

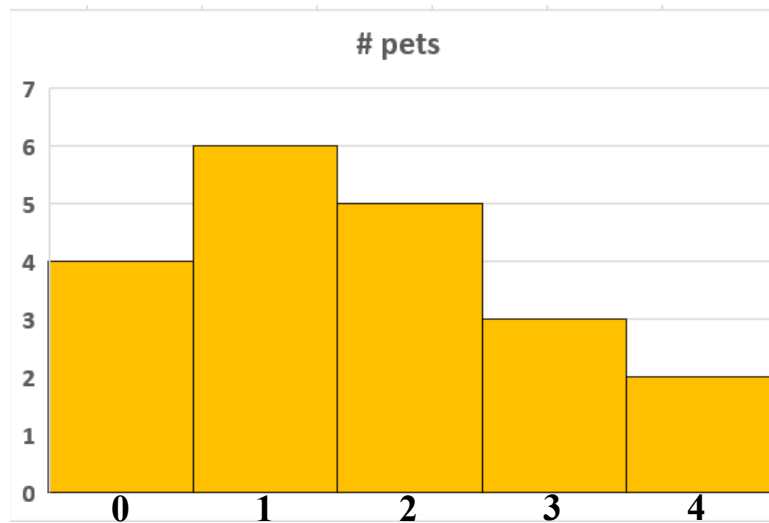
Histograms & Polygons

Histogram

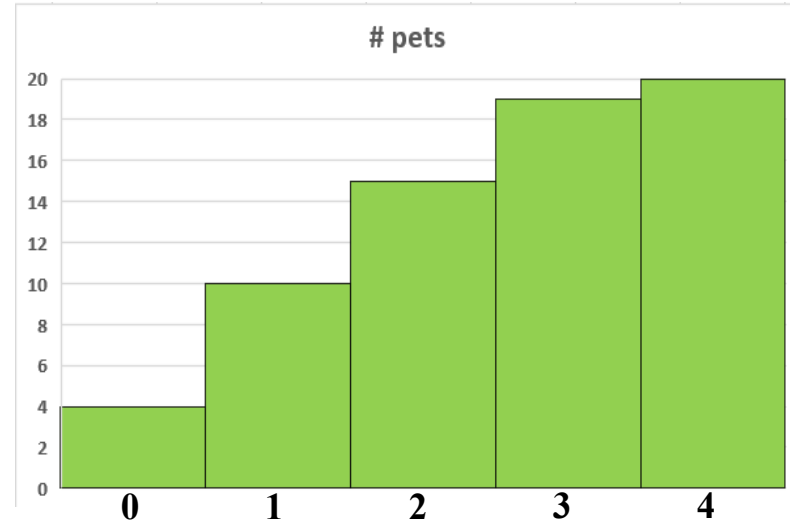
A histogram is a type of visual summary used in statistics. It uses columns that correspond to the scores in the frequency distribution table.

When constructing a histogram;

- * each column is centred on the score (these values are called the bins)
- * the columns join up with no gaps



