

Installing and Using the Histogram Function in Excel

First, you must install the Data Analysis ToolPak in Excel.

PC: Open Excel
 Select the File menu, then Options / Add-Ins
 Click Go...
 Check Analysis ToolPak then press OK

Mac: *Note: this does not work on Office 2011 and earlier*
 Open Excel
 Select the Tools menu
 Scroll to the bottom and select Add Ins
 Check Analysis ToolPak then press OK

Now, we need some data. Type the following into Excel:

-50 -50 15 20 21 25 30 31 32 33

We must then define our bins. Let's go with the following for now. We will interpret this in a moment.

0 10 20 30

Select the DATA menu / Data Analysis / Histogram

For your Input Range, select your data. For you Bin Range, select your bin values. (If you leave the Bin Range blank, the program will generate its own Bin Range.) Leave the rest of the fields blank. Click OK and your histogram will appear on a new sheet.

<i>Bin</i>	<i>Frequency</i>
0	2
10	0
20	2
30	3
More	3

Note how the ranges of the Bins are defined

$$-\infty < \text{Bin1} \leq 0 \rightarrow 2$$

$$0 < \text{Bin2} \leq 10 \rightarrow 0$$

$$10 < \text{Bin3} \leq 20 \rightarrow 2$$

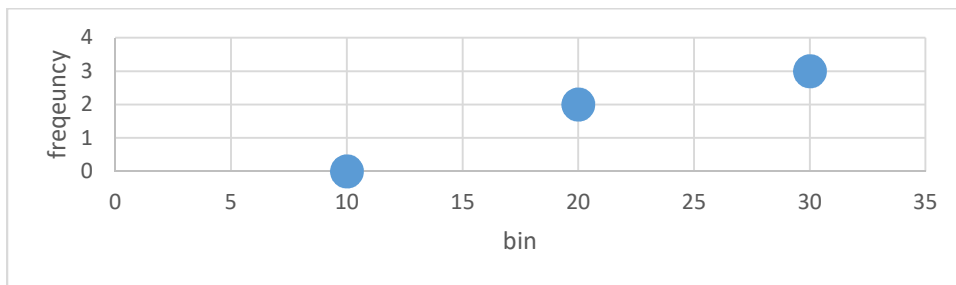
$$20 < \text{Bin4} \leq 30 \rightarrow 3$$

3 data points did not fit into any bin

Hence, Bin2 does not include 0 but does include 10.

We have two problems here. The first bin includes all the values up to and including 0. This is a ridiculously huge bin and is best ignored.

To see the second problem, consider the following plot of the histogram.



Since each bin is listed by its highest value, the histogram has been shifted to the right by half a bin width. It would be better to define each bin by its middle value: in this case, 5, 15, and 25. This would produce a more meaningful histogram.

