

MATHEMATICAL LITERACY

GRADE 12

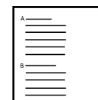
WINTER CLASSES

Topic: Data handling

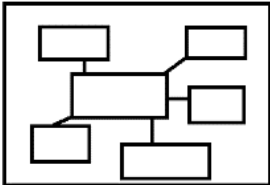



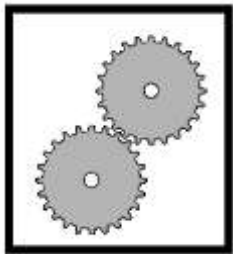


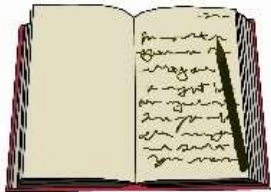
TEACHER AND LEARNER CONTENT MANUAL

MATHEMATICAL LITERACY PROGRAMME FOR WINTER CLASSES

STRUCTURE OF EXAMINATION			
PAPER	TOPICS	TOTAL MARKS	WEIGHTING
PAPER 1: (FINANCE & DATA HANDLING)	Finance	±90	± 60 %
	Data Handling	±53	± 35 %
	Probability	7	5%
TOTAL		± 4	± 32%
PAPER 2: (MEASUREMENT, MAPS, PLANS & SCALES)	Measurement	± 83	± 55%
	Maps, Plans and Scale	± 60	± 40%
	Probability	7	5%
TOTAL		150	±100%
Pre-test and Post-test to be administered since it's a revision of Term 1 & 2.			



GUIDELINES <ul style="list-style-type: none"> ○ Winter School Programme ○ Icon descriptions ○ Examination Guideline ○ Lesson Objectives ○ Important terms and definitions 	2 - 12
SECTION 1: Developing Data <ul style="list-style-type: none"> ○ Outlining the key concepts ○ Worked examples. ○ Activities 	13-16
SECTION 2: Collecting Data <ul style="list-style-type: none"> ○ Outlining the key concepts ○ Worked examples. ○ Activities 	16 - 18
SECTION 3: Classification & Organising Data <ul style="list-style-type: none"> ○ Outlining the key concepts ○ Worked examples. ○ Activities 	19 - 20
SECTION 4: Summarising Data <ul style="list-style-type: none"> ○ Outlining the key concepts ○ Worked examples. ○ Activities 	21 - 36
SECTION 5: Representing Data & Mixed Activities <ul style="list-style-type: none"> ○ Outlining the key concepts ○ Worked examples. ○ Activities 	37 – 43

ICON DESCRIPTION			
MIND MAP 	EXAMINATION GUIDELINE 	BIBLIOGRAPHY 	TERMINOLOGY 
WORKED EXAMPLES 	STEPS 	CONTENTS 	ACTIVITIES 

DATA HANDLING: EXAMINATION GUIDELINES



TOPIC: DATA HANDLING

Section	Level 1: Knowing	Level 2: Applying routine procedures	Level 3: Applying multi-step procedures in a variety of contexts	Level 4: Reasoning and reflecting
Developing Questions and Collecting Data	Read information directly from a given questionnaire/survey (e.g., the name of the organisation for which the questionnaire is being conducted). Complete a given questionnaire.	Conduct a given questionnaire/survey with a group of people.	Decide on appropriate questions to include on a questionnaire/survey, construct and then conduct the questionnaire/survey.	Critique the questions/layout of a questionnaire/survey
Classifying and Organising data	Sort data from smallest to biggest. Count the number of values in a dataset. State the difference between categorical data and numerical data; discrete and continuous data. Read information from frequency tables.	Sort data according to two categories (e.g., sort a set of data separately for females and males). Complete a given frequency table. Calculate percentage values to represent, the relative sizes of different categories of data.	When given a raw set of data, sort the data, decide on appropriate intervals (if necessary), and construct a frequency table to organise the data. If necessary, use the frequency table to draw an appropriate graph to represent, the data.	Make a deduction about whether collected information is biased or valid based on the structure of instrument. used to collect the data and the way in which the data was collected. Explain with justification whether data is discrete or continuous. Analyse data organised in tables and make deductions about trends in the data.

Measuring data/ Summarising data	<p>Identify the maximum and minimum values in a set of data.</p> <p>Identify the mode for arranged data.</p> <p>Identify the median for odd data that has already been arranged.</p>	<p>Calculate mean and range.</p> <p>Calculate the median for even data.</p> <p>Calculate the median if the data is not arranged.</p> <p>Calculate the quartile values for arranged data.</p> <p>Calculate the inter-quartile range when quartile values are given.</p>	<p>Calculate the mean, median and modal average for a set of data and decide with reasons which average provides the most accurate representation of the data.</p> <p>Use data presented on a graph to determine the mean, median, mode and range of a data set.</p> <p>Calculate the quartile values for data that is not arranged. Calculate the inter-quartile range when the quartile values are not given.</p>	<p>Analyse graphs and make deductions about trends in the data and predictions for the future.</p> <p>Identify and describe the use and misuse of statistics and make justified recommendations.</p>
Representing data	<p>Read values directly from the</p>	<p>Draw a specified graph date from a given table of data.</p> <p>Estimate values from given graphs.</p>	<p>Organise data using an appropriate table, decide on the most appropriate format for representing the data (that is, actual values or percentages), and decide on the most appropriate graph needed to represent the data.</p>	<p>Analyse graphs and make deductions about trends in the data and predictions for the future.</p> <p>Identify and describe the use and misuse of statistics and make justified recommendations.</p>

LESSON OBJECTIVES

Stage 1: Developing Questions

Learners must be able to:

- How to develop questions.
- Develop questions to collect multiple sets of data.

Stage 2: Collecting Data

Learners must be able to:

Develop and use instrument to collect multiple sets of data.

- Observation
- Interview
- Questionnaires:
 - Advantages and disadvantages of each type
 - Difference between sample and population
 - How to select an appropriate sample from a population
 - Impact of choice of sample on reliability of data

Stage 3: Classifying and Organising Data

Learners must be able to:

Classify collected data as a categorical or numerical.

- Numerical data classified as discrete or continuous.
- Sort collected numerical data in to two or more categories.
- Group collected data using intervals.
- Revise tally-tables and frequency tables

Stage 4: Summarising Data

Learners must be able to:

- Summarizing multiple sets of data using the mean, median, mode and range(central tendency and spread).
- Quartile and inter-quartile range values.
- Analyse calculated measures of central tendency.
- Percentiles(interpretation only).
- Understanding the function /purpose of values above.
- The role and impact of outliers.

Stage 5: Representing Data

Learners must be able to:

- Represent multiple sets of collected data using pie charts, histograms, single bar graph, line, and broken line graphs.
- Multiple bar graphs and compound/vertical stack graphs
- Scatter plot graphs.
- Box and whisker plot.
- Understand suitability of the types of graphs for different situations
- Understand the effect of the scale of the axes on the impression.
- Read/estimate information from graphs.

Stage 6: Analysing and Interpreting Data

NB: Analysing and Interpretation of data must be done whilst dealing with all other five (5) stages, i.e., it is integrated.

Learners must be able to:

- Identify/describe the trends (increase, decrease, stay the same and time)
- Ask questions about the way in which data was collected, organized, summarized, and presented to reveal possible sources of errors/bias. Indemnify & describe any misleading representations/summaries.
- Compare different representation of multiple sets of data and explain the differences.
- Investigate situations in which summarized/represented data is interpreted in different ways.
- Develop opposing arguments using the same summarized/represented data.

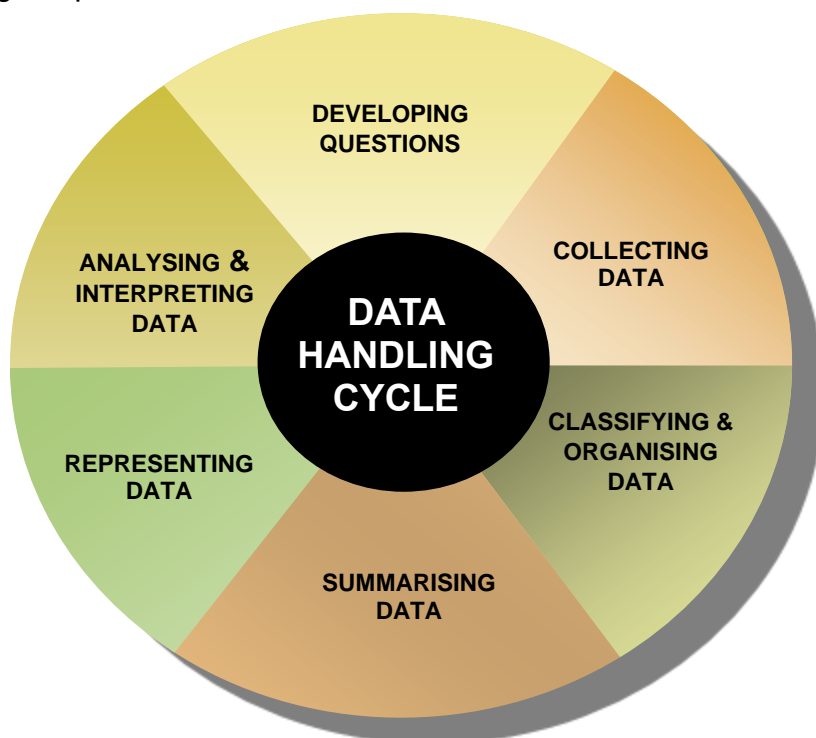
IMPORTANT TERMS AND DEFINITIONS	
DEVELOPING QUESTIONS AND COLLECTING DATA	
Data	Raw information that has been collected, without any organization of analysis.
Data Handling	Process of collecting, organizing, summarizing, representing, and analyzing information.
Qualitative data	Descriptive information (it describes something). It is referred to as categorical data
Quantitative data	Numerical data which involves numbers. e.g., measurements, length, height, area, volume, mass, etc. It can be further divided into Continuous and Discrete data.
Categorical data	The data that is given in the form of words, names, or labels. It is generally descriptive in nature, as data classified and organized into categories
Observation	Recording of data by watching someone or something closely.
Discrete data	Numerical (fixed numbers, like size of family, age); data that can have only certain values (quantities that can be counted)
Continuous data	Numerical data quantities that can be measured (e.g., measurements like weight, height, etc.)
Interview	List of questions that can be used to collect data. An instrument consisting of questions for the purpose of collecting data.
Survey	List of questions and obtain the data you need from the answers to those questions
Grouped data	The data given in the form of intervals.
Frequency	Number of times a data value is recorded.
Frequency table	Table showing frequencies in organized form, and summarizing the frequencies of all the data values in a data set

Tally	Mark/Tick that shows one count of a data value
Classify	Identify the type or class.
Inter-quartile range	The difference between quartile 3 and quartile 1 or the difference between largest quartile and the smallest quartile.
Population	The entire source of data involved in an investigation
Sample	The small group chosen from the population
Mean	Average of the values in a data set; sum of all the observed values divided by the number of observations.
Mode	Value or values appearing most often in a data set.
Median	Middle value in an ordered data set.
Range	Difference between the highest and lowest values in a data set.
Inter-Quartile Range (IQR)	The difference between largest quartile (Q3) and the smallest quartile (Q1).
Box-and-Whisker Plot	Diagram used to show the distribution of data along a number line divided into the quartiles.
Bias	To favour or oppose something or someone in a way which is considered unfair
Ascending Order	Arranging data from smallest to the biggest value.
Descending Order	Arranging data from biggest to the smallest value.

DATA HANDLING CYCLE

Data handling consist of six (6) inter-connected stages of statistical process shown below:

- The learners must know how to unpack each stage and remember that each stage depends on the other.



Important facts about six (6) stages of statistical process/data handling cycle

Stage	Statistical Process	Meaning
Stage 1	Developing a question or posing a problem	A question is posed regarding the problem and appropriate data sources are identified.
Stage 2	Collect data (information) related to the posed question	A data collection method and tool/technique are chosen and data is collected.
Stage 3	Classifying and Organising data	Process of organizing raw data, by classifying them into different categories.
Stage 4	Summarising Data	A large amount of data can be summarised using one or two numbers to represent all of it.
Stage 5	Representing Data	It is often very useful to represent the data visually. It can further quickly illustrate general behaviour or pattern of the data.
Stage 6	Analysing Data	Develop opposing arguments using the same summarised and/or represented data to conclude and bring about solutions to the posed question at stage 1.

- Each stage in the process is dependent on the stage that precedes it and directly impacts on the stage that follows.
- If the data that is collected is biased, then every subsequent stage will be flawed.
- If the data is summarised using an inappropriate or incorrect average, then the analysis of the data will be incorrect.
- In the diagram of the statistical process above, you will see that “analyse is linked to all stages. Data must be analysed throughout the statistical process, and it is not a separate step following one specific stage.

STAGE 1: DEVELOPING QUESTIONS

Statistical process starts with developing or posing a question. The phrasing of the question will determine the way in which the data is collected, organised, and analysed.

How to develop a questions

- Identify the main reason or aim for the research or investigation. Develop a main question with sub-questions which give more information or details.

➤ **Worked Example 1:**

Main question: *“What is the performance of Grade 12 performance in March test”?*

Sub-question:

1. *How many girls and boys passed the test?*
2. *What is the average and pass rate?*
3. *Which class that obtained more level 7?*



Activity 1 (Developing Questions)

(18 Marks; 20 Minutes)

1.1 Write down three (3) questions that will enable you to collect meaningful data, for each of the following topics below

1.1.1 Will the weekends extra classes improve learners' academic performance? (6)

1.1.2 What is the importance of exercising to the learners? (6)

1.2 Develop or pose a question that requires the collection of data on:

1.2.1 Test and Examination results (2)

1.2.2 School sport results (2)

1.2.3 Learners in your class or school (2)

(18)

Populations and Sample		
Concept	Definition	Worked Example
Population	Entire group or source of data involved in the research.	Identify the population and the sample from the question posed below: <u>Question:</u> <i>What percentage of Grade 12 learners in my school has cell phones?</i> <u>Answer:</u> <u>Population:</u> All the learners at the school <u>Sample:</u> Grade 12 learners
Sample	The small group chosen from the population to represent that population.	



Activity 1.3 (Population and Sample)

(12 Marks; 15 Minutes)

1.3 Read the statements below and identify the population and sample for each.

- 1.3.1 Which cell phone brand is popular amongst the Grade 12 learners of Nzimeni high school?

Population:

Sample:

(4)

- 1.3.2 Five rugby players of the Karoo team were born in Eastern Cape.

Population:

Sample:

(4)

- 1.3.3 Skinny jeans is one of the most preferred made of Just-Bongi's clothing store.

Population:

Sample:

(4)

(12)

STAGE 2: COLLECTING DATA

Data Collection: The process of gathering and analysing accurate data from various sources to find answers to research problems.

Below are methods of collecting data:

The advantages and disadvantages of the data collection methods:

Methods	Example	Instrument	Advantages	Disadvantages
Observation	Watch someone or something closely and record the data you need on a data collection sheet.	Recording sheet	Easy to record information. Participants do not fill forms.	Time consuming for the observer. Reliant on accuracy of the observer.
Interview	Talk to someone face to face or over the telephone	Recording sheet and Questionnaire. The interviewer can clarify questions and make follow-ups.	Data is obtained immediately. Interviewee can be asked to clarify responses.	Time consuming and expensive. Difficult to target large audience.
Survey	Hand or send out a list of questions and obtain the data you need from the answers to those questions.	Questionnaire	Can be completed in convenient time. Many people can complete it simultaneously.	People can be dishonest. Questions can be vague or ambiguous.
Database	Class register, performance academic results, employee's personal information, etc.	Electronic device. (e.g., Computer)	Data can be retrieved quickly. It is often accurate, consistent, and reliable.	It can be expensive. Needs computer skills.



Activity 2 (Collecting Data)

(20 Marks; 25 Minutes)

2.1 Which data collection method would be used to the following scenarios?

2.1.1 Whether there is a need for a pedestrian bridge in a local village. (2)

2.1.2 Collecting data whether a school should reward teachers for producing good results in Grade 12. (2)

2.1.3 Finding out about the number of cars passing near the mall in an hour. (2)

2.1.4 The water consumption by members of the household in a week. (2)

2.1.5 The 2021-2023 Grade 11 final examination performance. (2)

2.2 Choose the method or instrument from Column B that matches a description in Column A. Write only the letter next to the question number.

COLUMN A	COLUMN B
2.2.1 Printing the names of Grade 12 A	Interview
2.2.2 Looking at the behaviour of the birds.	Survey
2.2.3 A tool used in an interview.	Database
2.2.4 Radio presenter ask the President questions.	Questionnaire
2.2.5 Census employees counts people in the city.	Observation

(10)

(20)

STAGE 3: CLASSIFYING & ORGANISING DATA

Classifying Data

Once the data has been collected, it can be classified as qualitative or quantitative, where qualitative data is **categorical** and quantitative is **numerical**.