frequency histogram



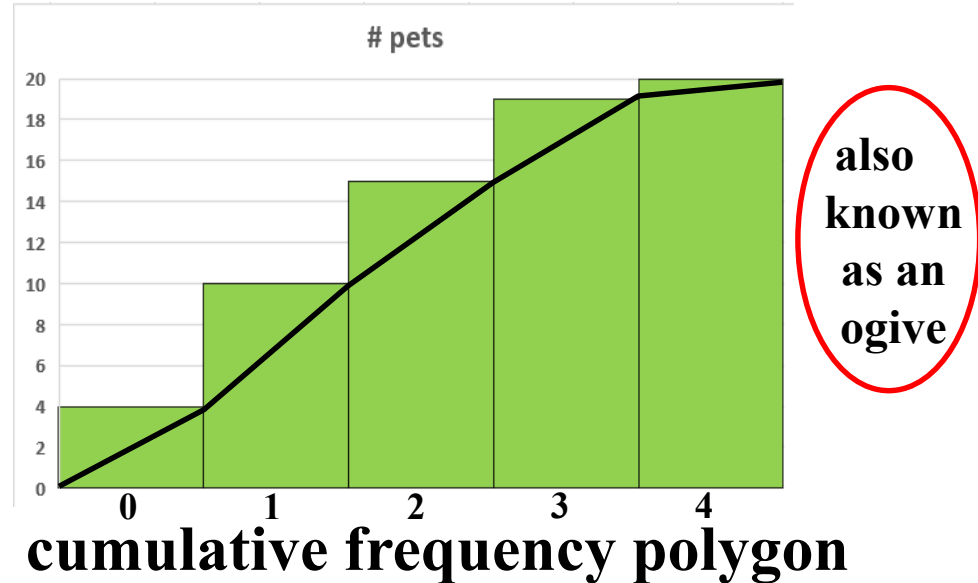
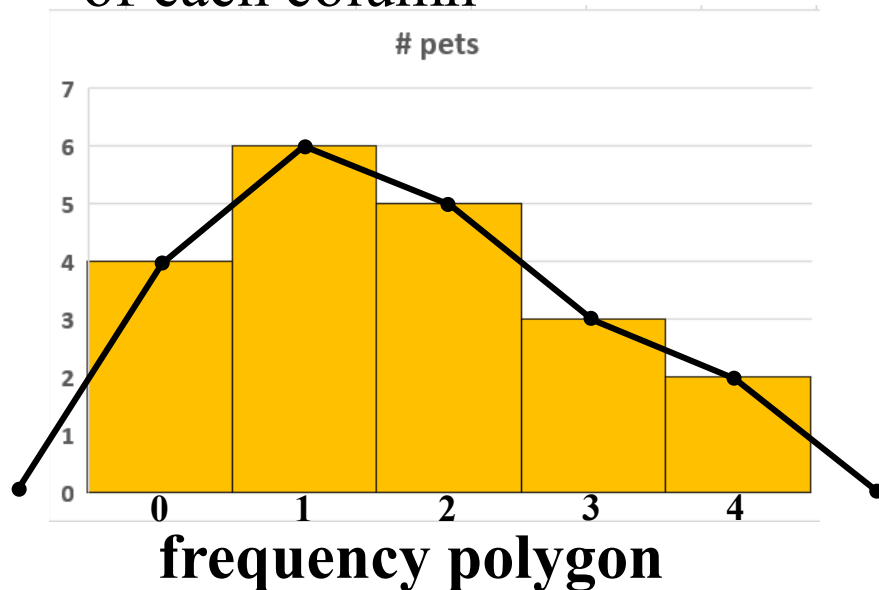
cumulative frequency histogram

Polygon

A polygon is the other common type of visual summary used in statistics. It is a line graph

When constructing a polygon;

- * the plotted points on a frequency polygon are at the centre top of each column
- * start the polygon at the previous bin and finish at what would be the next bin
- * A cumulative frequency polygon starts at the bottom left vertex of the first column and then continually connects to the top right vertex of each column



also known as an ogive

Grouped Data

When data is continuous or the number of different discrete pieces of data is large, it is useful to group the data into class intervals (classes).

Guide to grouping data

- * classes must be inclusive and non-overlapping
- * each score must belong to one, and only one class
- * any scores on a boundary, must be treated the same for all classes
i.e. always place in the lower class or always place in the upper class
and note how the data has been handled
- * determine the number of classes, generally it should be between 5 and 10
- * intervals should be the same width determined by;

$$\text{width} = \text{range} \div \text{number of intervals}$$

Note: both width and number of intervals should be rounded up

- * for calculations, each score takes the value of the class centre

e.g. a sample of the heights, in centimetres, of 300 students has been placed into eight classes

Height (cm)	120 - 130	130 - 140	140 - 150	150 - 160	160 - 170	170 - 180	180 - 190	190 - 200
# of people	16	24	59	100	41	31	19	10

