

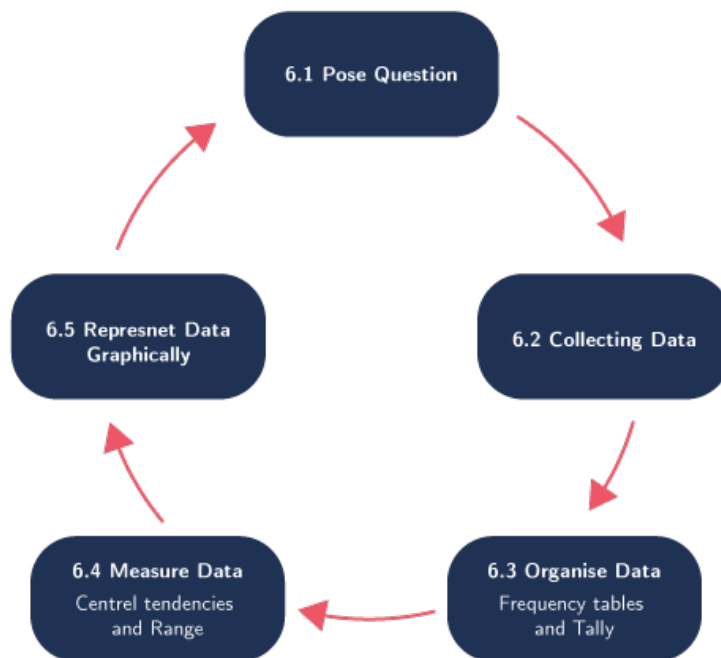
CHAPTER 6

Data Handling

CONTENTS

1	Pose questions	1
2	Collecting data	1
3	Classifying and Organising Data	2
3.1	Exercise 44: Frequency Table	2
4	Measuring Data	3
4.1	Exercise 45: Central tendencies and Range	3
5	Data	6
5.1	Exercise 46: Scatter Plots	7
5.2	Exercise 47: Histogram & Frequency Polygon	8
5.3	Exercise 48: Bar graph & Broken Line Graph	9
5.4	Exercise 49: Pie graph	10
5.5	Exercise 50: Pictograms	11
6	Interpreting and Analysing Data	13
6.1	Exercise 51: Misleading graphs	13
7	Answers for Exercises	15
7.1	Exercise 44	15
7.2	Exercise 45	15
7.3	Exercise 46	16
7.4	Exercise 47	16
7.5	Exercise 48	17
7.6	Exercise 49	18
7.7	Exercise 50	19
7.8	Exercise 51	19

April 20, 2021



1 POSE QUESTIONS

Discuss several questions on several topics.

2 COLLECTING DATA

"We need information about our changing environment, so that we can act effectively."

Nelson R. Mandela.

Discuss the difference between Categorical and Numerical Data.

Discuss the difference between Discrete and Continuous Data.

We can collect data from many different sources e.g. books, newspapers, journals, the internet, interviews with people and surveys. One can draw up a questionnaire.

3 CLASSIFYING AND ORGANISING DATA

3.1 Exercise 44: Frequency Table

1. Summarising Data

The weights (in kg) of each member of two teams (Under- 14 and Under-16) are recorded below:

Under-14 team	38	38	39	39	39	41	42	42	45	47	48	49	50
----------------------	----	----	----	----	----	----	----	----	----	----	----	----	----

Under-16 team	45	53	51	56	49	53	56	44	53	42	43	47
----------------------	----	----	----	----	----	----	----	----	----	----	----	----

1.1 Complete the Frequency Table for the Under-14 team:

Mass/ Weight in kg	Tally	Frequency
38 to 41		
42 to 45		
46 to 49		
50 to 53		
Total:		

1.2 Complete the Frequency Table for the Under-16 team:

Mass/ Weight in kg	Tally	Frequency
42 to 45		
46 to 49		
50 to 53		
54 to 57		
Total:		

1.3 How many under-16 players have a mass lower than 50 kg ?

1.4 Calculate what percentage of the under-16 team members has a mass of less than 50 kg ?

1.5 What is the probability of randomly selecting a player with a weight lower or equal to 53 kilogram, from the group of under-16 players?

