GRADE 10 MATHEMATICAL LITERACY

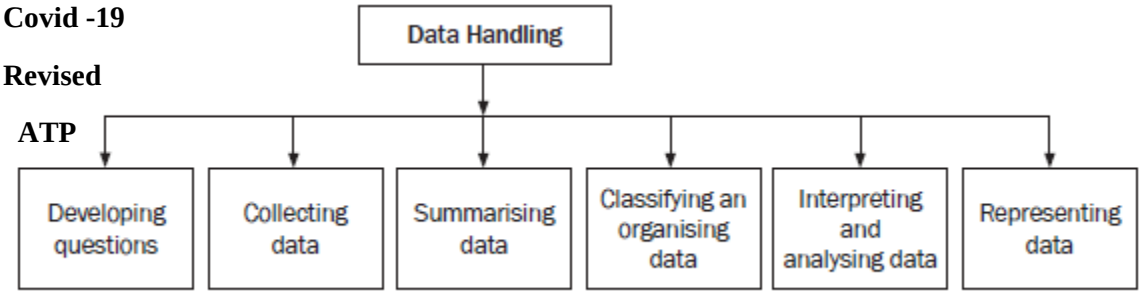
Topic: Data Handling 1 Compile by A Yankey for FHS gr. 10

NAME: _____ KEY: _____

Covid -19

Revised

ATP



DEFINITION:

Data is information that has been collected. Such information is used to answer questions in our daily lives.

TYPES OF DATA

Data can be **numerical or categorical**

1. Numerical Data- the items in this data are numbers (make up of numbers). E.g. the marks obtained in a test by learners.

Numerical data can be sub-divided into two:

- Discrete numerical data- Such data is obtained by **counting;**hence the values can only be whole numbers. E.g. The number of desks in each grade 8 class. (7 desks, 19 desks)
- Continuous numerical data – such data is obtained by measuring, hence include decimal or fractional values e.g. the numbers of hours athletes completed a marathon race. (2 hours, 1,5 hours; 3,15hours)

2. Categorical data-the items in this data are not numbers, hence can be put in categories. E.g. the colour of cars parked in a mall

Exercise 1

Indicate the type of data by ticking in the appropriate box.

		Categorical Data	Numerical data	
			Discrete	Continuous
e.g.	The colour of cars parked in a mall	√		
1	The number of cars sold over 2 years by a car dealer			
2	Mode of transport use by learners in a certain school			
3	The temperature of patients in a hospital			
4	Types of grant a government provides to its citizens			
5	The number of learners in each class that receive government grant			
6	Average height of players in a team			
7	The weight of new-born babes in a community			
8	The type of phones used by learners			
9	The number of learners in each grade that has Instagram account			
10	The number of homes in our community that has running waters			
11	Learner's race in each class			
12	The number of learners in the school per race			
13	Math literacy marks obtained by grade 11 learners .			

❖ DEVELOPING QUESTIONS

The first stage in any statistical process is to develop questions. These questions will provide a reason for collecting the data. The way the questions are phrased can impact on the data collected, and therefore on the findings of such investigation. The question will also inform the way in which the data is collected, organized, represented and measured.

The questions may be open-ended (the answer is usually the opinion of the respondent and the respondent can answer in their own words) or closed questions (options are given for respondents to choose from.

Example 1

Consider a teacher who wants to make sense of the results in his class. To do this he could pose the following questions relating to their marks:

- How did the class perform as a whole?
- Did the female students perform better than the male students?
- Did the class perform better than the other classes in the same grade?

These questions will inform the way in which the teacher collects information on the marks of his students, the way he organises those marks, and the best types of graphs and measures that he can use to make sense of the marks

Example 2

A house owner wants to reduce the amount of electricity that is used in the home and the amount of money that has to be paid for that electricity. The home owner may pose the following questions to help decide how to reduce the electricity consumption:

- How many units of electricity am I using, on average, per month?
- How much am I paying for electricity, on average, per month?
- Which appliances in my household use the most electricity?
- Which appliances can I use less often in order to reduce my electricity consumption and cost?

