

# What are the methods for identifying skewness in box plots?

Authored by  
**stats writer**

April 29, 2024

## RECOMMENDED CITATION

stats writer (2024). *What are the methods for identifying skewness in box plots?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=141129>

Box plots are a commonly used graphical representation for summarizing the distribution of numerical data. Skewness, or the lack of symmetry in the data, is an important aspect to consider when interpreting box plots. In order to accurately identify skewness in a box plot, there are several methods that can be used. These include visually inspecting the shape of the box plot, calculating the skewness coefficient, and comparing the position of the median and the mean. Additionally, the use of complementary statistical measures such as quartiles and interquartile range can also aid in detecting skewness. These methods can provide valuable insights into the distribution of the data, allowing for a more comprehensive understanding of the underlying patterns and trends.

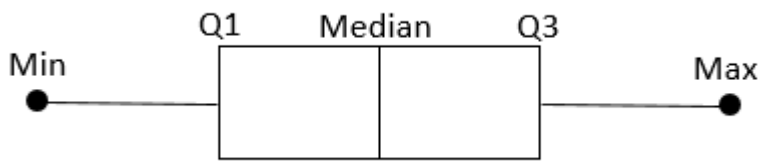
## Identify Skewness in Box Plots

**A box plot is a type of plot that displays the five number summary of a dataset, which includes:**

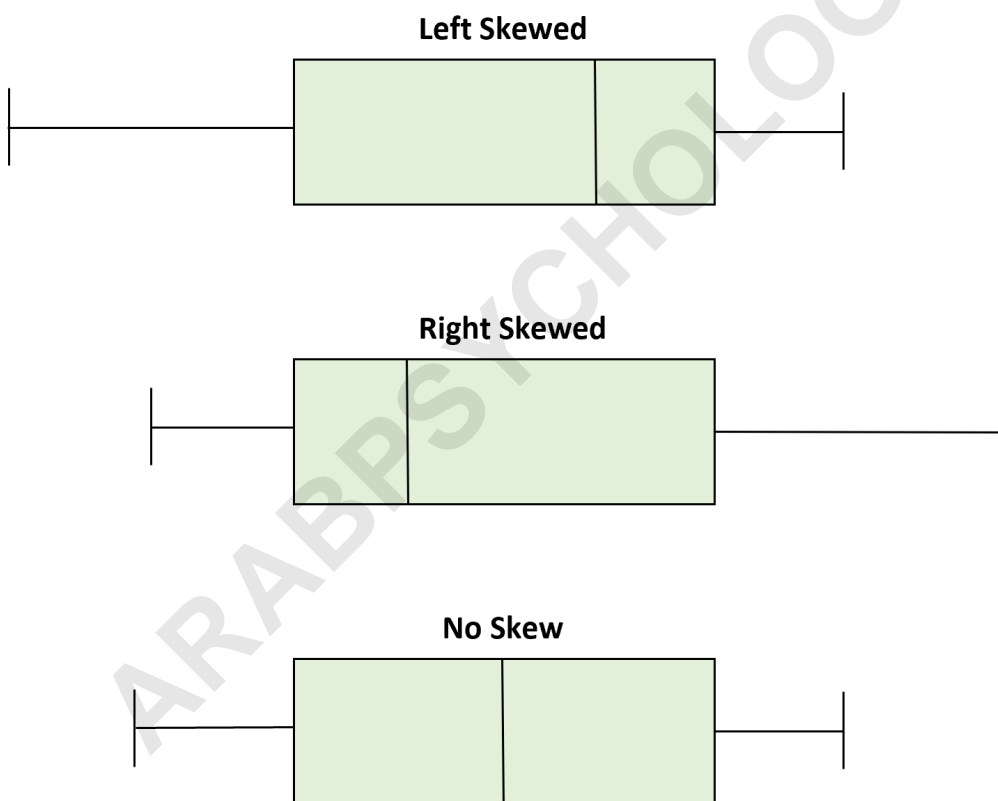
**The minimum value  
The first quartile (the 25th percentile)  
The median value  
The third quartile (the 75th percentile)  
The maximum value**

