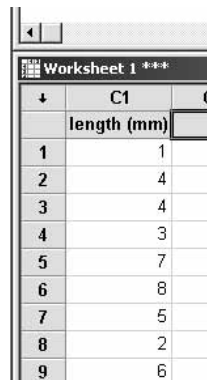


### 3.8.1. Are my data parametric?

**EXAMPLE 3.2.** Length (mm) of two-spot ladybirds (*Adalia bipunctata*)

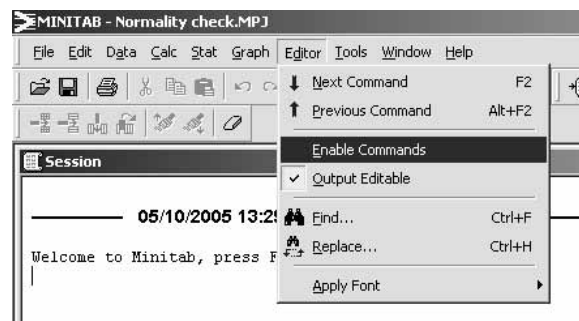
**BOX 3.2.** How to check if your data are normally distributed (parametric)

**Step 1.** Enter the data into the worksheet window of Minitab, giving the column(s) sensible headings.



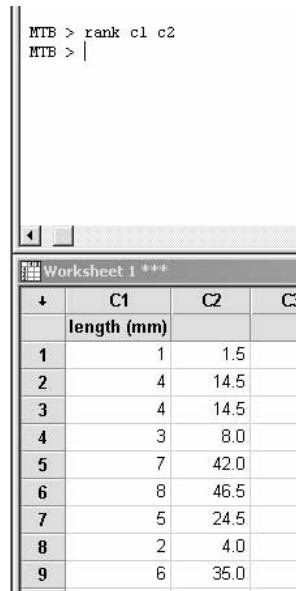
	C1	C2
	length (mm)	
1	1	
2	4	
3	4	
4	3	
5	7	
6	8	
7	5	
8	2	
9	6	

**Step 2.** Rank the data and calculate some descriptive statistics.  
To rank the data, make sure that commands are enabled by clicking in the session window and going to 'Editor', 'Enable Commands'.



At the 'MTB > prompt', type 'rank c1 c2'. This will generate ranks for each element in column 1, and place the ranks in column 2.

```
MTB > rank c1 c2
MTB > |
```

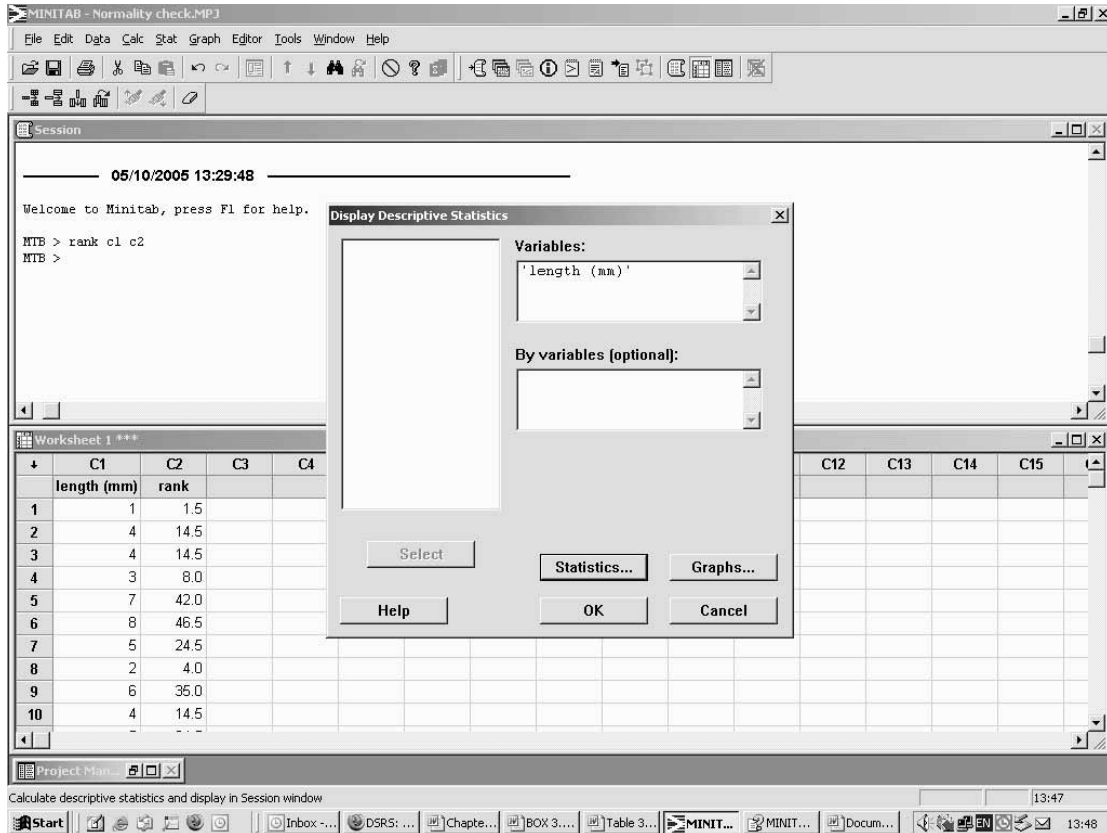


The screenshot shows the Minitab interface. At the top, the command window displays the command 'rank c1 c2' and a cursor. Below it, the worksheet 'Worksheet 1 \*\*\*' is visible, containing a table with three columns: C1, C2, and C3. The first row is a header for 'length (mm)'. The subsequent rows contain numerical data for each column.

