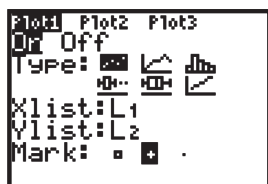


# 3

## Relationships Between Two Quantitative Variables

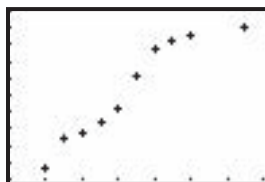
### Calculator Note 3A: Scatterplots

The TI-83 Plus and TI-84 Plus can display scatterplots of bivariate data contained in any two lists, say, lists L1 and L2. This particular example uses the mean net income for family practitioners from Display 3.17 on page 121 of the student book.



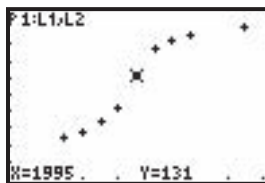
Go to the Stat Plot screen,  $\boxed{2\text{nd}}$  [STAT PLOT], and define a scatterplot. Scatterplots are the first option under Type.

Press  $\boxed{\text{ZOOM}}$  and select 9:ZOOMSTAT to fill the Graph window, or define an appropriate window.



[1988, 2002, 2, 100, 150, 5]

When you trace the scatterplot, you see the coordinates of the data points.

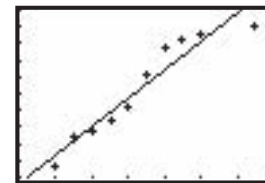
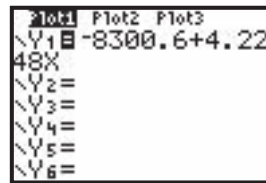
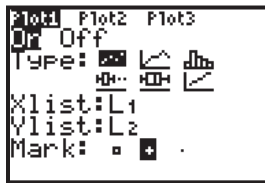


[1988, 2002, 2, 100, 150, 5]

It is important to keep in mind which is the independent variable (Xlist) and which is the dependent variable (Ylist).

### Calculator Note 3B: Graphing a Line on a Scatterplot

To graph a line on a scatterplot, you must separately define the scatterplot and enter the equation of the line into one of the functions on the Y= screen. This particular example uses the mean net income for family practitioners from the example on pages 121–122 of the student book.



[1988, 2002, 2, 100, 150, 5]

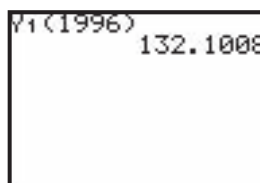
On a TI-84 Plus, you can use the Manual-Fit feature to place a line on the screen, and adjust it by changing the steepness and  $y$ -intercept manually on the Graph screen. You access this command by pressing **[STAT]**, arrowing over to **CALC**, and selecting **D: Manual-Fit**. If you press **[Y=]** while the line is on the screen, the equation is stored into  $Y_1$ .

## Calculator Note 3C: Finding Function Values

### Calculating Function Values on the Home Screen

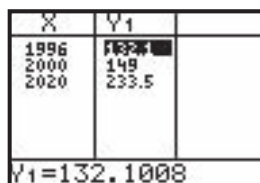
Once an equation is entered into the  $Y=$  screen, you can evaluate the function at any value of the independent variable on the Home screen. To access the function names, press **[VARS]**, arrow over to **Y-VARS**, and select **1:Function**.

For example, with  $Y_1 = -8300.6 + 4.2248X$  defined, you can predict the  $y$ -coordinate when  $x = 1996$  by entering  $Y_1(1996)$ .



### Using TABLE Features

Once an equation is entered into the  $Y=$  screen, you can evaluate the function at multiple values of the independent variable on the Table screen. Press **[2nd]** **[TBLSET]** to access the Table Setup screen. Set the independent variable to **Ask** so that you can enter any value for  $x$ . Press **[2nd]** **[TABLE]** to view the Table screen, and enter the desired values for  $x$ .



## Calculator Note 3D: Finding Residuals

The TI-83 Plus and TI-84 Plus will efficiently calculate residuals for any line of fit using the spreadsheet capabilities of the List Editor.

- First make sure that your data are entered into lists  $L_1$  and  $L_2$  and that the equation of your fitted line is entered into the  $Y=$  screen.
- Define list  $L_3$  as the predicted  $y$ -coordinates,  $Y_1(L_1)$ .