As with the teacher, these questions will inform the type of information that the homeowner collects with respect to electricity usage and the types of tools that they make use of in collecting, organising and making sense of the information.

Exercise 1

Pose a question that requires the collection of two sets of data (comparison) on:

1.1 Provincial sports results

_____

_____

1.2 Shoppers at a shopping Centre

_____

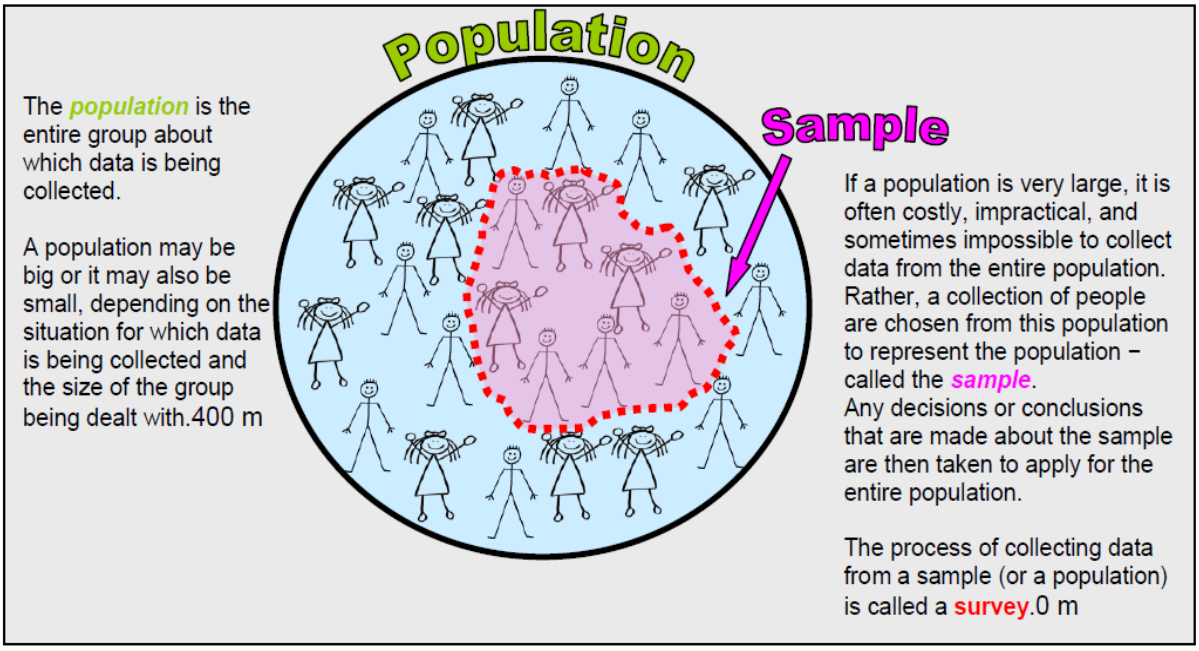
_____

1.3 Vehicles at a shopping Centre

_____

❖ COLLECTING DATA

Data can be obtained from differentsources, including people, animals, plants databases, etc. When collecting data from any of these sources, it is important to distinguished between the population and sample.
Look at the diagram below:



Example 1

If data is being collected on the most popular mode of transport used by learners of a school:

Population – All the learners of the school

Sample

A small number of learners comprising students from the different grades in the school and both male and female learners who can be interviewed about the mode of transport they use to travel to school.

Example 2

If data is being collected on the number of people in South African who have/have had malaria.

Population- the entire population of south Africa

Sample-collection of female and male South Africa citizens from each of the different provinces and different racial groups who could be interviewed about malaria

Exercise 1

1.1 Identify the **population** if the following questions are posed?

1.1.1 What percentage of learners in your community has cellphones?

_____

1.1.2 What percentage of learners in your province has cellphones?

_____

1.2 Identify a possible **sample** if the population is the following.

1.2.1 Learners in your community.

_____

1.2.2 Learners in your province.

_____

- **Considering bias**

When collecting data, it is important to ensure that the data collected is free of 'bias'(deviation from the truth). One of the ways in which this is achieved is by ensuring that a sample selected is **representative** of the population so that any decisions or conclusions made about the sample will accurately apply to the whole population:

- ✓ The sample must reflect the same features as the characteristics of the population.
- ✓ The sample must also be large enough.