Type of Data	Description	Examples
Categorical Data	<ul style="list-style-type: none"> It is descriptive in nature. Consists of words representing categories. Usually observed, not measured. 	<ul style="list-style-type: none"> Colour of the shoes Types of food Gender Names of the provinces in a country, etc.
Numerical Data	<ul style="list-style-type: none"> It is numerical. Consists of quantities. Can be further divided in two groups (<i>continuous and discrete</i>) <u>Continuous</u>: data which can be measured. <u>Discrete</u>: data which can be counted. 	<p>Continuous (measured):</p> <ul style="list-style-type: none"> Speed of the car Height of the building Mass of a cow Volume of the water <p>Discrete (counted):</p> <ul style="list-style-type: none"> Number of balls Number of calculators Quantity of money Number of water bottles

Organising Data

A raw data can be organised using the following:

- Ascending or Descending Order**
- Frequency Tables** (*which includes tallies, frequencies, and cumulative frequencies*)
- Where there is a large amount of data, it can be grouped into class intervals, e.g., 0-5; 10-15; 15-20; etc. depending on the size of intervals.
- The class intervals must be the same width i.e., each interval must have the same number of data points to be comparable

Number of data points to be compared							
Concept	Description	Examples					
Ascending order	Arranging data from lowest to highest value.	Arrange the following data in ascending order:					
		0,08	1,2	0,05	3,2	3,1	1,15
		2	0,01	4	2,75	4,05	1
		Possible Answer: 0,01; 0,05; 0,08; 1; 1,15; 1,2; 2; 2,75; 3,1; 3,2; 4; 4,05					
Descending Order	Arranging data from the highest to lowest value.	Arrange the following data in descending order:					
		0,08	1,2	0,05	3,2	3,1	1,15
		2	0,01	4	2,75	4,05	1
		Possible Answer: 0,01; 0,05; 0,08; 1; 1,15; 1,2; 2; 2,75; 3,1; 3,2; 4; 4,05					

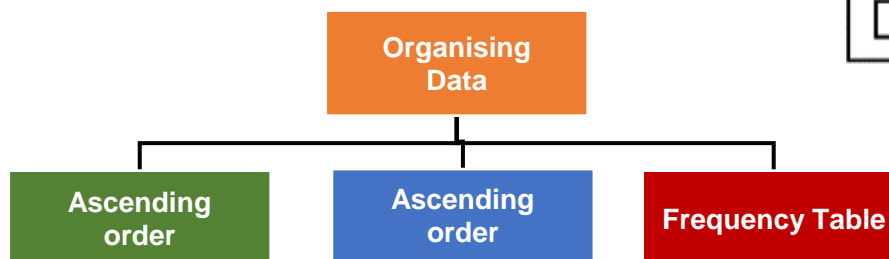
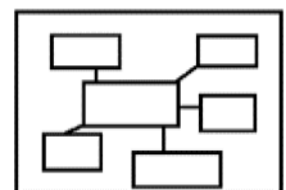
Frequency tables

Draw and complete a frequency table based on the shoe's sizes of Grade 12D class learners given below.

7	5	6	3	4	6
4	4	4	8	5	3
8	9	5	6	5	5
5	5	3	7	4	6

Possible Answer:

Size	Tally	Frequency	Cumulative Frequency
3		3	3
4		5	8
5		7	15
6		4	19
7		2	21
8		2	23
9		1	24
Total		24	





Activity 3 (Classifying data)

(20 Marks; 25 Minutes)

3.1 Classify the following types of data in either **categorical** or **numerical**:

3.1.1 Types of television brand. (2)

3.1.2 Names of the learners in Grade 12F class. (2)

3.1.3 The speed of Tom's motorbike. (2)

3.1.4 The capacity of fuel tank of the car. (2)

3.1.5 The mass of an aeroplane. (2)

3.2 Read the following data and state whether it is discrete or continuous.

3.2.1 The volume of Mr Monyane's Jojo tank is 2 000 litres. (2)

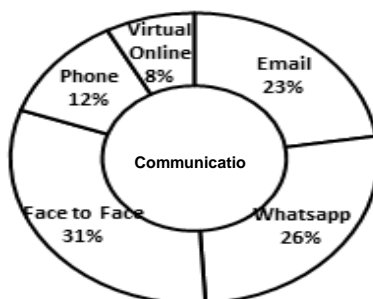
3.2.2 Sdwala call centre agents are given 5 000 calls per month. (2)

3.2.3 Below are the sizes of Frikkie Booysen's farms.

Farm 1	Farm 2	Farm 3
3 hectares	$2\frac{1}{2}$ hectares	4 hectares

(2)

3.2.4 The rate of communication method.



(2)

3.2.5 Only 6% of the South African population have university degrees, a new report by the Higher Education Department has revealed. (2)

(20)



Activity 3.2 (Organising data)

(20 Marks; 25 Minutes)

- 3.2.1 Below are the leave days taken by the cleaning employees of Tom's Clothing Textiles during the month of March 2023.

2; 4; 0; 3; 0; 0; 0; 5; 1; 2; 2; 3; 1; 0; 1; 2; 1; 0; 1; 2; 4

Use the above given data to answer the following:

- a) How many employees are working at Tom's Clothing Textiles? (2)
- b) Determine the number of employees who did not take a leave in March 2023. (2)
- c) Draw up a frequency table using the data given. (2)

- 3.2.2 The data below indicates the heights of 24 athletes of the Silithuli marathon team.

1,61	1,50	1,63	1,60	1,49	1,66	1,41	1,70	1,72	1,53	1,54	1,72
1,45	1,57	1,63	1,41	1,51	1,80	1,68	1,74	1,56	1,77	1,40	1,39

Complete the frequency table given below:

Height in metres (m)	Tally	Frequency
$1,30 < m \leq 1,35$		
$1,35 < m \leq 1,40$		
$1,40 < m \leq 1,45$		
$1,45 < m \leq 1,50$		
$1,50 < m \leq 1,55$		
$1,55 < m \leq 1,60$		
$1,60 < m \leq 1,65$		
$1,65 < m \leq 1,70$		
$1,70 < m \leq 1,75$		
$1,75 < m \leq 1,80$		
TOTAL		

STAGE 4: SUMMARISING DATA

Once the data has been organised, it can be summarised using measures of **central tendency** and measures of **spread**, to provide additional regarding the trends in data trends.

Measures of Central Tendency

There are three (3) types of measures of tendency namely:

- **Mean, Mode and Median**

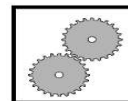
These three (3) measures of tendency all reflect some aspect of the data values which is **representative** of the whole data set, and further provides an indication of the “**middle**” or “**centre**” of the data.

Mean

- Commonly referred to as the “**average**”
- To calculate the **mean**, you add all the values of the data set and divide the sum by the total number of values in the data set.
- It can only be calculated if the data set is **numerical**.

$$\text{Mean} = \frac{\text{sum of all values in the data set}}{\text{total numbers of values in the data set}}$$

1. Worked Example (Mean)



The rugby team recorded all the scores of the matches they played as shown below.

Match 1	Match 2	Match 3	Match 4	Match 5	Match 6	Match 7	Match 8
23	34	18	45	28	39	12	9

Use the table of the scores given in the above table to answer the following questions.

- 1.1 How many matches did the rugby team play?

Possible Answer: They played 8 matches

- 1.2 Calculate the mean score.

$$\begin{aligned}\text{Mean} &= \frac{23 + 34 + 18 + 45 + 28 + 39 + 12 + 9}{8} \\ &= \frac{208}{8} = 26\end{aligned}$$

- 1.3 Paul, Nimmi, Nomvuyo, Sipho and Fifi write their university entrance competency test and compare their results. Paul achieved 89%;

Nimmi 98%; Nomvuyo 88%. Calculate Fifi's mark, if the mean value of all five (5) marks is 92%.

Possible Answer:

$$\text{Mean} = \frac{\text{sum of all values in the data set}}{\text{total numbers of values in the data set}}$$

$$92 = \frac{89 + 98 + 88 + 90 + \text{Fifi}}{5}$$

$$92 \times 5 = 365 + \text{Fifi}$$

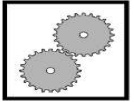
$$460 = 365 + \text{Fifi}$$

$$\text{Fifi} = 460 - 365 = 95$$

Fifi achieved 95% in the competency test.

Median

- This is the middle value of the data set when arranged in ascending order.
- If the total number of the data set is **odd** number, the middle value will be the median.
- If the total number of the data set is **even** number, there will be two values in the middle of the data set, and you will have to add these two values and divide them by two (2) to determine the median.
- It is important to know that the median can only be determined or calculated if the data set is **numerical**.



1. Worked Example (Median)

1.1

Below are the marks Babalwa got from five Mathematical Literacy tests she wrote. Each test had total marks of 50.

Test 1	Test 2	Test 3	Test 4	Test 5
29	31	22	40	20

Possible Answer:

Step 1: Arrange the data values in ascending order.

20; 22; 29; 32; 40

Step 2: Then find the middle value, which be the median

(3)

Therefore, the median is 29.

1.2

If Babalwa in 1.1 wrote test 6 and got 44 marks. Calculate the median of the new data set.



New marks: 20; 22; 29; 32; 40; 44

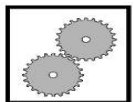
Possible Answer:

The median will now lie between 29 and 32. Therefore to calculate the median; add the two “middle” values and divide the sum by 2.

$$\text{Mean} = \frac{29 + 32}{2} = 30,5$$

Median

- It is the data value that occurs most frequently in a data set, i.e., the value with the highest frequency.
- In case where two data values occur most frequently, then the data set is referred as bimodal.
- Mode can be determined for both categorical and numerical data.



1. Worked Example (Mode)

1.1

Read the distances in kilometres which Unathi ran when he was exercising and determine the mode.

10 ; 15 ; 15 ; 7 ; 8

Possible Answer:

(2)

Mode = 15

1.2

If Unathi had an extra 10 km which he ran, determine the mode.

New data values:

10 ; 15 ; 15 ; 7 ; 8 ; 10

Possible Answer:

(2)

Mode = 10 and 15 (that is known as **bimodal** data set)

Outlier

- A value in a data set that is far away from most of the other values.

Example: 62 ; 56 ; 51 ; 58 ; **7** ; 65 ; 69

outlier
↖

Skewed Data

- It occurs when an outlier affects the measure of central tendency, thereby resulting in a middle value that is not necessarily representative of the middle of the data set.

When to use the Mean, Median and Mode (Measures of Central Tendency)

- Now that all three measures of central tendency describe the “middle” value of a set of data, you need to know the most suitable measure amongst the three measures of central tendency.

Limitations of Measures of Central Tendency

Mean:

- This averages the total of the data by the number of pieces of data and is the most used measurement, but it is strongly affected by outliers.

Median:

- This is the most accurate measure of the centre of the data, but it can be very difficult to calculate with a large dataset.

Mode:

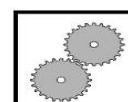
- This is often not a useful measure of the “average”. It is only useful when the data is categorical (e.g., shoe size or cell phone brand)

Analysis using Measures of Central Tendency

- Two sets of data can be compared by looking at their measures of central tendency (**mean, median and mode**), but they do not always give the complete picture or analysis. In Grade 12 a more and thorough analysis is required.

Advantages and Disadvantages of the Measures of Central Tendency

	Advantages	Disadvantages
Mean	<ul style="list-style-type: none">- Helpful when comparing two sets of data, e.g., <i>comparing test results of two different classes</i>.- Unique and it has only one answer.	<ul style="list-style-type: none">- It is affected by outliers.
Median	<ul style="list-style-type: none">- Useful when comparing sets of data.- Unique and it has only one answer.- Not affected by outliers as strongly as it is the case with mean.	<ul style="list-style-type: none">- The data needs to be arranged in ascending first.- The median is not always representative of the, especially when the data is not spread evenly about the median.
Mode	<ul style="list-style-type: none">- It is not affected by the outliers.	<ul style="list-style-type: none">- Not unique, answers maybe more than one.- In case of more than one mode, it is difficult to compare the data



Worked Example:

1. Class **A** and **B** below indicates the final examination results of two Grade 12 (matric) classes.

A	28	36	37	42	48	52	53	55	56	58	59	60	61	62	63	63	65	78	79	93	97
B	50	52	53	54	54	54	57	58	58	60	63	63	64	65	65	66	72	81			

- 1.1 Determine or analyse as to which class performed better in the examination using one of the measures of central tendency.

Possible Answer:

Measure of Central Tendency	Method	Class A	Class B
Mean	$\frac{\text{Sum of data}}{\text{No. of data}}$	$\frac{1245}{21}$ = 59,3%	$\frac{1089}{18}$ = 60,5%
Median	Middle value in an ordered data set	The value in the middle of the data: 59%	There are two values in the middle of the data, so we average them: $\frac{58 + 60}{2}$ = 59,0%
Mode	Most frequent value	63%	54%

Analysis	
Mean	According to the mean, Class B performed <i>slightly better</i> than Class A . However, it is not enough of a difference to say that they performed significantly better.
Median	Both Class A and Class B have the same median so this measure could not decide between them.
Mode	According to the mode, Class A performed better than Class B , but in type of data, mode is not useful measurement . There are too few of the modal data to declare it to be a good indicator.

Measures of Spread

There are four (4) measures of spread: range, quartiles, inter-quartile range and percentiles, namely:

- **Range**
- **Quartiles**
- **Inter-quartile Range**
- **Box and Whisker Plot**
- **Percentiles**
-

Range

- The difference between highest (maximum) value and lowest (minimum) value:

$$\text{Range} = \text{Highest value} - \text{Lowest value}$$

- If the range is small, the data is clustered together, and if the range is large, the data is more spread apart.

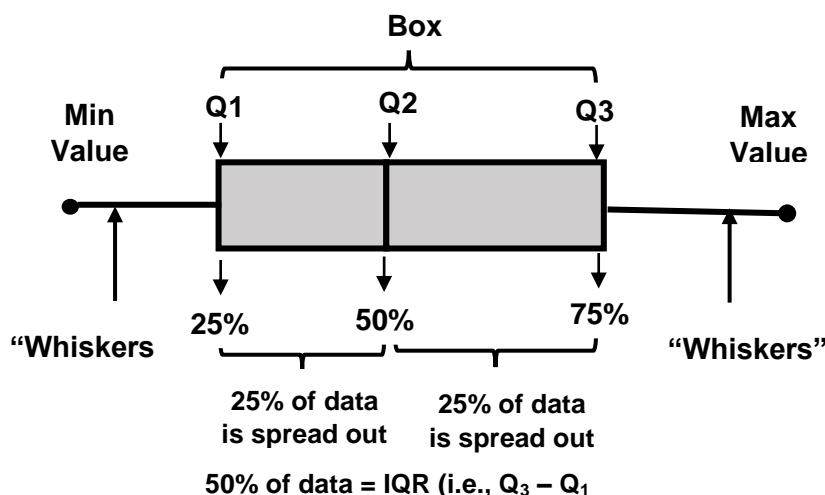
Interquartile Range (IQR)

- This is a measure of spread of the middle 50% of the data.

$$\text{IQR} = \text{Quartile 3 (Q}_3\text{)} - \text{Quartile 1 (Q}_1\text{)}$$

Box and Whisker Plot

- Also known as Box Plots
- The box-and-whisker plot is a visual representation of the “5-Number Summary” of a data set.

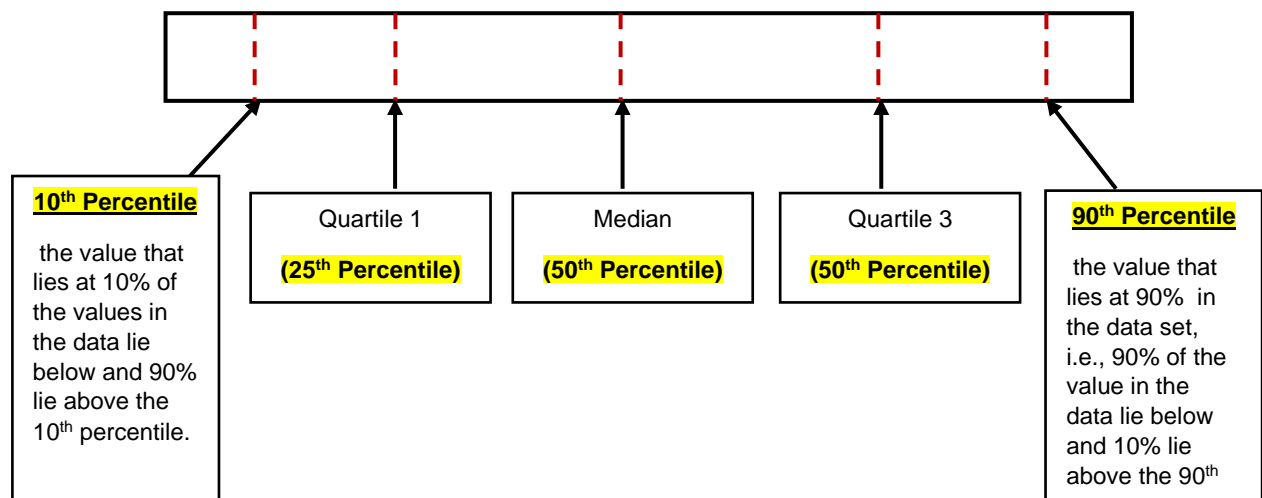


The 5-Number Summary	
No.	Value
1	Minimum Value
2	Lower Quartile (Q_1)
3	Median (Q_2)
4	Upper Quartile (Q_3)
5	Maximum Value

Percentiles

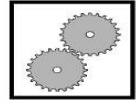
- The word “**percentile**” is derived from the word *percent* or *percentage*. Percentile divides the data into one hundred equal parts/segments, i.e., The 20th percentile is the value on or below which 20 percent of the data set lies.
- The **25th percentile** is **equivalent** to the **first quartile (Q₁)**; the **50th percentile** is **equivalent to the median or second quartile (Q₂)**; and the **75th percentile** is **equivalent to the third quartile (Q₃)**.