L1	L2	RES
1990	103	106.75
1991	112	110.98
1992	114	115.2
1993	117	119.43
1994	121	123.65
1995	131	127.88
1996	139	132.1

$L_3 = "Y_1(L_1)"$

- c. Then define list  $L_4$  as the residuals,  $L_2 - L_3$ .

L2	L3	#	L4	#
103	106.75		-3.752	
112	110.98		1.0232	
114	115.2		-1.202	
117	119.43		-2.426	
121	123.65		-2.651	
131	127.88		3.124	
139	132.1		6.8992	

L4 = "L2 - L3"

Alternatively, you could directly define list  $L_3$  as the residuals using the expression  $L_2 - Y_1(L_1)$ .

L1	L2	L3	#
1990	103	-3.752	
1991	112	1.0232	
1992	114	-1.202	
1993	117	-2.426	
1994	121	-2.651	
1995	131	3.124	
1996	139	6.8992	

L3(1) = -3.752

### Calculator Note 3E: Sum of Squared Errors (SSE)

If the residuals for a fitted line are contained in a list, say, list  $L_4$ , the sum of squared errors (SSE) can be calculated. (See Calculator Note 3D to learn how to calculate the residual.) The example that follows uses the passenger jet data from page 123 of the student book.

- a. Define list  $L_5$  as the squares of the residuals,  $L_4^2$ .

L4	#	L5	#	L6	#
-67		4489			
133		17689			
-67		4489			
---		---			
---		---			

L5 = "L4 ^ 2"

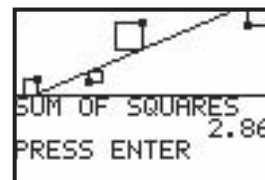
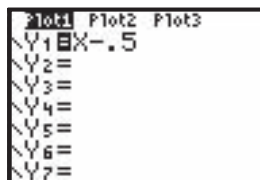
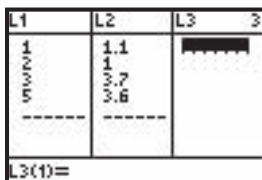
- b. On the Home screen, calculate the sum of list  $L_5$ . Find the `sum(` command by pressing `[2nd]` `[LIST]`, arrowing over to `MATH`, and selecting `5:sum(`. The result is the SSE.

SUM(L5)	26667
---------	-------

### Calculator Note 3F: Using the Calculator to Visualize the SSE—The SQUARES Program

To visualize the squared errors and calculate the sum of squared errors, use the SQUARES program. Enter your data into  $L_1$  and  $L_2$ , enter your line into  $Y_1$ , and set the window appropriately. To run the SQUARES program, press `[PRGM]`, select SQUARES, and press `[ENTER]`. Your screen will be split horizontally—the upper half of the screen will show the data and line, as well as squares with a vertical

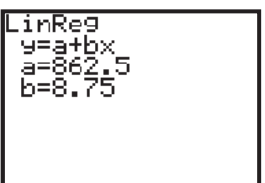
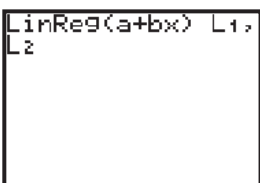
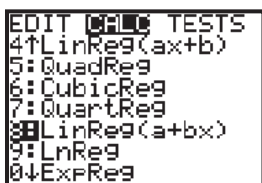
edge drawn from each data point to the corresponding point on the line. The lower half of the screen shows the sum of the squared errors, which is the sum of the areas of the squares. Press **ENTER** to end the program and restore the screen to full size.



### Calculator Note 3G: Using the Calculator to Find the Least Squares Regression Line

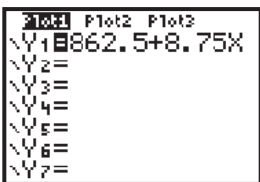
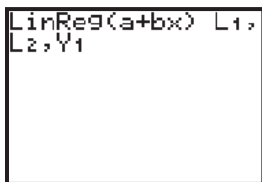
The TI-83 Plus and TI-84 Plus provide two forms of the least squares regression line:  $y = ax + b$  and  $y = a + bx$ . The **LinReg(ax+b)** and **LinReg(a+bx)** commands, as well as other regression techniques, are found by pressing **STAT** and arrowing over to **CALC**. As explained on pages 96–97 of the *Statistics in Action Instructor's Guide, Volume 1*, the form  $y = a + bx$  has been accepted as the standard notation in statistics.

