WHAT IS A HISTOGRAM?

A frequency distribution shows how often each different value in a set of data occurs. A histogram is the most commonly used graph to show frequency distributions. It looks very much like a bar chart, but there are important differences between them. This helpful [data collection and analysis tool](https://asq.org/quality-resources/data-collection-analysis-tools) is considered one of the [seven basic quality tools](https://asq.org/quality-resources/seven-basic-quality-tools).

WHEN TO USE A HISTOGRAM

Use a histogram when:

* The data are numerical
* You want to see the shape of the data’s distribution, especially when determining whether the output of a process is distributed approximately normally
* Analyzing whether a process can meet the customer’s requirements
* Analyzing what the output from a supplier’s process looks like
* Seeing whether a process change has occurred from one time period to another
* Determining whether the outputs of two or more processes are different
* You wish to communicate the distribution of data quickly and easily to others



**Histogram Example**

HOW TO CREATE A HISTOGRAM

1. Collect at least 50 consecutive data points from a process.
2. Use the [histogram worksheet](https://asq.org/quality-resources/histogram#Worksheet) to set up the histogram. It will help you determine the number of bars, the range of numbers that go into each bar, and the labels for the bar edges. After calculating *W*in Step 2 of the worksheet, use your judgment to adjust it to a convenient number. For example, you might decide to round 0.9 to an even 1.0. The value for *W*must not have more decimal places than the numbers you will be graphing.
3. Draw x- and y-axes on graph paper. Mark and label the y-axis for counting data values. Mark and label the x-axis with the *L*values from the worksheet. The spaces between these numbers will be the bars of the histogram. Do not allow for spaces between bars.
4. For each data point, mark off one count above the appropriate bar with an X or by shading that portion of the bar.

HISTOGRAM ANALYSIS

* Before drawing any conclusions from your histogram, be sure that the process was operating normally during the time period being studied. If any unusual events affected the process during the time period of the histogram, your analysis of the histogram shape likely cannot be generalized to all time periods.
* Analyze the meaning of your histogram's shape. Typical histogram [shapes and what they mean](https://asq.org/quality-resources/histogram#Shapes) are covered below.

TYPICAL HISTOGRAM SHAPES AND WHAT THEY MEAN

**Normal Distribution**

A common pattern is the bell-shaped curve known as the "normal distribution." In a normal or "typical" distribution, points are as likely to occur on one side of the average as on the other. Note that other distributions look similar to the normal distribution. Statistical calculations must be used to prove a normal distribution.

It's important to note that "normal" refers to the typical distribution for a particular process. For example, many processes have a natural limit on one side and will produce skewed distributions. This is normal—meaning typical—for those processes, even if the distribution isn’t considered "normal."



**Skewed Distribution**

The skewed distribution is asymmetrical because a natural limit prevents outcomes on one side. The distribution’s peak is off center toward the limit and a tail stretches away from it. For example, a distribution of analyses of a very pure product would be skewed, because the product cannot be more than 100 percent pure. Other examples of natural limits are holes that cannot be smaller than the diameter of the drill bit or call-handling times that cannot be less than zero. These distributions are called right- or left-skewed according to the direction of the tail.



**Double-Peaked or Bimodal**

The bimodal distribution looks like the back of a two-humped camel. The outcomes of two processes with different distributions are combined in one set of data. For example, a distribution of production data from a two-shift operation might be bimodal, if each shift produces a different distribution of results. Stratification often reveals this problem.



**Plateau or Multimodal Distribution**

The plateau might be called a “multimodal distribution.” Several processes with normal distributions are combined. Because there are many peaks close together, the top of the distribution resembles a plateau.