(i) Using a frequency distribution table, find the mean, median class, modal class, variance and standard deviation

x	125	135	145	155	165	175	185	195	Σ
f	16	24	59	100	41	31	19	10	300
fx	2000	3240	7316	15500	6765	5425	3515	1950	46950
fx^2	250000	437400	1240475	2402500	1116225	949375	650275	380250	7426500

need to divide by n as we are using frequencies instead of probabilities

$$\begin{aligned} \bar{x} &= \frac{\Sigma fx}{n} \\ &= \frac{46950}{300} \\ &= \underline{156.5} \end{aligned}$$

$$\begin{aligned} s^2 &= \frac{\Sigma fx^2}{n} - (\bar{x})^2 \\ &= \frac{7426500}{300} - (156.5)^2 \\ &= \underline{262.75} \end{aligned}$$

$$\begin{aligned} s &= \sqrt{262.75} \\ &= \underline{16.21} \text{ (to 2dp)} \end{aligned}$$

x	125	135	145	155	165	175	185	195	Σ
f	16	24	59	100	41	31	19	10	300
cf	16	40	99	199	240	271	290	300	

Thus mean = 156.5

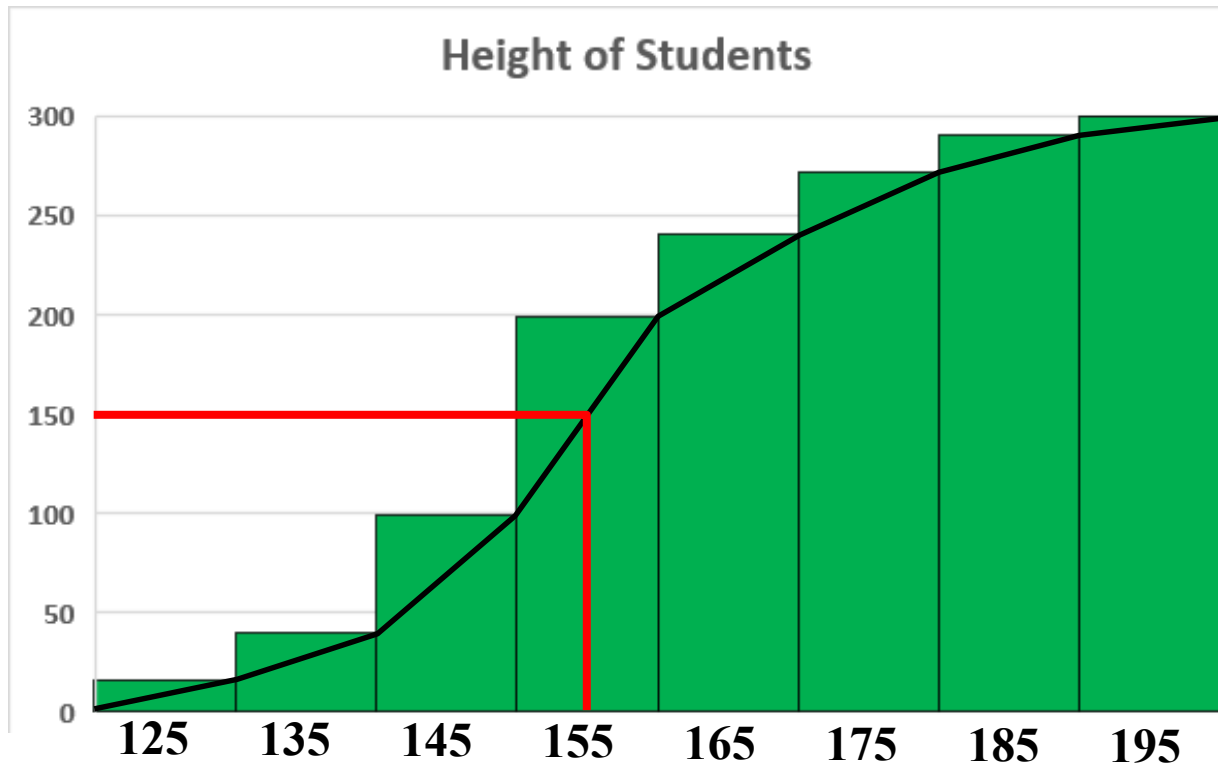
variance = 262.75

standard deviation = 16.21

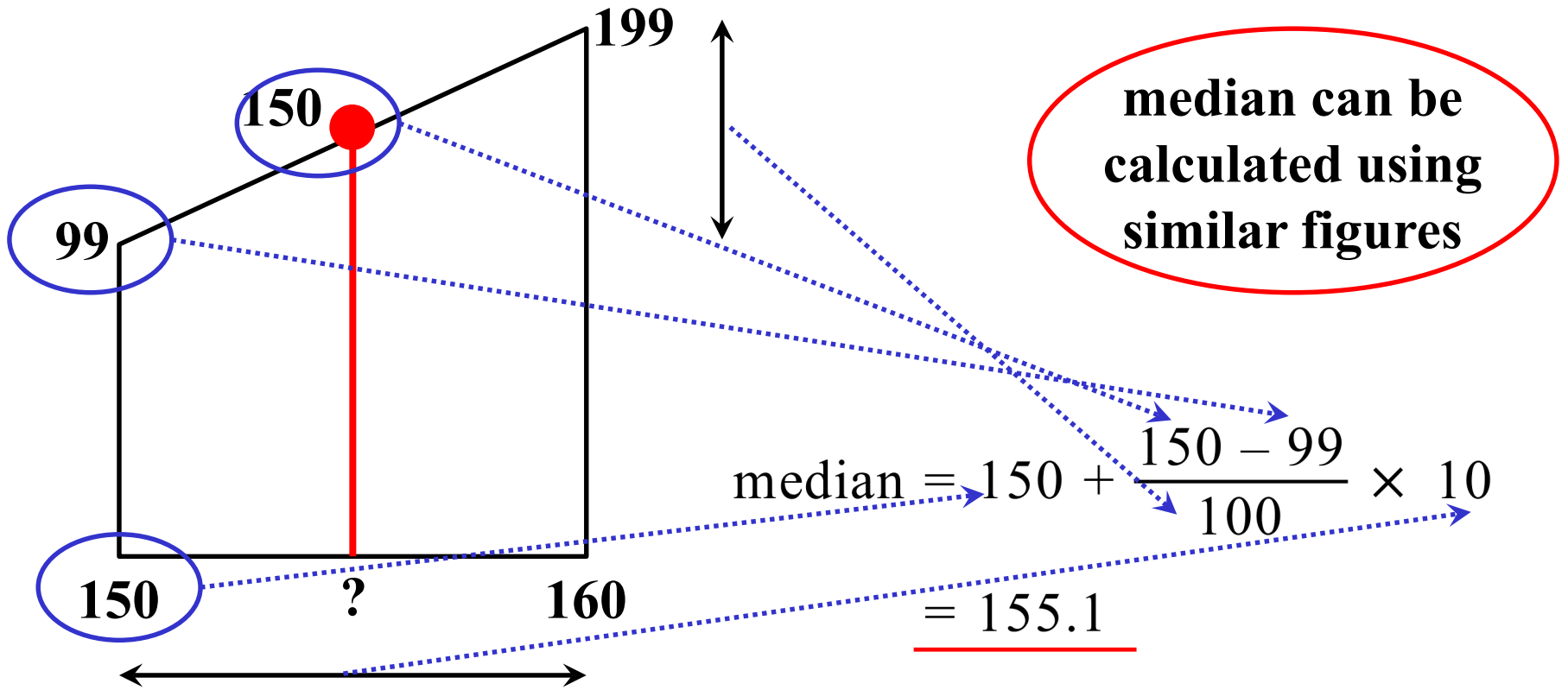
median class = 150 – 160

modal class = 150 – 160

(ii) Use the ogive to calculate the median



The median height corresponds to a cumulative frequency of 50% (150)



Exercise 15B; 1b, 3bd, 4, 5, 6, 7