsisfun.com/data/pie-charts.html	Understanding how to draw a pie chart.
	http://www.edplace.com/worksheet_preview.php?eId=1975&type=nc	How to read and understand a pie chart.
	http://www.mathgoodies.com/lessons/graphs/line_graph_part2.html	Understanding line graphs.
	http://www.bbc.co.uk/bitesize/ks3/maths/handling_data/representing_data/revision/3/	Broken line graphs and trends.
6. Manipulating Graphs	http://www.youtube.com/watch?v=oGGYIw_pIj8&feature=related	A YouTube clip with multiple statistics which will provide for interesting debate in class.
	http://en.wikipedia.org/wiki/Sampling_bias	Sampling bias: This offers advanced enrichment for a teacher and not recommended for learners.
	http://www.mathsisfun.com/data/survey-conducting.html	How to do a survey: Site also has interactive games and printable worksheets for all junior grades.

Task

Question 1

Look at the following questions and state if they are reliable or not. If not, rewrite the question to make it reliable.

1.1 Do you own a laptop?
1.2 If you do have a laptop, do you use it for schoolwork?
1.3 Do you drink alcohol?
1.4 Have you ever drunk alcohol while underage at school?
1.5 How old are you?
1.6 What age should it be made legal for learners to smoke at school?
1.7 Do you play a sport at school or with a club?

Question 2

For each of the following scenarios, decide which method of collecting data would be the best choice. Provide a reason for your answer.

2.1 Gathering data about the choice of sports learners in Grade 10 do at your local school.
2.2 Gathering data about the vehicles that drive past the spaza shop between the times of 14:00 and 17:00 every Monday, Wednesday and Friday.
2.3 Gathering data about why learners do not like attending assembly at school.

Question 3

3.1 Data can be classified into two main types. What are these two types?
3.2 What is the difference(s) between discrete and continuous data?
3.3 Classify each of the following statements below as categorical, discrete or continuous data.
    3.3.1   The number of SMSs you send per day.
    3.3.2   The number of Grade 11's who play soccer at your school.
    3.3.3   The temperature in Pretoria over the months of June and July.
    3.3.4   The brands of cellphones that the learners in your school own.

Question 4

The Grade 11's with their Life Science teacher conducted a study using fifty bean plants. These plants were measured in centimetres to see how tall they were. The heights were rounded to the nearest centimetre. After each plant was measured, the data was recorded in groups and a frequency table was constructed. The frequency table is shown below.

Height of bean plants (cm)	Number of plants
20 - 24	2
25 - 29	6
30 - 34	14
35 - 39	16
40 - 44	10
45 - 49	2
Total	50

Study the frequency table and answer the following questions.

4.1 Complete the following table by adding in the tally.

| Height of bean plants (cm)	Tally	Number of plants
20 - 24		2
25 - 29		6
30 - 34		14
35 - 39		16
40 - 44		10
45 - 49		2
Total		**50**

4.2 How many bean plants had the height of less than 35cm?

4.3 How many bean plants had the height of 19cm or less?

4.4 How many bean plants had the height of between 25 – 44cm?

4.5 What conclusions could you draw from this frequency table and the data?

Question 5

Your school principal is very interested in what types of food the learners are buying at the school tuck shop. He believes that a healthier approach to school lunches is needed.

Here is a pie chart that shows the type of lunches bought by a random sample of 50 learners.

TUCKSHOP MEALS SOLD
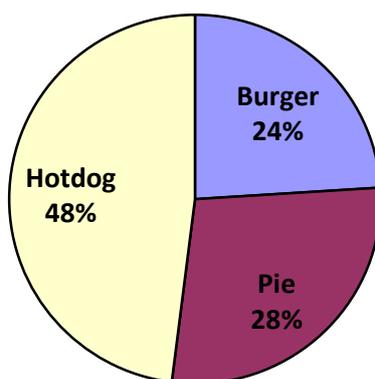
Use the pie chart to answer the following questions.

5.1 How many learners completed the survey?

5.2 How many of the learners bought pies?

5.3 Estimate what fraction of learners bought burgers.

5.4 Estimate what fraction of learners bought hotdogs.

5.5 What was the most and least popular choice of tuck shop food?

5.6 How many degrees is represented by each sector of the pie chart?

Question 6

The data set below shows the number of hours in a week twenty-five teenagers spent watching television. Their times have been rounded off and are correct to the nearest hour.

17	14	18	19	12
6	15	13	12	10
11	16	10	15	30
22	7	12	24	5
25	8	9	16	27

6.1 Arrange the data in ascending order.
6.2 Calculate the mean (average) time teenagers spent watching television. Round your answer off to the nearest hour.
6.3 What does the mean tell us about the time teenagers spent watching television?
6.4 Calculate the median.
6.5 Why should we use the mean to discuss this data, instead of the median or mode?

Question 7

Study the following graph, and answer the following questions.