	C1	C2	C3
	length (mm)		
1	1	1.5	
2	4	14.5	
3	4	14.5	
4	3	8.0	
5	7	42.0	
6	8	46.5	
7	5	24.5	
8	2	4.0	
9	6	35.0	

Give column 2 a name – 'rank' will do.

Generate some descriptive statistics – we will need the median, mean, and standard deviation. Go to ‘Stat’, ‘Basic Statistics’, ‘Display Descriptive Statistics’. Select column 1 by clicking on it to highlight it, then clicking on ‘Select’ to transfer it to the ‘Variables’ window.



The default options for 'Statistics' include what we want, so click on 'Graphs' and select 'Histogram of data'.

The screenshot shows the Minitab interface. The 'Session' window contains the following text:

```

05/10/2005 13:29:48
Welcome to Minitab, press F1 for help.
MTB > rank c1 c2
MTB >

```

The 'Worksheet 1' window shows the following data:

	C1	C2	C3	C4
	length (mm)	rank		
1	1	1.5		
2	4	14.5		
3	4	14.5		
4	3	8.0		
5	7	42.0		
6	8	46.5		
7	5	24.5		
8	2	4.0		
9	6	35.0		
10	4	14.5		

The 'Display Descriptive Statistics' dialog box is open, showing the following options:

- Variables: length (mm)
- Display Descriptive Statistics - Graphs:
  - Histogram of data
  - Histogram of data, with normal curve
  - Individual value plot
  - Boxplot of data

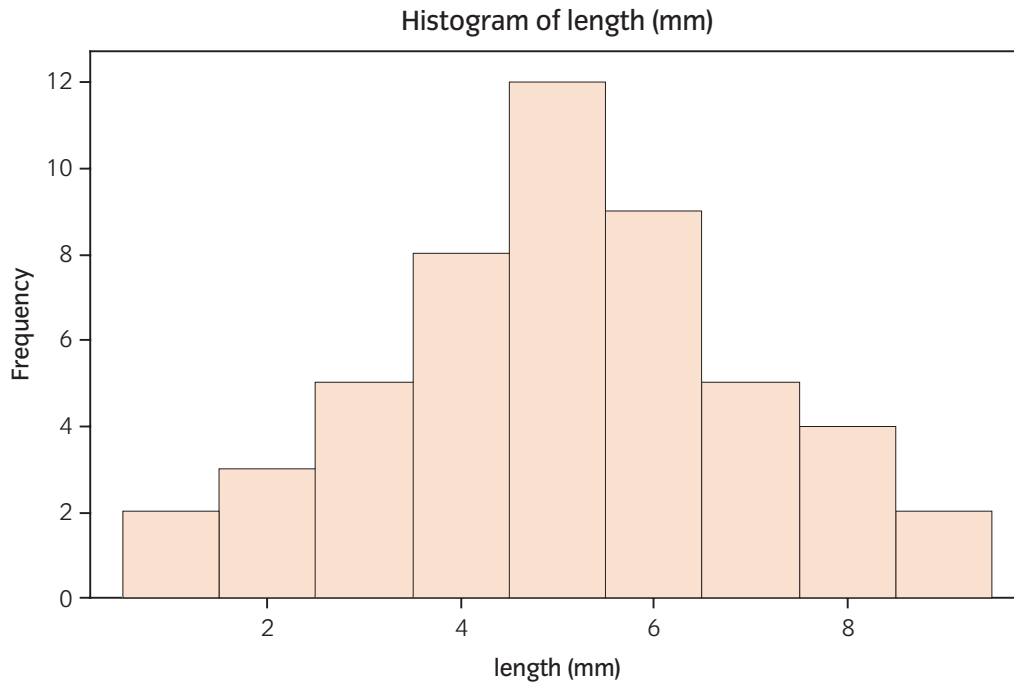
Click on 'OK', and on 'OK' again. The descriptive statistics appear in the 'Session' window, and the histogram appears in a new window.

### Descriptive Statistics: length (mm)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
length (mm)	50	0	5.080	0.274	1.936	1.000	4.000	5.000	6.000

Variable	Maximum
length (mm)	9.000

From the descriptive statistics, we note that the mean is 5.080 mm, the median is 5.000 mm, and the standard deviation is 1.936 mm.



The graph tells us that the mode is 5.0 mm.

**Step 3.** Go through the individual tests.

- Are the data on an interval scale? Yes – they are measured in mm.
- Does the distribution appear to be a ‘bell-shaped’ curve? Looking at the graph, the answer is ‘yes’.
- Do about 68% of your observations fall within the range  $\bar{x} \pm 1 s$ ?

The mean is 5.080 mm, and the standard deviation is 1.936 mm.

$$\bar{x} + 1 s = 5.080 + 1.936 = 7.016$$

$$\bar{x} - 1 s = 5.080 - 1.936 = 3.144$$

The measurements between these are 4, 5, 6, and 7 mm. From the ranking table in column 2, these correspond to ranks of 14.5, 24.5, 35.0, and 42.0. Looking at the table (or reading from the graph) there are 8, 12, 9, and 5 ladybirds in these categories, a total of 34 ladybirds. This is  $(34/50) \times 100 = 68\%$  of the total. Thus about 68% of our observations lie in the range  $\bar{x} \pm 1 s$ .

(d) Does median = mode = mean?

Median = 5.0 mm, mode = 5.0 mm (from the graph), and mean = 5.080 mm. Thus they are all close together.

(e) For the goodness of fit chi-squared test, see Box 5.2.

From the information we can obtain at this stage, it looks as if the data are normally distributed.