

What is a Residuals vs. Leverage Plot?

Authored by
stats writer

May 6, 2024

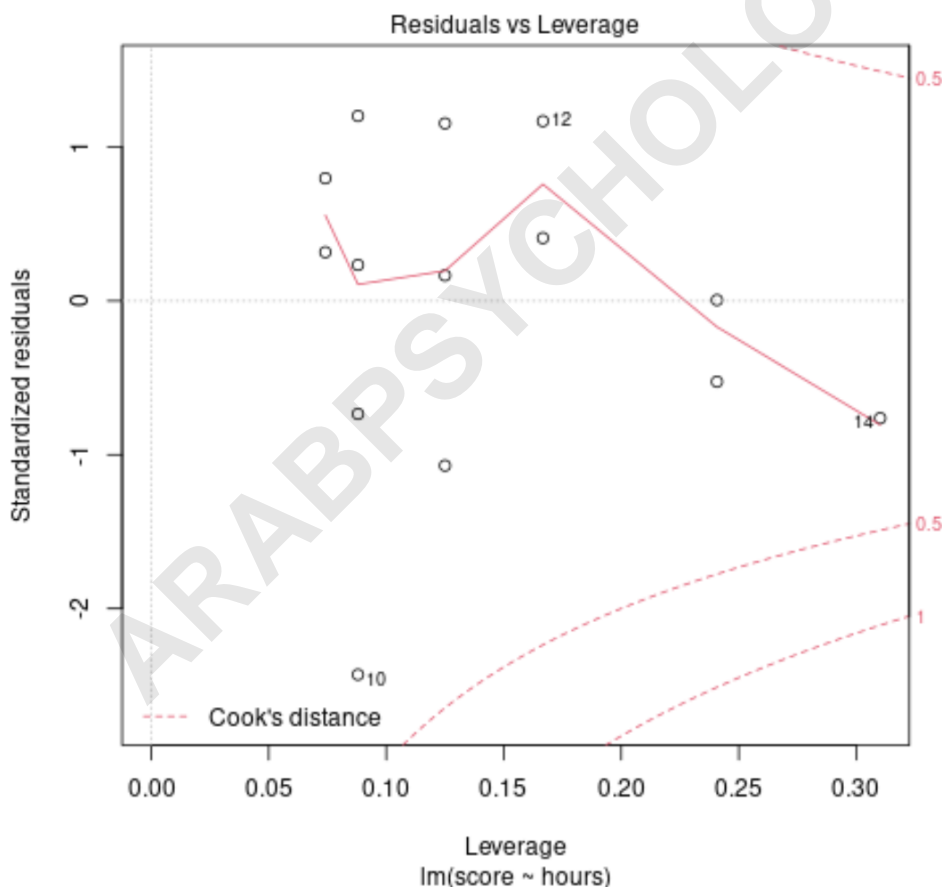
RECOMMENDED CITATION

stats writer (2024). *What is a Residuals vs. Leverage Plot?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=143545>

What is a Residuals vs. Leverage Plot? (Definition & Example)

A residuals vs. leverage plot is a type of plot that allows us to identify influential observations in a regression model.

Here is how this type of plot appears in the statistical programming language R:



Each observation from the dataset is shown as a single

point within the plot. The x-axis shows the leverage of each point and the y-axis shows the standardized residual of each point.

Leverage refers to the extent to which the coefficients in the regression model would change if a particular observation was removed from the dataset.

Observations with high leverage have a strong influence on the coefficients in the regression model. If we remove these observations, the coefficients of the model would change noticeably.

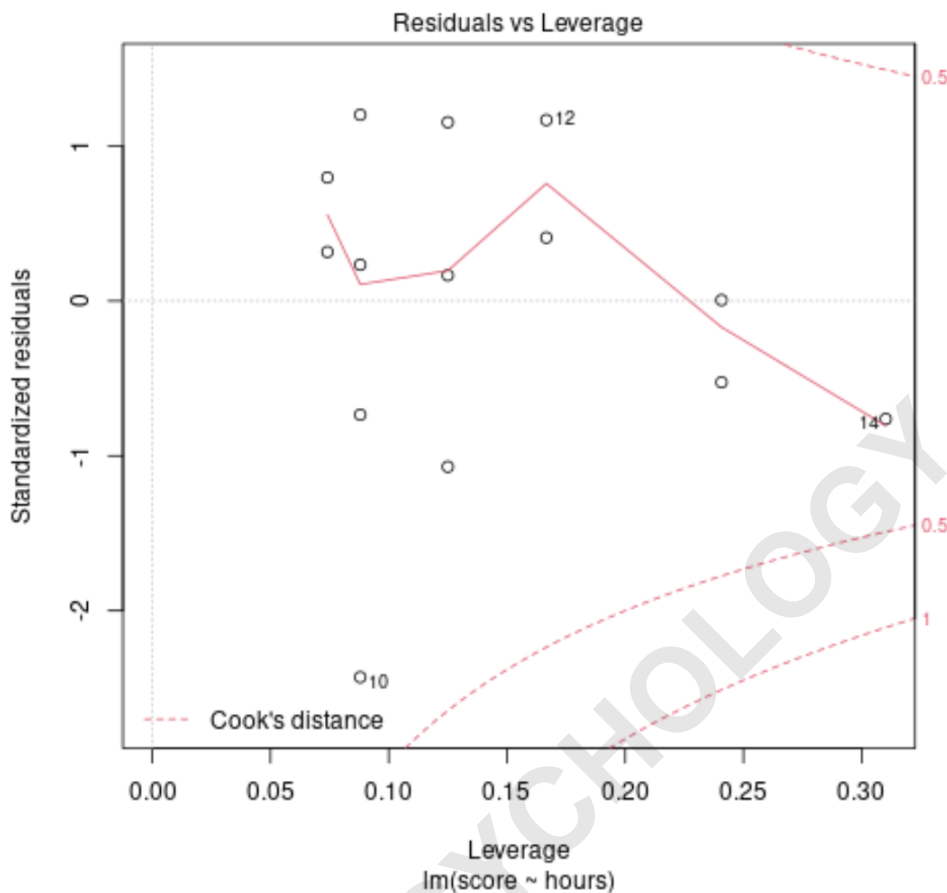
Standardized residuals refer to the standardized difference between a predicted value for an observation and the actual value of the observation.

