Statistics Kingdom

<u>Home</u> > <u>Regression</u> > Linear Regression

Linear Regression Calculator

Linear regression calculator and prediction interval calculator with step-by-step solution.

Simple Linear regression Multiple Linear regression Logistic regression Multinomial logistic regression

How to do with R?

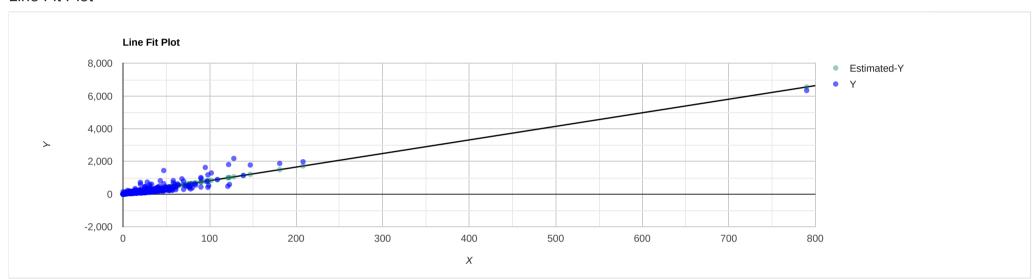
Regression line equation

$\hat{\mathbf{Y}} = -12.6294 + 8.3074\mathbf{X}$

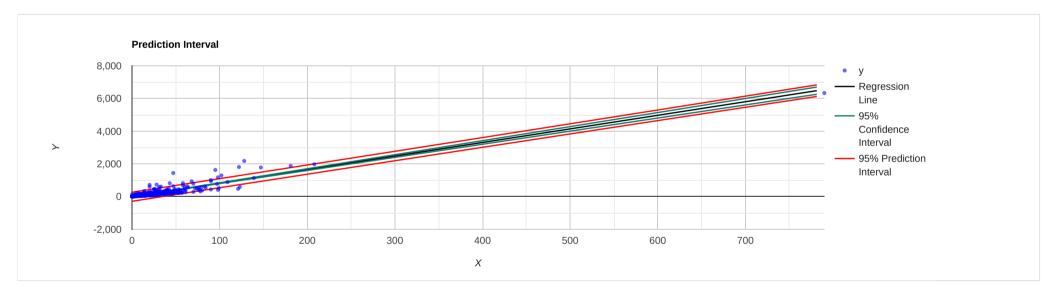
Reporting linear regression in APA style

X predicted Y, R^2 = .86, F(1,498) = 3124.72, p < .001. β = 8.31, p < .001, α = -12.63, p = .092.

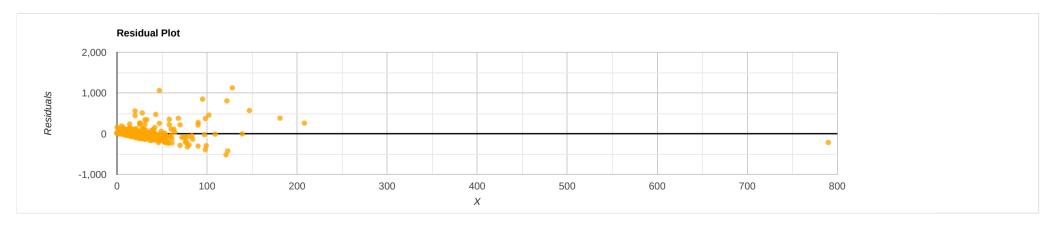
Line Fit Plot



Prediction online



Residual Plot



Prediction

Interpretation of the results



Hover over the cells to see the formulas.

Source	DF	Sum of Square	Mean Square	F Statistic (df ₁ ,df ₂)	P-value
Regression (between \hat{y}_i and \bar{y})	1	62226688.0363	62226688.0363	3124.7159 (1,498)	0
	498	9917346.7637	19914.3509		
Total (between y_i and \bar{y})	499	72144034.8	144577.224		

1. Y and X relationship

R-Squared (R²) equals **0.8625.** This means that 86.3% of the variability of Y is explained by X.

