

Why am I getting the error “na/nan/inf in ‘y’” when using lm.fit() with an offset in R?

Authored by
stats writer

June 29, 2024

RECOMMENDED CITATION

stats writer (2024). *Why am I getting the error “na/nan/inf in ‘y’” when using lm.fit() with an offset in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=159990>

The error "na/nan/inf in 'y'" occurs when using the `lm.fit()` function in R with an offset. This error indicates that there are missing, infinite, or undefined values in the dependent variable (y) being used for the linear regression model. This can happen if the offset is not properly specified or if there are issues with the data being used. It is important to carefully check the data and adjust the offset appropriately in order to avoid this error.

Fix: error in `lm.fit(x, y, offset = offset, ...)` : na/nan/inf in 'y'

One error you may encounter when using R is:

**Error in `lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...)` :
NA/NaN/Inf in 'y'**

This error occurs when you attempt to use the function to fit a linear regression model in R, but either the predictor or response variable contains NaN or Inf values.

The following example shows how to fix this error in practice.

How to Reproduce the Error

Suppose we have the following data frame in R that contains information about minutes played and points scored for various basketball players:

```
#create data frame with some NA, NaN, Inf values  
df <- data.frame(minutes=c(4, NA, 28, 12, 30, 21, 14),  
points=c(12, NaN, 30, Inf, 43, 25, 17))
```

```
#view data frame
```

```
df
```

```
minutes points
```

```
1 4 12
```

```
2 NA NaN
```

```
3 28 30
```

```
4 12 Inf
```

```
5 30 43
```

```
6 21 25
```

```
7 14 17
```

Notice that the data frame contains some NaN and Inf values.

Now suppose we attempt to fit a linear regression model using "minutes" as the predictor variable and "points" as the response variable:

```
#attempt to fit regression model
```

```
lm(points ~ minutes, data=df)
```

**Error in lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :
NA/NaN/Inf in 'y'**

We receive an error because there are NaN and Inf values present in the data frame.

How to Fix the Error

It's worth noting that the NA values in the data frame are not an issue. In fact, R simply ignores the NA values when fitting the linear regression model.

The real issue is caused by the NaN and Inf values.

The easiest way to resolve this issue is to replace the NaN and Inf values with NA values:

```
#Replace NaN & Inf with NA
```

```
df = NA
```

```
#view updated data frame
```

```
df
```

```
minutes points
```

```
1 4 12
```

2 NA NA

3 28 30

4 12 NA

5 30 43

6 21 25

7 14 17

Now we can fit the regression model:

```
#fit regression model
```

```
lm(points ~ minutes, data=df)
```

Call:

```
lm(formula = points ~ minutes, data = df)
```

Coefficients:

```
(Intercept) minutes
```

```
5.062 1.048
```

The output shows the of the regression model.

Notice that we don't receive any error since we replaced the NaN and Inf values in the data frame.

Additional Resources

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

ARABPSYCHOLOGY.COM