It's worth noting that an observation can have a high absolute value for a standardized residual, yet have a low value for leverage.

How to Interpret a Residuals vs. Leverage Plot

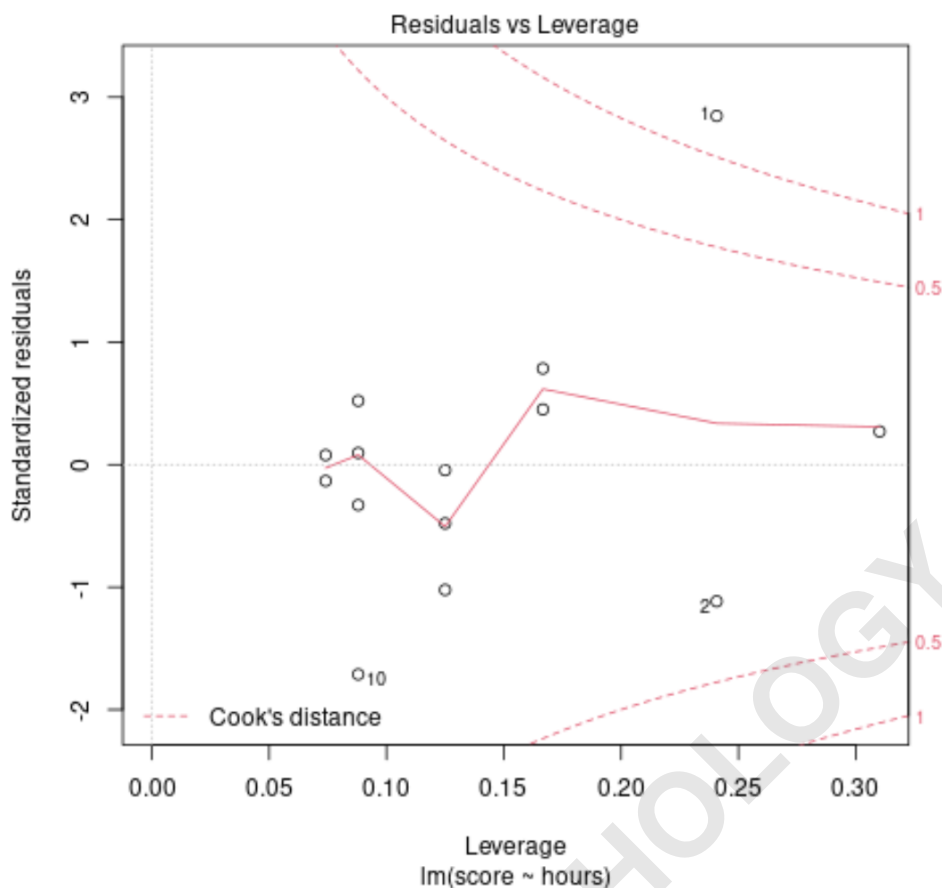
If any point in this plot falls outside of Cook's distance (the red dashed lines) then it is considered to be an influential observation.

Let's refer to the residuals vs. leverage plot from earlier:



In the example above, we can see that observation #10 lies closest to the border of Cook's distance, but it doesn't fall outside of the dashed line. This means there are not any influential points in our regression model.

However, suppose we had the following residuals vs. leverage plot:



We can see that observation #1 in the top right corner falls outside of the red dashed lines. This indicates that it is an influential point.

This means that if we removed this observation from our dataset and fit the regression model again, the coefficients of the model would change significantly.

How to Handle Influential Observations

If you create a residuals vs. leverage plot for a model and you find that one or more observations are

identified as influential, there are a few things you can do:

1. Verify that the observation is not an error.

Before you take any action, you should first verify that the influential observation(s) are not a result of a data entry error or some other odd occurrence.

2. Attempt to fit another regression model.

Influential observations could indicate that the model you specified does not provide a good fit to the data. In this case, you may try a or a nonlinear model.

3. Remove the influential observations.

Lastly, you may decide to simply remove the influential observations if the model you specified seems to fit the data well except for the one or two influential observations.

The following tutorials provide additional information on how to use residuals to assess the fit of regression models.