

# How can we plot the distribution of column values in R?

Authored by  
**stats writer**

June 25, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can we plot the distribution of column values in R?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=152021>

In order to plot the distribution of column values in R, we can utilize the built-in functions and packages available in the software. First, we need to import the data set containing the desired column values into R. Then, we can use functions such as "hist()" or "density()" to create a histogram or density plot respectively. These functions allow us to visualize the frequency and spread of the column values. Additionally, we can use the "boxplot()" function to create a box plot, which displays the median, quartiles, and outliers of the column values. By utilizing these functions and adjusting parameters such as bin size or number of bins, we can effectively plot the distribution of column values in R.

## Plot Distribution of Column Values in R

You can use the following methods to plot a distribution of column values in R:

### Method 1: Plot Distribution of Values Using Density Plot

```
plot(density(df$my_column))
```

### Method 2: Plot Distribution of Values Using Histogram

```
hist(df$my_column)
```

The following examples show how to use each method in practice with the following data frame

```
#create data frame
```

```
df = data.frame(team=rep(c('A', 'B'), each=10),  
points=c(3, 3, 4, 5, 4, 7, 7, 7, 10, 11, 8,
```

7, 8, 9, 12, 12, 12, 14, 15, 17))

**#view data frame**

**df**

**team points**

**1 A 3**

**2 A 3**

**3 A 4**

**4 A 5**

**5 A 4**

**6 A 7**

**7 A 7**

**8 A 7**

**9 A 10**

**10 A 11**

**11 B 8**

**12 B 7**

**13 B 8**

**14 B 9**

**15 B 12**

**16 B 12**

**17 B 12**

**18 B 14**

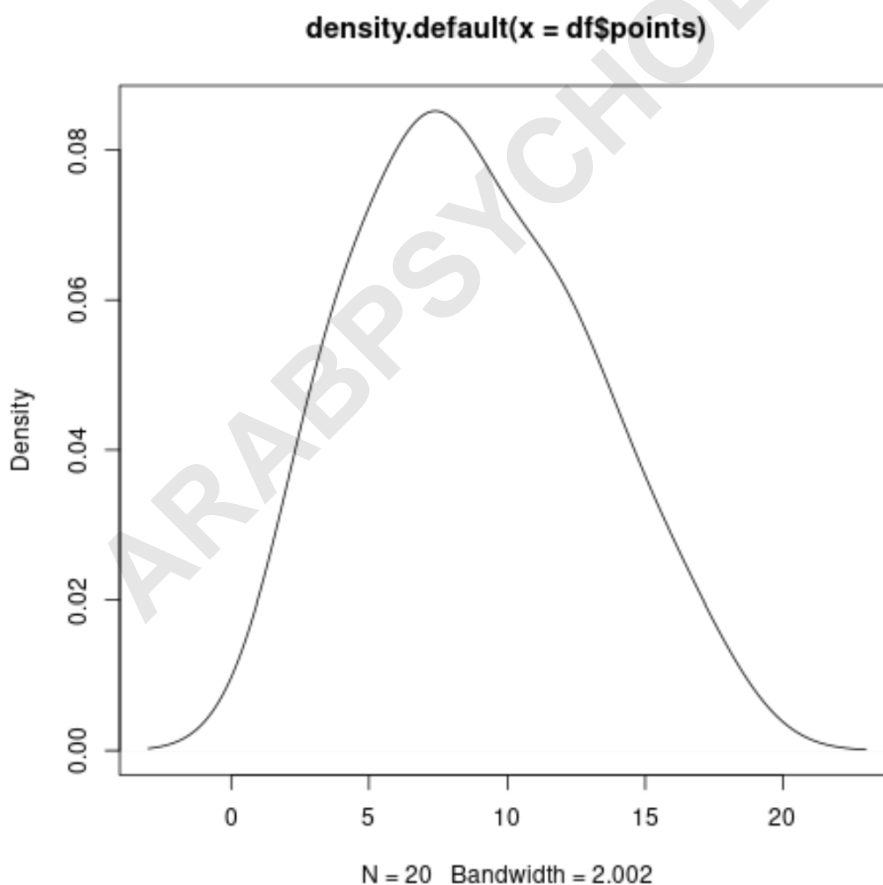
**19 B 15**

**20 B 17**

### Example 1: Plot Distribution of Values Using Density Plot

The following code shows how to plot the distribution of values in the points column using a :

