Frequency Distribution

Frequency tables

Class	Number of weeks
20 to 39	14
40 to 59	19
60 to 79	13
80 to 99	6

- We have already met a frequency table of the form shown here. This divides weekly sales into a number of distinct classes and shows the number of weeks where demand fell in each class.
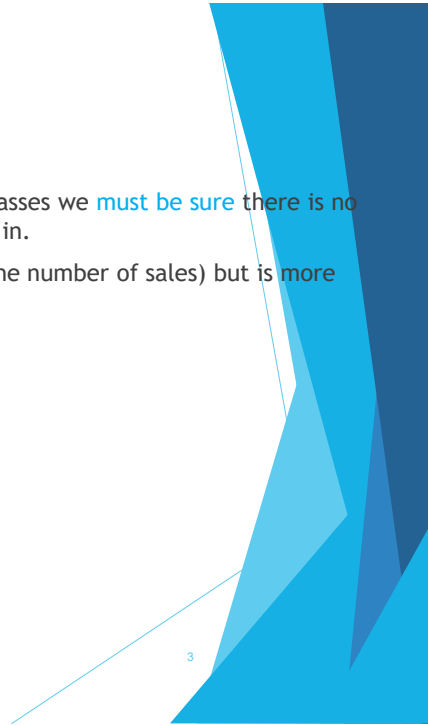
▶ The result is called a **frequency distribution**.

Frequency tables

▶ When defining the boundaries between classes we must be sure there is no doubt about which class an observation is in.

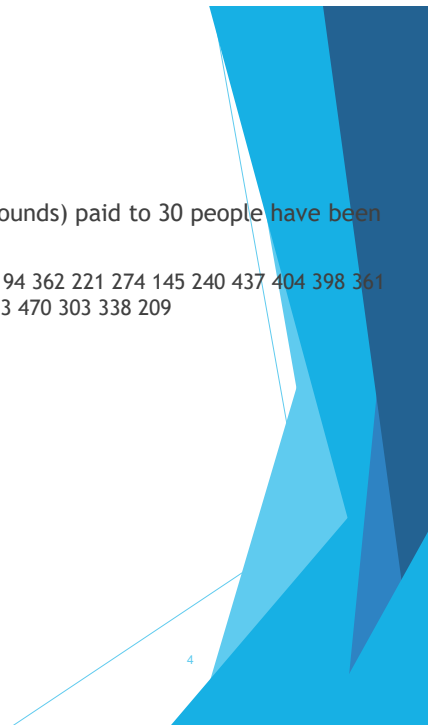▶ This is easy if data are discrete (such as the number of sales) but is more difficult with continuous data.

Frequency tables

▶ During a particular period the wages (in pounds) paid to 30 people have been recorded as follows:

202 457 310 176 480 277 87 391 325 120 554 94 362 221 274 145 240 437 404 398 361 144 429 216 282 153 470 303 338 209

▶ Draw a frequency table of these data.

Frequency tables

- The first decision concerns the number of classes.
- Too few classes (say, three) does not allow patterns to be highlighted; too many classes (say, 20) is confusing and too detailed.
- In this example we shall look for about six classes.

Frequency tables

Class	Frequency
less than £100	2
£100 or more, but less than £200	5
£200 or more, but less than £300	8
£300 or more, but less than £400	9
£400 or more, but less than £500	5
£500 or more	1

Frequency tables

- Frequency distributions show the actual number of observations in each class.
- A useful extension is a **percentage frequency distribution**, which shows the percentage of observations in each class.
- The results are presented in exactly the same way as in standard frequency tables.

L6

7

Percentage frequency distribution

Class	Frequency	Percentage frequency
Less than £100	2	6.67%
£101 or more, but less than £200	5	16.67%
£201 or more, but less than £300	8	26.67%
£301 or more, but less than £400	9	30.00%
£401 or more, but less than £500	5	16.67%
£501 or more, but less than £600	1	3.33%
Total	30	100.00%

L6

8

Cumulative percentage
~~frequency distribution~~

Class	Frequency	Cumulative frequency	Percentage frequency	Cumulative percentage frequency
<100	2	2	6.67%	6.67%
100≥　<200	5	7	16.67%	23.33%
200≥　<300	8	15	26.67%	50.00%
300≥　<400	9	24	30.00%	80.00%
400≥　<500	5	29	16.67%	96.67%
500≥	1	30	3.33%	100.00%
Total	30		100.00%	

IN SUMMARY

▶ Frequency tables show the number of observations that fall into different classes.

