

How to construct an accurate histogram in Excel

This procedure uses the scatter diagram with lines between points option to build a 'dot-to-dot' histogram.

1. Create a frequency distribution of the data, using separate cells for the lower and upper limits of the cell 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, using the formula:

$$\text{Frequency density} = \frac{\text{Frequency}}{(\text{Upper class limit} - \text{lower class limit} + 1)}$$
 for each cell.
 The '+1' is there to include the half unit above and below the class limits that make up the class boundaries.

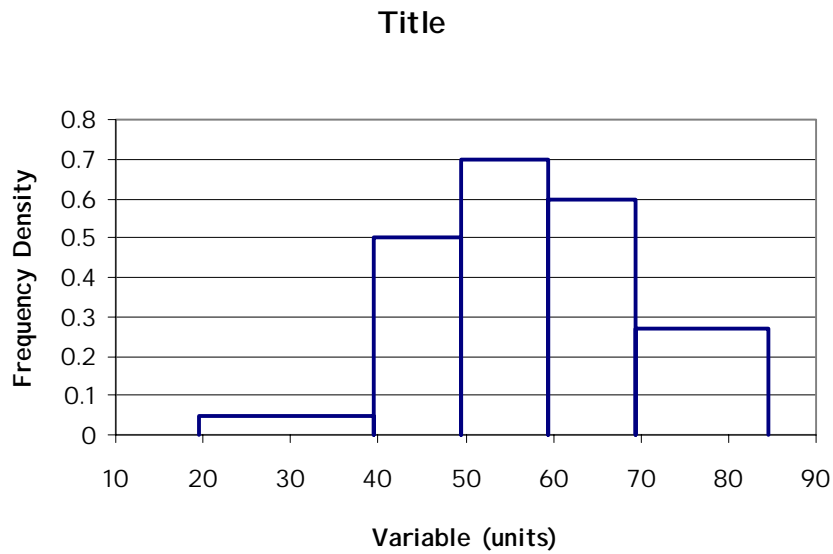
Classes			Frequency	Frequency density
20	to	39	1	0.05
40	to	49	5	0.50
50	to	59	7	0.70
60	to	69	6	0.60
70	to	84	4	0.27

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 adjacent to the frequency distribution like this:
5. Now highlight all the cells and click on Chart Wizard. Go for xy(scatter) bottom left option, then NEXT and NEXT. Put in chart title and x-axis and y-axis labels. Uncheck legend box. Click NEXT and choose either option (as new sheet or object within the sheet).
6. Double click on Plot Area and under Patterns Area choose None
7. Double click over Series point. On Patterns, thicken the line and in markers check NONE.
8. Double click over the title, axes and labels in turn and ensure that the font is the one you want and the size is at least 12 if not 14 - when you paste it into WORD, you want to be able to see it! Also in the value axis, make sure only the relevant part of the scale is displayed.

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



The histogram looks like this:



9. If you want a frequency polygon, a few adjustments can produce one. If the class widths are unequal, then an assessment needs to be made for the x-coordinates of the end points. One method is to double the class width of the end classes. Hence the midpoint (i.e. the x-coordinate) is one class width backwards/ forwards. If however the class widths are equal, then the empty classes at the end are assumed to be the same as all the others and thus the end points are $\frac{1}{2}$ a class width. In this case the classes are unequal and so the zero point classes have widths of 40 and 30 respectively and the midpoints are thus 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. And the frequency polygon looks like this, after adjustments to the scale, start and finish points, line thickness and background have been made:

