

# How can I create a residual plot in Google Sheets?

Authored by  
**stats writer**

June 30, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I create a residual plot in Google Sheets?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=162059>

To create a residual plot in Google Sheets, first create a scatter plot with the independent variable (x-axis) and dependent variable (y-axis) data. Then, add a trendline to the scatter plot by right-clicking on a data point and selecting "Add Trendline." In the trendline options, choose "Linear" as the type and check the box for "Display equation on chart." This will add a line of best fit and its equation to the scatter plot. Next, create a new column next to the dependent variable data and label it "Residuals." In this column, use the equation of the trendline to calculate the difference between the actual and predicted values for each data point. Finally, create a scatter plot with the independent variable data on the x-axis and the residuals on the y-axis. This will create a residual plot, which can be used to identify any patterns or trends in the errors of the original scatter plot.

## Create a Residual Plot in Google Sheets

**A residual plot is a type of plot that displays the fitted values against the residual values for a regression model.**

**This type of plot is often used to assess whether or not a linear regression model is appropriate for a given dataset and to check for of residuals.**

**The following step-by-step example explains how to create a residual plot for a simple linear regression model in Google Sheets.**

**Step 1: Enter the Data**

**First, let's enter the following values for a dataset:**

	A	B	C	D
1	<b>x</b>	<b>y</b>		
2	8	41		
3	12	42		
4	12	39		
5	13	37		
6	14	35		
7	16	39		
8	17	45		
9	22	46		
10	24	39		
11	26	49		
12	29	55		
13	30	57		
14				
15				
16				
17				
18				

## Step 2: Calculate the Equation of the Regression Model

Next, we'll use the **SLOPE** and **INTERCEPT** functions to calculate the equation of the simple linear regression model for this dataset:

	A	B	C	D
1	<b>x</b>	<b>y</b>		
2	8	41		
3	12	42		
4	12	39		
5	13	37		
6	14	35		
7	16	39		
8	17	45		
9	22	46		
10	24	39		
11	26	49		
12	29	55		
13	30	57		
14				
15	<b>Slope</b>	0.7552878554	=SLOPE(B2:B13, A2:A13)	
16	<b>Intercept</b>	29.63090069	=INTERCEPT(B2:B13, A2:A13)	
17				
18				
19				

Using these values, we can write the following simple linear regression equation:

$$y = 29.631 + 0.755x$$

Step 3: Calculate the Predicted Values

Next, we can use the regression equation to calculate the predicted values for each observation.

We'll type the following formula into cell C2:

$$= \$B\$16 + \$B\$15 * A2$$

We can then copy and paste this formula down to every remaining cell in column C:

	A	B	C	D
1	<b>x</b>	<b>y</b>	<b>Predicted Value</b>	
2	8	41	35.7	
3	12	42	38.7	
4	12	39	38.7	
5	13	37	39.4	
6	14	35	40.2	
7	16	39	41.7	
8	17	45	42.5	
9	22	46	46.2	
10	24	39	47.8	
11	26	49	49.3	
12	29	55	51.5	
13	30	57	52.3	
14				
15	<b>Slope</b>	0.7552878554		
16	<b>Intercept</b>	29.63090069		
17				
18				
19				

#### Step 4: Calculate the Residuals

A is the difference between an observed value and a predicted value in a regression model.

It is calculated as:

**Residual = Observed value - Predicted value**

To calculate the residual for each observation in our

dataset, we can type the following formula into cell D2:

**=B2-C2**

We can then copy and paste this formula down to every remaining cell in column D:

	A	B	C	D	E
D2				$=B2-C2$	
1	<b>x</b>	<b>y</b>	<b>Predicted Value</b>	<b>Residual</b>	
2	8	41	35.7	5.3	
3	12	42	38.7	3.3	
4	12	39	38.7	0.3	
5	13	37	39.4	-2.4	
6	14	35	40.2	-5.2	
7	16	39	41.7	-2.7	
8	17	45	42.5	2.5	
9	22	46	46.2	-0.2	
10	24	39	47.8	-8.8	
11	26	49	49.3	-0.3	
12	29	55	51.5	3.5	
13	30	57	52.3	4.7	
14					
15	<b>Slope</b>	0.7552878554			
16	<b>Intercept</b>	29.63090069			
17					
18					
19					
20					

