

Support Sheet for the 'SUGAR GUZZLERS' activity.

Data: This support sheet will use the follow data sets for examples given

Dataset A = {4, 5, 1, 7, 12, 8, 2, 5, 6} 9 values, an odd number of data values

Dataset B = {1, 25, 3, 3, 6, 2, 2, 7, 5, 4} 10 values, an even number of data values

The Mean

To find this average, count up all the values and get a total. Divide this total by number of values.

Example 1. The mean of dataset 'A' is $\frac{4+5+1+7+12+8+2+5+6}{9} = \frac{50}{9} = 5.56$ (2 d.p.)

Example 2. The mean of dataset 'B' is $\frac{1+25+3+3+6+2+2+7+5+4}{10} = \frac{58}{10} = 5.8$

Advantages

- + It is the most commonly used 'average' and can be easily worked out
- + All the data values are used and represented
- + It is used in many higher-level statistical procedures (standard deviation for example)

Disadvantages

- The mean is unduly affected by extreme values in the dataset

The Mode

The value that occurs most often in a data set. There may be more than one modal value.

Example 3. The mode of dataset 'A' is 5 as '5' appears the most frequently in the dataset.

Example 4. The mode of dataset 'B' is 2 and 3 as '2' and '3' are most frequent.

Advantages

- + Can easily be found from any dataset
- + Easy to understand
- + Not affected by outliers
- + Can be obtained for qualitative data
E.g. From the table water was consumed most often

Disadvantages

- Not all datasets have a modal value
- The mode lacks useful mathematical properties

The Median

The middle value when the data is lined up in numerical order.

Example 5. To find the median of dataset 'A'

Line the data up in numerical order

1, 2, 4, 5, 5, 6, 7, 8, 12 the middle value is 5 (coloured blue)

This is the median

Example 6. To find the median of dataset 'B'

Line the data up in numerical order

1, 2, 2, 3, 3, 4, 5, 6, 7, 25 the middle value lies between 3 and 4. The median is 3.5 or $3\frac{1}{2}$

Advantages

- + Concept is easy to understand
- + Can be determined in any dataset
- + Is not unduly influenced by extreme value in the dataset

Disadvantages

- The data must be ordered (ascending or descending)
- The median lacks useful mathematical properties