Correlation (R) equals 0.9287. This means that there is a very strong direct relationship between X and Y.

The Standard deviation of the residuals (S_{res}) equals **141.1182**.

The slope: b_1 =8.3074 CI[8.0154, 8.5994] means that when you increase X by 1, the value of Y increases by 8.3074.

The y-intercept: b_0 =-12.6294 CI[-27.3314, 2.0726] means that when X equals 0, the prediction of Y's value is -12.6294.

The x-intercept equals 1.5203.

2. Goodness of fit

Overall regression: right-tailed, F(1,498) = 3124.7159, p-value = 0. Since p-value < α (0.05), we reject H_0 .

The linear regression model, $Y = b_0 + b_1 X + \epsilon$, provides a better fit than the model without the independent variable resulting in $Y = b_0 + \epsilon$.

The slope (b_1) : two-tailed, T(498)=55.8992, p-value = 0. For one predictor it is the same as the p-value for the overall model.

The y-intercept (b_0): two-tailed, T(498) = -1.6878, p-value = 0.09208. Hence, b_0 is not significantly different from zero. It is still most likely recommended not to force b_0 to be zero.

3. Residual normality

The linear regression model assumes normality for residual errors. The Shapiro-Wilk p-value equals 0. It is assumed that the data is not normally distributed, But since the sample size is large, it should not adversely affect the regression model.

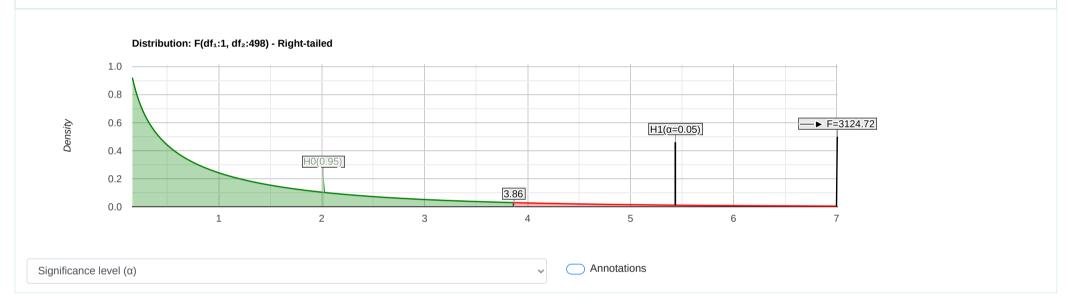
4. Outliers

Outliers may affect the regression line.

In this case, the distribution of the residuals is normal. Therefore, the probability of detecting 11 valid outliers or more is 1, (outliers:

-520.5705,445.4807,559.4807,510.0212,473.4096,455.2708,1061.1799,849.4229,568.4362,805.1221,1126.2775).

You should only remove outliers if you identify them as errors!



Residuals normality



Calculation

Step-by-step solution $\hat{Y} = b_0 + b_1 X$ $b_1 = \frac{SP_{xy} - \sum_{x} (x_i - \bar{x})(y_i - \bar{y})}{SS_x - \sum_{x} (x_i - \bar{x})^2}$ $b_1 = \frac{7490479.76}{901659.542} = 8.3074$ $b_0 = \bar{y} - b_1 \bar{x}$ $\bar{x} = 27.054$

$$\bar{y} = 212.12$$

b₀ = 212.12-8.3074*27.054 = -12.6294

$$R^{2} = \frac{SS_{Regression}}{SS_{total}} = \frac{\Sigma(\ \hat{y_{i}} - \bar{y})^{2}}{\Sigma(\ y_{i} - \bar{y})^{2}} = \frac{62226688.0363}{72144034.8} = 0.8625$$

The standard deviation of the residuals is:

$$MS_{residual} = S_{res}^2 = \frac{\sum (y_i - \hat{y})^2}{n - 2}$$

Residual outliers

 $S_{res} = \sqrt{MSE} = \sqrt{19914.3509} = 141.1182.$

The average of the residuals is always zero.

The thresholds used to calculate the outliers are: $\pm k*S_{res}$.