Important: *You are ONLY expected to INTERPRET percentiles. You will NOT be expected to calculate them from a data set.*



Advantages and Disadvantages of the Measures of Spread

	Advantages	Disadvantages
Range	Quick and easy to calculate	Affected by outliers
Interquartile range	Gives you an indication of the spread of 50% of the data values. Not affected by outliers.	The data must be first arranged in ascending order. Time consuming to calculate, as you first need to determine Q ₁ and Q ₃ .



Worked Example:

Use the performance of **Class A** and **Class B** to determine the following:

1. Range of both classes
2. Find Quartiles for both **Class A** and **Class B** and interpret the two data set.

A	28	36	37	42	48	52	53	55	56	58	59	60	61	62	63	63	65	78	79	93	97
B	50	52	53	54	54	54	57	58	58	60	63	63	64	65	65	66	72	81			

Possible Answer:

1. Range of **Class A** = Maximum Value – Minimum Value

$$= 97 - 28$$

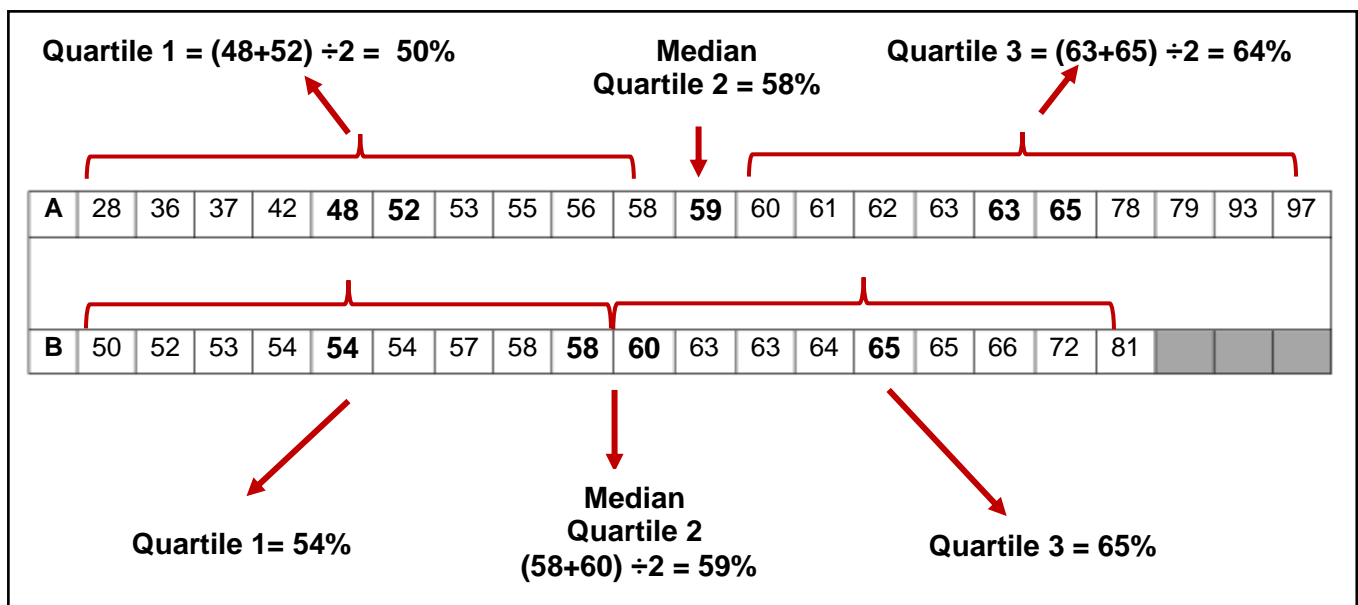
$$= 69$$

- Range of **Class A** = Maximum Value – Minimum Value

$$= 81 - 50$$

$$= 31$$

2. Quartiles of **Class A** and **Class B**

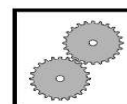
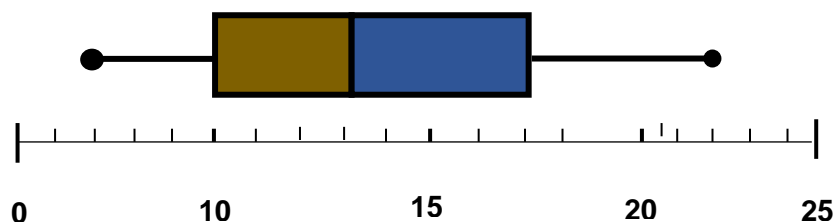


Analysis of the data:

	Class A	Class B	Analysis
Bottom 25% of the data	28% to 50%	50% to 54%	The bottom quarter of Class A scored at most 50% , while the bottom quarter of Class B scored at least 50% , but at most 54% .
Top 25% of the data	64% to 97%	65% to 81%	The top quarter of Class A scored at least 64% which is almost the same as Class B at 65% , but Class A results were as high as 97% while Class B's results peaked at 81% .

1. Worked Example (Range, Interquartile range, box and whisker plot and percentiles)

- 1.1 Study the box-and-whisker plot below and interpret all the key “5-Number Summary” values with respect to the spread of data.



Possible Answer:

- “Five (5) Number Summary” values:

1. Minimum value = 2
2. Q_1 (lower quartile) = 10
3. Q_2 (median) = 13
4. Q_3 (upper quartile) = 17
5. Maximum value = 22

- **Range** = $22 - 2 = 20$

- **Interquartile Range (IQR)** = $Q_3 - Q_1 = 17 - 10 = 7$ (IQR is the lower, it means the data values are clustered/close together)

- 1.2 A student earned a mark of 73% for an exam and it is in the 92nd percentile of the grade.

Possible Answer:

This means that 92% of the grade obtained a mark of 73% or less for the exam and only 8% of the grade achieved a mark that was higher than (or equal) to 73%.

Activity 1 Developing Questions, Collecting Data and Classifying & Collecting Data

1.1 For each of the main questions given below, formulate 3 sub-questions that will enable meaning data collection

1.1.1 How can the South African matric pass be improved? (3)

1.1.2 What is the importance of matric dance? (3)

1.1.3 How can we decrease the rate of school drop-out in South Africa (3)

1.2 Decide on the best method of collecting data for the following investigations.

1.2.1 Types of motorcycles for Kumba Biking Club (2)

1.2.2 The dissatisfaction of the community about poor service delivery (2)

1.2.3 Finding a suitable candidate to be an administration clerk (2)

1.2.4 The number of trucks passing in an hour on the National Road (2)

1.3 Sizwe attends the schools with the total of 1 700 learners and he is in Grade 12 class which has 37 learners. Identify the following from this given information.

1.3.1 Population (2)

1.3.2 Sample (2)

1.4 State which data is Categorical or Numerical

1.4.1 Quantity of boxes of books delivered to the schools. (2)

1.4.2 The country of birth. (2)

1.4.3 Yvonne has a choice of either installing wooden or glass door at her house. (2)

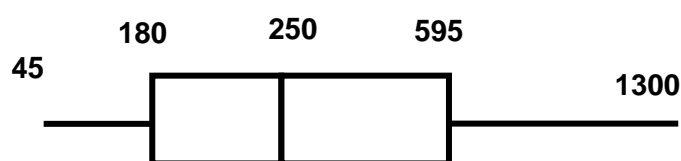
1.5 Determine whether the data represented below is discrete or continuous.

1.5.1 The rate at which water flows from the tap (2)

1.5.2 The participants of South African School Sports awards. (2)

1.5.3 The floor size of Ms Jasi's double storey house. (2)

1.5.4 Money (in rands) saved matric class for the market day at their school.



1.6 The incomplete frequency table below shows the Mathematical Literacy marks of a group of 67 learners.

INTERVAL As a %	TALLIES	FREQUENCY	CUMULATIVE FREQUENCY
90–100		0	0
80–89	III	3	3
70–79
60–69	HHH HH II	12	21
50–59	HHH II	7	28
40–49	HHH HH HHH	15	43
30–39	HHH HH HHH II	17	60
20–29	IIII	4	64
10–19	II	2	66
0–9	I	1	67

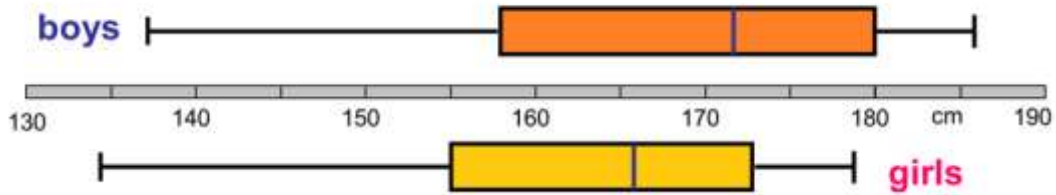
Use the frequency table above to answer the questions that follow.

