Measures of Spread

The measures of central tendency indicate the central values of a set of data. Often, we will also want to know how closely the data cluster around these centres. In other words, we will want to know how spread out the data is. The Range, the IQR (interquartile range), the variance and the standard deviation are measures of spread.

Quartiles are values that divide a body of data into four equal parts.

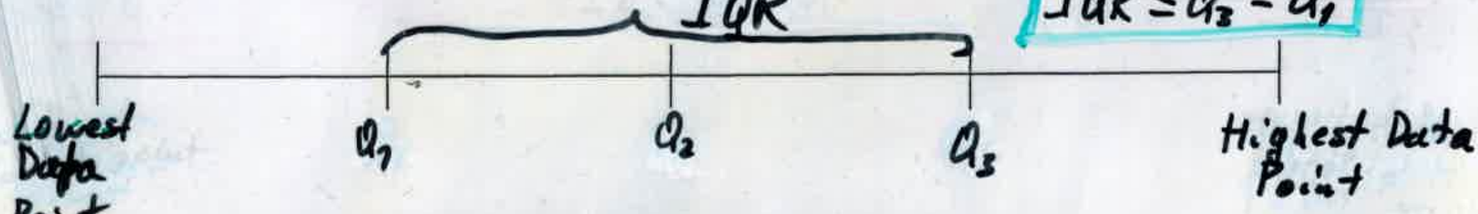
Q_1 (first quartile) - The median of the lower-half of the data

Q_2 (second quartile) - The median of the data set (to be found first)

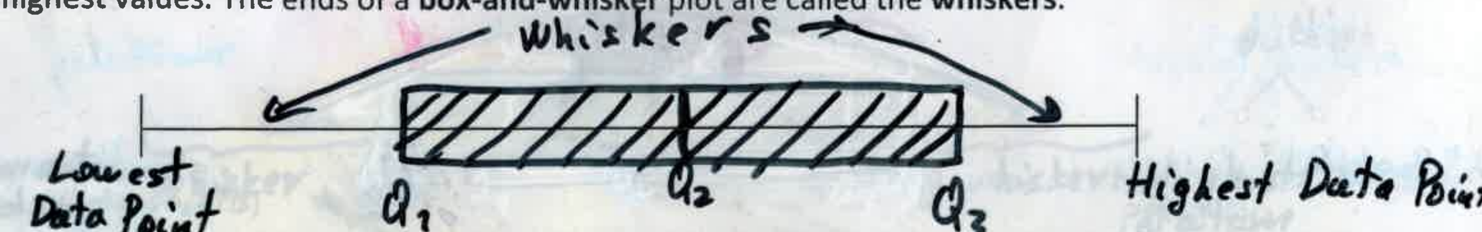
Q_3 (third quartile) - The median of the upper-half of the data

IQR (Interquartile range) is the difference between the first and third quartile.

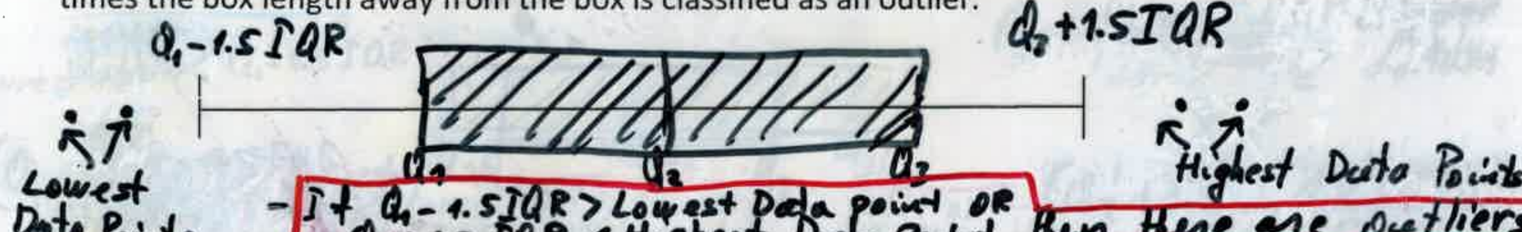
$\boxed{IQR = Q_3 - Q_1}$



A **box-and-whisker** plot shows the first quartile, the median, the third quartile, and the lowest and highest values. The ends of a **box-and-whisker** plot are called the **whiskers**.