In this case, the thresholds are $\pm 3*141.1182 = \pm 423.3546$.

We tagged the outliers with an arrow (\Leftarrow) at the 'Residual' column.

SS_x and SP_{xy}

$ x-\bar{x} \rangle^2$	$(x-\bar{x})(y-\bar{y})$
\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	(* ^)(J J)

10/2	23/25, 5:38 PM	Linear regression ca	alculator - calculates the linear regression equation, draws the pred	iction interva
	-27.054	-212.12	731.9189	5738.6945
	-27.054	-212.12	731.9189	5738.6945
	-27.054	-212.12	731.9189	5738.6945
	-27.054	-212.12	731.9189	5738.6945
	-27.054	-212.12	731.9189	5738.6945
	-27.054	-212.12	731.9189	5738.6945
	-27.054	-210.12	731.9189	5684.5865
	-27.054	-210.12	731.9189	5684.5865
	-27.054	-210.12	731.9189	5684.5865
	-27.054	-209.12	731.9189	5657.5325
	-27.054	-209.12	731.9189	5657.5325
	-25.054	-208.12	627.7029	5214.2385
	-27.054	-208.12	731.9189	5630.4785
	-26.054	-207.12	678.8109	5396.3045
	-26.054	-207.12	678.8109	5396.3045
	-25.054	-206.12	627.7029	5164.1305
	-25.054	-206.12	627.7029	5164.1305
	-27.054	-206.12	731.9189	5576.3705
	-27.054	-205.12	731.9189	5549.3165
	-26.054	-205.12	678.8109	5344.1965
	-26.054	-205.12	678.8109	5344.1965
	-25.054	-205.12	627.7029	5139.0765
	-25.054	-204.12	627.7029	5114.0225
	-25.054	-203.12	627.7029	5088.9685
	-25.054	-203.12	627.7029	5088.9685
	-26.054	-202.12	678.8109	5266.0345
	-25.054	-202.12	627.7029	5063.9145
	-24.054	-202.12	578.5949	4861.7945
	-25.054	-202.12	627.7029	5063.9145
	-26.054	-200.12	678.8109	5213.9265
	-26.054	-198.12	678.8109	5161.8185
	-25.054	-198.12	627.7029	4963.6985
	-24.054	-197.12	578.5949	4741.5245
	-26.054	-197.12	678.8109	5135.7645
	-23.054	-195.12	531.4869	4498.2965
	-25.054	-195.12	627.7029	4888.5365
	-26.054	-195.12	678.8109	5083.6565
	-26.054	-195.12	678.8109	5083.6565
	-25.054	-194.12	627.7029	4863.4825
	-25.054	-194.12	627.7029	4863.4825
	-24.054	-194.12	578.5949	4669.3625
	-26.054	-191.12	678.8109	4979.4405
	-22.054	-188.12	486.3789	4148.7985
	-25.054	-187.12	627.7029	4688.1045
	-24.054	-187.12	578.5949	4500.9845
	-18.054	-184.12	325.9469	3324.1025
	-23.054	-184.12	531.4869	4244.7025
	-21.054	-183.12	443.2709	3855.4085
	-22.054	-182.12	486.3789	4016.4745
	-19.054	-180.12	363.0549	3432.0065
	-25.054	-180.12	627.7029	4512.7265
	-23.054	-178.12	531.4869	4106.3785
	-23.054	-178.12	531.4869	4106.3785
	-21.054	-178.12	443.2709	3750.1385
	-16.054	-177.12	257.7309	2843.4845
	-24.054	-176.12	578.5949	4236.3905
	-20.054	-176.12	402.1629	3531.9105
	-20.054	-174.12	402.1629	3491.8025
	-21.054	-174.12	443.2709	3665.9225
	-21.054	-174.12	443.2709	3665.9225
	-16.054	-173.12	257.7309	2779.2685
	-21.054	-173.12	443.2709	3644.8685
	-24.054	-173.12	578.5949	4164.2285
	-23.054	-173.12	531.4869	3991.1085
	-16.054	-173.12	257.7309	2779.2685
	-20.054	-173.12	402.1629	3471.7485
	-22.054	-172.12	486.3789	3795.9345
	-22.054	-171.12	486.3789	3773.8805
	-19.054	-170.12	363.0549	3241.4665
	-24.054	-170.12	578.5949	4092.0665
	-24.054	-168.12	578.5949	4043.9585
	-20.054	-168.12	402.1629	3371.4785
	-22.054	-167.12	486.3789	3685.6645
	-17.054	-167.12	290.8389	2850.0645
	-17.054	-166.12	290.8389	2833.0105
	-11.054	-164.12	122.1909	1814.1825
	-19.054	-163.12	363.0549	3108.0885
	-12.054	-163.12	145.2989	1966.2485
	-16.054	-162.12	257.7309	2602.6745
	-14.054	-159.12	197.5149	2236.2725
	-18.054	-159.12	325.9469	2872.7525
	-23.054	-159.12	531.4869	3668.3525
	-13.054	-159.12	170.4069	2077.1525
	-15.054	-158.12	226.6229	2380.3385
	-17.054	-158.12	290.8389	2696.5785
	-18.054	-158.12	325.9469	2854.6985
	-20.054	-157.12	402.1629	3150.8845
	-17.054	-157.12	290.8389	2679.5245
	-16.054	-156.12	257.7309	2506.3505
	-19.054	-156.12	363.0549	2974.7105
	-26.054	-155.12	~~~ ^~~	**** ****
	-16.054	-154.12		
	-18.054	-154.12		
	-21.054	-154.12		
. '	1	r regression salsulatorhtml		