Although the regression commands default to lists  $L_1$  and  $L_2$ , it is strongly recommended that you *always* include the names of the lists—independent, then dependent—separated by commas.

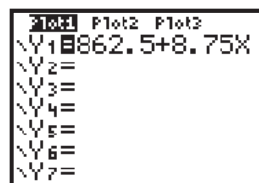
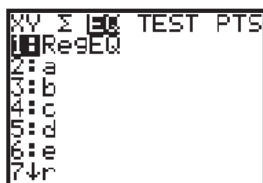
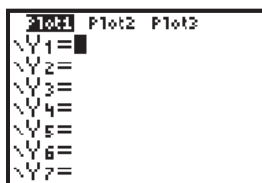


### Pasting the Equation of the Least Squares Line into the Y= Screen

You may paste the equation of the least squares line into the **Y=** screen in one of two ways. First, when using the regression command (see above), you can specify a function after specifying the lists, for example, **LinReg(a+bx) L1, L2, Y1**. The calculator will automatically put the equation into the **Y=** screen.

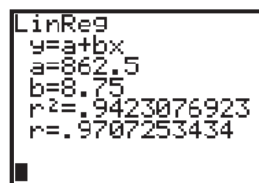
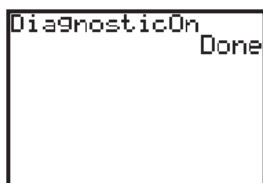
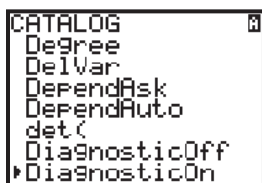


Second, you can paste the equation into the **Y=** screen after running the regression command. Press **Y=** and position the cursor beside one of the available functions. Press **VARS**, select 5:Statistics, arrow over to **EQ**, and select 1:RegEq to paste the equation in place.



### Calculator Note 3H: Finding $r$ and $r^2$

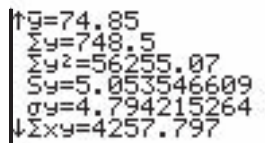
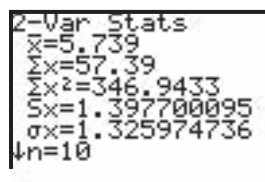
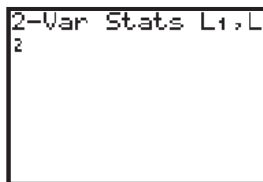
If you'd like the calculator to calculate and display the correlation,  $r$ , and coefficient of variation,  $r^2$ , you must first turn Diagnostics on. Press  $\text{2nd}$ [CATALOG], press  $\text{ALPHA}$  [D], arrow down to DiagnosticOn, and press  $\text{ENTER}$   $\text{ENTER}$ . Now, when you perform a regression, the values of  $r^2$  and  $r$  are shown.



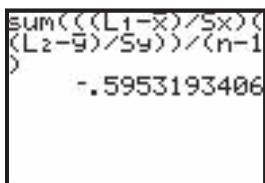
### A Formula for the Correlation, $r$

If desired, the spreadsheet capabilities of the List Editor can aid the calculation of the correlation. This process resembles Display 3.42 on page 142 of the student book.

- Enter the bivariate data into two lists, say, lists  $L_1$  and  $L_2$ .
- Press  $\text{STAT}$ , arrow over to CALC, and select 2:2-VarStats and specify the lists. Performing a 2-Var Stats calculates and stores the means,  $\bar{x}$  and  $\bar{y}$ , the standard deviations,  $s_x$  and  $s_y$ , and the number of data values,  $n$ . All of these variables can now be accessed by pressing  $\text{VAR}$ , and selecting 5:Statistics.



- Press  $\text{STAT}$  and select 1:Edit to return to the List Editor screen.
- Define list  $L_3$  with the expression  $(L_1 - \bar{X})/s_x$ .
- Define list  $L_4$  with the expression  $(L_2 - \bar{Y})/s_y$ .
- Define list  $L_5$  with the expression  $L_3 * L_4$ .
- On the Home screen, enter  $\text{sum}(L_5)Y(n-1)$ . The result is the correlation,  $r$ . Find  $\text{sum}$ ( by pressing  $\text{2nd}$  [LIST], arrowing over to MATH, and selecting 5:sum(.