1.6.1 Complete the tally for the 70–79% interval. (2)

1.6.2 Write down the frequency for the 70–79% interval. (2)

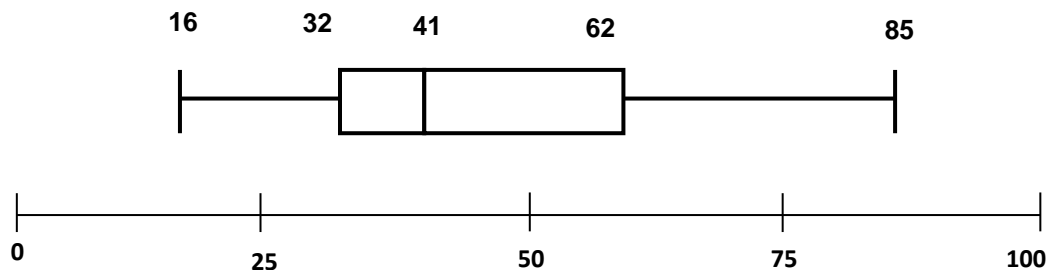
1.6.3 Show, by means of calculations, how the cumulative frequency for the 30–39% interval was determined. (2)

- 1.2 The box and whisker plots below represent the heights of boys and girls in Grade 12 class. Interpret the box and whisker to answer the following questions:



State whether the following statements are True or False

- 1.2.1 The tallest person is a boy. (2)
- 1.2.2 The shortest person is a girl. (2)
- 1.2.3 The girls show less spread in height. (2)
- 1.2.4 Half the boys are over 172 cm tall. (2)
- 1.2.5 The girls are taller on average. (2)
- 1.3 Below is a box-and-whisker diagram of the results (in %) of a Mathematical Literacy test written by 96 Grade 12 learners in a school. Use the diagram to answer the questions that follow:

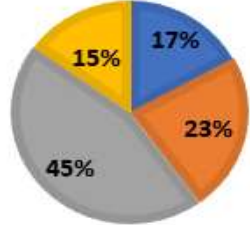
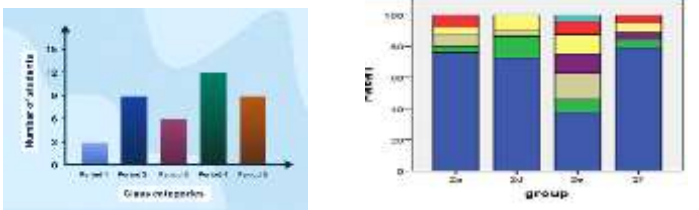
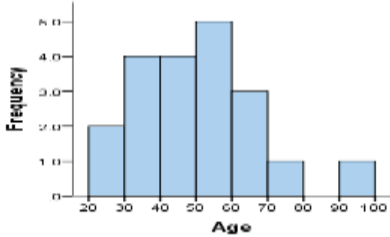
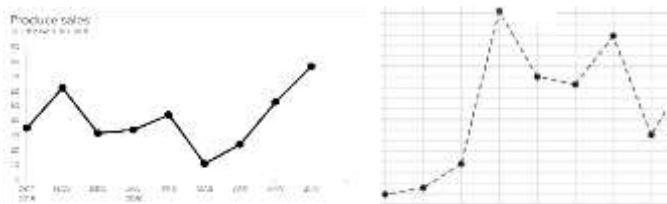
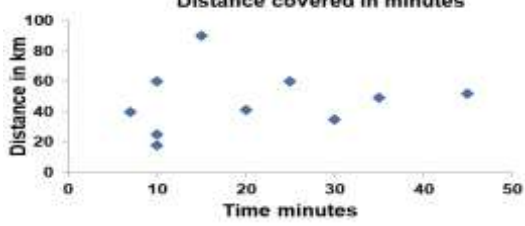
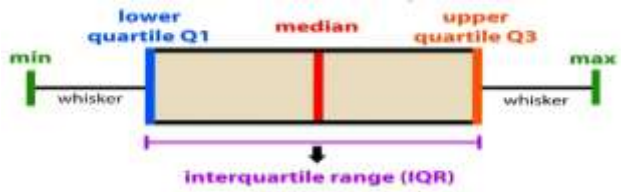


- 1.2.1 What is the highest mark obtained? (2)
- 1.2.2 What percentage of learners obtained more than 41%? (2)
- 1.2.3 How many learners obtained more than 62%? (2)
- 1.2.4 Determine the inter-quartile range. (2)
- 1.2.5 The lower quartile is 32%. Say in your own words what does this mean. (2)

STAGE 5: REPRESENTING DATA

Once the data has been summarised, it is important to represent it visually.

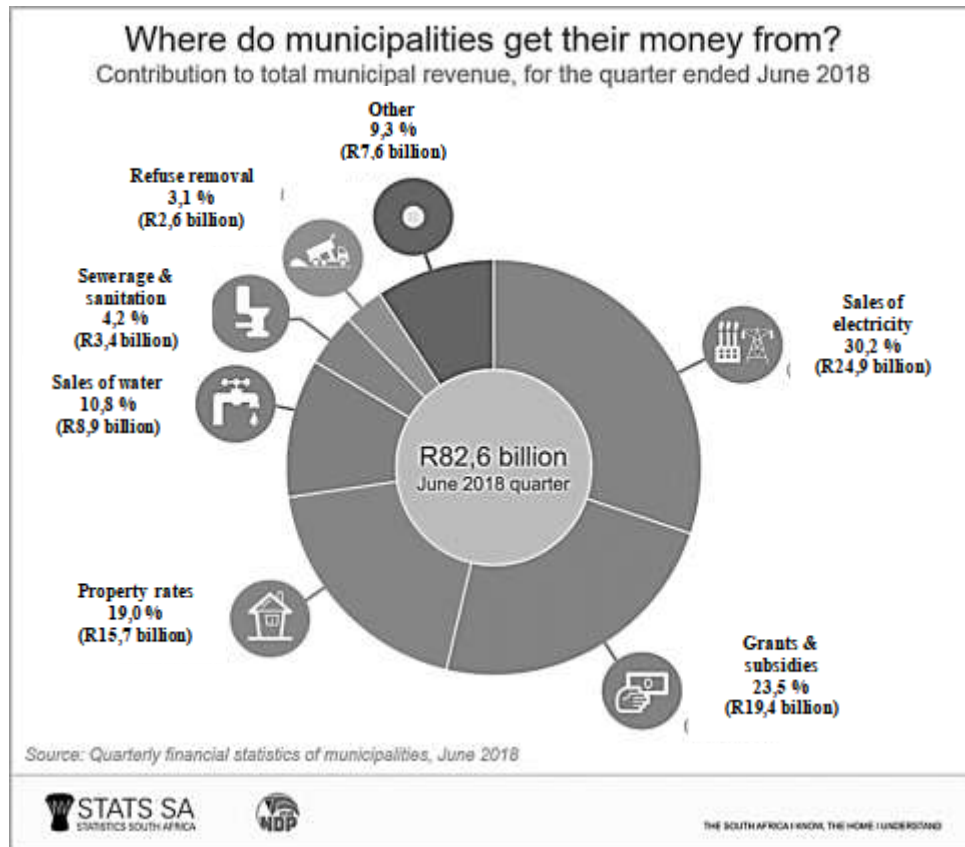
- The following representations of data can be drawn:

<p>Pie Charts</p> <ul style="list-style-type: none"> - Each sector/slice represents data value. - Commonly used for categorical data <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> <p>NB: You are not required to draw it, just read, and interpret values.</p> </div>	<p style="text-align: center;"> ■ Twitter ■ Facebook ■ Tiktok ■ Instagram </p> 
<p>Single bar graphs and compound bar graphs (stacked and multiple)</p> <ul style="list-style-type: none"> - Show frequency of each data - Used for discrete categorical data. - Spaces between the bars indicate discrete nature of the data 	
<p>Histograms</p> <ul style="list-style-type: none"> - Represents continuous data. - Data grouped into intervals. - Class intervals on the horizontal axis - Bars on the same width. 	
<p>Line and Broken line graphs</p> <ul style="list-style-type: none"> - Show the trend between plotted points of continuous data. - Broken line shows the trend between plotted points of discrete data. - Points are joined to show discrete nature of data. 	
<p>Scatter plots</p> <ul style="list-style-type: none"> - Increasing straight line pattern gives positive correlation. - Decreasing straight line pattern gives negative correlation. - Scattered random points without pattern gives no correlation. 	<p style="text-align: center;">Distance covered in minutes</p> 
<p>Box-and-whisker plots</p> <ul style="list-style-type: none"> - Only expected to interpret, not to draw. 	

