

# How does the function `stat_bin()` handle the warning of using 30 bins?

Authored by  
**stats writer**

June 25, 2024

## RECOMMENDED CITATION

stats writer (2024). *How does the function `stat_bin()` handle the warning of using 30 bins?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=152873>

The function `stat_bin()` is a statistical tool used for creating histograms in data analysis. It is designed to automatically determine the number of bins needed to accurately represent the data. However, when the number of bins exceeds 30, a warning is generated. This is because using too many bins can result in overfitting and misleading representations of the data. To address this warning, `stat_bin()` automatically limits the number of bins to 30, ensuring that the histogram remains informative and accurate. This approach allows for a helpful visual representation of the data without compromising its integrity.

## Handle R Warning: `stat_bin()` using `bins = 30`

One common warning you may encounter in R is:

``stat_bin()` using `bins = 30`. Pick better value with `binwidth`.`

This warning appears when you use `geom_histogram()` to create a histogram in `ggplot2` and fail to specify the number of bins to use in the histogram.

To avoid this warning, you can use the `bins` argument to specify the number of bins to use:

```
ggplot(df, aes(x=my_variable)) +  
geom_histogram(bins=10)
```

The following example shows how to avoid this warning in practice.

## Example: How to Avoid Warning `stat_bin()` using `bins = 30`

Suppose we use the `geom_histogram()` function from `ggplot2` to create a histogram for some variable in a data frame:

```
library(ggplot2)
```

```
#make this example reproducible  
set.seed(0)
```

```
#create data frame
```

```
df <- data.frame(my_values = rnorm(1000))
```

```
#view head of data frame
```

```
head(df)
```

```
my_values
```

```
1 1.2629543
```

```
2 -0.3262334
```

```
3 1.3297993
```

```
4 1.2724293
```

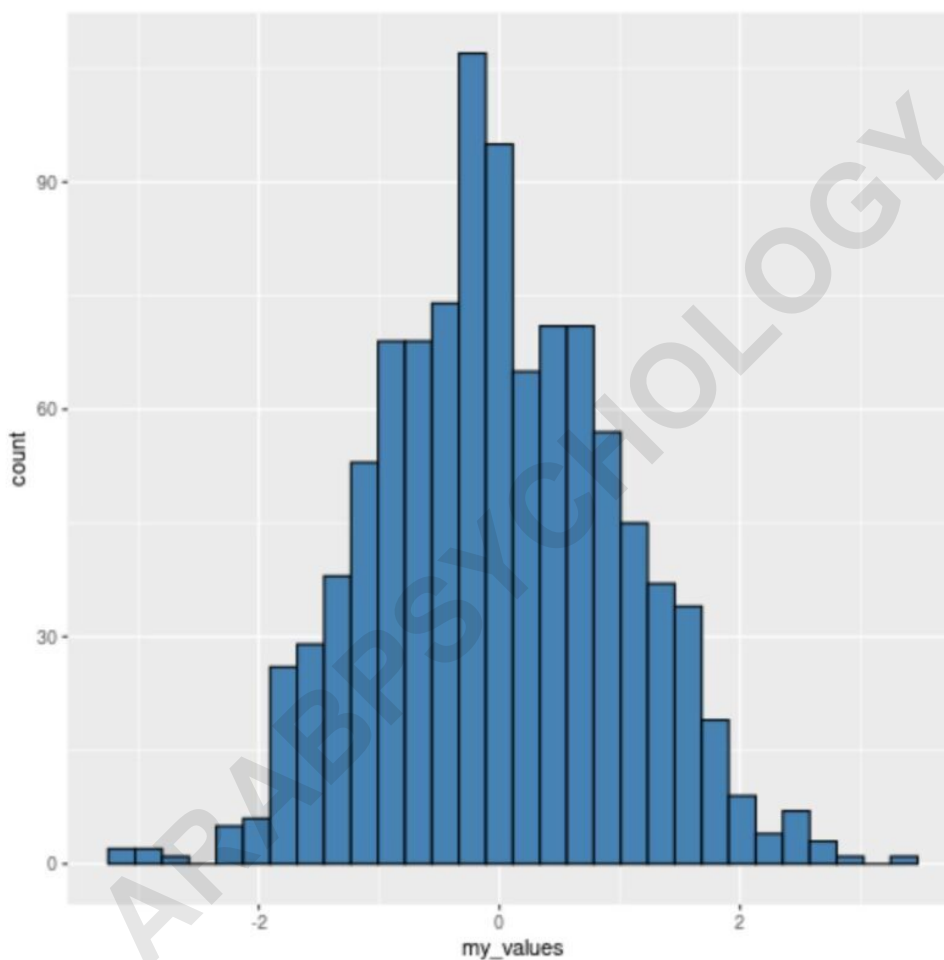
```
5 0.4146414
```

```
6 -1.5399500
```

```
#create histogram
```

```
ggplot(df, aes(x=my_values)) +  
geom_histogram(col='black', fill='steelblue')
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.



Notice that we receive the warning `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

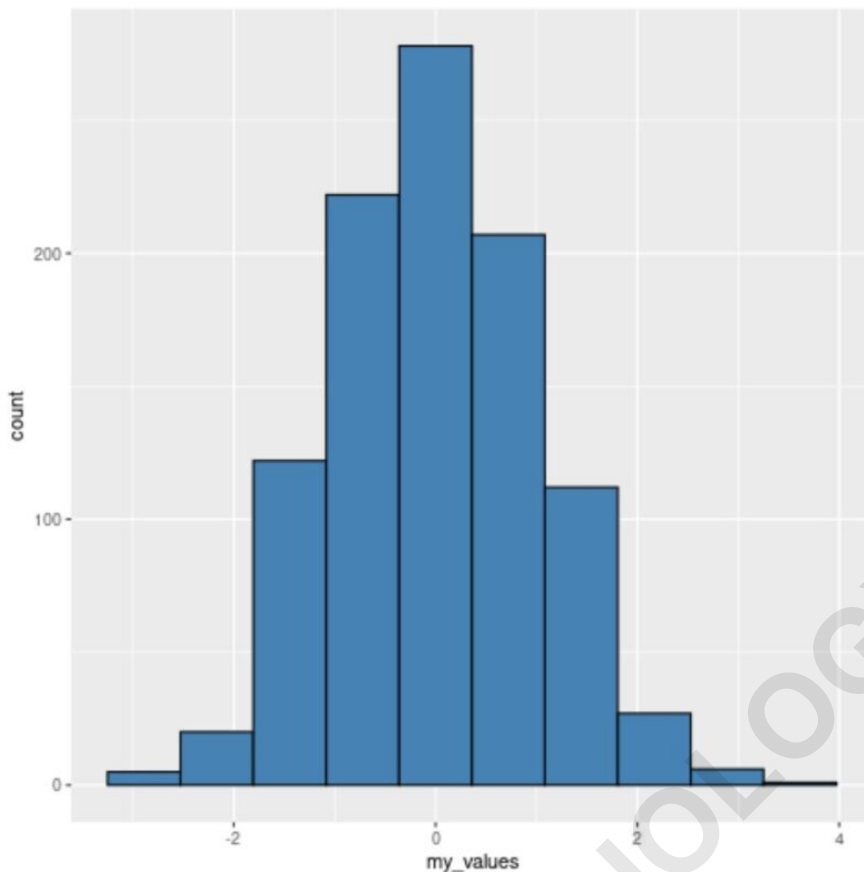
It's worth noting that this is just a warning and the

**histogram is still created by `ggplot2`.**

**However, we can avoid the warning altogether by using the `bins` argument within `geom_histogram()` to specify the number of bins to use in the histogram.**

**For example, we can use the following code to use 10 bins for the histogram:**

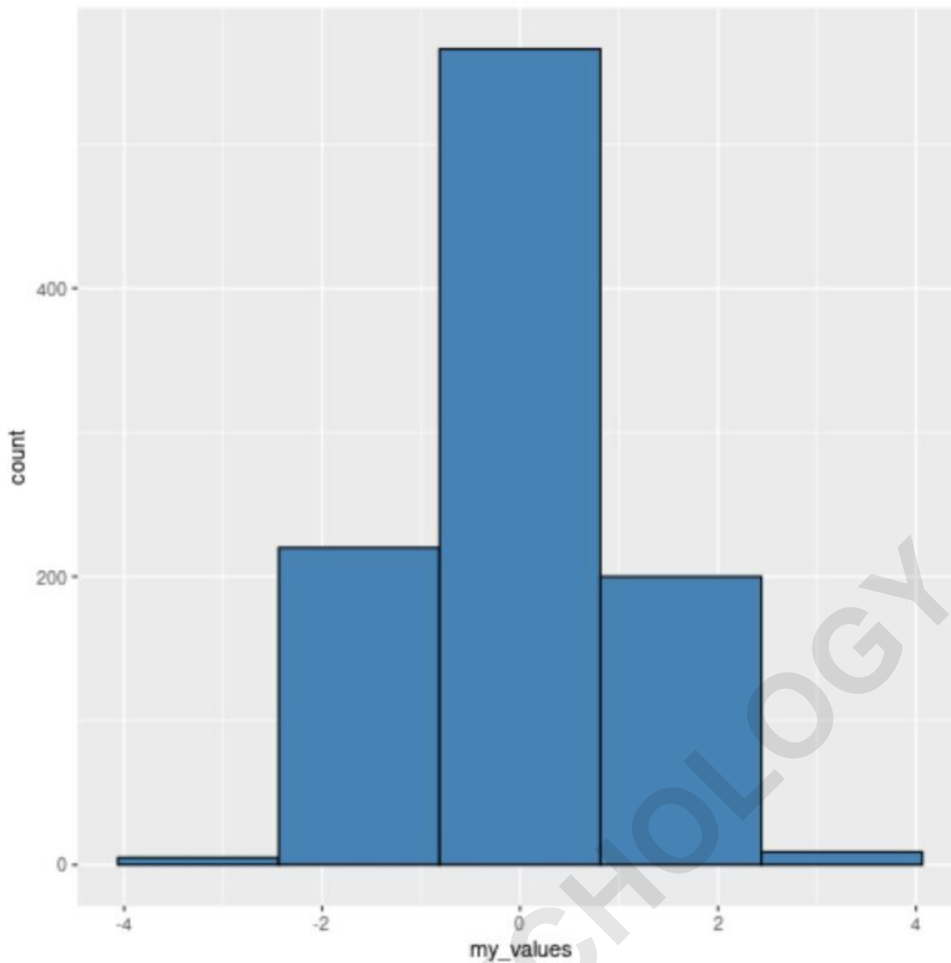
```
#create histogram with 10 bins  
ggplot(df, aes(x=my_values)) +  
geom_histogram(col='black', fill='steelblue', bins=10)
```



Notice that we don't receive any warning message this time and the histogram contains exactly 10 bins.

For example, we could instead use 5 bins:

```
#create histogram with 5 bins  
ggplot(df, aes(x=my_values)) +  
geom_histogram(col='black', fill='steelblue', bins=5)
```



**Notice that there are fewer but wider bins in this histogram.**

**Feel free to use as many bins as you'd like in your histogram.**

**The following tutorials explain how to troubleshoot other common errors in R:**

**[How to Fix in R: NAs Introduced by Coercion](#)**