  If the sample is not representative of the population, then that sample will provide a skewed or biased impression of the features or characteristics of the population

Example 3

Consider a scenario where data is being collected on alcohol abuse amongst high school learners at a school.

The population -is all the learners in the school

A representative sample of this population will include:

- ✓ Both female and male learners;
- ✓ Leaners from every grade in the school;

- ✓ Learners from all of the different racial background in the school in the correct proportion. If the school consist primarily only black learners, then more black learners must be included in the sample.
- ✓ Sample should be big enough. For example, for a school of about 980 learners, a sample of 150 or 200 learners will be big enough.

Exercise 2

Philemon, a grade 8 learner, is conducting a survey in his school on the number of learners that smoke. In collecting information, Philemon interviews 20 learners,all in grade 8.

2.1 What is the population of this data?

_____

2.2 What is the sample of this data?

_____

2.3 Why would you consider this sample to be biased?

_____

2.4 How could Philemon improve his sample to deal with bias?

_____

Exercise 3

Zipho and her friends are given a project to collect data and investigate the most popular make of car in their town. To do this, they walk to Pick n' pay and spend an hour counting the various makes of cars parked in the parking lot. The table below contains the data that they collected and the deduction that they reached .

Make of car	Number of cars
Ford	38
Mazda	59
BMW	25
Toyota	66
Mercedes	8
Peugeot	1
Hyundai	7
Nissan	52
Deduction: *From the data that we have collected we concluded that Toyota is the most popular car in our town.*	

3.1 What is the population of the data?

_____

3.2 What is the sample of the data?

_____

3.3 Do you think the deduction Zipho and her friends made is correct, or do you think is biased? Explain your answer.

_____

3.4 How could they improve their sample?

_____

_____

_____

_____

_____

Exercise 4

Xolani reads in the newspaper that 21,1% of high school leaners smoke cigarettes regularly. He is skeptical about this statistic and decides to conduct his own survey at his school. The table below contains the results of his survey.

Smoking Survey					
Number of people surveyed	200	Males	100	Black	50
		Females	100	Coloured	50
				Asian	50
				White	50
Total number of smokers	26		Percentage of smokers		13%
Total number of non-smokers	174		Percentage of non-smokers		87%
Conclusion: *Based on the information collected in my survey, I have concluded that the statistic given in the newspaper is unrealistically high and that only about 13% of high school learners smoke cigarettes regularly.*					

Is Xolani's conclusion valid, or did he collect a biased sample? Explain

_____

_____

_____

_____

_____

- **Data collection tools/instruments**

It is important to consider the type of data collection tool/instrument that will be most effective for collecting the data. There are three methods of collecting data:

	Some advantage(s)	Some disadvantage(s)
Observation Data is collected by watching someone or something closely and record the data you need on a collection sheet.	-You don't have to interact with people to get response. - You get firsthand information	-People may behave/act differently when they know they are being observed -Not all social occurrences can be observed - requires skills
Questionnaire Hand or send out list of questions	-Is a quick and easy way or getting results	-Respondents may not be truthful

and obtain the data you need from the answers to those questions.	-Can maintain anonymity ad privacy -can complete in own time	-Some questions may be left unanswered or incomplete -Possible low response -Expensive
Interview Talk to someone face to face or telephonically and record the data you need o a recording sheet.	-The interviewer/interviewee may ask a further question/clarity if the response/question is vague. - Response rate likely to be high	-Can be costly and time consuming -Respondents cannot maintain anonymity hence may not be open up to issues that are personal

Exercise 4

4.1 Indicate the type of method that would be appropriate for collecting each of the following, by placing a tick in the appropriate box.