Activity 1



- 1.1 The following graph was released by STATSSA indicating the main sources of income for municipalities in South Africa.



Source: [www.statssa.gov.za]

Study the diagram above and answer the questions that follow.

- 1.1.1 State whether the sources of income are qualitative or quantitative data. (2)
 - 1.1.2 Determine whether the data is discrete or continuous data. (2)
 - 1.1.3 What is the name of the graph used above. (2)
 - 1.1.4 What is the probability that one of the main sources of income can be less than 1,5 million? (2)
 - 1.1.5 Show how the amount from sales of electricity of R24,9 billion was calculated (3)
- (9)**

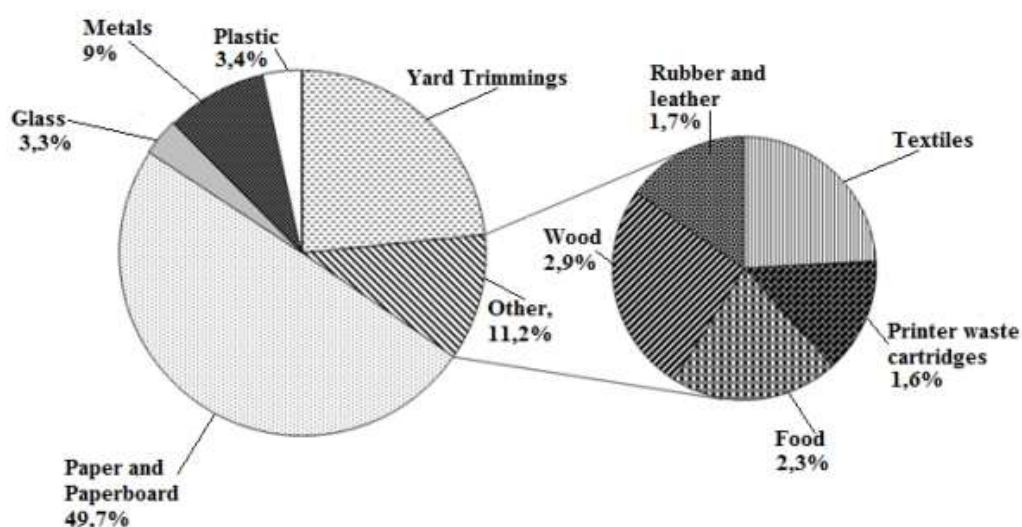


Activity 2

- 2.1 The pie chart below shows the total recycling and composting of material for 2015.

TOTAL RECYCLING AND COMPOSTING BY TYPES OF MATERIAL FOR 2015

(91,16 million tons)



Study the pie charts above and answer the questions that follow.

- 2.1.1 Write down ONE possible data collection method that was used to gather the above data. (2)
- 2.1.2 Calculate the percentage allocated for yard trimmings. (3)
- 2.1.3 Determine the percentage allocated for textiles. (2)
- 2.1.4 Calculate (in tons) the total amount of plastic recycled in 2015.
- 2.1.5 Give ONE possible example of a product that could be recycled under the metals category. (2)
- 2.1.6 State another type of graph that could be used to represent the data above. (2)
- 2.1.7 Determine, as a decimal, the probability of randomly selecting a material in the 'other' category that is not textiles. (4)

(15)



Activity 3

- 3.1 The pie chart below shows the total recycling and composting of material for 2015.

TOTAL RECYCLING AND COMPOSTING BY TYPES OF MATERIAL FOR 2015

(91,16 million tons)

PARTY	DELEGATE TYPE	PROVINCE									TOTAL
		EC	FS	GP	KZN	LP	MP	NW	NC	WC	
ANC	Permanent	4	4	3	4	4	4	4	4	2	33
	Special	3	3	2	3	4	4	3	3	2	27
DA	Permanent	1	1	2	1	1	1	1	1	4	13
	Special	1	1	2					1	2	7
EFF	Permanent		1	1		1	1	1	1		6
	Special							1			1
IFP	Permanent				1						1
NFP	Special				1						1
UDM	Permanent	1									1
TOTAL											90

[Source: www.wikipedia.org]

African National Congress	ANC	Inkatha Freedom Party	IFP
Democratic Alliance	DA	National Freedom Party	NFP
Economic Freedom Fighters	EFF	United Democratic Movement	UDM

Study the pie charts above and answer the questions that follow.

- 3.1.1 State the number of KZN delegates in the NCOP. (2)
- 3.1.2 Write down (in simplified form) the ratio of the total number of permanent seats to special seats for the ANC in the NCOP. (3)
- 3.1.3 Identify ONE party that has no permanent seat in the NCOP. (2)
- 3.1.4 An incomplete bar graph showing the different types of delegates representing each party in the NCOP, is drawn on the ANSWER SHEET BELOW

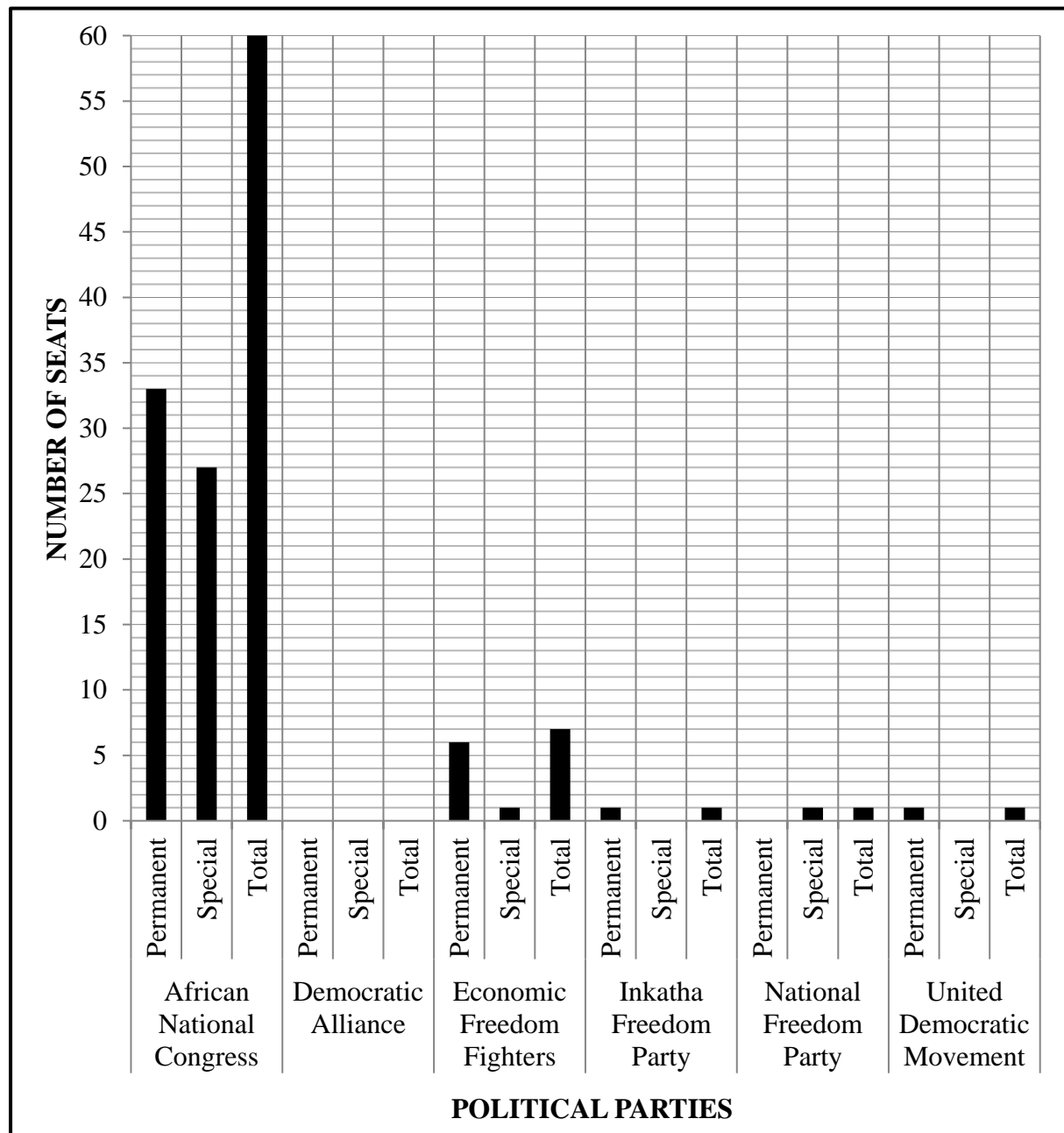
On the same ANSWER SHEET complete the bar graph for the Democratic Alliance (DA). (3)

(10)