10/23/25, 5:38 PM	Linear regression o	calculator - calculates the linear regression equation, draws the prec	diction interval
12.946	19.88	167.5989	257.3665
5.946	19.88	35.3549	118.2065
19.946	20.88	397.8429	416.4725
6.946	21.88	48.2469	151.9785
	21.88		
0.946		0.8949	20.6985
21.946	22.88	481.6269	502.1245
5.946	23.88	35.3549	141.9905
-13.054	24.88	170.4069	-324.7835
0.946	25.88	0.8949	24.4825
0.946	25.88	0.8949	24.4825
9.946	25.88	98.9229	257.4025
3.946	29.88	15.5709	117.9065
13.946	29.88	194.4909	416.7065
6.946	30.88	48.2469	214.4925
-8.054	30.88	64.8669	-248.7075
-5.054	32.88	25.5429	-166.1755
-12.054	36.88	145.2989	-444.5515
4.946	38.88	24.4629	192.3005
-8.054	39.88	64.8669	
			-321.1935
29.946	39.88	896.7629	1194.2465
7.946	39.88	63.1389	316.8865
8.946	41.88	80.0309	374.6585
11.946	42.88	142.7069	512.2445
27.946	44.88	780.9789	1254.2165
11.946	49.88	142.7069	595.8665
7.946	50.88	63.1389	404.2925
33.946	51.88	1152.3309	1761.1185
8.946	52.88	80.0309	473.0645
-1.054	53.88	1.1109	-56.7895
3.946	56.88	15.5709	224.4485
	56.88		1134.5285
19.946		397.8429 526.5180	
22.946	58.88	526.5189	1351.0605
16.946	58.88	287.1669	997.7805
0.946	58.88	0.8949	55.7005
16.946	58.88	287.1669	997.7805
16.946	60.88	287.1669	1031.6725
14.946	61.88	223.3829	924.8585
12.946	63.88	167.5989	826.9905
-0.054	66.88	0.002916	-3.6115
6.946	67.88	48.2469	471.4945
4.946	69.88	24.4629	345.6265
42.946	69.88	1844.3589	3001.0665
15.946	72.88	254.2749	1162.1445
-6.054	74.88	36.6509	-453.3235
13.946	75.88	194.4909	1058.2225
15.946	77.88	254.2749	1241.8745
17.946	81.88	322.0589	1469.4185
23.946	82.88	573.4109	1984.6445
14.946	86.88	223.3829	1298.5085
23.946	87.88	573.4109	2104.3745
8.946	92.88	80.0309	830.9045
12.946	96.88	167.5989	1254.2085
10.946	98.88	119.8149	1082.3405
50.946	98.88	2595.4949	5037.5405
16.946	100.88	287.1669	1709.5125
8.946	100.88	80.0309	902.4725
2.946	103.88	8.6789	306.0305
15.946	117.88	254.2749	1879.7145
26.946	122.88	726.0869	3311.1245
0.946	124.88	0.8949	118.1365
-13.054	126.88	170.4069	-1656.2915
4.946	129.88	24.4629	642.3865
19.946	142.88	397.8429	2849.8845
1.946	150.88	3.7869	293.6125
13.946	154.88	194.4909	2159.9565
10.946	163.88	119.8149	1793.8305
33.946	165.88	115.3309	5630.9625
52.946	165.88	2803.2789	8782.6825
27.946	174.88	780.9789	4887.1965
30.946	177.88	957.6549	5504.6745
21.946	180.88	481.6269	3969.5925
26.946	191.88	726.0869	5170.3985
12.946	196.88	167.5989	2548.8085
48.946	196.88	2395.7109	9636.4885
70.946	197.88	5033.3349	14038.7945
12.946	207.88	167.5989	2691.2145
49.946	208.88	2494.6029	10432.7205
-2.054	216.88	4.2189	-445.4715
62.946	217.88	3962.1989	13714.6745
25.946	228.88	673.1949	5938.5205
32.946	234.88	1085.4389	7738.3565
21.946	241.88	481.6269	5308.2985
-2.054	243.88	4.2189	-500.9295
-1.054	245.88	1.1109	-259.1575
3.946	257.88	15.5709	1017.5945
93.946	259.88	8825.8509	24414.6865
14.946	277.88	223.3829	4153.1945
44.946	281.88	2020.1429	12669.3785
46.946	292.88	2203.9269	13749.5445
71.946	302.88	5176.2269	21791.0045
48.946	315.88	2205 7100	15/61 0625
56.946	337.88		
36.946	344.88		
51.946	359.88		
https://www.statskingdom.com/line	ar-regression-calculator.html		

