Chapter 5: Averaging

Extra questions

1. You have recently been given the expense claims of five sales representatives making the same journey:

£78.00   £56.84   £78.00   £64.76   £68.98

Determine the mean, mean and median.

2. A group of seven employees have been asked how much they typically spend each day in the canteen:

£2.40   £2.50   £0   £0   £2.50   £0   £1.60

Determine the mean, mean and median. Comment on your results.

3. You have been given the following list of numbers.

18	20	19	19	21	20	20	22
21	20	18	21	22	19	20	21
20	21	20	18	20	21	18	19
22	18	21	19	20	21	22	20
19	22	19	20	21	21	19	20

Determine the mean, median and mode.

4. A supermarket has now decided to look at the amount requested in the form of 'cash back'.   They have the following records on the last 20 customers:

0	0	£30	0	0
0	0	0	0	£30
0	£40	0	0	0
0	0	0	£30	£40

Using the basic statistics that you are familiar with, describe the data given in this table.

5. You have been given the following table giving the number of defective items found in boxes of components. Determine the mean, median and mode.

Number of defective components	Number of boxes
0	60
1	4
2	2
3	1
4	1

6. A sample of motorists was asked what they had paid for a recent car service. Their replies were summarised in the following table

Cost of car service	Number of motorists
under £100	8
£100 but under £150	14
£150 but under £200	17
£200 but under £300	12
£300 and over	5

Determine the mean, median and mode. Comment on the construction of this table.

7. You have been given the following survey results on the average age of customer for a particular product:

Age in years	Number
Under 20	4
20 but under 30	23
30 but under 40	45
40 but under 50	47
50 or more	1

Determine the mean, median and mode.

8. Prepare a response to the comment 'that you only need to calculate the mean because that's the only one people really use'.

Extra answers

1. The calculations are shown below:

$$\bar{x} = \frac{78.00 + 56.84 + 78.00 + 64.76 + 68.98}{5} = \frac{346.58}{5} = £69.316$$

The median is the middle value of the ordered list:

£56.84   £64.76   £68.98   £78.00   £78.00

The median position is found using $(n + 1)/2 = 6/2 = 3$. In this case the median is the value of the 3rd observation within this ordered list. The median is therefore £68.98.

The mode (most frequent) is £78.00.

This data is very limited and there is little we can say. It could be that the two submitted expense claims for £78.00 came from sales representatives making the same journey. To comment on the variation in the data, we would need to know more about the travel e.g. mode of travel, time of day, day of the week.

2. The three summary measures of location are shown below:

$$\bar{x} = \frac{2.40 + 2.50 + 0 + 0 + 2.50 + 0 + 1.60}{7} = \frac{9}{7} = £1.29$$

The ordered values are £0 £0 £0 £1.60 £2.40 £2.50 £2.50

The median is £1.60.

The mode is £0

In this case we have included all seven employees even though three out of the seven do not typically spend money in the canteen each day. Looking at the data would suggest that the employees using the canteen would typically spend between £2.40 and £2.50. As with all work of this kind it is important to define the group of interest and relate the calculated statistics to this.

3. You could calculate the mean, median and mode directly from the list of numbers given. However, as these lists become longer it can be easier (depends also on the technology) to work with tabulations. In this case, it would also provide you with the information needed to produce a bar chart.

Number	Frequency	fx	Cum freq
18	5	90	5
19	8	152	13
20	12	240	25
21	10	210	35

22		5	110	40
		40	802	

$$\bar{x} = \frac{802}{40} = 20.05$$

The average is 20 years if we round to the nearest whole number.

Note

If these numbers referred to age we would need to be particularly careful with the calculations. Age usually means age last birthday. A quoted age of 18 years for example would '18 but under 19 years'. In this case, we would need to consider the age interval and work with mid-points.

The median corresponds to the $20^{1}/_{2}$th ordered observation $(40 + 1)/2$. The median lies in the row with 25 or less observations and is therefore 20.

