Worksheet: Describing a Set of Data

This worksheet explains how to describe a set of data. When describing data there are two aspects of the data you should describe. The first is where the data is centered and the second is how the data is spread out. There are three measures of center that tell where data is centered. The five number summary describes the spread of the data.

The Measures of Center

There are three measures of center: the mean, median, and mode.

Mean: This is also called the average. To find the mean, add all the data values then divide by the number of data points.

Median: This is the middle number in an ordered list. To find the median, first list the data points from smallest to largest. The data is now in an ordered list. Now find the middle data point, that data point is the median. If the middle is between two data points, then find the average of those two data points.

Mode: This is the data value that occurs most often. There can be more than one mode if data points tie in occurring the most. If all the data points occur the same number of times, there is no mode.

Example 1: Find the three measures of center for the data {2, 3, 6, 7, 6, 4, 8}.

Mean = (2+3+6+7+6+4+8)/7 = 36/7 = 5.14

Median: First order the list: 2, 3, 4, 6, 6, 7, 8. The middle number is 6. The median = 6

Mode: Mode = 6 The data value that occurs the most is the 6, it occurs twice.

Example 2: Find the three measures of center for the data {3, 7, 2, 10, 12, 5}.

Mean = (3+7+2+10+12+5)/6 = 39/6 = 6.5

Median: First order the list: 2, 3, 5, 7, 10, 12. The middle is between the 5 and 7. Median = (5+7)/2 = 12/2 =6.

Mode: There is no mode. All the numbers occur once.
The Five Number Summary

The five number summary describes the spread of the data in two ways. It gives the range and the interquartile range of the data. It consists of five numbers:

**Minimum:** The smallest data value.
**Lower Quartile:** The middle value in the lower half of an ordered list.
**Median:** The middle value in an ordered list.
**Upper Quartile:** The middle value in the upper half of an ordered list.
**Maximum:** The largest data value.

The minimum and maximum values show the range of the data. The **range** is the distance from the smallest to the largest data value. The median is the center of the data; it is at the 50% mark. The lower quartile is the 25% mark and the upper quartile is the 75% mark. The distance from the lower to the upper quartile is called the **interquartile range**. The interquartile range shows where half (50%) of the data is located.

To find the lower and upper quartiles, first find the median. Then using just the data points below the median, find the middle value of those points. This is the lower quartile. If the middle is between two values, average the values. To find the upper quartile use the data values above the median, and find the middle value. The lower and upper quartiles are just the medians of the lower and upper halves of the data.

*Example 1:* Find the five number summary for the data {2, 3, 6, 7, 6, 4, 8}.

First order the list 2, 3, 4, 6, 6, 7, 8.

Find the median, median = 6, note it occurs at the first 6 data point. The lower quartile is the middle of 2, 3, 4, the lower quartile = 3. The upper quartile is the middle of 6, 7, 8, the upper quartile = 7.

Minimum = 2
Lower Quartile = 3
Median = 6
Upper Quartile = 7
Maximum = 8
**Example 2:** Find the five number summary for the data 
\{3, 7, 2, 10, 12, 5\}.

First order the list 2, 3, 5, 7, 10, 12.

Find the median, median = (5+7)/2 =6. The median occurs between the 5 and 7.
The lower quartile is the middle of 2, 3, 5, the lower quartile = 3.
The upper quartile is the middle of 7, 10, 12, the upper quartile = 10.

Minimum = 2
Lower Quartile = 3
Median = 6
Upper Quartile = 10
Maximum = 12