4.2 Hence explain your choice of method in **4.1**

		DATA COLLECTION TOOL/METHOD		
	QUESTION	**INTERVIEW**	**OBERVATION**	**QUESTIONNAIRE**
1	The number of pens in each learner in the class has			
	Reason:			
2	The weight of all learners in your class			
3	The number of hours each learner in your class slept last night			
	Reason:			
4	Customer's opinion on a new design of a shop			
	Reason:			
5	The most popular make of car in your neighborhood			
	Reason:			

6	Number of childheaded families in your community			
	Reason:			
7	The opinion of girls in your school on dating			
	Reason:			
8	Unemployment rate for small community			
	Reason:			
9	Price history data for grocery items			
	Reason:			

Exercise 5

Develop two or three interview questions or questionnaire you can use to collect information about whether a learner experienced any form of abuse during the level 5 lockdown.

Note: Consider the sensivity of the issue when phrasing your question

1._____

_____

2._____

_____

3._____

_____

EXERCISE 6

Design a data collectio sheet to use if you have to observe the gender of adult shoppers leaving a shoppig centre and whether they are:

6.1 Carrying their purchases by hand or in a trolley.

6.2 Carrying their purchases in plastic or reusable bags or both.

❖ CLASSIFYING AND ORGANISING DATA

Organising data is taking information and **arranging/sorting** it into some kind of form that is manageable. This involves sorting, categories, tallies and tables. Categorical data can be sorted into categories such as males/females, fruit/vegetables, or in alphabetical order from A to Z or Z to A. Numerical data can be arranged I **ascending** order (from the smallest value to the largest) or in **descending** order (from the largest to the smallest value).

Classifying data: Once data has been sorted, it can be organized into frequency table (a table that shows a record of how often each value/set of values occurs in a set of data. Tallies can be used to help count the frequency of each value in a set of data. Sometimes it becomes impractical to construct a frequency table that includes every possible value. In such a case, we group the data by making using of class intervals to help organize the data more efficiently.

Tally table: include tallies and frequency. Frequency table: No tallies. Can be for grouped (using class interval) or ungrouped data

Exercise 1

Here are the ages(in years) of the teachers of Christ the King international school:

34, 58, 28, 36, 44, 29, 51, 36, 49, 54, 43, 59, 45, 37, 29, 48, 57, 29, 35, 43, 53, 47, 31

1.1 Sort the ages of the teachers in ascending order.

_____

1.2 Sort the ages of the teachers in descending order.

_____

1.3 Use the sorted values in 1.1 to complete the frequency table below.

Age	Frequency	Age	Frequency	Age	Frequency	Age	Frequency	Age	Frequency
28				42		49		56	
29								57	
30	0	37						58	
31								59	1
32				46		53			
33									
34									

1.4 Use the sorted values in 1.1 to group the ages of the teachers into the frequency table below.

Age	**Frequency**
20 -29	
60-69	0

Exercise2

An Audi sales person has ordered cars from their plant in Germany. The table below shows the number of cars they received.

Colour / **Model**	**Red**	**White**	**Silver**	**Black**
A3	2	5	4	3
A4	3	2	3	6
S3	4	3	5	5
Q7	1	4	4	3
R8	2	3	1	4

Use the information provided in the table to construct a tally table for Audi A3 cars.

SOLUTION

COLOUR	**FREQUENCY**	**TALLY**
Red		
White		
Silver		
Black		
TOTAL		

Exercise 3

The Geography examination marks, expressed as a percentage, of 52 learners were recorded as follows:

54	67	83	34	49	56	78	89	90	79	20	49	50
70	89	57	27	48	56	65	70	22	98	89	29	56
47	95	49	67	89	48	46	89	63	75	45	50	58
73	67	45	76	70	38	46	37	47	36	38	99	100

3.1 Use the above exam result to complete the frequency table below to show the performance of the learners according to performance level.

FREQUENCY TABLE: LEARNER PERFORMANCE IN GEOGRAPHY			
PERFORMANCE LEVEL	PERCENTAGE RANGE	FREQUENCY	TALLY
1	0 to 29		
2	30 to 39		
3	40 to 49		
	50 to 59		
5			
	70 to 79		
7			
TOTAL			

3.2 How many learners wrote the Geography examination?_____

3.3 How many learners scored less than 50%?

_____

3.4 How many learners scored at least 70%?

_____

3.5 What percentage of the learners scored more than 39%?

_____

❖ SUMMARISING DATA

After data has been collected, classified and organised it is not always possible to mention every piece of data in a report. Instead we summarise data by describing the whole data set using just a few numbers. Summarising data also makes it easier to analyse the data later. Data can be summaried by using measures of central tendency or measure of spread.