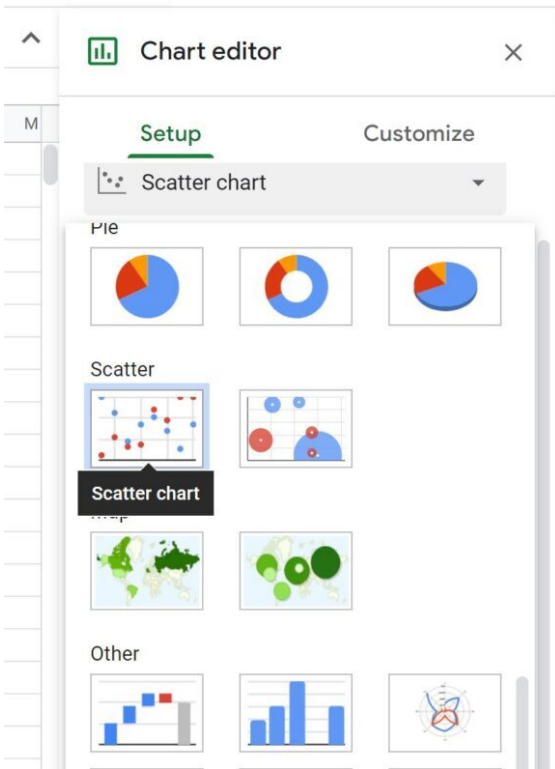
### Step 5: Create the Residual Plot

To create the residual plot, we can highlight the values in the range A2:A13, then hold the "Ctrl" key and highlight the values in the range D2:D13.

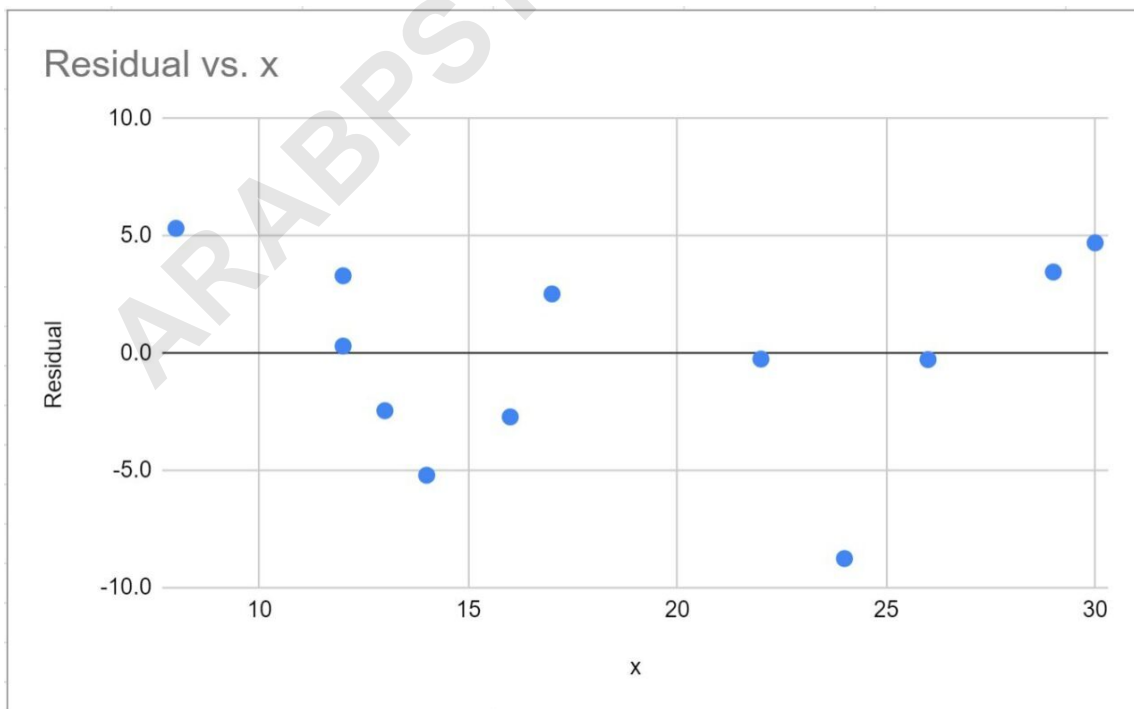
	A	B	C	D
1	<b>x</b>	<b>y</b>	<b>Predicted Value</b>	<b>Residual</b>
2	8	41	35.7	5.3
3	12	42	38.7	3.3
4	12	39	38.7	0.3
5	13	37	39.4	-2.4
6	14	35	40.2	-5.2
7	16	39	41.7	-2.7
8	17	45	42.5	2.5
9	22	46	46.2	-0.2
10	24	39	47.8	-8.8
11	26	49	49.3	-0.3
12	29	55	51.5	3.5
13	30	57	52.3	4.7
14				
15	<b>Slope</b>	0.7552878554		
16	<b>Intercept</b>	29.63090069		
17				
18				
19				
20				

Then click the **Insert** tab, then click **Chart** in the dropdown menu.

In the **Chart** editor panel that appears on the right side of the screen, choose **Scatter chart** as the **Chart type**:



The following residual plot will automatically appear:



The x-axis displays the values for the predictor variable in our regression model and the y-axis displays the residuals.

One of linear regression is that the residuals have constant variance at every level of x, so we often use a residual plot to determine if this assumption is met.

If the residuals are roughly evenly scattered around zero in the plot with no clear pattern, then we typically say the assumption of constant variance is met.

In our residual plot above we can see that the points in the plot seem to be randomly scattered around zero with no clear pattern, thus we would conclude that the constant variance assumption is met for this particular regression model.

#### Additional Resources

The following tutorials explain how to perform other common tasks in Google Sheets: