

SECTION 13

Making a table of data, or tabulating data, is a very important skill. If information is presented in a purely haphazard fashion, or data as a random hotch-potch of numbers, it is very difficult to make sense of it.

Putting the information into the form of a table is often the first step taken in organising it before either trying to assimilate it yourself, or presenting it to someone else. A table is virtually essential before any kind of graph or chart can be drawn.

Contingency Tables

Example 13a

In a survey 87 people, 48 of them women, were asked if they wore seat belts when driving their car. 11 of the women and 14 of the men admitted to not wearing seat belts. Present this information in an organised fashion.

Solution:

A table is clearly required here. We are dealing with two 'variables' – gender (women and men) and seat belts (wear or not) so a simple two-way table will suffice. Tables of this type are called contingency tables, because they cover (or at least try to cover) every eventuality.

We start by making up the skeleton table and putting in the data which we know:

	Wear Belts	Don't Wear Belts	Totals
Men		14	
Women		11	48
Totals			87

Then, if necessary, we fill in any missing numbers.

	Wear Belts	Don't Wear Belts	Totals
Men	25	14	39
Women	37	11	48
Totals	62	25	87

Example 13b

Manufacturers of three brands of soap powder launch different advertising campaigns to see how the powders appealed to various socio-economic groups. Of group A/B, 10 bought brand X and 43 thought about using it; 6 bought brand Y and 14 thought about it; 24 bought brand Z and 75 considered doing so. As regards socio-economic group C, the figures for brand X were 14 bought and 42 considered buying it; brand Y 5 and 17 respectively; brand Z 23 and 63 respectively. For socio-economic class D/E the corresponding results were X 17 (50); Y 0 (12) and Z 17 (61).

Present these results in an organised way.

Solution:

Notice first that there are three variables to be considered: the brand (X, Y or Z), the socio-economic class (A/B, C or D/E) and the response (bought or only considered buying). Thus the table is not as simple as in the previous example.

There are actually several different ways of sub-dividing the rows or columns of the table and any way which is sensible is OK.

Brand of Powder	Actually Bought the Powder			Considered Buying the Powder			Totals
	A/B	C	D/E	A/B	C	D/E	
Brand X	10	14	17	43	42	50	176
Brand Y	6	5	0	14	17	12	54
Brand Z	24	23	17	75	63	61	263
Totals	40	42	34	132	122	123	493

Brand of Powder	Class A/B		Class C		Class D/E		Totals
	Bought	Considered	Bought	Considered	Bought	Considered	
Brand X	10	43	14	42	17	50	176
Brand Y	6	14	5	17	0	12	54
Brand Z	24	75	23	63	17	61	263
Totals	40	132	42	122	34	123	493

Brand of Powder	Consumer Response	Socio-Economic Class			
		A/B	C	D/E	Totals
Brand X	Bought	10	14	17	41
	Considered	43	42	50	135
Brand Y	Bought	6	5	0	11
	Considered	14	17	12	43
Brand Z	Bought	24	23	17	64
	Considered	75	63	61	199
Totals		172	164	157	493

All of the above show the same information in a slightly different way, grouping the data in a variety of ways to emphasise different aspects of it.

Frequency Tables

A frequency table is a way of displaying numerical data in an organised way in order to carry out certain statistical analyses. If you look back at page 32 of this pack you will see a frequency table from which an ogive will be drawn. Page 20 (Example 7b) shows one from which a histogram will be drawn.

Simple Frequency Table

Example 13c

Suppose we have raw data (i.e. data as it arrives, with no attempt to organise it in any way) showing the ages of a group of students:


17 19 20 25 19 28 26 26 20 21 19 24 19 22 18 26 19 23 21 19 19

We could organise it using either a dot plot or a stem leaf chart. Here is a third way.

The smallest number is 17 and the largest is 28 so we write these in a vertical list. Next, we put a mark against each of these as we go along the line of raw data, perhaps crossing each one off as we go to make sure we don't lose our place. (I know it's very easy to do this with just as small group of numbers here, but imagine if you had a couple of hundred numbers to work with!)

Below is the frequency table with the first five numbers of the raw data list entered and ticked off on the list. We call this process of entering marks making a tally.

\checkmark \checkmark \checkmark \checkmark \checkmark
 17 19 20 25 19 28 26 26 20 21 19 24 19 22 18 26 19 23 21 19 19

17 18 19 20 21 22 23 24 25 26 27 28	This is the table with the first five numbers entered on it.	<table border="1"> <thead> <tr> <th>Age</th> <th>Tally</th> <th>Frequency</th> </tr> </thead> <tbody> <tr><td>17</td><td> </td><td>1</td></tr> <tr><td>18</td><td> </td><td>1</td></tr> <tr><td>19</td><td> </td><td>7</td></tr> <tr><td>20</td><td> </td><td>2</td></tr> <tr><td>21</td><td> </td><td>2</td></tr> <tr><td>22</td><td> </td><td>1</td></tr> <tr><td>23</td><td> </td><td>1</td></tr> <tr><td>24</td><td> </td><td>1</td></tr> <tr><td>25</td><td> </td><td>1</td></tr> <tr><td>26</td><td> </td><td>3</td></tr> <tr><td>27</td><td> </td><td>0</td></tr> <tr><td>28</td><td> </td><td>1</td></tr> </tbody> </table>	Age	Tally	Frequency	17		1	18		1	19		7	20		2	21		2	22		1	23		1	24		1	25		1	26		3	27		0	28		1	This is the table with the complete tally on it, followed by a frequency column. Notice that when we get 5 tally marks we cross them off to make a bundle of 5 like this:  This makes them easier to add up
Age	Tally	Frequency																																								
17		1																																								
18		1																																								
19		7																																								
20		2																																								
21		2																																								
22		1																																								
23		1																																								
24		1																																								
25		1																																								
26		3																																								
27		0																																								
28		1																																								

Grouped Frequency Table

Example 13d

Here is a record of the bonuses earned by a group of workers:

77	59	84	73	51	43	50	81	61	53	69	37
58	63	67	61	90	61	50	60	84	56	77	57
42	43	41	49	37	21	24	35	34	50	11	52
38	22	22	49	46	50	64	23	73	73	48	26
30	16	33	67	87	64	47	59	37	92	88	30
51	51	85	71	57	45						

Show these data in a frequency table.

Solution:

As you see, they are spread about quite a bit.

The **range** = largest – smallest = 92 – 11 = 81 so a simple frequency table as in Example 13c is out of the question.

We will have to group the numbers in bundles. This will lead to some loss of detail, but it will make analysis much easier.

But how do we bundle them? The lowest number is 11, so we could make our first bundle (or class interval, to use the correct term) 10-15 and the next one 15-20 and so on, making the width of each interval 5.

But this would, all in all, give us 17 intervals (to get us up to the highest bonus of £92), which is a bit too many; we still have a lot of work to do and the effort of grouping hardly seems worth it.

Suppose we bundled the numbers in 20s, the first being 10-30. The last then has to be 90-110 in order to fit in the £92. (We try to ensure that all intervals are of the same width. Remember the problem of drawing a histogram with unequal intervals?) But this gives us only 5 intervals; we're now losing too much definition.

So bundle the numbers in tens, the first bundle being 10-20, the second being 20-30 (so a bonus of £20 will appear in the second interval, not the first) and the last 90-100. This gives us 9 intervals which is OK.

In theory, statisticians tell us that between 7 and 15 class intervals strikes a happy medium between losing too much detail and giving ourselves too much work. But don't ask me why!

We now make up a tally chart as before. I know there is a temptation to hunt through the entire list of raw data and pick out all the numbers between 10 and 19 inclusive and so on but you are almost guaranteed to make a mistake this way.

Doing a tally may seem longer, and perhaps a bit childish, but it is more efficient in the long run.

If we go down the columns in the raw data list instead of along the rows (i.e. tally the *short* lines of data rather than the long lines) we can find our place more easily if we are distracted during the task.

Here is the raw data list again for easy reference:

77	59	84	73	51	43	50	81	61	53	69	37
58	63	67	61	90	61	50	60	84	56	77	57
42	43	41	49	37	21	24	35	34	50	11	52
38	22	22	49	46	50	64	23	73	73	48	26
30	16	33	67	87	64	47	59	37	92	88	30
51	51	85	71	57	45						

This is the tally chart after the first column has been entered:

Bonus (£)	Tally
10 - 20	
20 - 30	
30 - 40	
40 - 50	
50 - 60	
60 - 70	
70 - 80	
80 - 90	
90 - 100	

This is the tally chart after the second column has been entered:

Bonus (£)	Tally
10 - 20	
20 - 30	
30 - 40	
40 - 50	
50 - 60	
60 - 70	
70 - 80	
80 - 90	
90 - 100	

And this is the final frequency table.

Notice how the frequencies have been automatically totalled to 66.

This allows us to keep a check of sorts on the tally.

The original data list has 66 numbers on it (5 rows of 12 each and a last row of 6.)

Bonus (£)	Tally	Frequency
10 - 20		2
20 - 30		6
30 - 40		9
40 - 50		10
50 - 60		15
60 - 70		10
70 - 80		6
80 - 90		6
90 - 100		2
TOTAL		<u>66</u>

? 13

Make up a table or frequency table (as appropriate) for each of the following sets data.

- Three companies produced figures for their turnover and pre-tax profits for the same financial year, and also the size of their workforce. ABC plc had a turnover of £23.4 million with profits of £4.6 million and a workforce of 102. PQR Ltd employed 35 fewer people, but made a profit of £7.3 million on a turnover of £25.8 million. Finally, XYZ Bros had profits of £2.7 million on a turnover of £10.9 million, with 43 employees.

- The data shows the time taken (in hours to the nearest hour) to finish each of 48 houses on a building site.

911	902	900	867	897	915	945	917	883	874	880	932
919	899	903	872	901	874	925	886	928	917	906	925
913	898	888	912	896	921	908	933	903	920	885	901
892	931	902	893	940	887	928	907	926	895	907	864

- The total number of employees in a company is exactly 1,000. For the purposes of analysis they have been divided into three areas of employment: Production, Administration and Sales, subdivided further into male/female and adult/juvenile. In Production, there were 500 adults (350 of them male) and 100 juveniles (half of them male). Administration had a total of 300 employees, of whom 210 were adults (10 more women than men) and the rest juveniles (10 fewer women than men). The rest of the company's employees worked in the Sales division, with 10 juvenile males, 20 juvenile females, 30 adult females and the rest adult males.

- The following is a record of the weights (in grams) of a sample of components:

105.7	108.3	119.4	111.6	106.2	115.8
114.3	100.2	107.4	112.8	107.3	101.9
114.4	116.8	105.7	108.4	116.1	116.0
109.3	100.0	105.3	119.6	115.2	104.8
104.0	116.6	119.1	117.4	114.2	116.1
117.8	118.4	117.9	116.0	115.7	110.4
117.3	118.3	112.4	119.0	116.4	113.2