V. ii. Location ← find the data

eg I want to know a data that is in 60th percentile.

Steps: 1. Sort the datas

2. Compute the location L with k, where k is the percentile and n is the total number of values.

$$L = \frac{k}{100} \cdot n$$ ← reverse of percentile

A. If L is an integer, then count it.

B. If L is decimal, take $\frac{L\text{th value} + (L+1)^{th}\text{ value}}{2}$ average

eg $L = 7.8$, data $= \frac{7\text{th value} + 8^{th}\text{ value}}{2}$

eg 1, 3, 5, 7, 10 (1st, 2nd)

If $L = 2$, data is $\boxed{3}$

eg 4, 6, 10, 15 (2nd, 3rd)

If $L = 2.3$, data is $\frac{6+10}{2} = \boxed{8}$

eg 1, 2, 3, 5, 7, 9, 10 (6th, 7th)

If $L = 6.4$, data is $\frac{9+10}{2} = \boxed{9.5}$

Eg. Find the data value of the 18th percentile.

7th 8th

19	19	20	20	20	20	22	22	22	22
23	23	23	23	23	23	23	24	24	24
24	24	25	25	25	25	25	25	25	26
26	26	26	26	26	27	27	28	28	30

← 40 of them

S: $n = 40$, $k = 18^{th}$

$L = 0.18 \cdot 40 = 7.2$ ← decimal

$P_{18} = \frac{22+22}{2}$

$= \boxed{22}$ chips ← data value

Eg. Find the data value of the 75th percentile.

19	19	20	20	20	20	22	22	22	22
23	23	23	23	23	23	23	24	24	24
24	24	25	25	25	25	25	25	25	26
26	26	26	26	26	27	27	28	28	30

← 30th

S: $n = 40$, $k = 75^{th}$

$L = 0.75 \cdot 40 = 30$

$P_{75} = \boxed{26}$ chips ← data value

Eg. Find the data value of the 48th percentile.

19	19	20	20	20	20	22	22	22	22
23	23	23	23	23	23	23	24	24 (19th)	24 (20th)
24	24	25	25	25	25	25	25	25	26
26	26	26	26	26	27	27	28	28	30

S: $n = 40$, $k = 48^{th}$

$L = 0.48 \cdot 40 = 19.2$

$Data = \frac{24 + 24}{2}$

$= \boxed{24}$ chips

Eg. Given the sorted list of 20 values:

3 5 6 7 8 9 11 12 (8th) 14 18 20 24 26

27 28 30 33 34 35 40.

Find the value of P_{40}.

S: $n = 20$, $k = 40^{th}$

$L = 0.4 \cdot 20 = 8$

$P_{40} = \boxed{12}$

Eg. Given the sorted list of 20 values: (9th) (10th)

3 5 6 7 8 9 11 12 14 18 20 24 26

27 28 30 33 34 35 40.

Find the value of P_{47}.

S: $n = 20$, $k = 47^{th}$

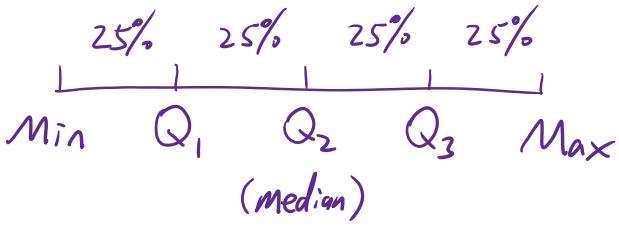
$$L = 0.47 \cdot 20 = 9.4$$

$$P_{47} = \frac{14 + 18}{2}$$

$$= \boxed{16}$$

iii. Quartiles

It is a special measure of position for the n number of data values. It divides into 4 intervals.



We have 5-number summary:

The 5-number summary is Minimum, first quartile Q_1, second quartile Q_2 (median), third quartile Q_3 and Maximum.

They are for plotting the boxplot.

