Lesson 11

11.1 Histograms #2

Having looked at some simple histograms in Lesson 10, the moment has come to look at some of the more subtle aspects of their underlying mathematics. Firstly, note that only continuous data should be represented by a histogram. Secondly, recall that the key idea behind a histogram is that Area is directly proportional to frequency.

$$Area \propto Frequency$$

$$\therefore \ Area \ = \ k \ \times \ Frequency$$

The constant of this proportionality need not be 1 which can give rise to some quite tricky problems.
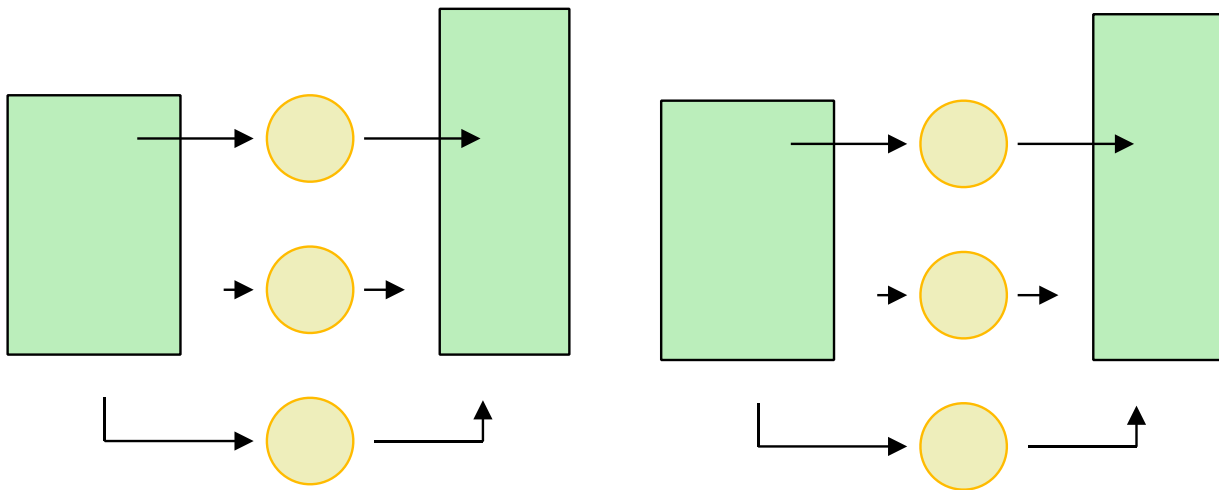
11.2 Example

The number of hours of sunshine each day, h, over a seven week period in the summer was recorded at Heathrow Airport;

Hours	$0 \leqslant h < 5$	$5 \leqslant h < 8$	$8 \leqslant h < 11$	$11 \leqslant h < 12$	$12 \leqslant h < 14$
Frequency	18	9	12	6	4

A histogram was drawn to represent these data.

The $8 \leqslant h < 11$ group was represented by a bar of width 1.5 cm and height 12 cm.

Find the width and height of the $0 \leqslant h < 5$ group.

11.3 Exercise

Question 1

The DMG (daily maximum gust) is measured (in knots) at Hurn (Bournmouth Airport) throughout May and June 2019.

The data is summarised in the table;

DMG, g	$10 \leqslant g < 15$	$15 \leqslant g < 18$	$18 \leqslant g < 20$	$20 \leqslant g < 25$	$25 \leqslant g < 50$
Frequency	3	9	9	20	16

Marta draws a histogram to represent this data.