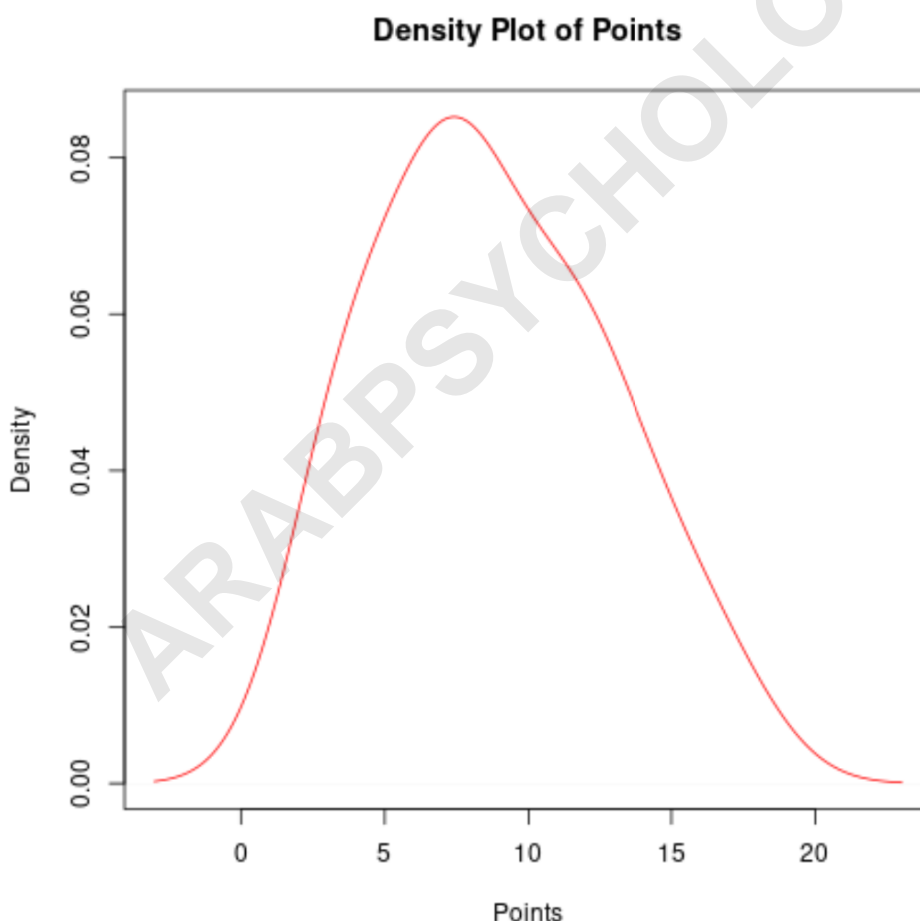
```
#plot distribution of values in points column  
plot(density(df$points))
```



This syntax produces a smooth curve that summarizes the distribution of values for a variable.

Note that we can also modify the title, axis labels, and color of the line in the density plot if we'd like:

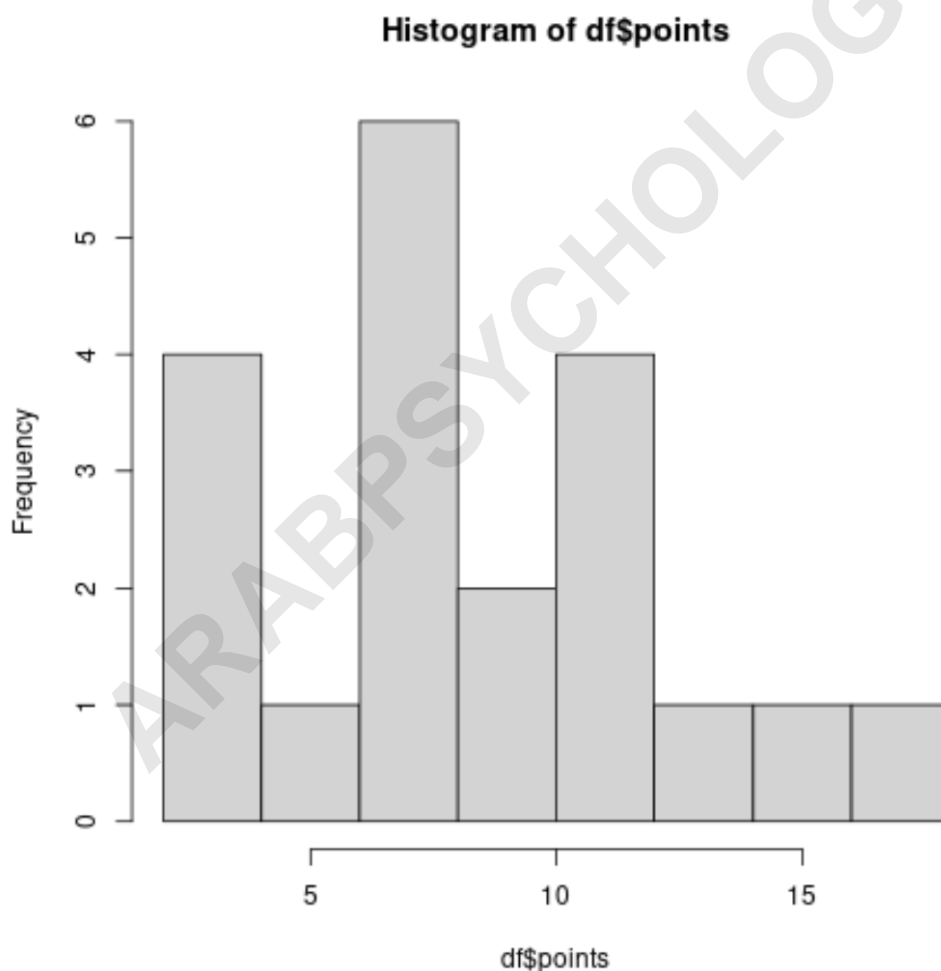
```
#plot distribution of values in points column  
plot(density(df$points), col='red', main='Density Plot of  
Points', xlab='Points')
```