ANSWER SHEET

QUESTION 3.1.4

**NUMBER OF SEATS FOR DIFFERENT PARTIES IN THE
NATIONAL COUNCIL OF PROVINCES (NCOP)**



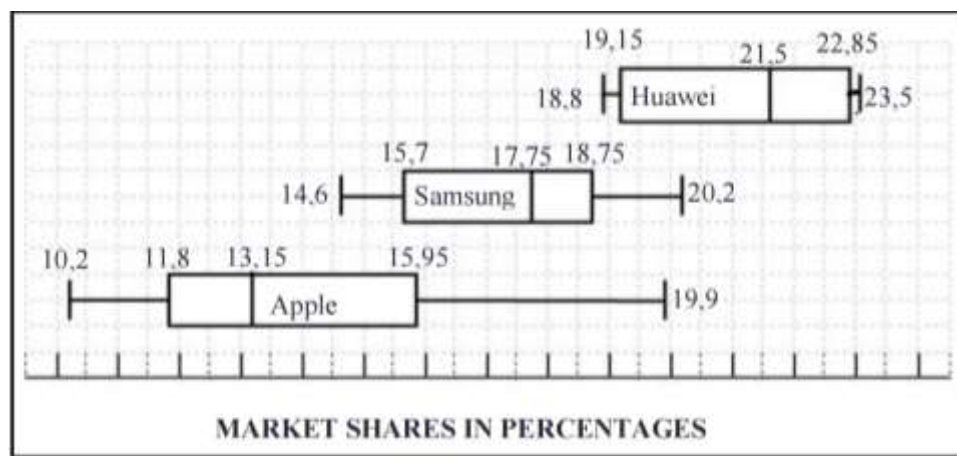
Activity 4



- 4.1 Marius wants to buy a new laptop. He studied the available data for the three most popular brands of laptops sold in South Africa.

The box and whisker plots below show (in percentages) the spread of the market share of the Samsung, Apple, and Huawei laptop brands from the fourth quarter in 2019 to the third quarter in 2020.

BOX-AND-WHISKER PLOTS SHOWING MARKET SHARE OF THREE LAPTOP BRANDS



Use the box and whisker plots above to answer the questions that follow.

- 4.1.1 Write down the name of the most popular laptop brand. (2)
- 4.1.2 Hence, write down the 50th percentile of the brand identified in QUESTION 5.2.1(a) (2)
- 4.1.3 Calculate the interquartile range (IQR) of the Samsung brand. (4)
- 4.1.4 A data analyst claims that 75% of the dataset of Apple was less than 16%. (2)

Explain whether his statement is valid.

(2)

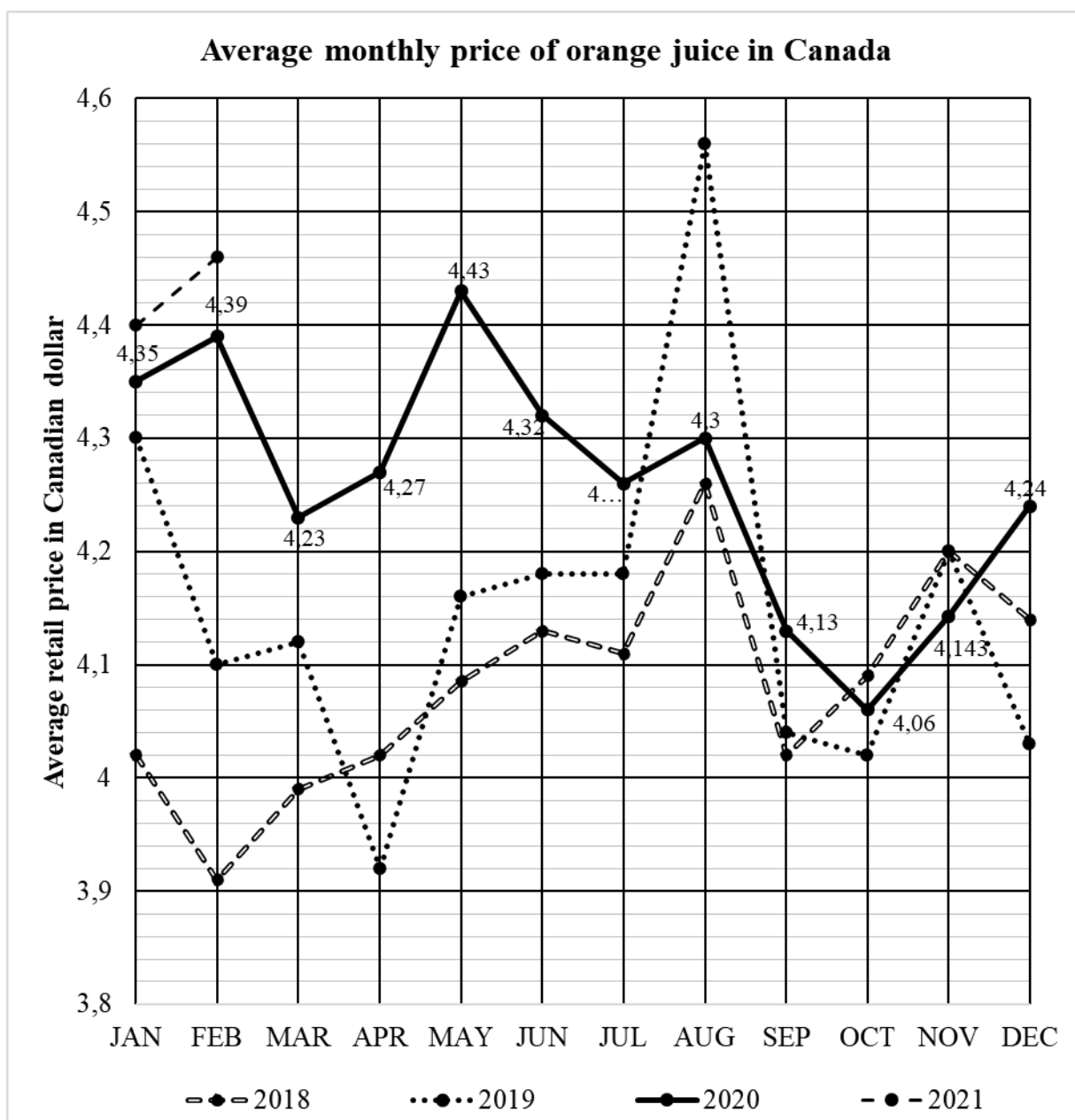


Activity 5

- 5.1 The analyst's prediction for the price of orange juice for the rest of the year 2021 is shown in TABLE 3 below.

Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
4,25	4,28	4,3	4,05	4,35	4,2	4,15	4,27	4,2





















- 5.1.1 Use the ANSWER SHEET to complete the line graph for March 2021 to December 2021. (4)
- December 2021.



Activity 6

The Republic of South Africa (RSA) conducts household censuses to collect information. The next census will take place in 2021.

Census information regarding household size is shown below.

HOUSEHOLD SIZE	CENSUS 1996	CENSUS 2001	CENSUS 2011
 One	 16%	 19%	 27%
 Two	 17%	 18%	 19%
 Three	 15%	 15%	 15%
 Four	 15%	 15%	 14%
 Five or more	 36%	 33%	 25%
Total number of households	8,7 million	10,8 million	14,5 million

NOTE:

[Source: statssa.gov.za]

- A census is an official count or survey.

According to Stats SA, a household consists of a single person or a group of people who live together for at least four nights a week, who eat together and who share resources.

Use the data above to answer the questions that follow.

- 6.1.1 Determine the percentage increase in the total number of households from 2001 to 2011. (2)
- 6.1.2 State which household size matches EACH of the following trends:
- (a) Increased every year, but only by a small percentage (2)
- (b) Remained constant in every census from 1996 to 2011 (2)
- 6.1.3 It was stated that the percentage of households with five or more persons decreased from 2001 to 2011, therefore the number of households with five or more persons decreased by 0,060 million.
- Verify, showing ALL calculations, whether this statement is CORRECT. (5)
- 6.1.4 Explain why the percentages for the 1996 census do not add up to 100% (2)
- 6.1.5 Write down the probability of randomly choosing a household from the 2011 census with a household size fewer than four persons. (3)

Bibliography

Books and Journals

1. R. Allen, *Mathematics* (2011).
2. P. F. Howard, *E. Howard, "Mathematics in the Curriculum"*, 2011, pp. 1-10.

Electronic Resources

1. J. Allen, *Mathematics*, <http://www.maths.org> (2011).

BIBLIOGRAPHY/REFERENCES

1. 2021 Mathematical Literacy Grade 12 Examination Guideline
2. Curriculum Assessment Policy Statement (CAPS)
3. DBE NSC Examination November 2019
4. DBE NSC Examination November 2021
5. DBE SC/NSC Examination May/June 2022
6. Grade 12 Answer Series 3-in-1
7. JENN Consolidated Winter Booklet 2017
8. FS Grade 12 Daily Activities