

Graphing Calculator Lab

Lines of Regression

Casio FX-9750G

You can use a Casio FX-9750G graphing calculator to find a function that best fits a set of data. The graph of a linear function that models a set of data is called a **regression line** or **line of best fit**. You can also use the calculator to draw scatter plots and make predictions.

ACTIVITY

INCOME The table shows the median income of U.S. families for the period 1970–2002.

Year	1970	1980	1985	1990	1995	1998	2000	2002
Income (\$)	9867	21,023	27,735	35,353	40,611	46,737	50,732	51,680

Source: U.S. Census Bureau

Make a scatter plot of the data. Find a function and graph a regression line. Then use the function to predict the median income in 2015.

Step 1 Make a scatter plot.

- Enter the years in L1 and the income in L2.

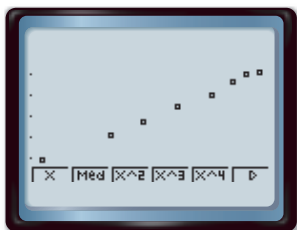
KEYSTROKES: **MENU** 2 1970 **EXE** 1980
EXE ...

- Set the viewing window to fit the data.

KEYSTROKES: **SHIFT** **F3** 1965 **EXE** 2015
EXE 5 **EXE** 0 **EXE** 55000
EXE 10000 **EXE** **EXIT**

- Use STAT PLOT to graph a scatter plot.

KEYSTROKES: **F1** **F1**

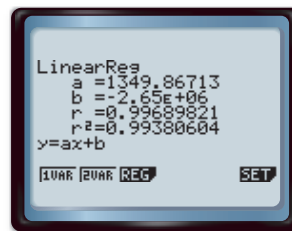


[1965, 2015] scl: 5 by [0, 55,000] scl: 10,000

Step 2 Find the equation of a regression line.

- Find the regression equation by selecting LinReg(ax + b) on the STAT CALC menu.

KEYSTROKES: **EXIT** **EXIT** **F2** **F3** **F1**

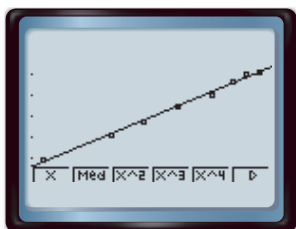


The regression equation is about $y = 1349.87x - 2,650,768.34$. The slope indicates that family incomes were increasing at a rate of about \$1350 per year.

The number r is called the **linear correlation coefficient**. The closer the value of r is to 1 or -1 , the closer the data points are to the line. In this case, r is very close to 1 so the line fits the data well. *If the values of r^2 and r are not displayed, use DiagnosticOn from the CATALOG menu.*

Step 3 Graph the regression equation.

KEYSTROKES: **EXIT** **EXIT** **F1** **F1** **F1**
F6



[1965, 2015] scl: 5 by [0, 55,000] scl: 10,000

The graph of the line will be displayed with the scatter plot. Notice that the regression line seems to pass through only one of the data points, but comes close to all of them. As the correlation coefficient indicated, the line fits the data very well.

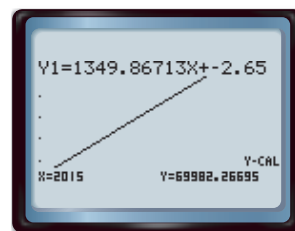
Step 4 Predict using the function.

Find y when $x = 2015$. Copy the equation to the $Y=$ list.

KEYSTROKES: **F5** **EXE** **MENU** 5 **F6**

Use $Y=$ cal on the $G-Solv$ menu.

KEYSTROKES: **SHIFT** **G-solve** **F6** **F1**
2015 **EXE**



According to the function, the median family income in 2015 will be about \$69,982. Because the function is a very good fit to the data, the prediction should be quite accurate.

EXERCISES

BASEBALL For Exercises 1–3, use the table at the right that shows the total attendance for minor league baseball in some recent years.

1. Make a scatter plot of the data.
2. Find a regression equation for the data.
3. Predict the attendance in 2010.

Year	Attendance (millions)
1985	18.4
1990	25.2
1995	33.1
2000	37.6

Source: National Association of Professional Baseball Leagues

GOVERNMENT For Exercises 4–6, use the table below that shows the population and the number of representatives in Congress for the most populous states.

State	CA	TX	NY	FL	IL	PA	OH
Population (millions)	35.5	22.1	19.2	17.0	12.7	12.4	11.4
Representatives	53	32	29	25	19	19	18

Source: *World Almanac*

4. Make a scatter plot of the data.
5. Find a regression equation for the data.
6. Predict the number of representatives for South Carolina, which has a population of about 4.1 million.

MUSIC For Exercises 7–11, use the table at the right that shows the percent of music sales that were made in record stores in the United States for the period 1995–2004.

7. Make a scatter plot of the data. Is the correlation of the data positive or negative? Explain.
8. Find a regression equation for the data.
9. According to the regression equation, what was the average rate of change of record store sales during the period?
10. Use the function to predict the percent of sales made in record stores in 2015.
11. How accurate do you think your prediction is? Explain.

Record Store Sales	
Year	Sales (percent)
1995	52
1996	49.9
1997	51.8
1998	50.8
1999	44.5
2000	42.4
2001	42.5
2002	36.8
2003	33.2
2004	32.5

Source: Recording Industry Association of America

RECREATION For Exercises 12–16, use the table at the right that shows the amount of money spent on sporting footwear in some recent years.

12. Find a regression equation for the data.
13. Use the regression equation to predict the sales in 2010.
14. Delete the outlier (1999, 12,546) from the data set and find a new regression equation for the data.
15. Use the new regression equation to predict the sales in 2010.
16. Compare the correlation coefficients for the two regression equations. Which function fits the data better? Which prediction would you expect to be more accurate?

Sporting Footwear Sales	
Year	Sales (\$ millions)
1998	13,068
1999	12,546
2000	13,026
2001	13,814
2002	14,144
2003	14,446
2004	14,752

Source: National Sporting Goods Association

EXTENSION

For Exercises 17–20, design and complete your own data analysis.

17. Write a question that could be answered by examining data. For example, you might estimate the number of students who will attend your school 5 years from now or predict the future cost of a piece of electronic equipment.
18. Collect and organize the data you need to answer the question you wrote. You may need to research your topic on the Internet or conduct a survey to collect the data you need.
19. Make a scatter plot and find a regression equation for your data. Then use the regression equation to answer the question.
20. Analyze your results. How accurate do you think your model is? Explain your reasoning.