MEASURE OF CENTRAL TENDENCY AND MEASURE OF SPREAD

A measure of central tendency is a single value that attempts a central position of aset of data. There are three types of measure :mean, mode and median.

MODE

Mode or modal value is the data value that appears most in a data set (value with the highest frequecy). It is possible for the mode to be more than one value or no mode in the data set. For a group data , the class interval with the highest frequency is the modal class.

MEDIAN

The median is the number in the middle of the data set after the data has been arrange in ascending order. If there are two numbers in the middle (i.e.if the data contains an even number of data values) add the two middle numbers and divide by 2. Fifty % of the data lies above the median and another fifty % lies below the median, hence the median is also called the 50th percentile.

MEAN

Is the most common measure of central tendency that is used,but can easily be influence by very low or high numbers **(outlier**) in the data set. The mean is commonly referred to as the **average**. To calculate the mean, add all the values together and divide by the number of values in the data set.

MEASURE OF SPREAD

The spead of the data provides us with information about whether the values in a data set are grouped closely together or spread far apart. Commonly used measure of spread is range.

Range = Highest value – lowest value

Is affected by outlier. If there are outliers the range will appear to be unrealistically too small or big.

COMPARING MEASURES

To compare, calculate the mean, range, median and mode and then compare each measure to the other values in the data set to see which measure provides the best or most reflective indication of the majority of the values in the data set.Usually a higher mean, a higher median and low range will reflect the best set of data.

EXERCISE 1

Zandile employs 10 workers on her construction sites. Below are the weekly wages of the 10 Workers.

R250 ; R275 ; R200 ; R520 ; R260 ; R250 ; R250 ; R260 ; R200 ; R240

1.1 Calculate the mean, median and modal weekly wage of the workers.

_____

_____

_____

_____

1.2 How much more did the highest paid worker get?

_____

1.3 Which indicator(s) provides the most realistic average weekly income of the workers? Explain.

_____

_____

1.4 Which indicator(s) provides an unrealistic average weekly income of the workers? Explain your answer.

_____

EXERCISE 2

The table below contains the test scores of two grade 11 clases in recent mathematical literacy test. The test was out of 50 marks.

Class 1														
44	25	18	37	48	22	15	33	38	36	29	30	42	24	35
26	29	41	32	19	15	38	43	20	26	25	33	38	40	36

Class 2														
17	15	38	45	41	29	37	36	28	47	49	25	19	47	22
13	42	28	37	31	45	29	40	35	32	28	25	27	30	43

2.1 Determine the median and range for the two classes.

_____

_____

2.2 Determine the mean test scores of both classes as a percentage.

_____

2.3 Which class performed the best in the test? Why?

_____

EXERCISE 3

Province	Teachers
Eastern Cape	2 116 426
Free State	686 155
Gauteng	1 661 817
KwaZulu-Natal	2 783 051
Limpopo	1 816 852
Mpumalanga	914 739

North West	891 036
Northern Cape	202 010
Western Cape	956 836

3.1 which provincehas the highest number of teachers?

_____

3.2 What is the range of the number of teachers per proviince

_____

3.3 Calculate the median number of teachers per province for all the provinces.

_____

3.4 Calculate the mean number of teachers per province for all the provinces.

_____

3.6 Which indicator provides the best indication of the general distributions of teachers per province. Explain your answers.

_____

_____

EXERCISE 4

The table below shows the annual average school fees per province for government schools in south Africa

Province	**Average annual school fees (R)**
Eastern cape	220
Free State	**A**
Gauteng	830
Kwazulu-Natal	**B**
Limpopo	95
Mpumalanga	150
North West	120
Northern Cape	370
Western Csape	700

4.1 Free state has the lowest annual school fees. If the range of the school fees per province is R740, determine the value of **A**.