Alternatively, after running the 2-Var Stats, you can calculate  $r$  directly on the Home screen by entering  $\text{sum}(((L1-\bar{x})/S_x)((L2-\bar{y})/S_y))/(n-1)$ .

### Calculator Note 3I: Residual Plots

Each time a TI-83 Plus or TI-84 Plus performs a regression, the calculator automatically computes the residuals and stores them in a list called RESID. Access this list by pressing  $\boxed{2\text{nd}}$  [LIST] and looking in the NAMES menu.

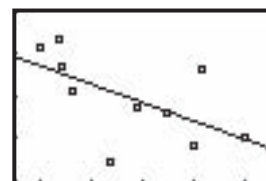
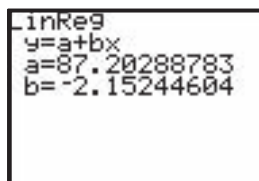


Note that you can easily verify a necessary condition of residuals—that they sum to zero—by summing the values in RESID.

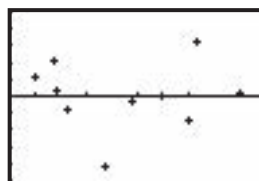
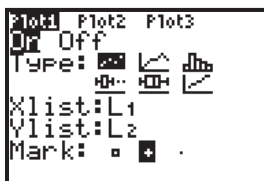
You can easily make a residual plot by creating a scatterplot with RESID as the Ylist. Here are the original plot and the residual plot for the data on percentage of on-time arrivals versus mishandled baggage from Display 3.68 on page 169 of the student book.

L1	L2	L3	1
4	80.9		
4.36	81.9		
4.42	78.4		
4.62	75.7		
5.5	63.3		
6.5	70.8		
6.5	79.1		

L1(4)=4



[3.5, 8.5, 1, 65, 85, 5]



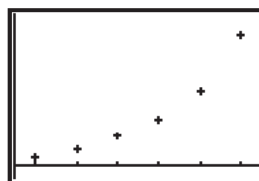
[3.5, 8.5, 1, -10, 10, 2]

### Calculator Note 3J: Shape-Changing Transformations

The TI-83 Plus and TI-84 Plus can accomplish all of the transformations that are discussed in Section 3.5. However, the calculator always returns the coefficients of an equation in the form  $y = a + bx$ . It is up to you to note which transformations were used on which variables. For example, consider this small hypothetical data set that appears to exhibit an exponential relationship. Assume this data set relates days and growth, or data points in the form  $(d, g)$ .

L1	L2	L3	2
1	10		
1.5	15		
2	20		
3	30		
4	45		
5	60		
6	90		
7	135		
8	200		

L2(7) =



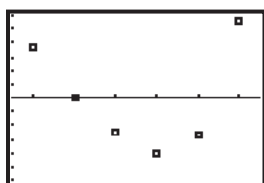
[0.5, 6.5, 1, -15.5, 185.5, 1]

A least squares regression on the original data gives the equation  $y = -36.333 + 28x$ . You should interpret this as  $\hat{g} = -36.333 + 28d$ .

```
LinReg(a+bx) L1,
L2, Y1
```

```
LinReg
y=a+bx
a=-36.33333333
b=28
r2=.8776119403
r=.9368094472
```

A residual plot confirms that a linear equation is a bad fit.



[0.5, 6.5, 1, -30, 30, 5]

Now use the List Editor screen to take the logarithm (base 10) of each of the growth values (list L2). Store these transformed values in list L3.

L1	L2	L3	# 2
1	10	1	
2	20	1.301	
3	35	1.5441	
4	55	1.7404	
5	90	1.9542	
6	160	2.2041	

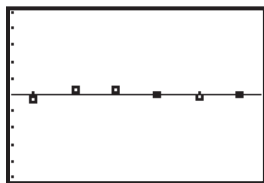
L3="log(L2)"

Performing a least squares regression on the transformed data,  $(d, \log(g))$ , gives the equation  $y = 0.806 + 0.234x$ . You should interpret this as  $\widehat{\log(g)} = 0.806 + 0.234d$ .

```
LinReg(a+bx) L1,
L3, Y2
```

```
LinReg
y=a+bx
a=.806317327
b=.2336152029
r2=.9956910377
r=.997843193
```

A residual plot shows that the new equation is a better fit.



[0.5, 6.5, 1, -0.5, 0.5, 0.1]