

Chapter 2 Notes:

Median - middle value

The position of the median in a dataset is found by $\frac{n+1}{2}$ $n =$ number of data values.

If the median position falls between 2 numbers, find the mean of the 2 numbers

e.g. 13 numbers in data set ; median occurs at $\frac{13+1}{2} = 7$

12 numbers in data set ; median occurs at $\frac{12+1}{2} = 6.5$

so then $\frac{6\text{th value} + 7\text{th value}}{2} = \text{median}$

Range = maximum value - minimum value.

IQR = Interquartile range (this covers the middle 50% of data)

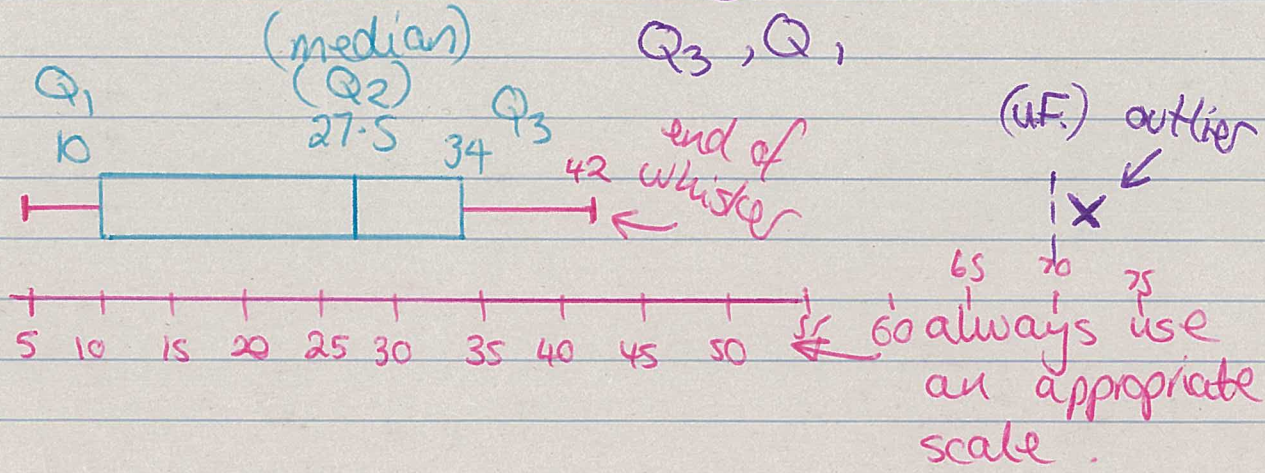
= $Q_3 - Q_1$ ← median of lower half.

↑ median of upper half

* Q_1 bottom 25% of results are below this

* Q_3 top 25% of results are above this

5 figure summary - min, max, median, Q_3, Q_1



Outliers

If a data value is OUTSIDE

$$Q_1 - 1.5 \times IQR \text{ (lower fence)}$$

$$\text{or } Q_3 + 1.5 \times IQR \text{ (upper fence)}$$

} then it is an outlier

So in above example if there was a value of 72, this would be an outlier.

$$IQR = 34 - 10$$

$$= 24$$

$$1.5 \times IQR = 36$$

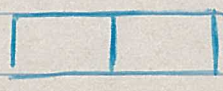
$$Q_3 + 36 = 70$$

72 is above the upper fence of 70 so is an outlier

The whiskers stop at the actual data values that are inside the fences. In this case, that value is 42.

Distributions

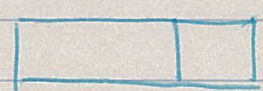
same descriptions as for histograms.



symmetrical



positive skew



negative skew,

Comparing Results / Interpretation

Students tend to find this harder than constructing them.

In general you need to compare and contrast

- 5 figure summary
- spread (range, IQR)
- distribution