_____

_____

4.2 Hence determine the value of **B**(the annual fees for Kwazulu-Natal), if the mean of the school fees is **R307,78**

_____

_____

_____

EXERCISE 5

Below are the weights of 12 athletes (in kg):

65; 59; 62,7; 72; 60; 74; 88; 79; **p**; 69,8; 85; **p**

5.1 Determie the value of p, if the mean weight of the athletes is79,9

_____

_____

_____

5.2 Hence determine the median weight.

_____

EXERCISE 7

The South African rugby team (Springboks) played against New Zealand rugby team (All Blacks) during the World Cup. Study the sample of Springboks and All Blacks players' test statistics provided in ANNEXURE A and answer the questions that follow.

ANNEXURE A . . South African (Springbok) rugby players test statistics

PLAYERS		Height (cm)	Weight (kg)	Matches	Points
1.	Tendai Mtawarira	183	116	72	10
2.	Bismack du Plessis	190	115	76	55
3.	Eben Etzebeth	203	117	41	5
4.	Lood de Jager	205	125	16	20
5.	Schalk Burger	193	110	83	80
6.	Duane Vermeulen	193	108	32	10
7.	Handré Pollard	189	98	17	141
8.	Bryan Habana	179	94	114	320
9.	Jesse Kriel	186	96	8	15
10.	JP Pietersen	191	102	63	110
11.	Adriaan Strauss	184	112	51	30
12.	Trevor Nyakane	178	120	20	5
13.	Willem Alberts	192	120	35	35
14.	Ruan Pienaar	186	92	87	135

7.1 Calculate the average weight for the springbok players.

_____

_____

New Zealand (All Blacks) rugby players test statistics

PLAYERS		Height (cm)	Weight (kg)	Matches	Points
1.	Joe Moody	188	112	9	5
2.	Brodie Retallick	204	120	45	10
3.	Richie McCaw	187	105	146	135
4.	Daniel Carter	178	96	110	1 569
5.	Conrad Smith	186	95	92	130
6.	Keven Mealamu	181	109	130	60
7.	Ben Franks	183	117	45	10
8.	Sam Cane	189	105	29	50
9.	Tawera Kerr-Barlow	187	117	45	10
10.	Dane Coles	184	110	34	25
11.	Samuel Whitelock	202	116	71	20
12.	Kieran Read	193	110	82	100
13.	Julian Savea	192	106	39	190
14.	Ben Smith	186	94	46	95

[Adapted from the *Daily Dispatch*, 24 October 2015]

7.2 Identify the value of an outlier from the All Blacks players's points.

_____

7.3 7.3 Determine the modal heights for both teams.

_____

_____

7.4 Complete the frequency table for the number of matches played by both All Blacks and SpringBok players.

Interval	Tally	Frequency
0 – 30		
31 – 60		
61 – 90		
91 – 120		
121 – 150		

7.5 Explain the term probability.

_____

_____

7.6 Determine the probability of randomly selecting a player from both teams with a name starting with J. Write your answer as a percentage.

_____

_____

_____

EXERCISE 8

Patty High School has 22 Grade 11B learners. Learners were asked to measure their heights and weigh themselves using the bathroom scale provided by the school for Mathematical literacy Assignment. They recorded the data as indicated in the table below.

TABLE A: Boys

Mass (kg)	56	72	65	73	54	56	51	55	63	56	58	71
Height (m)	1,5	1,5	1,2	1,4	1,5	1,5	1,6	1,6	1,4	1,5	1,54	1,6

TABLE B: Girls

Mass (kg)	62	63	64	63	52	59	46	66	58	52
Height (m)	1,43	1,55	1,56	1,51	1,62	1,54	1,48	1,6	1,63	1,6

8.1 Arrage the data in TABLE A from the lowest mass value to the highest value.

_____

_____

8.2 Calculate the mean of the mass values in TABLE A.

_____

_____

8.3 Determine the median mass of the girls.

_____

_____

8.4 Determine the modal weight of thw whole class.

_____

_____

8.5 Express the probability of getting a boy with a mass of 56 kg as a percentage.

_____