Using the Histogram tool in the Data Analysis Toolpak

First, install the Data Analysis Tookpak in Excel. Instructions via QR code.

After installing the Toolpak in Excel, the histogram function is found by selecting . . .

Data/Data Analysis

You first must define your data. You can either select the specific cells of the data or the entire column.

You may then define your bins. If you do not define your bins, they will be automatically generated. This may or may not work well. You can define your own bins with a column of numbers. Consider the example below:

data	bins			
-55	20			
-10	30			
-5	40			
2	50			
10				
15				
20				
25				
30				
50				
55				
100				
110				

We can now apply the histogram tool. Select . . .

DATA/Data Analysis/

Choose **Histogram** from the list of **Analysis Tools**.

Select your Input and Bin ranges, and where you want the output to appear on the spreadsheet. Then click **OK**. You should see the following:

| data	bins		Bin	Frequency
-55	20		20	7
-10	30		30	2
-5	40		40	0
2	50		50	1
10			More	3
15				
20				
25				
30				
50				
55				
100				
110				

The first bin is the number of values less than and including 20: 7

The second bin is the number of values from 20 (exclusive) to 30 (inclusive): 2

The third bin is the number of values from 30 (exclusive) to 40 (inclusive): 0

The fourth bin is the number of values from 40 (exclusive) to 50 (inclusive): 1

"More" is the number of values that did not fit into any bin: 3
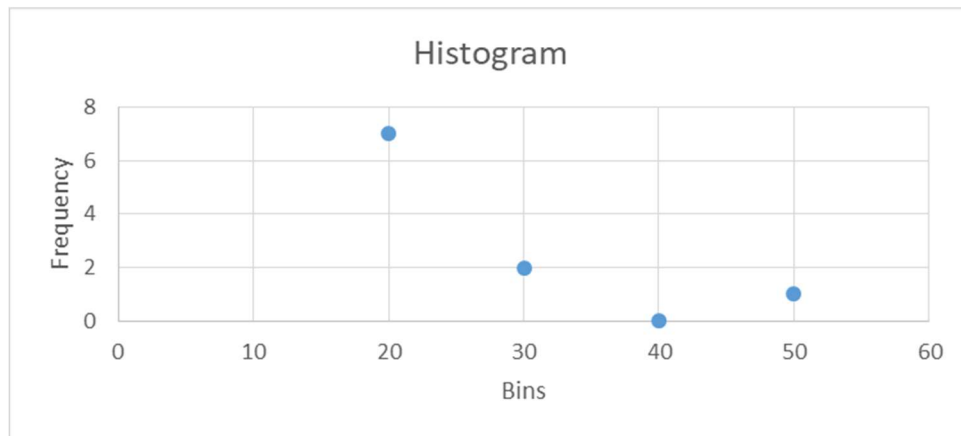
We have two problems here. The first bin includes all the values up to and including 0. This is a ridiculously huge bin and is best ignored.

To see the second problem, look at the histogram plot. Since each bin is listed by its highest value, this histogram has been shifted to the right by half a bid width. It would be better to define

each bin by its middle value: in this case, 25, 35, and 45. This will produce a more meaningful histogram.