Eg. Find the 5-number summary for the following:

Min → 19	19	20	20	20	20	22	22	22	22
23	23	23	23	23	23	23	24	24	24
24	24	25	25	25	25	25	25	25	26
26	26	26	26	26	27	27	28	28	30 ← Max

S: $n = 40$ here, we need 25th, 50th, 75th

$L_{25} = 0.25 \cdot 40 = 10$,

$Q_1 = \boxed{22}$

$L_{50} = 0.5 \cdot 40 = 20$

$Q_2 = \boxed{24}$

$L_{75} = 0.75 \cdot 40 = 30$

$Q_3 = \boxed{26}$

The 5-number summary: $\boxed{19, 22, 24, 26, 30}$

Eg. Find the 5-number summary for the following datas:

20	15	56	18	89	56	43	38	76

S: Sort: 15 (min) 18 20 38 43 56 56 76 89 (max)

$n = 9$, find L_{25}, L_{50}, L_{75}

$L_{25} = 0.25 \cdot 9 = 2.25$,

$Q_1 = \frac{18 + 20}{2} = \boxed{19}$

$L_{50} = 0.5 \cdot 9 = 4.5$

$Q_2 = \frac{38 + 43}{2} = \boxed{40.5}$

$L_{75} = 0.75 \cdot 9 = 6.75$

$Q_3 = \frac{56 + 56}{2} = \boxed{56}$

The 5-number summary: $\boxed{15, 19, 40.5, 56, 89}$

$Q_3 = \frac{\quad}{2} = 56$

The 5-number summary: 15, 19, 40.5, 56, 89

Eg. Find the 5-number summary for the following datas:

32	27	26	20	17	39	41	23	24	27

S: Sort: 17 20 23 24 26 27 27 32 39 41 ← 10 of them

$L_{25} = 0.25 \cdot 10 = 2.5$

$Q_1 = \frac{20+23}{2} = 21.5$

$L_{50} = 0.5 \cdot 10 = 5$

$Q_2 = 26$

$L_{75} = 0.75 \cdot 10 = 7.5$

$Q_3 = \frac{27+32}{2} = 29.5$

The 5-number summary: 17, 21.5, 26, 29.5, 41

VI. Boxplot

It gives a sense of distribution, which it displays the distance for a 5-number summary.

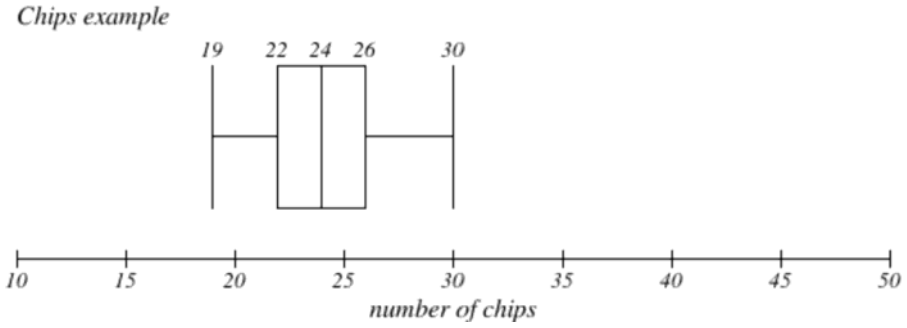
https://www.imathas.com/stattools/boxplot.html

Eg. Plot a boxplot for 19, 22, 24, 26, 30.

Boxplot 1 title: Chips example
Min: 19 Q1: 22 Median: 24 Q3: 26 Max: 30

Boxplot limits:
Overall min: 10 Overall max: 50 Distance between tick marks: 5
Axis Title: number of chips

Display Numbers on Boxplot: ☑
Image Size: Width= 550 Height= 200
Draw here



eg. A random SAT subject score data base:

Min = 230, Q_1 = 330, Q_2 = 520, Q_3 = 670, Max = 800.

Construct a boxplot.

S:

Boxplot 1 title: SAT score data base
Min: 230 Q1: 330 Median: 520 Q3: 670 Max: 800

Boxplot limits:
Overall min: 200 Overall max: 800 Distance between tick marks: 100
Axis Title: SAT scores

Display Numbers on Boxplot: ☑
Image Size: Width= 550 Height= 200
Draw here