## Example 2: Plot Distribution of Values Using Histogram

The following code shows how to plot the distribution of values in the points column using a histogram:

```
#plot distribution of values in points column using histogram  
hist(df$points)
```



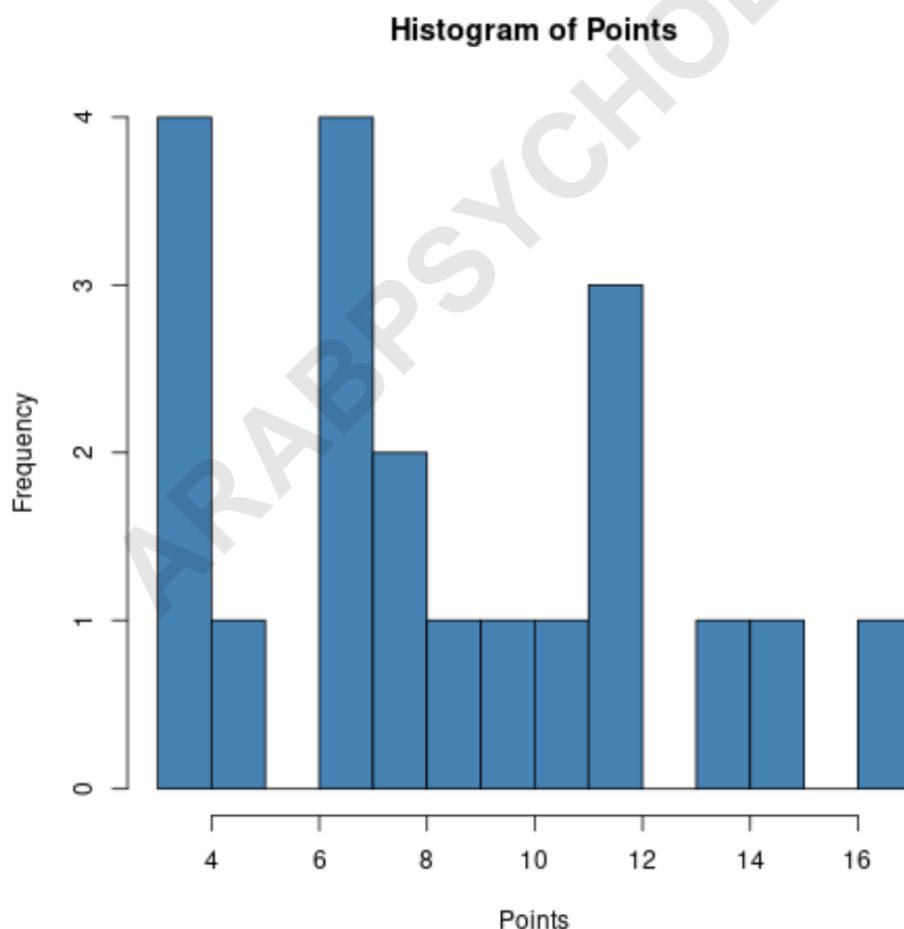
A histogram uses bars to represent frequencies of

values in the points column as opposed to a smooth line that summarizes the shape of the distribution.

Note that we can also modify the title, axis labels, color, and number of breaks used in the histogram:

**#plot distribution of values in points column using histogram**

```
hist(df$points, main='Histogram of Points',  
xlab='Points', col='steelblue', breaks=12)
```



**Note: The larger the value you choose for the breaks argument, the more bars there will be in the histogram.**

ARABPSYCHOLOGY.COM