0	0	901659.542 (SS _x)	7490479.76 (SP _{xy})
762.946	6121.88	582086.5989	4670663.8585
100.946	1964.88	10190.0949	198346.7765
180.946	1761.88	32741.4549	318805.1385
153.946	1660.88	23699.3709	255685.8325
94.946	1593.88	9014.7429	151332.5305
119.946	1564.88	14387.0429	187701.0965
67.946	1413.88	4616.6589	96067.4905
19.946	1226.88	397.8429	24471.3485
74.946	1077.88	5616.9029	80782.7945
70.946	961.88	5033.3349	68241.5385
111.946	923.88	12531.9069	103424.6705
62.946	796.88	3962.1989	50160.4085
62.946	735.88	3962.1989	46320.7025
40.946	718.88	1676.5749	29435.2605
81.946	668.88	6715.1469	54812.0405
30.946	610.88	957.6549	18904.2925
15.946	605.88	254.2749	9661.3625
42.946	574.88	1844.3589	24688.7965
69.946	558.88	4892.4429	39091.4205
0.946	517.88	0.8949	489.9145
-7.054	500.88	49.7589	-3533.2075
30.946	478.88	957.6549	14819.4205
19.946	420.88	397.8429	8394.8725
55.946	411.88	3129.9549	23043.0385
35.946	409.88	1292.1149	14733.5465
5.946	392.88	35.3549	2336.0645
-7.054	386.88	49.7589	-2729.0515
32.946	383.88	1085.4389	12647.3105
3.946 95.946	374.88 374.88	15.5709 9205.6349	35968.2365

Linear regression calculator

The linear regression calculator generates the linear regression equation. It also draws: a linear regression line, a histogram, a residuals QQ-plot, a residuals x-plot, and a distribution chart. It calculates the R-squared, the R, and the outliers, then testing the fit of the linear model to the data and checking the residuals' normality assumption and the priori power.

What is linear regression?

The linear regression is the linear equation that best fits the points.