▶ They can be used for both continuous and discrete data, and can be extended to show percentage frequency distributions and cumulative distributions.
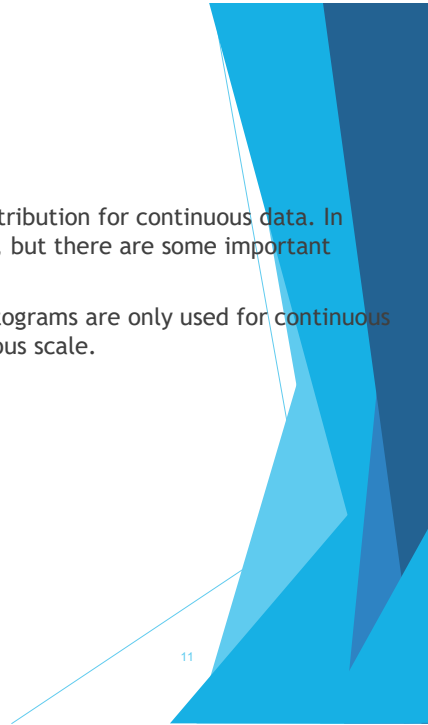
Histograms

▶ **Histograms** are diagrams of frequency distribution for continuous data. In appearance they are similar to bar charts, but there are some important differences.

▶ The most important difference is that histograms are only used for continuous data, so the horizontal axis has a continuous scale.
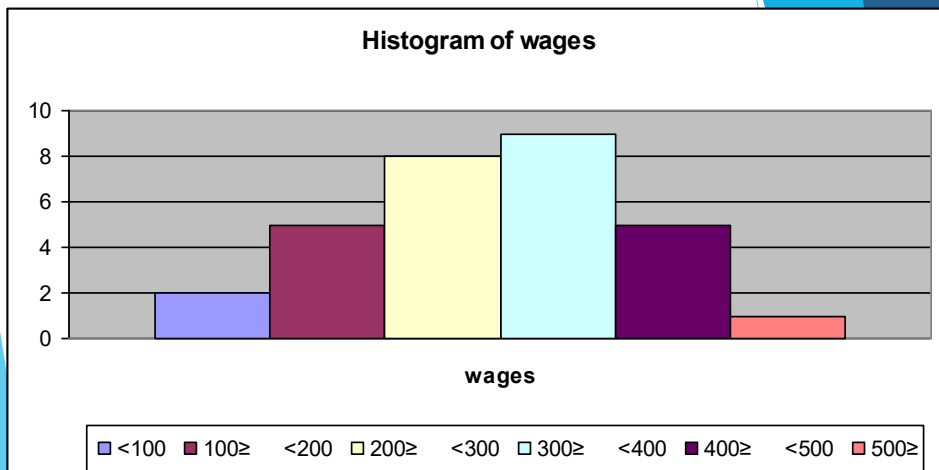
Histograms

▶ Bars are drawn on this scale, so their width, as well as their height, has a definite meaning. This is an important point: in bar charts it is only the height of the bar that is important, but in histograms it is both the width and the height, or in effect the area. We can show this by drawing a histogram of the continuous data for wages shown above.

Histogram of wages

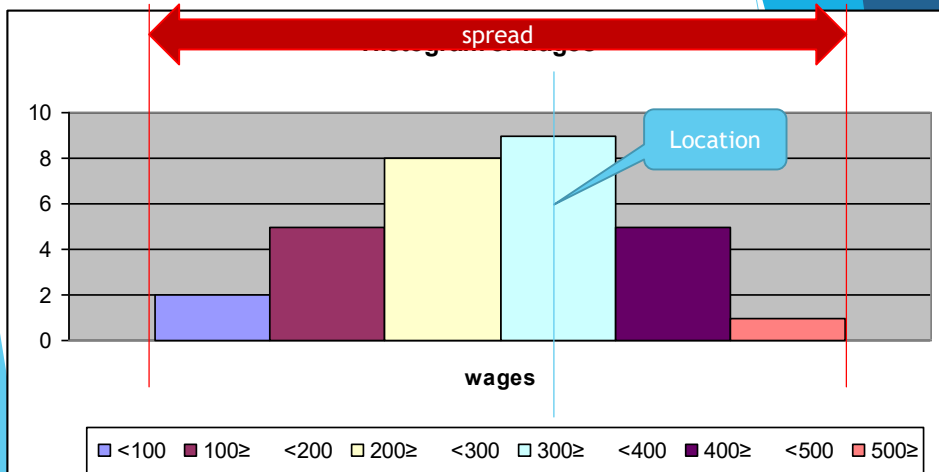Legend: <100   100≥   <200   200≥   <300   300≥   <400   400≥   <500   500≥

Histograms

Two **measures for data**

▶ a measure of location to show where the centre of the data is: one suggestion for this would be an average value

▶ a measure of spread to show how spread out the data is around the centre: one suggestion for this would be the range of values

Histograms

▶ If we take a histogram of observations, a measure of location would show where this histogram lies on the x axis, while a measure of spread would show how dispersed the data are along the axis.

▶ In practice, these two measures are the most important, but others are used to a lesser extent. We could, for example, measure the spread of data in relation to frequency.