**We use the following process to draw a box plot:**

**Draw a box from the first quartile (Q1) to the third quartile (Q3)  
Then draw a line inside the box at the median  
Then draw "whiskers" from the quartiles to the minimum and maximum values**



**We can determine whether or not a distribution is skewed based on the location of the median value in the box plot.**



**When the median is closer to the bottom of the box and the whisker is shorter on the lower end of the box, the distribution is right-skewed (or "positively" skewed).**

**When the median is closer to the top of the box and the whisker is shorter on the upper end of the box, the distribution is left-skewed (or "negatively" skewed).**

**When the median is in the middle of the box and the whiskers are roughly equal on each side, the distribution is symmetrical (or "no" skew).**

**The following examples illustrate how to use box plots to determine if a distribution is right-skewed, left-skewed, or has no skew.**

#### **Example 1: Right-Skewed Distribution**

**The distribution of annual household incomes in the United States is right-skewed. Most households earn between \$40k and \$80k per year but there's a long right tail on the distribution that represents households that earn much more.**

**If we created a box plot to visualize the distribution of household incomes, it would look something like this:**

### Distribution of Household Incomes



Notice that the vertical line inside the box that represents the median is much closer to the first quartile than the third quartile, which means the distribution is right-skewed.

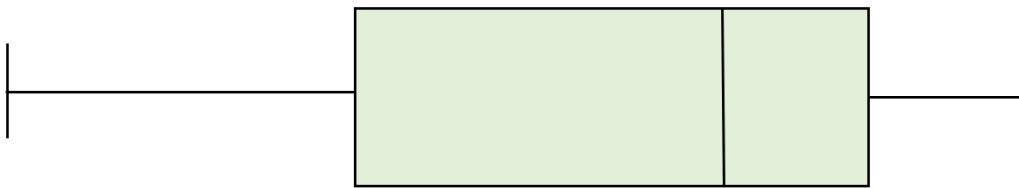
### Distribution of Household Incomes



### Example 2: Left-Skewed Distribution

The distribution of the age of deaths in most populations is left-skewed. Most people live to be between 70 and 80 years old, with fewer and fewer living less than this age.

If we created a box plot to visualize the distribution of the age of deaths, it would look something like this:

**Distribution of Age of Death**

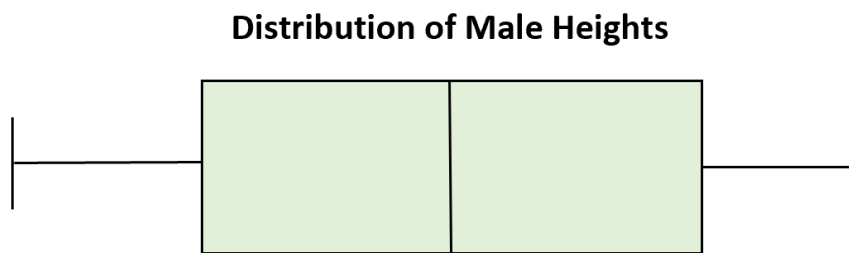
Notice that the vertical line inside the box that represents the median is much closer to the third quartile than the first quartile, which means the distribution is left-skewed.

**Distribution of Age of Death**

### Example 3: Symmetrical Distribution

The distribution of the height of males is roughly and has no skew. For example, the average height of a male in the United States is roughly 69.1 inches. The distribution of heights is roughly symmetrical, with some being shorter and some being taller.

If we created a box plot to visualize the distribution of the height of males in the United States, it would look something like this:



Notice that the vertical line inside the box that represents the median is equally close to the first quartile and the third quartile, which means the distribution is symmetrical and has no skew.