4 MEASURING DATA

4.1 Exercise 45: Central tendencies and Range

Measures of central tendencies

Definition: Median

The **Median** is the value in the middle of an ordered (ranked) set of data values:

- If there is an **ODD** number of values in the data set, then the median is the middle number, and will be one of the given data values.
- If there is an **EVEN** number of values in the ranked data set, then the Median is calculated by taking the average of the middle two data values. Here the median will not be one of the data values.

Indicate which one is the median data item. Suppose this is arranged data.



Definition: Mode

The number that occur the most. With grouped data, the modal class is the class with the highest frequency.

Definition: Mean

Another word for average.

The mean is calculated as follows:

$$\text{Mean} = \frac{\text{The sum off all the data items}}{\text{The number of data items}}$$

Definition: Range

The highest value of the data minus the lowest. It calculates the spread of the data, in other words the distance between the highest and lowest data item.

1. Give the **MODE, MEDIAN, MEAN & RANGE** of the following data:

1.1 Data:

4 7 8 13 13 15 17 20 22 26 32

Median:

Mode:

Mean:

Range:

1.2 Data:

4 7 8 13 14 15 17 17 20 26 30 32

Median:

Mode:

Mean:

Range:

1.3 Data:

3 9 10 11 14 18 25 27 33 33 33 35 35

Median:

Mode:

Mean:

Range:

1.4 Data:

1 2 2 2 3 4 5 7 10 11 15 16 17 17 17 18

Median:

Mode:

Mean:

Range:

1.5 Data:

1 2 2 2 3 4 5 7 10 11 15 16 17 17 17 18 19 19

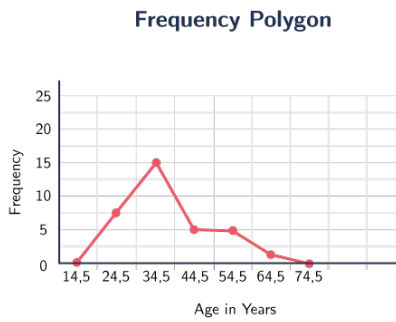
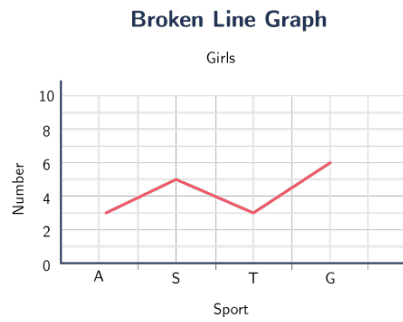
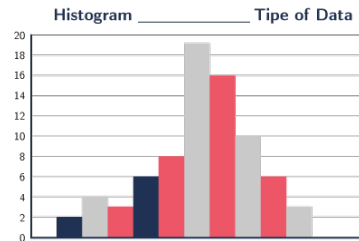
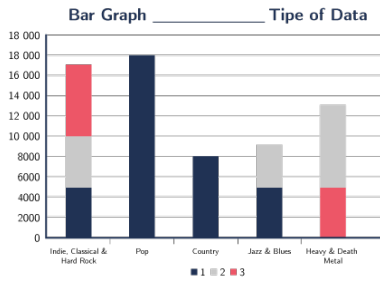
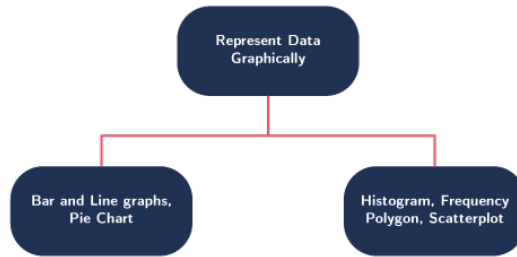
Median:

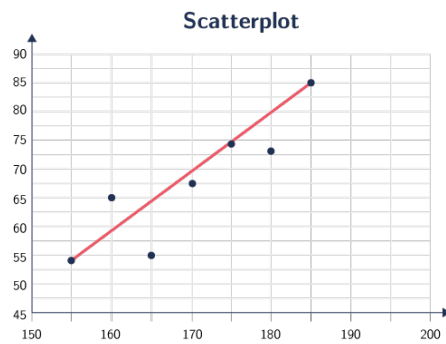
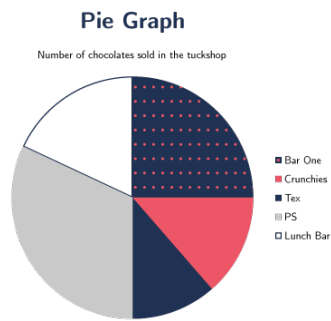
Mode:

Mean:

Range:

5 DATA





5.1 Exercise 46: Scatter Plots

Definition: Scatter plots

Scatter plots are graphs that show the relationship between two data sets (variables).

This line shows a negative correlation. Points lie in a perfect downward line.



No correlation. Points lie in a horizontal line.



This line shows a positive correlation. Points lie in a perfect rising line.



Outliers are values in a data set that are completely "out of line".

1. A doctor recorded his daily number of house visits for nine days.

Day	1	2	3	4	5	6	7	8	9
Number of visits	10	7	4	4	12	10	3	2	15

1.1 Plot these results on a graph. Draw a line of best fit if there is one.

1.2 Do you think there is a positive or a negative correlation or no correlation?

5.2 Exercise 47: Histogram & Frequency Polygon

1. A survey among the Gr 10 learners showed that they received the following amount of extra pocket money per month.

47 31 69 33 52 59 28 62 68 65
51 30 43 64 54 49 39 53

1.1 What is the range of the data?

1.2 Arrange the data

1.3 Determine the following:

- Arithmetic mean:
- The median:
- The mode:

1.4 Complete the following frequency table:

INTERVAL	TALLY	FREQUENCY
20 – 29		
30 – 39		
40 – 49		
50 – 59		
60 – 69		
TOTAL		

1.5 Draw a histogram of the data on the table.

1.6 Draw a frequency polygon of the data in the table.

The midpoint of a class interval is exactly halfway between the two values of a class interval:

Example

For the class interval (20 – 29) the midpoint is:

$$= \left[\frac{20+29}{2} = \frac{49}{2} \right]$$

$$= 24,5$$

Determine the midpoint of each interval.

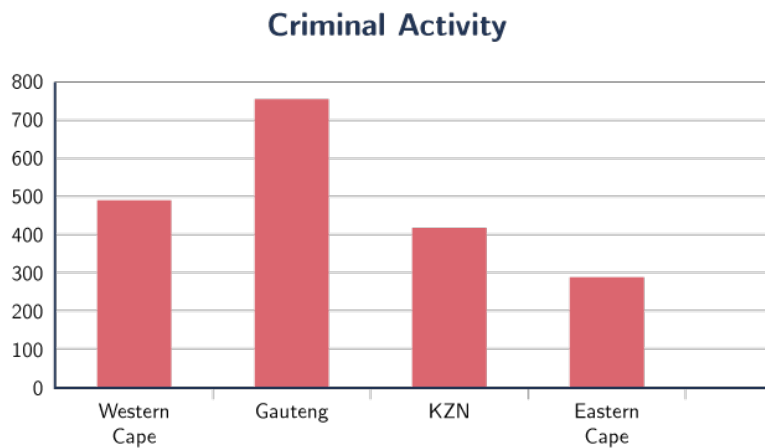
Complete the tally below:

Interval Midpoint	Tally	Frequency
	Total	

5.3 Exercise 48: Bar graph & Broken Line Graph

- The bar graph shows the number of criminal activities reported (in thousands) from April to March 2003 in the Western Cape, Gauteng, KwaZulu-Natal and Eastern Cape.