The mode is also 20.

4. The construction of the table includes all those who did not want cash back:

Cash back £'s	Requests	fx	Cum freq
0	15	0	15
10	0	0	15
20	0	0	15
30	3	90	18
40	2	80	20
	20	170	

The mean is £8.50 (£170/20), the median is £0 (75% of customers in this sample did not request any cash back) and the mode is £0.

In this case, we would have to question what these descriptive statistics could usefully tell us. We have calculated a mean of £8.50 but most customers have taken £0, three have taken £30 and two have taken £40. Essentially we are looking at two distinct groups, those that have taken cash back and those that have not. It would be more useful to say that in this small sample (it really is too small), only 25% of customers requested cash back, of those asking for cash back, three out of five wanted £30 and two out of five wanted £40.

5. To determine the mean and median, columns have been added to give totals (*fx*) and cumulative frequency.

Number of defective components (*x*)	Number of boxes (*f*)	fx	Cumulative frequency
0	60	0	60
1	4	4	64
2	2	4	66
3	1	3	67
4	1	4	68
	68	15	

The mean is 0.2706 ($\bar{x}$ = 15/162)

The median corresponds to the $34^1/_2$ ordered observation $(68 + 1)/2$. The median lies with 60 or less observations and is therefore 0.

The mode is 0.

All the statistics are indicating the most likely number of defectives is 0. It could still be more useful to say that 88% of boxes contain no defectives, 6% of boxes contain 1 defective and so on.

6. To determine the mean and median, columns have been added to give totals (*fx*) and cumulative frequency.

Cost of car service	Number of motorists (*f*)	Mid-point (*x*)	fx	Cumulative frequency
under £100	8	50*	400	8
£100 but under £150	14	125	1750	22
£150 but under £200	17	175	2975	39
£200 but under £300	12	250	3000	51
£300 and over	5	350*	1750	56
	56		9875	

* assumed mid-point value

$$\bar{x} = \frac{9875}{56} = £176.34$$

(rounded to the nearest pence)

$$median = 150 + 50\frac{(28 - 22)}{17} = £167.65$$

mode (to be determined from constructing the histogram) = £168.75

The results are reasonably close in this case, suggesting only a small positive skew. It should be noted that the mid-point values assumed for the first and last groups can make a significant difference to the calculated mean. The mid-point values assumed for the first and last groups make no difference to the median and mode.

7.

* If you do have open-ended groups you need to make an assumption about the mid-points. In most research you will have some knowledge that will be helpful. The answer you get will, of course, depend on the mid-points chosen. The important thing is to state assumptions as you make them and justify what you have done. In this case, it might be reasonable to assume no customers under the age of 16 years (this may be a legal requirement) and you may know that there are very few customers over the age of 60 years.

Age in years (x)	Number of respondents(f)	Mid-point (x)	fx	Cumulative frequency
Under 20	4	18*	72	4
20 but less than 30	23	25	575	27
30 but less than 40	45	35	1575	72
40 but less than 50	47	45	2115	119
50 or more	1	55*	55	120
	120		4392	

$$\bar{x} = \frac{4392}{120} = 36.6$$

$$median = 30 + 10\frac{(60-27)}{45} = 37.33$$

mode (to be determined from the construction of the histogram) = 40.42

Note
The answers have come out as decimals but you will need to convert if you want to give your answer in years and months or years, months and days.

8. It is true that most people you are likely to meet will understand the mean and the mean is a particularly important statistic. However, in a wide range of occupations, a critical understanding of statistics is valued and will be seen as a useful competence. The availability of software packages, such as Excel, make the generation of statistics fast and easy.

The mean will only give you one kind of information (a weighted centre of location). The median and mode give different kinds of information. If you are concerned about relative position in the pay league, the median will tell you whether you are in the top

or bottom 50%. If you want to know the most popular shirt or skirt size then the mode and other frequency measures will be of interest.