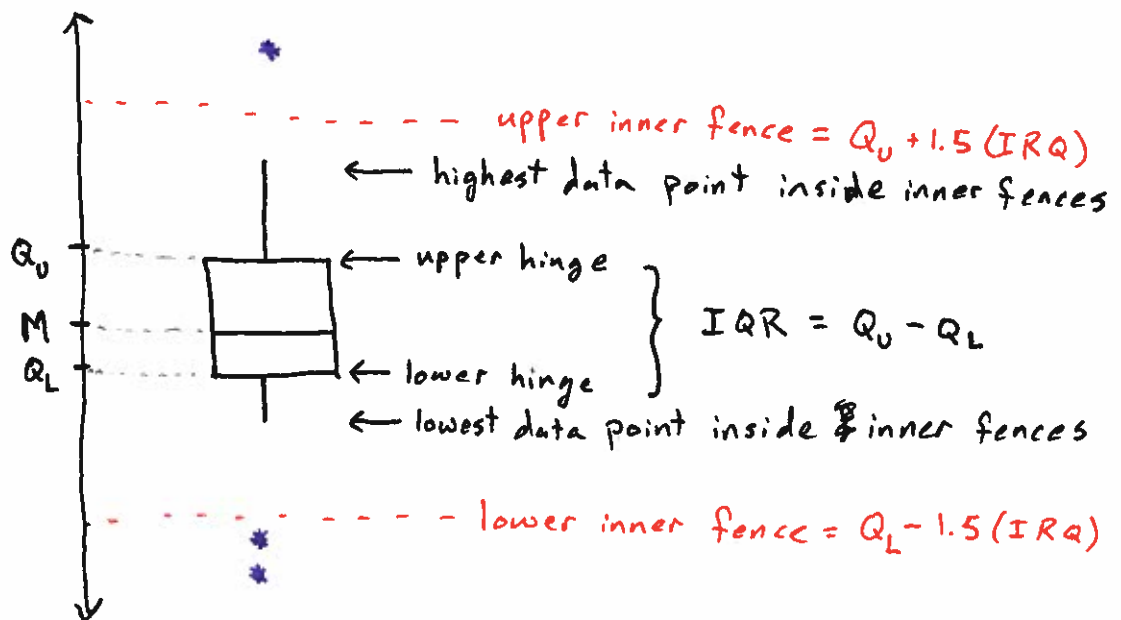


## § 2.7 Methods for Detecting Outliers: Box Plots

### Box Plot



Definition the interquartile range (IQR) is the distance from the lower quartile to the upper quartile.

$$IQR = Q_U - Q_L$$

Upper and lower inner fences do not appear on the box plot.

\* represent possible outliers.

Note: 50% of the data is "in the box".

The lines extending from the box are called whiskers.

A potential outlier is a data point that may have been a mistake and can possibly be omitted from the data. A mistake could occur because of data entry errors, miscommunication, errors in measurement, etc...

Example Draw a Box Plot for the following data

Stem	Leaf
15	89
14	
13	15
12	05 44 81
11	13 33 56 72
10	07 25   47 68 82
9	23 31 71
8	41 99
7	34 62
6	29
5	
4	55
3	
2	03

$$n = 24$$

$$M = \frac{10.25 + 10.47}{2} = 10.36$$

$$Q_L = \frac{8.41 + 8.99}{2} = 8.7$$

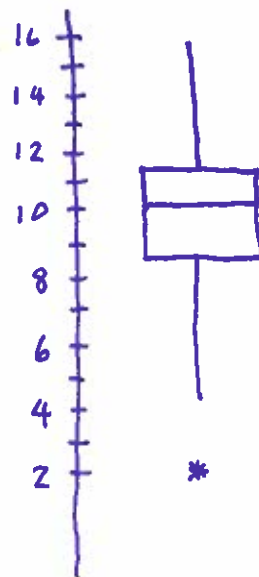
$$Q_U = \frac{11.56 + 11.72}{2} = 11.64$$

$$IQR = 11.64 - 8.7 = 2.94$$

$$LIF = 8.7 - 1.5(2.94) = 4.29$$

$$UIF = 11.64 + 1.5(2.94) = 16.05$$

Box Plot



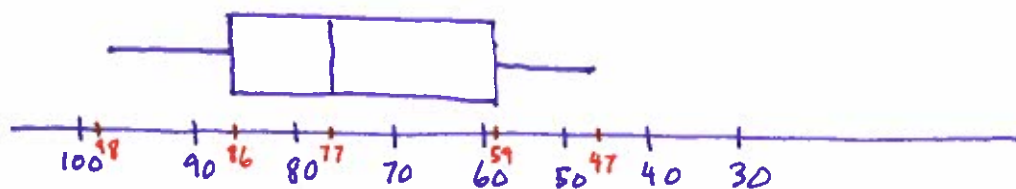
Possible outlier at 2.03. Lower whisker at 4.55  
 Upper whisker at 15.89. 50% of data is in the interval  
 [8.7, 11.64].

Example Draw a Box Plot for the following exam scores: 56, 76, 64, 55, 84, 92, 62, 88, 77, 96, 86, 63, 73, 82, 53, 52, 59, 76, 79, 98, 52, 78, 78, 47, 96, 84.

First put the data into ascending order:

47, 52, 52, 53, 55, 56, 59, 62, 63, 64, 73, 76, ~~76~~, 77, 78, 78, 79, 82, 84, 84, 86, 87, 88, 92, 96, 96, 98. *never mind*

$n = 27$ ,  $M = 77$ ,  $Q_L = 59$ ,  $Q_U = 86$ ,  $IQR = 86 - 59 = 27$   
 $LIF = 59 - 1.5(27) = 18.5$ ,  $UIF = 86 + 1.5(27) = 126.5$



No outliers. 50% of data is in the interval  $[59, 86]$ .  
 Lower whisker at 47, upper whisker at 98.