

A Histogram is a graphical representation of the distribution of numerical data. It is widely used in Statistical Analysis. Karl Pearson introduced histogram.

In Excel, you can create a Histogram from the Analysis ToolPak that comes as an add-in with Excel. However, in such a case, when the data is updated, Histogram will not reflect the changed data unless it is modified through Analysis ToolPak again.

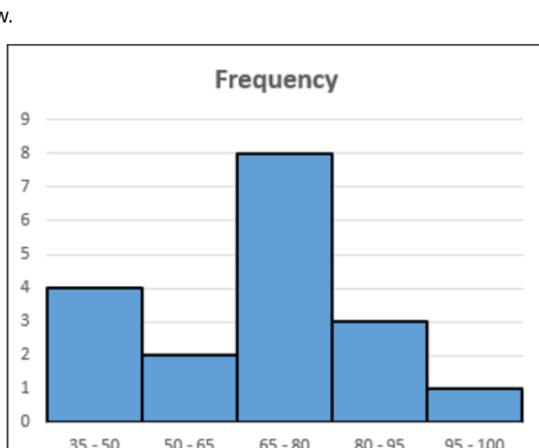
In this chapter, you will learn how to create a Histogram from a Column chart. In this case, when the source data is updated the chart also gets refreshed.

## What is a Histogram?

A Histogram is represented by rectangles with lengths corresponding to the number of occurrences of a variable in successive numerical intervals. The numerical intervals are called bins and the number of occurrences is called frequency.

The bins are usually specified as consecutive, non-overlapping intervals of the variable. The bins must be adjacent and are of equal size. A rectangle over a bin with height proportional to the frequency of the bin depicts the number of cases in that bin. Thus, the horizontal axis represents the bins whereas the vertical axis represents the frequency. The rectangles are colored or shaded.

A Histogram will be as shown below.



## Advantages of Histograms

Histogram is used to inspect the data for its underlying distribution, outliers, skewness, etc. For example, Histogram can be used in statistical analysis in the following scenarios –

- A census of a country to obtain the people of various age groups.
- A survey focused on the demography of a country to obtain the literacy levels.
- A study on the effect of tropical diseases during a season across different regions in a state.

## Preparation of Data

Consider the data given below.

	A	B
1		
2		Marks %
3		48
4		69
5		76
6		79
7		68
8		70
9		71
10		93
11		49
12		52
13		72
14		87
15		93
16		48
17		96
18		61
19		51
20		69

Create bins and calculate the number of values in each bin from the above data as shown below –

	A	B	C	D	E
1					
2		Marks %	Bins	No. of Values	
3		48	35 - 50	=COUNTIFS(\$B\$3:\$B\$20,">=35",\$B\$3:\$B\$20,"<50")	
4		69	50 - 65	=COUNTIFS(\$B\$3:\$B\$20,">=50",\$B\$3:\$B\$20,"<65")	
5		76	65 - 80	=COUNTIFS(\$B\$3:\$B\$20,">=65",\$B\$3:\$B\$20,"<80")	
6		79	80 - 95	=COUNTIFS(\$B\$3:\$B\$20,">=80",\$B\$3:\$B\$20,"<95")	
7		68	95 - 100	=COUNTIFS(\$B\$3:\$B\$20,">=95")	
8		70			
9		71			
10		93			
11		49			
12		52			
13		72			
14		87			
15		93			
16		48			
17		96			
18		61			
19		51			
20		69			

The number of values in a bin is referred to as the frequency of that bin.

Bins	Frequency
35 - 50	4
50 - 65	2
65 - 80	8
80 - 95	3
95 - 100	1

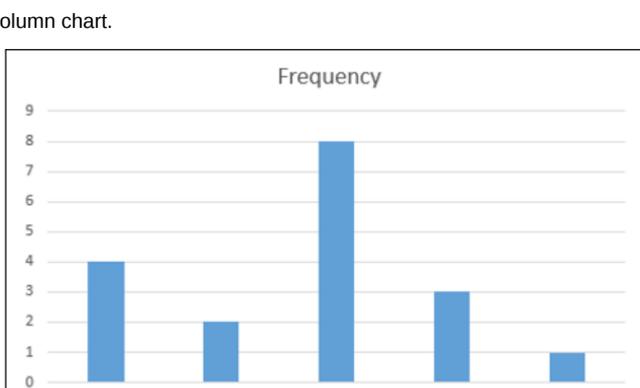
This table is called a Frequency table and we will use it to create the Histogram.

