

## Data Handling

### Part 1: Measuring Data

#### Key Concepts

In this session, we will focus on summarising what you need to know about:

- Measures of central tendency
- Measures of spread

#### Terminology & definitions

- **Data** is a computer term meaning information.
- When we **collect information**, particularly numeric information about a topic, we say that we are **collecting data**.

#### X-ample 1

Collecting the marks for a cycle test in a class.

#### X-ample 2

Collecting the heights of all the people in a school.

Once we have collected data, we need to process the data so that we can interpret the information in a meaningful way.

#### Concept: Measures of Central Tendency

Mean: This value is the average of all the values collected.

Median: This is the middle value of all the values collected. It tells us that 50% of the sample is less than or equal to this value and 50% of the sample is greater than or equal to this value.

Mode: This is the value that appears the most number of times of all the values collected.

#### X-ample 1

Sipho and his friends went fishing and caught 12 fish altogether. What was the average length of the fish caught? Find the median and mode for this sample.

12cm, 15cm, 10cm, 22cm, 6cm, 13cm, 5cm, 8cm, 8cm, 17cm, 15cm, 11cm

Mean =

Median =

Mode =

There are other values that we are interested in when sampling.

## Concept: Measures of Spread

- Range: The range is the spread of the values we have collected. It gives the difference between the highest and lowest value in our sample.
- 1<sup>st</sup> quartile: This is the value at which 25% of the sample is below or equal to this value.
- 3<sup>rd</sup> quartile: This is the value at which 75% of the sample is below or equal to this value.
- Interquartile range: The spread of values from the 1<sup>st</sup> quartile to the 3<sup>rd</sup> quartile. It is the difference between the 1<sup>st</sup> and 3<sup>rd</sup> quartile values.

### X-ample 2

In the Grade 12 class at Siyafunda Secondary School a sample was taken of the shoe sizes of various pupils. Here are the sample sizes.

12, 5, 6, 5, 14, 11, 11, 10, 7, 7, 10, 9, 9, 10, 3, 6, 8, 10

Calculate the Mean, Median, Mode, Range, 1<sup>st</sup> and 3<sup>rd</sup> Quartiles, as well as the inter-quartile range for this sample.

Order the sample:

3, 5, 5, 6, 6, 7, 7, 8, 9, 9, 10, 10, 10, 10, 11, 11, 12, 14

Mean =

Median =

Mode =

Range =

1<sup>st</sup> Quartile =

3<sup>rd</sup> Quartile =

Inter-quartile range =

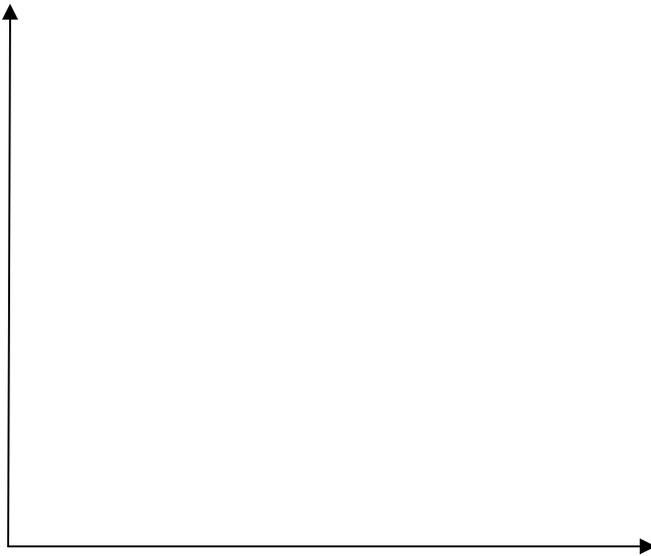
Standard Deviation: Indicates the degree that values are spread out around the mean. A wider spread indicates a larger standard deviation.

Variance: Also measures spread of data and is found by squaring the standard deviation.

A larger standard deviation means that the values in a sample are more spread out. A smaller standard deviation means that the values in the sample are more closely spaced.

Outliers: Values that are far removed from the bulk of the values in the sample.

Standard Deviation: A measure of spread that gives an 'average' of how far all the values in the sample are removed from the mean.



Measuring one, two and three standard deviations.

### X-ample 3

A group of athletes is doing high jump and their results are recorded. The mean of the results comes out at 1,61m. The standard deviation for the sample was 0,12m.

1. Calculate the range of values for one standard deviation.
2. What % of the sample would fall within the range of one standard deviation?
3. What would the range of the sample be for 95% of the sample?
4. What would the range of the sample be for 99% of the sample?

### X-ample 4

We collect the following heights of the children in the Grade 11 Maths Literacy class.

165, 168, 155, 190, 187, 165, 195, 170, 168, 184,  
181, 177, 198, 170, 175, 172, 163, 170, 171

Step 1: Organise the data in sequence from lowest to highest.

155, 163, 165, 165, 168, 168, 170, 170, 170, 171,  
172, 175, 177, 181, 184, 187, 190, 195, 198

Step 2: Check that you have all the data samples

Mean = Sum of Samples  
Total No. Of Samples  
Sum = 3324  
Total No. of Samples =  
Mean =



- c. What measurement did 75% of the sample fall under?
  - d. What is the modal number of the fish caught?
  - e. What is the inter-quartile range?
  - f. Describe in words what the inter-quartile range indicates about the sample.
3. After survey, a shop finds that the mean size for men's feet is a size 9 with a standard deviation of 1,5 ( $\sigma = 1,5$ ). They stock shoes from size 4 to size 14.
- a. What is the range of the shoe sizes available?
  - b. What is the range of sizes that fall within two standard deviations?
  - c. Explain why the shop never seems to sell the smallest and biggest sizes of the shoes as quickly as the others?
  - d. What would the modal number of the shoe sizes be that the shop sells?

### X-ercise Answers

- 1.
- a. 33,3 loaves
  - b. 30
  - c. 27
  - d. 44.
  - e. It can help them estimate their future sales
  - f. It can help them estimate the minimum and maximum number of loaves to bake in a day.
- 2.
- a. 8,6 cm
  - b. 5 cm
  - c. 10,5 cm
  - d. 5 cm
  - e. 5,5cm
  - f. 50% of the sample falls within the range 5-10,5cm.
- 3.
- a. 10
  - b. 6 – from size 6 to 12
  - c. Two standard deviations covers 95% of the people surveyed. i.e. 95% of the people have shoe sizes in the range Size 6 – 12.
  - d. Size 9.