

How to construct an accurate histogram in Excel

This procedure uses a scatter diagram to build a 'dot-to-dot' histogram.

1. Create a frequency distribution of the data.
Use separate cells for the lower and upper values as well as for the symbol implying from/to.

Classes			Frequency
20	to	39	1
40	to	49	5
50	to	59	7
60	to	69	6
70	to	84	4

2. Add a column for frequency density.

For each cell use the formula:

$$\text{Frequency density} = \frac{\text{Frequency}}{\text{Upperclasslimit} - \text{lowerclasslimit} + 1}$$

The '+ 1' is needed in this example because the lower class boundary for each group is actually a half unit below the limit shown in the table and the upper class boundary is half a unit above that shown.

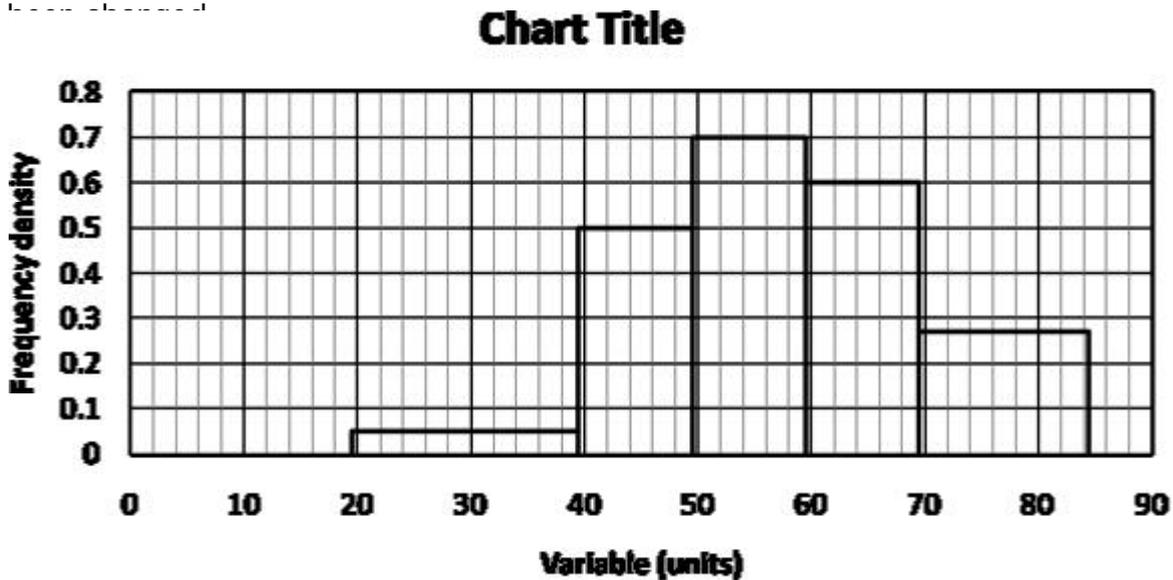
Classes			Frequency	Frequency density
20	to	39	1	0.05
40	to	49	5	0.50
50	to	59	7	0.70
60	to	69	6	0.60
70	to	84	4	0.27

3. To construct the histogram we need points starting at (19.5,0) to (19.5, 0.05) to (39.5, 0.05) to (39.5, 0) etc
4. Put these in a list of paired values on the spreadsheet next to the frequency distribution (as shown here).
5. Now highlight all these paired values and click on Insert, Scatter and the last option (Scatter with straight lines),
6. Click on Chart Layout 1 to add a title and axis labels. Right clicking on the title or labels allows you to change them in a variety of ways. Ensure that the font is the one you want and the size is at least 12 - when you paste it into WORD, you want to be able to see it!
7. Delete the legend box (key).
8. Right clicking on a part of the chart usually allows you to change it. For example, right click on the horizontal axis to add vertical gridlines. You can also use Format Axis to change the scale.

19.5	0
19.5	0.05
39.5	0.05
39.5	0
39.5	0.50
49.5	0.50
49.5	0
49.5	0.70
59.5	0.70
59.5	0
59.5	0.60
69.5	0.60
69.5	0
69.5	0.27
84.5	0.27
84.5	0



In the histogram below the colour and thickness of the axes gridlines have also



9. You can produce a frequency polygon by making a few changes to the histogram. First you need to decide values for the x-coordinates of the end points. When the class widths are equal the empty classes at each end are usually assumed to be the same width as all of the others. So the end points (i.e. the x-coordinates of the ends of the polygon) are $\frac{1}{2}$ of a class width from the upper and lower boundaries of the histogram. However in this example the classes are unequal. One method of dealing with this is to double the class width of the end classes. In this example this means that the empty classes have widths of 40 and 30 respectively and the midpoints are 20 before the lowest lower class boundary and 15 above the highest upper class boundary. The 'dot-to-dot' list therefore looks like this:

-0.5	0
29.5	0.05
44.5	0.50
54.5	0.70
64.5	0.60
77	0.27
99.5	0

10. After changing the data points and the scale on the horizontal axis, the frequency polygon looks like this:

