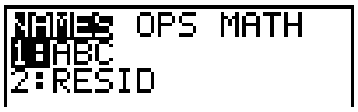
Regression Model Features

STAT CALC menu items **3** through **C** are regression models (page 12-24). The automatic residual list and automatic regression equation features apply to all regression models. Diagnostics display mode applies to some regression models.

Automatic Residual List

When you execute a regression model, the automatic residual list feature computes and stores the residuals to the list name **RESID**. **RESID** becomes an item on the LIST NAMES menu (Chapter 11).

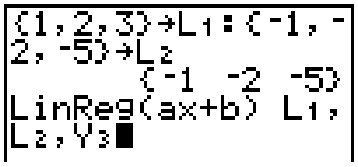


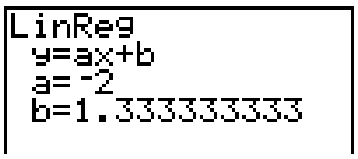
The TI-82 STATS uses the formula below to compute **RESID** list elements. The next section describes the variable **RegEQ**.

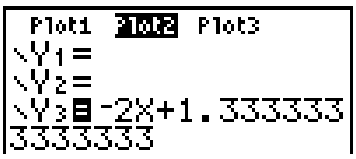
RESID = *Ylistname* − **RegEQ**(*Xlistname*)

Automatic Regression Equation

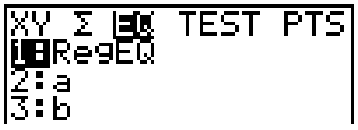
Each regression model has an optional argument, *regequ*, for which you can specify a Y= variable such as **Y1**. Upon execution, the regression equation is stored automatically to the specified Y= variable and the Y= function is selected.







Regardless of whether you specify a Y= variable for *regequ*, the regression equation always is stored to the TI-82 STATS variable **RegEQ**, which is item **1** on the VARS Statistics EQ secondary menu.



Note: For the regression equation, you can use the fixed-decimal mode setting to control the number of digits stored after the decimal point (Chapter 1). However, limiting the number of digits to a small number could affect the accuracy of the fit.

Diagnostics Display Mode

When you execute some regression models, the TI-82 STATS computes and stores diagnostics values for **r** (correlation coefficient) and $\mathbf{r^2}$ (coefficient of determination) or for $\mathbf{R^2}$ (coefficient of determination).

r and $\mathbf{r^2}$ are computed and stored for these regression models.

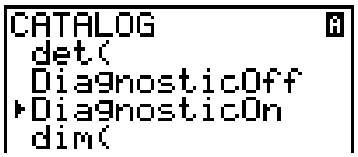
LinReg(ax+b)	**LnReg**	**PwrReg**
LinReg(a+bx)	**ExpReg**	

$\mathbf{R^2}$ is computed and stored for these regression models.

QuadReg	**CubicReg**	**QuartReg**

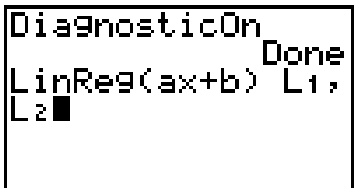
The **r** and $\mathbf{r^2}$ that are computed for **LnReg**, **ExpReg**, and **PwrReg** are based on the linearly transformed data. For example, for **ExpReg** (y=ab^x), **r** and $\mathbf{r^2}$ are computed on ln y=ln a+x(ln b).

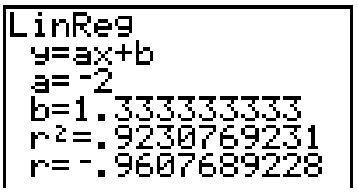
By default, these values are not displayed with the results of a regression model when you execute it. However, you can set the diagnostics display mode by executing the **DiagnosticOn** or **DiagnosticOff** instruction. Each instruction is in the CATALOG (Chapter 15).



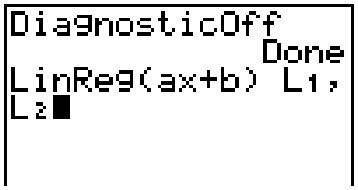
Note: To set **DiagnosticOn** or **DiagnosticOff** from the home screen, press 2nd [CATALOG], and then select the instruction for the mode you want. The instruction is pasted to the home screen. Press ENTER to set the mode.

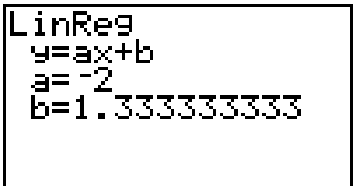
When **DiagnosticOn** is set, diagnostics are displayed with the results when you execute a regression model.





When **DiagnosticOff** is set, diagnostics are not displayed with the results when you execute a regression model.





STAT CALC Menu

To display the STAT CALC menu, press [STAT] [▶].