She draws the bar representing the $10 \leqslant g < 15$ class is 2.5 cm wide and 1.8 cm high.
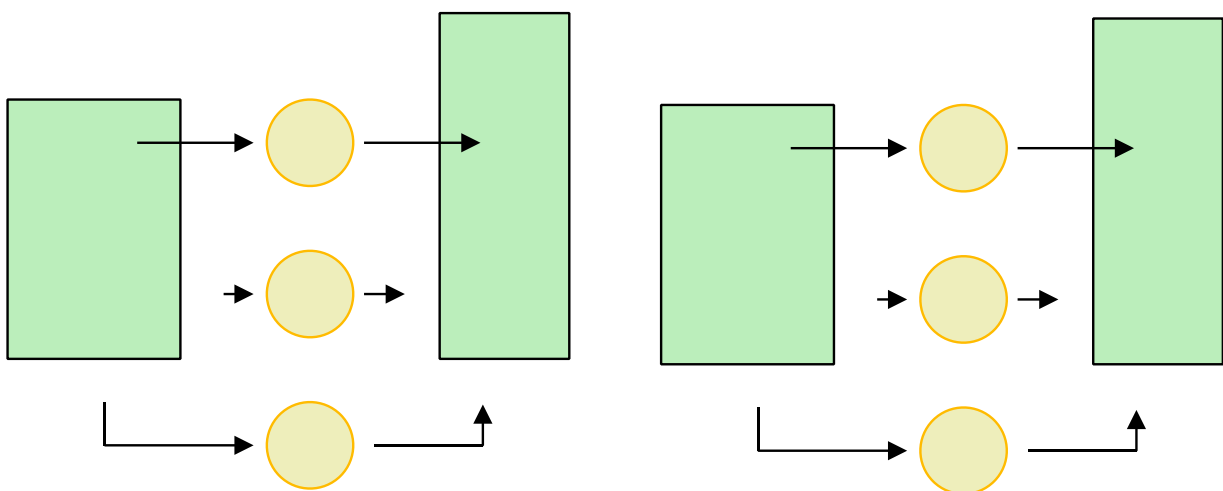
(i) Calculate the width and height of the bar representing the $18 \leqslant g < 20$ class.
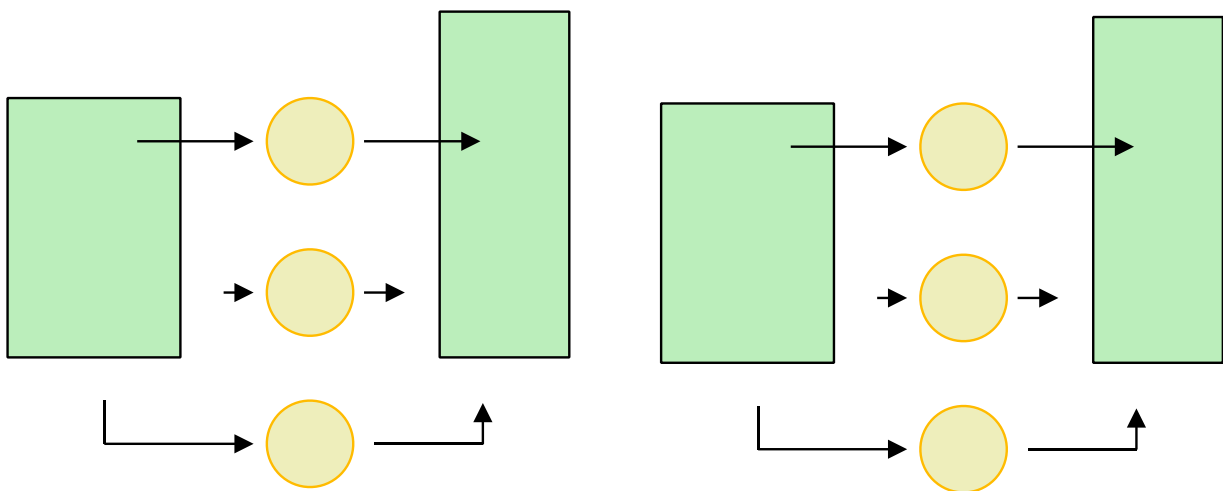
Cathy also draws a histogram to represent the same data.

She draws the bar representing the $10 \leqslant g < 15$ class is 2 cm wide and 3 cm high.

(ii) Calculate the width and height of the bar representing the $18 \leqslant g < 20$ class.

Question 2

The lifetime in thousand hours, L, of 36 LED lightbulbs is given in the following table;

Lifetime (1000 Hours)	$0 \leqslant L < 45$	$45 \leqslant L < 55$	$55 \leqslant L < 80$
Frequency	9	12	15

A histogram was drawn to represent these data.

The $45 \leqslant L < 55$ group was represented by a bar of width 2 cm and height 9 cm.

(i)   Find the width and height of the $55 \leqslant L < 80$ group.

(ii)   Find the width and height of the $0 \leqslant L < 45$ group.

Question 3

A survey of 100 households in Shrewsbury, Autumn 2019, gave the following results for weekly income £y

Income y (£)	Frequency f
$0 \leqslant y < 400$	12
$400 \leqslant y < 480$	28
$480 \leqslant y < 640$	22
$640 \leqslant y < 800$	18
$800 \leqslant y < 1200$	12
$1200 \leqslant y < 1600$	8

(i)   An IT consultant who took part in the survey has an annual salary of £54000. In which class interval was her income placed ?

A histogram represented $400 \leqslant y < 480$ by a rectangle of width 2cm and height 7 cm

(ii)   Calculate the width and the height of the rectangle representing $640 \leqslant y < 800$

(iii)   Use linear interpolation to determine the median annual income of a household in Shrewsbury, Autumn 2019.
Assume 52.143 weeks in a year, and give your answer to the nearest pound.

(iv)   Toby claims that a histogram should not be drawn for this data.
Do you agree with him ?
Justify your answer.

Question 4

S1 A-level Examination Question from June 2017, Q2 (edited)

An estate agent is studying the cost of office space in London.

He takes a random sample of 90 offices and calculates the cost, £x per square foot.

His results are given in the table below.

Cost (£x)	Frequency	Midpoint (£y)
$20 \leqslant x < 40$	12	30
$40 \leqslant x < 45$	13	42.5
$45 \leqslant x < 50$	25	47.5
$50 \leqslant x < 60$	32	55
$60 \leqslant x < 80$	8	70

The discrete data (money) is modelled as continuous data and a histogram is drawn.

The bar representing $50 \leqslant x < 60$ is 2 cm wide and 8 cm high.

(a) Calculate the width and height of the bar representing $20 \leqslant x < 40$

[3 marks]

(b) Use linear interpolation to estimate the median cost.

[2 marks]

(c)   Estime the mean cost of office space for these data.

<div align="right">

[2 marks]

</div>

(d)   Estimate the standard deviation for these data.
You may use the fact that,

$$\sum fy^2 \; = \; 226\,687.5$$

(e)   Describe, giving a reason, the skewness.

<div align="right">

[1 mark]

</div>