## Creating a Histogram

Following are the steps to create a Histogram.

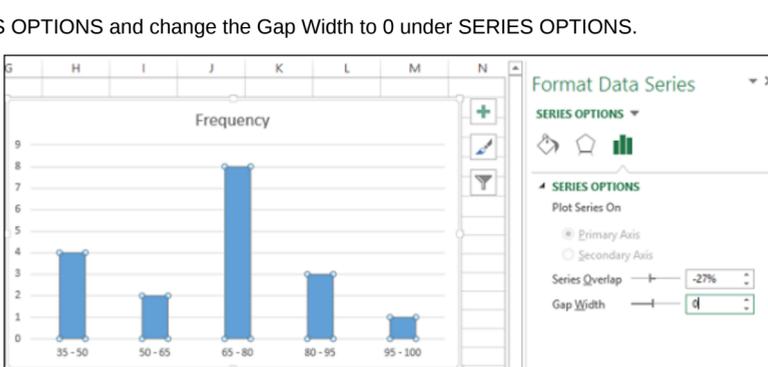
**Step 1** – Select the data in the Frequency table.

**Step 2** – Insert a Clustered Column chart.



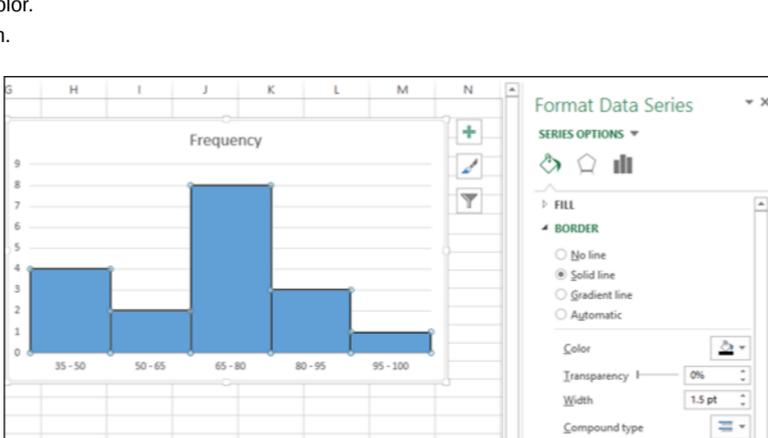
**Step 3** – Right click on the Columns and select Format Data Series from the dropdown list.

**Step 4** – Click SERIES OPTIONS and change the Gap Width to 0 under SERIES OPTIONS.

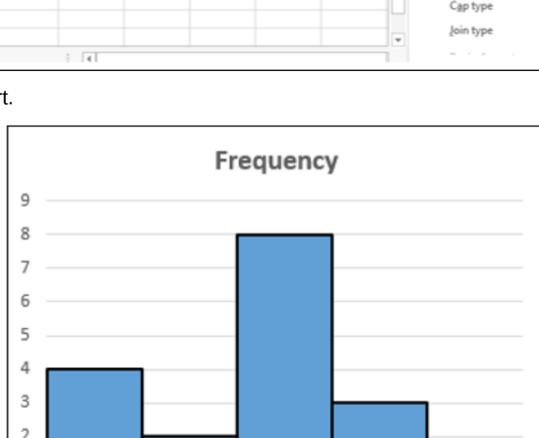


**Step 5** – Format the chart as follows.

- Click on Fill & Line.
- Click on Solid Line under Border.
- Select black for color.
- Type 1.5 for Width.



**Step 6** – Adjust the size of the chart.



Your Histogram is ready. As you can observe, the length of each column corresponds to the frequency of that particular bin.