Maths Marks for Term 1
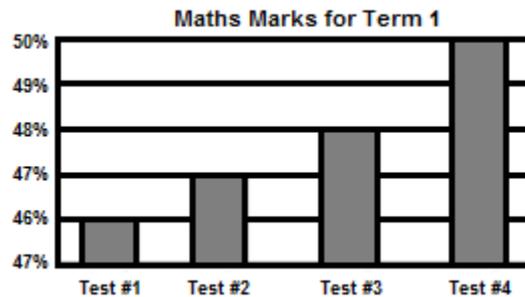
7.1 What are your first impressions of seeing this graph? Describe what do you see the graph is portraying, as if you were explaining this to your parents.
7.2 Do you think that the graph is misleading? If yes, in what ways is this graph misleading?

Task Answers

Question 1

1.1 Reliable

1.2 Unreliable – not all learners would have a laptop, and could have answered "no" to the previous question. Rather ask, "do you use a computer to help you with your school work?"

1.3 This question could be reliable, as it is straight forward, however, one needs to know the context behind this question.

1.4 Unreliable – the question is judgemental and learners could lie in their answers. Rather ask, "Have you ever drunk alcohol?"

1.5 Reliable

1.6 It is a reliable question, but whether it is an appropriate question to ask school learners or not is the issue. Choice answers might need to be provided.

1.7 Reliable

Question 2

2.1 Questionnaire. It is an excellent way of collecting data concerning a specific point from large amounts of people. It is quick and easy to do.

2.2 Observation. In this scenario, you would be collecting data about either the colour or make of the cars. It would not be possible (or safe) to ask the drivers of the cars for this information. By observing, you are able to collect this data from a safe distance.

2.3 Interviews. By speaking face to face with the learners, you would be able to collect accurate, personal information. The learners should give a variety of responses instead of just choosing a readymade multiple - choice answer.

Question 3

3.1 Categorical and numerical data.

3.2 Discrete data usually contain certain whole individual numbers and no values in between, where continuous data can contain fractions and can be any value.

3.3.1 Discrete

3.3.2 Discrete

3.3.3 Continuous

3.3.4 Categorical

Question 4

4.1

Height of bean plants (cm)	Tally	Number of plants
20 - 24	||	2
25 - 29	ЖГ |	6
30 - 34	ЖГ ЖГ ||||	14
35 - 39	ЖГ ЖГ ЖГ |	16
40 - 44	ЖГ ЖГ	10
45 - 49	||	2
Total		50

4.2 14 + 6 + 2 = 22

4.3 0

4.4 6 + 14 + 16 + 10 = 46

4.5 One can see that the majority of the bean plants grew to a height of between 35-39cm. The least amount of plants grew to heights of between 20-24cm and 45-49cm.

Question 5

5.1 50 learners

5.2 $\frac{28}{100} \times 50 = 14$ learners

5.3 $\frac{24}{100} \approx \frac{12}{50} \approx \frac{6}{25}$

5.4 $\frac{48}{100} \approx \frac{12}{25}$

5.5 Most popular: hot dogs. Least popular: Burgers

5.6 Pies: $\frac{28}{100} \times 360 = 100{,}8° \approx 101°$

Burgers: $\frac{24}{100} \times 360 = 86{,}4° \approx 86°$

Hotdogs: $\frac{48}{100} \times 360 = 172{,}8° \approx 173°$

Therefore, $101° + 86° + 173° = 360°$

Question 6

6.1 $5, 6, 7, 8, 9, 10, 10, 11, 12, 12, 12, 12, 13, 14, 15, 15, 16, 16, 17, 18, 19, 22, 24, 25, 27, 30$

6.2 $\frac{353}{25} = 14{,}12 \, hours \, \approx 14 \, hours$

6.3 The mean tells us the average amount of time the all 25 teenagers spent watching TV in one week, and approximately 14 hours is a lot of time spent watching TV.

6.4 Median = 14.

6.5 We should use the mean instead of the median and mode because the mean uses all of the data, and does not leave out data like the median and mode.

Question 7

7.1 This graph could be used by a learner to indicate how his or her Math's marks are improving over the term.

7.2 Yes, it is misleading. There are two problems with the vertical axis: the first is that the highest mark on the axis is 50%, which could draw our attention away from the fact that the learner's marks are so poor. The axis should extend to 100%, because that is the highest mark achievable for each test. The second is that the scale of the axis has been enlarged, so it exaggerates the increase in the percentages. In reality, the marks range from 46% to 50%, which is not a big change in this context.

Additional Activities

Activity 1

The schools in your local district conducted survey with certain learners about which mode of transport learners used to get to school, and back home again. The survey consisted of five questions. One of the questions was: "What mode of transport do you use to get to and from school?" Here are the responses:

walk	bicycle	taxi	walk	car
bicycle	bicycle	walk	walk	car
Walk	bus	walk	taxi	walk
car	bus	bus	bus	bus
car	walk	bus	skateboard	taxi
taxi	motorcycle	bus	motorcycle	skateboard
walk	bicycle	walk	bus	car
walk	taxi	bicycle	walk	bus

1.1 How many people took part in the survey?
1.2 Construct a frequency table that represents this data.
1.3 What mode of transport was the least used? Give a reason for your answer as to why do you think so.
1.4 What conclusions could you draw from this frequency table and the data?

Activity 2

A Grade 11 learner has conducted a survey to determine which colour of car was the most popular. As her time was limited, she conducted the survey on the first 200 cars that passed a certain point on the main road. She calculated the following data:

Colour	Percentage of vehicles	Number of vehicles	The angle of segment
Red	11		
Green	3		
White	44		
Blue	14		
Silver	11		
Other	17		

Use the above table and information to answer the following questions. Remember to show all of your working out.

2.1 Calculate the number of vehicles there were per colour.
2.2 Calculate the number degrees per car colour segment.
2.3 Using the table, construct an accurate pie chart. Measure each angle to see if the pie chart was drawn accurately.
2.4 How would you know if your pie chart was accurately drawn?

Activity 3:

Two schools were invited to compete against each other in a Mathematics competition. School A was from the Western Cape and School B was from Gauteng. Each school has six Grade 11 classes. The table below shows the mean averages of each class, per school, after taking part in the competition.

School A	27	33	33	45	47	51
School B	25	29	37	42	42	51

3.1 Find the measures of central tendencies for both school A and B. Show all calculations.

3.2 Calculate the range (measure of dispersion) for each school. Show all calculations.

3.3 Is range a good tool to compare the two schools? Explain your answer.

Additional Activities Answers

Activity 1

walk	bicycle	taxi	walk	car
bicycle	bicycle	walk	walk	car
Walk	bus	walk	taxi	walk
car	bus	bus	bus	bus
car	walk	bus	skateboard	taxi
taxi	motorcycle	bus	motorcycle	skateboard
walk	bicycle	walk	bus	car
walk	taxi	bicycle	walk	bus

1.1 40 people

1.2

Mode of Transport	Frequency
Walking	12
Bus	9
Taxi	5
Motorcycle	2
Car	5
Skateboard	2
Bicycle	5
TOTAL	**40**

1.3 There are two modes – motorcycle and skateboard. Possible reasons for this could be that a motorcycle is expensive to buy and run.

1.4 The most popular form of transport to and from school is walking, closely followed by taking the bus. However, the least popular is using a motorcycle or a skateboard.
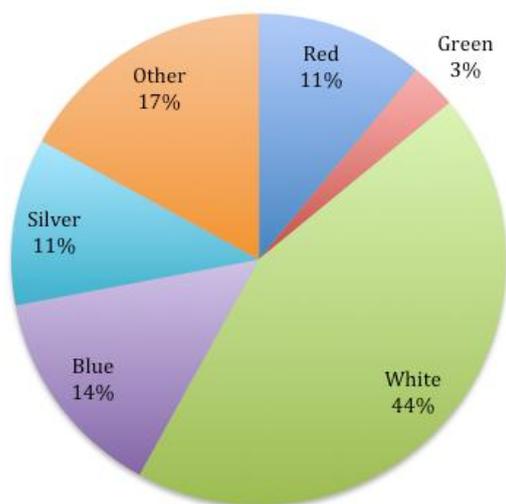
Activity 2

Colour	Percentage of vehicles	Number of vehicles	The angle of segment
Red	11	22	40
Green	3	6	11
White	44	88	158
Blue	14	28	50
Silver	11	22	40
Other	17	34	61

2.1 See table

2.2 See table

Number of Vehicles per Colour

2.3
2.4   All of the angles would add up to 360 degrees.

Activity 3

Two schools were invited to compete against each other in a Mathematics competition. School A was from the Western Cape and School B was from Gauteng. Each school has six Grade 11 classes. The table below shows the mean averages of each class, per school, after taking part in the competition.

School A	27	33	33	45	47	51
School B	25	29	37	42	42	51

3.1   School A:

Mean = $\frac{236}{6} = 39,33 \approx 39$

Median = $33 + 45 = 78 ; \frac{78}{2} = 39$

Mode = 33

School B:

Mean = $\frac{226}{6} = 37,66 \approx 38$

Median = $37 + 42 = 79 ; \frac{79}{2} = 39,5$

Mode = 42

3.2   School A = $51 - 27 = 24$
School B = $51 - 25 = 26$

3.3   No, it is not. The range only tells us how spread out or close together the data is. It does not tell us any useful information like the measures of central tendencies.

Acknowledgements

Mindset Learn Executive Head	Dylan Busa
Content Manager Classroom Resources	Jenny Lamont
Content Coordinator Classroom Resources	Helen Robertson
Content Administrator	Agness Munthali
Content Developer	Meryl Diab
Content Reviewer	Bernard Donkor

Produced for Mindset Learn by Traffic

Facilities Coordinator	Cezanne Scheepers
Production Manager	Belinda Renney
Director:	Alriette Gibbs
Editor:	Nonhlanhla Nxumalo
Presenter:	Talitha Ndima
Studio Crew:	Abram Tjale
	James Tselapedi
	Wilson Mthembu
Graphics:	Wayne Sanderson.