EDIT CALC TESTS	
1: 1-Var Stats	Calculates 1-variable statistics.
2: 2-Var Stats	Calculates 2-variable statistics.
3: Med-Med	Calculates a median-median line.
4: LinReg(ax+b)	Fits a linear model to data.
5: QuadReg	Fits a quadratic model to data.
6: CubicReg	Fits a cubic model to data.
7: QuartReg	Fits a quartic model to data.
8: LinReg(a+bx)	Fits a linear model to data.
9: LnReg	Fits a logarithmic model to data.
0: ExpReg	Fits an exponential model to data.
A: PwrReg	Fits a power model to data.
B: Logistic	Fits a logistic model to data.
C: SinReg	Fits a sinusoidal model to data.

For each STAT CALC menu item, if neither *Xlistname* nor *Ylistname* is specified, then the default list names are **L1** and **L2**. If you do not specify *freqlist*, then the default is **1** occurrence of each list element.

Frequency of Occurrence for Data Points

For most STAT CALC menu items, you can specify a list of data occurrences, or frequencies (*freqlist*).

Each element in *freqlist* indicates how many times the corresponding data point or data pair occurs in the data set you are analyzing.

For example, if **L1={15,12,9,14}** and **ʟFREQ={1,4,1,3}**, then the TI-82 STATS interprets the instruction **1-Var Stats L1, ʟFREQ** to mean that **15** occurs once, **12** occurs four times, **9** occurs once, and **14** occurs three times.

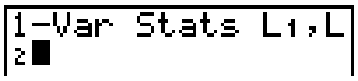
Each element in *freqlist* must be ≥ 0, and at least one element must be > 0.

Noninteger *freqlist* elements are valid. This is useful when entering frequencies expressed as percentages or parts that add up to 1. However, if *freqlist* contains noninteger frequencies, **Sx** and **Sy** are undefined; values are not displayed for **Sx** and **Sy** in the statistical results.

1-Var Stats

1-Var Stats (one-variable statistics) analyzes data with one measured variable. Each element in *freqlist* is the frequency of occurrence for each corresponding data point in *Xlistname*. *freqlist* elements must be real numbers > 0.

1-Var Stats [*Xlistname*,*freqlist*]



2-Var Stats

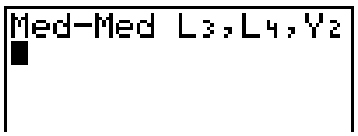
2-Var Stats (two-variable statistics) analyzes paired data. *Xlistname* is the independent variable. *Ylistname* is the dependent variable. Each element in *freqlist* is the frequency of occurrence for each data pair (*Xlistname,Ylistname*).

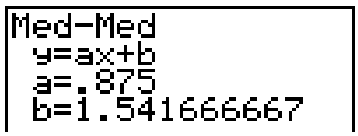
2-Var Stats [*Xlistname*,*Ylistname*,*freqlist*]

Med-Med (ax+b)

Med-Med (median-median) fits the model equation y=ax+b to the data using the median-median line (resistant line) technique, calculating the summary points **x1**, **y1**, **x2**, **y2**, **x3**, and **y3**. **Med-Med** displays values for **a** (slope) and **b** (y-intercept).

Med-Med [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]





LinReg (ax+b)

LinReg(ax+b) (linear regression) fits the model equation y=ax+b to the data using a least-squares fit. It displays values for **a** (slope) and **b** (y-intercept); when **DiagnosticOn** is set, it also displays values for $\mathbf{r^2}$ and **r**.

LinReg(ax+b) [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

QuadReg (ax2+bx+c)

QuadReg (quadratic regression) fits the second-degree polynomial $y=ax^2+bx+c$ to the data. It displays values for **a**, **b**, and **c**; when **DiagnosticOn** is set, it also displays a value for $\mathbf{R^2}$. For three data points, the equation is a polynomial fit; for four or more, it is a polynomial regression. At least three data points are required.

QuadReg [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

CubicReg (ax^3+bx^2+cx+d)

CubicReg (cubic regression) fits the third-degree polynomial $y=ax^3+bx^2+cx+d$ to the data. It displays values for **a**, **b**, **c**, and **d**; when **DiagnosticOn** is set, it also displays a value for R^2. For four points, the equation is a polynomial fit; for five or more, it is a polynomial regression. At least four points are required.

CubicReg [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

QuartReg ($ax^4+bx^3+cx^2+dx+e$)

QuartReg (quartic regression) fits the fourth-degree polynomial $y=ax^4+bx^3+cx^2+dx+e$ to the data. It displays values for **a**, **b**, **c**, **d**, and **e**; when **DiagnosticOn** is set, it also displays a value for R^2. For five points, the equation is a polynomial fit; for six or more, it is a polynomial regression. At least five points are required.

QuartReg [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

LinReg (a+bx)

LinReg(a+bx) (linear regression) fits the model equation y=a+bx to the data using a least-squares fit. It displays values for **a** (y-intercept) and **b** (slope); when **DiagnosticOn** is set, it also displays values for r^2 and **r**.