There is no one way to choose the best fit ting line, the most common one is the ordinary least squares (OLS). The linear regression describes the relationship between the dependent variable (Y) and the independent variables (X).

The linear regression model calculates the dependent variable (DV) based on the independent variables (IV, predictors).

What is "ordinary least squares"?

The ordinary least squares method chooses the line parameters that minimize the sum of squares of the differences between the observed dependent variables (Y) and the estimated value by the linear regression (Ŷ).

Why do you need linear regression?

We may use linear regression when we want to do one of the following

- Predict the dependent variable (\hat{Y}) .
- Estimate the effect of each independent variable (X) on the dependent variable (Y).
- Calculate the correlation between the dependent variable and the independent variables.
- Test the linear model significance level.

How to calculate linear regression?

Following the linear regression formula:

$$\hat{\mathbf{Y}} = \mathbf{b}_0 + \mathbf{b}_1 \mathbf{x}$$

 $\ensuremath{b_0}$ - the y-intercept, where the line crosses the y-axis.

b₁ - the slope, describes the line's direction and incline.

$$b_1 = \frac{SP_{xy}}{SS_x} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2}$$

$$b_0=\bar{y}-b_1\bar{x}$$

linear regression prediction

The prediction calculator uses the linear regrssion to predict the depdendent variable based on the independent value. The calculator also creates the confidence interval, and the prediction interval.

Confidence interval of the prediction

The prediction interval for the $\boldsymbol{mean\ value}$ of the dependent variable.

This is the interval for the equation line, the true value equation will be in this interval. If we would know the true equation then the width of this interval would be zero.

If you would calculate the confidence interval over an infinite number of regressions with the same sample size, 95% (confidence level) of the calculated confidence intervals will contain the mean's true value. Since this interval is for the mean, the standard error is smaller and the the range is narrower than the range of the prediction interval.

$$\begin{split} MS_{residual} &= S^2_{residual} = \frac{\Sigma (y_i - \hat{y})^2}{n - 2} \\ S.E^2_{ci} &= S^2_{residual} \Big(\frac{1}{n} + \frac{(x_0 - \hat{x})^2}{SS_x} \Big) \\ \hat{Y} &\pm T_{1-\alpha/2} (n-2)^* S.E_{ci} \end{split}$$

Prediction Interval

The prediction interval for a particular observation of the dependent variable.

This is the interval for any single value.

The prediction inteval takes into consideration the fact that you don't know the true equatio, and the fact the the liner regression explaned only part of the variance (the part is R-squared). Even if we would know the true equation then the width of this interval would be gre

Since this interval is for a single observation, the standa

$$S.E_{prediction}^2 = S_{residual}^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_x} \right)$$

 $\hat{Y} \pm T_{1\text{-}\alpha/2} (n\text{-}2) * S. E_{prediction}$

How to calculate R squares?

R squares is the percentage of the variance explain by the regression (SS_{Regression}) from the overall variance (SS_{Total}).

$$R^2 = \frac{SS_{Regression}}{SS_{Total}}$$

Linear regression in calculator

This online calculator supports all the basic functionality and more.

The right-tailed F test checks if the entire regression model is statistically significant. Why only right tail?

For Multiple regression calculator with the stepwise method and assumptions validations: multiple regression calculator

The following statistic checks if the linear regression model supports better results than the average of Y.

Hypotheses

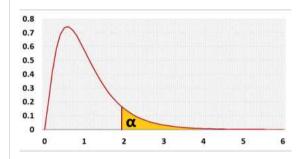
 H_0 : Y = b_0

 H_1 : $Y = b_0 + b_1 X$

Test statistic

 $F = \frac{MS(regression)}{1}$ MS (residual)

F distribution



R Code

The following R code should produce similar results

rm(list = ls())

if(!"car" %in% installed.packages()){install.packages("car")}

library(car)

x10 <-

x11 <-

x1 <- c(x10,x11)

y10 <-

y1 <- c(y10,y11) model1 = Im(y1~x1)

summary(model1)

What is linear regression?

Tutorial

Calculators

Correlation

Regression sample size