(www.saps.gov.za)



- 1.1 Relatively seen, which province has the highest crime rate?
- 1.2 Approximately how many crimes were reported in Gauteng?
- 1.3 Which provinces have almost the same reported crimes?
- 1.4 Name the other five provinces not shown in this analysis.

2. Consider the table given that shows the temperature of a male patient, recorded at 3 different times over 3 days and answer the following questions:

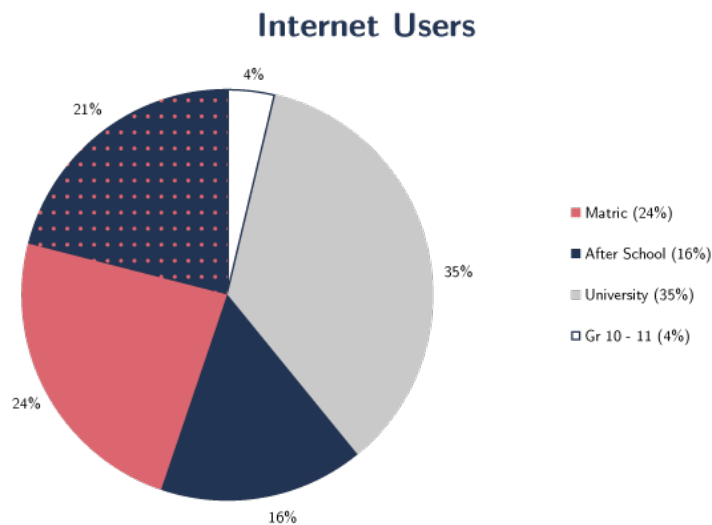
Day 1	a.m.	37°C
	p.m.	36°C
Day 2	a.m.	37,5°C
	p.m.	37,5°C
Day 3	a.m.	40°C
	p.m.	38°C

- 2.1 Draw TWO Broken Line Graphs to show the temperature of a male patient, recorded at different times over 3 days. Draw one graph for A.M and one graph for P.M. on the same set of axes.
- 2.2 What was his average temperature on the second day?
- 2.3 What is normal body temperature (in degrees Celsius)?

5.4 Exercise 49: Pie graph

1. Internet users and education levels.

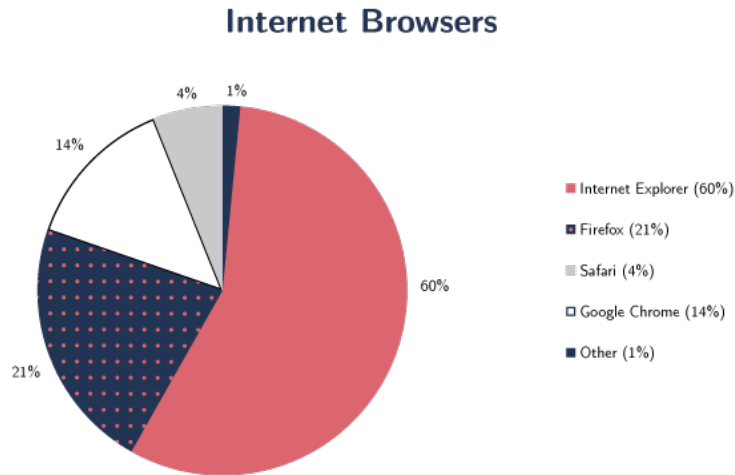
Study the pie chart and answer the questions below:



- 1.1 How large is the technician section?
- 1.2 1 500 people were interviewed as part of the survey. How many of them were still in matric?

2. Popular Web Browsers.

Study the pie chart and answer the questions below.



1.1 See if you can identify the different sectors.


1.2 3 000 people were interviewed as part of the survey. How many of them used "Google Chrome"?

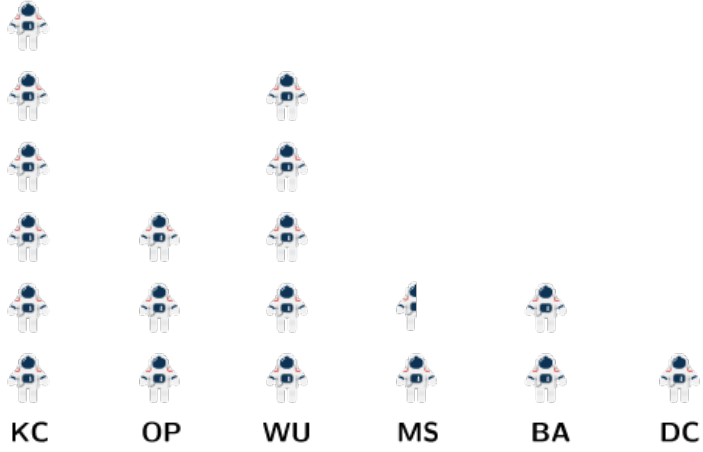
5.5 Exercise 50: Pictograms

Zani and Annette wanted to find out which is the most popular soccer team. They recruited one researcher in each class to get this information from the class. Then they summarised the data in a table and drew a pictogram.

Complete the table from the Pictogram:

Team	Votes
KC	300
OP	
WU	
MS	75
BA	
DC	50

 = 50 votes

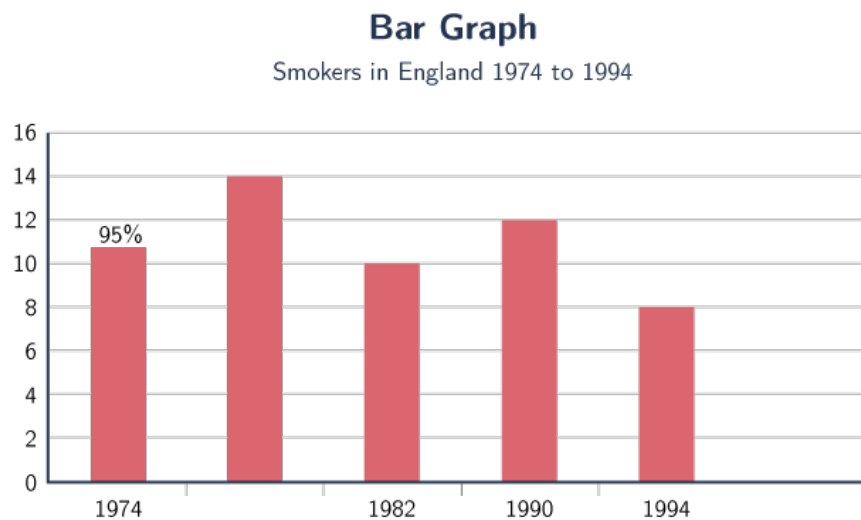


6 INTERPRETING AND ANALYSING DATA

6.1 Exercise 51: Misleading graphs

Statistical data can mislead people. These are techniques used to create misleading graphs:

- Bars are not the same size
- Gaps are not even
- Headings are not the same
- Bright colours are used to compare with dull colours
- Calibration is inconsistent
- Headings are not the same and in the same place
- The horizontal and vertical axes are swapped around
- The axes do not start at zero (Broken axes)
- The intervals on the axes are not evenly spaced
- Turn around chronological sequence (months, days, time years)
- Intervals on the axes are not evenly spaced



1. Study the graph and discuss the ways in which people can be misled.

1.1 Discuss more misleading ways.

1.2 Why, do you think, can the thickness of the bars confuse you?

1.3 Give reasons why there are misleading graphs.

7 ANSWERS FOR EXERCISES

7.1 Exercise 44

1.1

Mass/ Weight in kg	Frequency
38 to 41	5
42 to 45	4
46 to 49	3
50 to 53	1
Total:	13

1.2

Mass/ Weight in kg	Frequency
42 to 45	4
46 to 49	2
50 to 53	4
54 to 57	2
Total:	12

1.3 6

1.4 50%

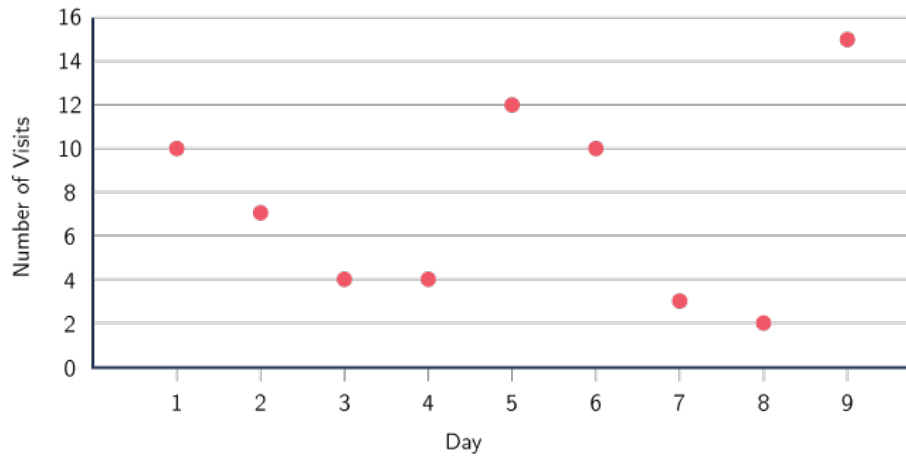
1.5 $\frac{5}{6}$

7.2 Exercise 45

- 1.1 Median=15 , Mode=13 , Mean= $\frac{177}{11}$, Range= 28
- 1.2 Median=16 , Mode=17 , Mean= $\frac{203}{12}$, Range= 28
- 1.3 Median=25 , Mode=33 , Mean= $\frac{256}{13}$, Range= 32
- 1.4 Median=8, 5 , Mode=17 , Mean= $\frac{147}{16}$, Range= 17
- 1.5 Median=10, 5 , Mode=17 , Mean=10, 28 , Range= 18

7.3 Exercise 46

1. 1.1



1.2 No correlation.

7.4 Exercise 47

1. 1.1 41

1.2 28; 30; 31; 33; 39; 43; 47; 49; 51; 52; 53; 54; 59; 62; 64; 65; 68; 69

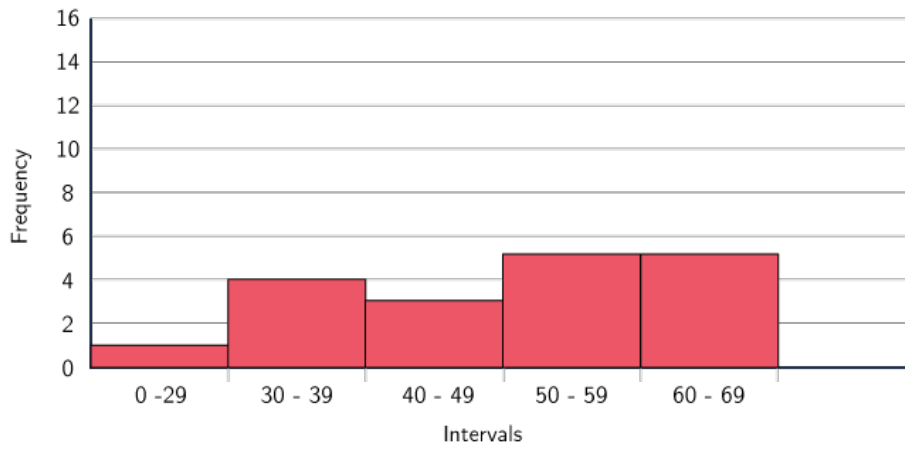
1.3 • $\frac{897}{18}$

• 51, 5

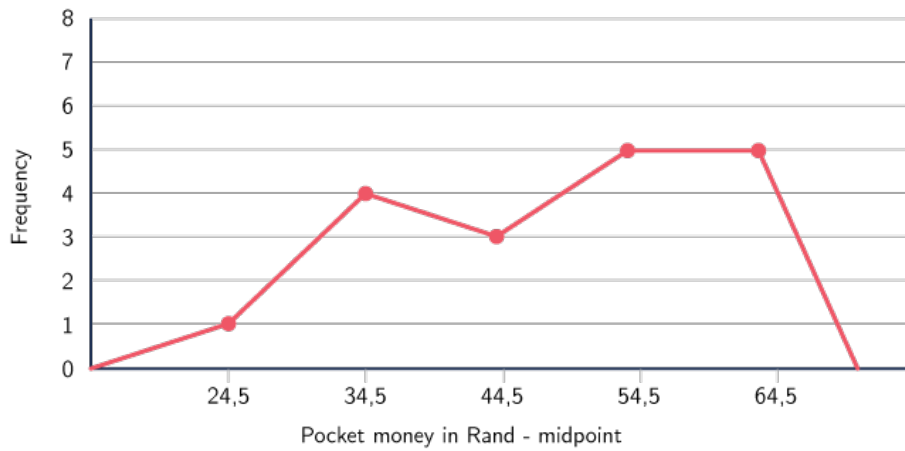
• There is no mode

INTERVAL	FREQUENCY
20 – 29	1
30 – 39	4
40 – 49	3
50 – 59	5
60 – 69	5
TOTAL	18

1.5



1.6



7.5 Exercise 48

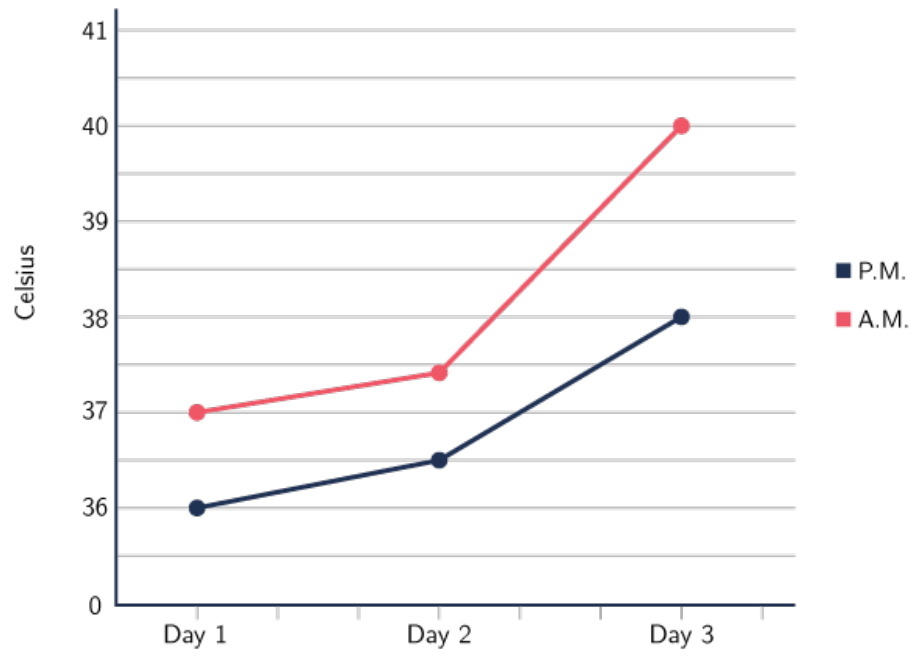
1.1 Gauteng

1.2 745

1.3 Western Cape KZN

1.4 Limpopo, Mpumalanga, Free State, Northern Cape, North West

2.1



2.2 37°C

2.3 37°C

7.6 Exercise 49

1.1 21%

1.2 360

2.1 60

2.2 420

7.7 Exercise 50

Team	Votes
KC	300
OP	50
1. WU	250
MS	75
BA	100
DC	50

7.8 Exercise 51

- 1.1 Above graph is mathematically inconsistent, making it impossible to reach a correct mathematical conclusion.
- 1.2 Since graphs are a visual representation, your eyes is drawn to the thicker graph and it puts more emphasis to the thicker graph and it does not necessarily create the correct mathematical conclusion.
- 1.3 Many situations require visual representation and the wrong emphasis is put to advance role players own situation. One can easily convey your own message through wrong emphasis.