LinReg(a+bx) [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

LnReg (a+b ln(x))

LnReg (logarithmic regression) fits the model equation y=a+b ln(x) to the data using a least-squares fit and transformed values ln(x) and y. It displays values for **a** and **b**; when **DiagnosticOn** is set, it also displays values for r^2 and **r**.

LnReg [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

ExpReg (ab^x)

ExpReg (exponential regression) fits the model equation $y=ab^x$ to the data using a least-squares fit and transformed values x and ln(y). It displays values for **a** and **b**; when **DiagnosticOn** is set, it also displays values for r^2 and **r**.

ExpReg [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

PwrReg
(ax^b)

PwrReg (power regression) fits the model equation $y=ax^b$ to the data using a least-squares fit and transformed values ln(x) and ln(y). It displays values for **a** and **b**; when **DiagnosticOn** is set, it also displays values for r^2 and **r**.

PwrReg [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

Logistic
$c/(1+a*e^{-bx})$

Logistic fits the model equation $y=c/(1+a*e^{-bx})$ to the data using an iterative least-squares fit. It displays values for **a**, **b**, and **c**.

Logistic [*Xlistname*,*Ylistname*,*freqlist*,*regequ*]

SinReg
a sin(bx+c)+d

SinReg (sinusoidal regression) fits the model equation y=a sin(bx+c)+d to the data using an iterative least-squares fit. It displays values for **a**, **b**, **c**, and **d**. At least four data points are required. At least two data points per cycle are required in order to avoid aliased frequency estimates.

SinReg [*iterations*,*Xlistname*,*Ylistname*,*period*,*regequ*]

iterations is the maximum number of times the algorithm will iterate to find a solution. The value for *iterations* can be an integer ≥ 1 and ≤ 16; if not specified, the default is 3. The algorithm may find a solution before *iterations* is reached. Typically, larger values for *iterations* result in longer execution times and better accuracy for **SinReg**, and vice versa.

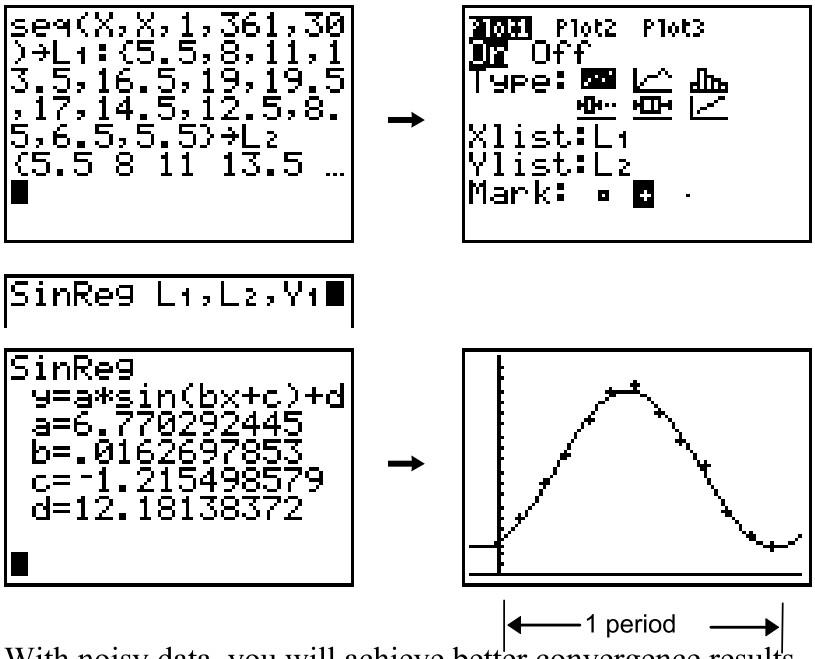
A *period* guess is optional. If you do not specify *period*, the difference between time values in *Xlistname* must be equal and the time values must be ordered in ascending sequential order. If you specify *period*, the algorithm may find a solution more quickly, or it may find a solution when it would not have found one if you had omitted a value for *period*. If you specify *period*, the differences between time values in *Xlistname* can be unequal.

Note: The output of **SinReg** is always in radians, regardless of the **Radian/Degree** mode setting.

A **SinReg** example is shown on the next page.

SinReg Example: Daylight Hours in Alaska for One Year

Compute the regression model for the number of hours of daylight in Alaska during one year.



With noisy data, you will achieve better convergence results when you specify an accurate estimate for *period*. You can obtain a *period* guess in either of two ways.

- Plot the data and trace to determine the x-distance between the beginning and end of one complete period, or cycle. The illustration above and to the right graphically depicts a complete period, or cycle.
- Plot the data and trace to determine the x-distance between the beginning and end of N complete periods, or cycles. Then divide the total distance by N.

After your first attempt to use **SinReg** and the default value for *iterations* to fit the data, you may find the fit to be approximately correct, but not optimal. For an optimal fit, execute **SinReg 16**,*Xlistname*,*Ylistname*,**2π / *b*** where *b* is the value obtained from the previous **SinReg** execution.