

# When Should You Use a Box Plot?

Authored by  
**stats writer**

May 12, 2024

## RECOMMENDED CITATION

stats writer (2024). *When Should You Use a Box Plot?*. PSYCHOLOGICAL SCALES.  
Retrieved from <https://scales.arabpsychology.com/?p=143997>

A box plot is a graphical representation of numerical data that displays the distribution, variability, and median values of a dataset. It consists of a box with whiskers extending from both ends, which represents the range of the data. The use of a box plot is appropriate in various scenarios, including when comparing data from multiple groups, analyzing data with outliers, and visualizing the distribution of a continuous variable. In these situations, a box plot provides a concise and informative summary of the data, making it a useful tool for data analysis and decision making. By utilizing a box plot, you can easily identify patterns, outliers, and differences between groups, making it a valuable addition to any statistical analysis.

## **When Should You Use a Box Plot? (3 Scenarios)**

**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)**

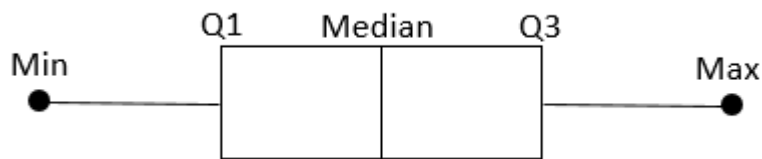
**The maximum value**

**We use three simple steps to create a box plot for any dataset:**

**1. Draw a box from the first to the third quartile**

**2. Draw a vertical line at the median**

**3. Draw "whiskers" from the quartiles to the minimum and maximum value**



**We typically create box plots in one of three scenarios:**

**Scenario 1: To visualize the distribution of values in a dataset.**

A box plot allows us to quickly visualize the distribution of values in a dataset and see where the five number summary values are located.

**Scenario 2: To compare two or more distributions.**

Side-by-side box plots allow us to visualize the differences between two or more distributions and compare the median values and the spread of values between distributions.

**Scenario 3: To identify outliers.**

In box plots, outliers are typically represented by tiny circles that extend beyond either whisker. An observation is defined to be an outlier if it meets one of

**the following criteria:**

**An observation is less than  $Q1 - 1.5 \times (\text{Interquartile range})$**

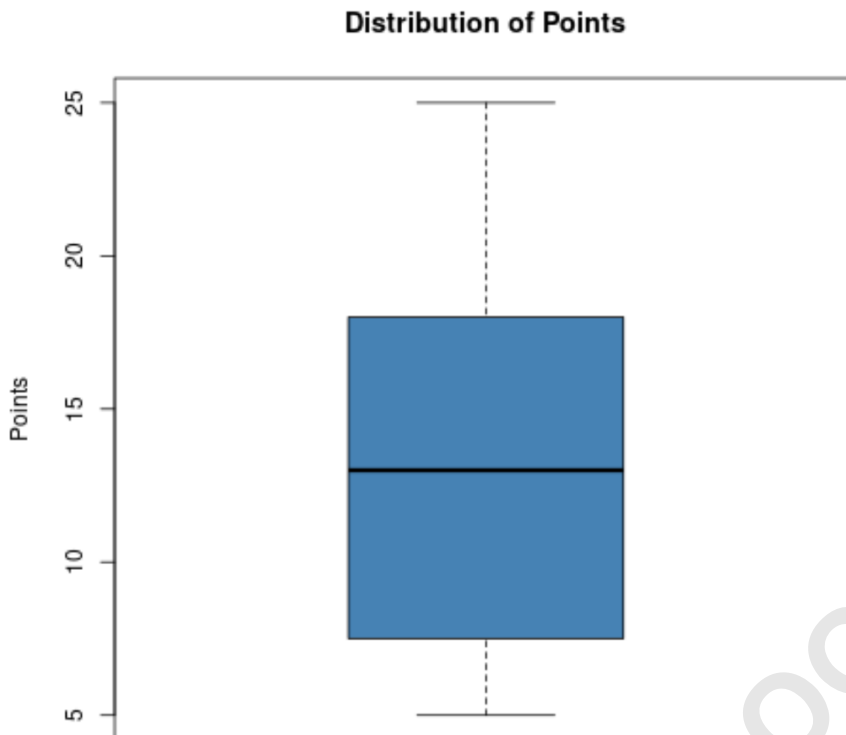
**An observation is greater than  $Q3 + 1.5 \times (\text{Interquartile range})$**

**By creating a box plot, we can quickly see whether or not a distribution has any outliers.**

**The following examples show how we would use a box plot in each scenario.**

**Scenario 1: Visualize the Distribution of Values in a Dataset**

**Suppose a basketball coach wants to visualize the distribution of points scored by players on his team so he creates the following box plot:**



Based on this box plot, he can quickly see the following values:

**Minimum: 5**

**Q1 (First Quartile): About 8**

**Median: About 13**

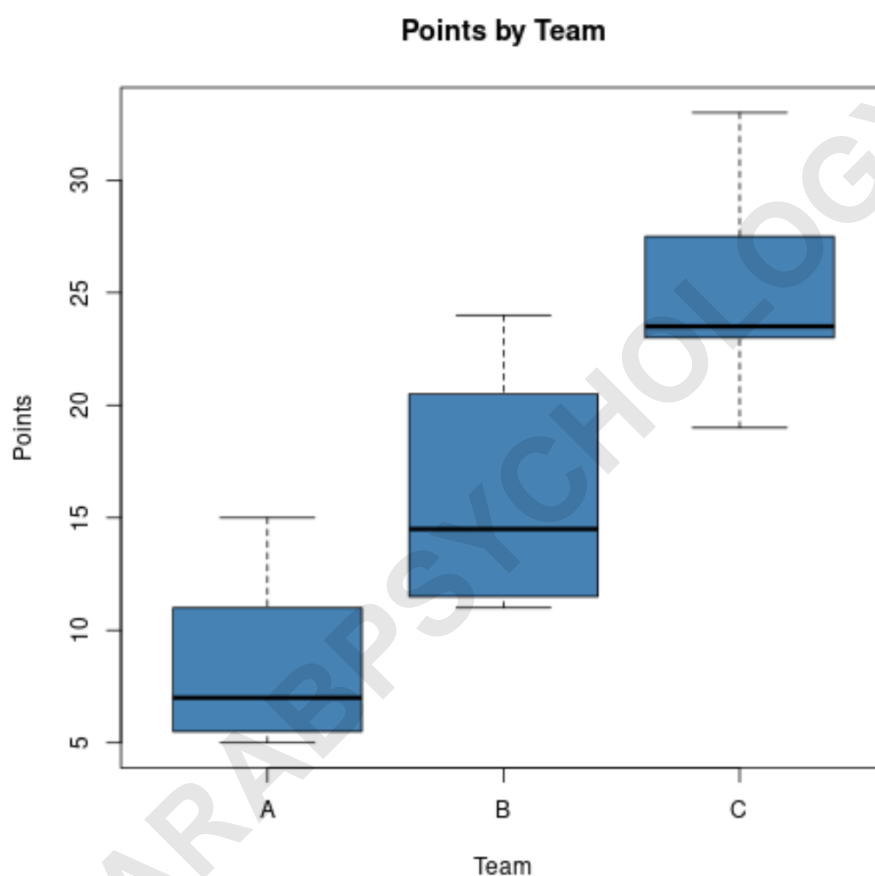
**Q3 (Third Quartile): About 18**

**Maximum: 25**

This allows the coach to quickly see that the points scored by players ranges from 5 to 25, the median points scored is about 13, and 50% of his players score between about 8 and 18 points per game.

## Scenario 2: Compare Two or More Distributions

Suppose a sports analyst wants to compare the distribution of points scored by basketball players on three different teams so he creates the following box plots:



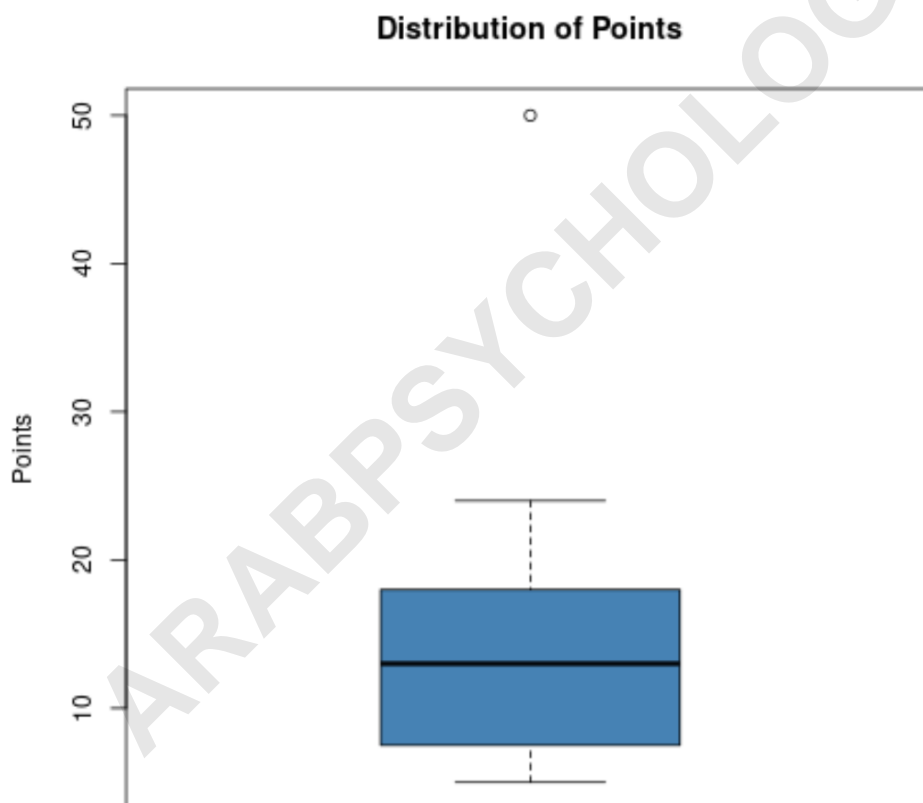
Using these plots, he can quickly see that Team C has the highest median points scored and Team A has the lowest median points scored.

He can also quickly see that Team B has the highest

**spread of points scored since the box plot for Team B has the longest box.**

### Scenario 3: Identify Outliers

**Suppose a basketball coach wants to know if any of his players are outliers in terms of points scored. He decides to create the following box plot to visualize the distribution of points scored by his players:**



**Using this plot, the coach can see that the tiny dot at the top of the plot indicates an outlier.**

**Specifically, one of the players scored about 50 points which is considered an outlier compared to all of the other points scored.**

**The following tutorials offer in-depth explanations of how to use box plots in practice:**

**The following tutorials explain how to create box plots in different statistical software:**

ARABPSYCHOLOGY.COM