When the data contain outliers, a modified **box-and-whisker** plot is used, which shows the outliers as separate points instead of including them in the whiskers. By convention, any point that is at least 1.5 times the box length away from the box is classified as an outlier.



- If $Q_1 - 1.5IQR >$ Lowest Data point OR $Q_3 + 1.5IQR <$ Highest Data Point, then there are outliers.

Example 1: Make a box-and-whisker plot for the following data marks:

~~88~~ ~~56~~ ~~72~~ ~~67~~ ~~59~~ ~~48~~ ~~81~~ ~~62~~ 90 ~~75~~ ~~75~~ ~~43~~ ~~71~~ ~~64~~ ~~78~~

Example 1 Ans: Rearrange the data marks.

Example 2: Do the previous example in Fathom.

Example 3: If n=18, determine position of Q_1, Q_2, Q_3.

Pg. 168: #1,2,5 (do not evaluate standard deviation),9

Example 1 Ans: Rearrange the data marks. $n = 15$

$$\underbrace{43\quad 48\quad 56\quad 59\quad 62\quad 64\quad 67}_{\text{lower-half}}\quad \underset{\substack{\uparrow \\ Q_2}}{71}\quad \underbrace{72\quad 75\quad 75\quad 78\quad 81\quad 88\quad 90}_{\text{upper-half}}$$

$Q_2 = t_8$
$= 71$

- Find Q_1, next

$$43\quad 48\quad 56\quad \underset{\substack{\uparrow \\ Q_1}}{59}\quad 62\quad 64\quad 67$$

$Q_1 = 59$

- Find Q_3: $\quad 72\quad 75\quad 75\quad \underset{\substack{\uparrow \\ Q_3}}{78}\quad 81\quad 88\quad 90$

$Q_3 = 78$

so, $IQR = Q_3 - Q_1$
$= 78 - 59$
$= 19$

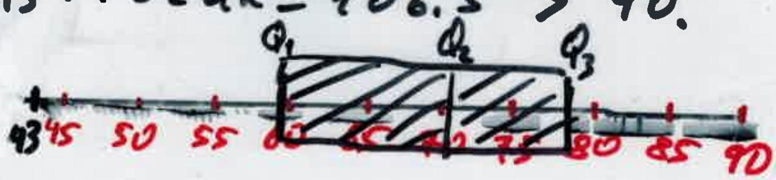
$$\therefore \boxed{\begin{aligned} Q_1 &= 59 \\ Q_2 &= 71 \\ Q_3 &= 78 \\ IQR &= 19 \end{aligned}}$$

Check for outliers:

$Q_1 - 1.5\, IQR = 59 - (1.5)(19)$
$= 59 - 28.5$
$= 30.5$

$Q_3 + 1.5\, IQR = 78 + (1.5)(19)$
$= 78 + 28.5$
$= 106.5$

so, there are no outliers since $Q_1 - 1.5\, IQR = 30.5 < 43$ and $Q_3 + 1.5\, IQR = 106.5 > 90$.



Example 3 Ans: $n = 18$

$$\underbrace{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9}, \overbrace{t_{10}, t_{11}, t_{12}, t_{13}, t_{14}, t_{15}, t_{16}, t_{17}, t_{18}}$$

$$\boxed{\begin{aligned} Q_2 &= \frac{t_9 + t_{10}}{2} \\ Q_1 &= t_5 \\ Q_3 &= t_{14} \end{aligned}}$$