5 Figure Summary

Definitions: – A Reminder

Mean Average – Add up all the data and divide by the number of data items.
e.g. $\{x_1, x_2, x_3, \ldots\ldots, x_n\}$

then: $\bar{x} = \frac{\sum x}{n}$. Note that we use the symbol $\bar{x}$ to denote the mean of *x*

and that the symbol Σ is shorthand for "the sum of all of the"

Median The data item in the middle, **when they are arranged in order**.

Note: if there are an odd number of data then it is the one in the middle
e.g. a, b, c, d, e then it is item '*c*'

if there are an even number of data then it is in the middle
(or the average of) the two middle items.
e.g. a, b, c, d, e, f then it is the average of items '*c*' and '*d*'

Mode The data item that occurs the most, or most frequently occurring.
e.g. 1, 3, 7, 8, 8, 8, 9, 11 then the mode would be 8
Note there may be more than one mode.
If there are two modes, then the data set is said to be **bimodal**.

Range The highest – lowest. (The spread of the data)
e.g. 12, 15, 19, 21, 23, 35 Range is: 35 – 12 = 23

Quartiles The median divides the data set into two halves.

There will be a data item, that divides the lower half into two equal parts.
This is known as the **Lower quartile** or Q_1

Similarly,

There will be a data item, that divides the upper half into two equal parts.
This is known as the **Upper quartile** or Q_3

The median is also known as the **mid–Quartile** or Q_2

In all the above cases, the **data MUST be arranged in order**.

e.g. 1, 3, 6, 7, 9, 11, 15, 16, 16, 18, 20, 21, 23

Q_2
(median)

1, 3, 6, 7, 9, 11, 15, 16, 16, 18, 20, 21, 23

Q_1 (lower quartile) Q_2 (median) Q_3 (lower quartile)

6.5 **15** **19**

5 Figure Summary

Inter–Quartile Range Upper quartile – lower quartile

Semi Inter–Quartile Range ½ (Upper quartile – lower quartile)

Abbreviations: We use:

LQ or Q_1 for lower quartile

Q_2 for the median

UQ or Q_3 for upper quartile

L for the lowest data item (the first when arranged in order)
H for the highest data item (the last when arranged in order)
SIR for the semi–interquartile range.

Five Figure Summary For any set of data, by arranging in order, we can easily obtain the:

Lowest, L
Lower quartile LQ, Q_1
Median Q_2
Upper quartile UQ, Q_3
Highest, H

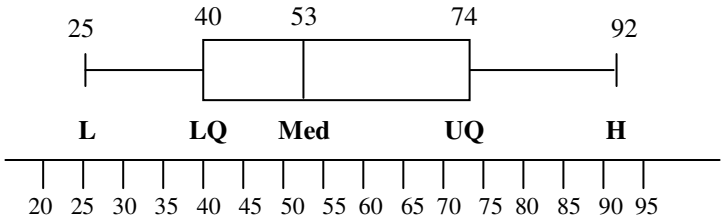
This is known as the Five Figure Summary and can be illustrated in a **Box Plot.**

This is a ***sketch*** which can readily demonstrate the spread of the data.

Box Plot

For data which has the following 5–figure summary

L = 25, Q_1 = 40, Q_2 = 53, Q_3 = 74, H = 92



Dot Plot

Data can also be illustrated with a Dot Plot.

e.g. 1, 2, 2, 2, 4, 5, 6, 6, 7, 7, 7, 7, 8, 8, 9, 10, 10