Note: to submit the work click on Proj #1 Due Date 10/5

Focus on Modeling - Fitting Lines to Data

1-Read the sections posted in your textbook pages 139-146.
2- Practice using the calculator on how to plot a regression line (Example 1 –pg 140)
3- Use the information from pages 139-146 to work as a group of four to solve
2 page 144.
6 page 145
12 page 146
Once you solve these problems, write a short introduction to the topic and a conclusion by critically analyzing your results and how well do they apply to real life situations.

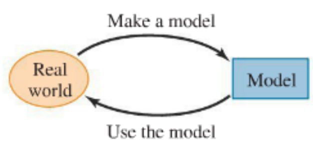
Open attached sheet.

Directions:

Don't wait until last minute to form your group. Submit your work as an attachment right here by clicking on the hyperlink
Form a group of at most four students to work on your end of the chapter # 1 project.
Post a call on the discussion board and from a coherent group of four. Work together by meeting online or offline depending on your group availability. Share the work by assigning individuals to each part and meet to discuss the items! Write up the project and submit as an attachment through the link inside this item.
Important note: each student should submit the same write-up through his own link the project so I can have a slot for posting your own grade an make sure to have the same write-up with all the names at the top.
The project takes time. Don’t wait until the last minute.
Note: Submit your project right here by clicking on the hyper link beside Proj# 1 and send it as an attachment through BB in order to have your grade posted in your online grade book. Don't send it to my email. Each student should submit through the link a unified copy of the group work with the list of names on the top of the document.

Final Note: Alternative to a calculator
Excel can be used as well if you prefer to, instead of a calculator to draw the regression line.
You can watch this video:
https://www.youtube.com/watch?v=ExfknNCvBYg

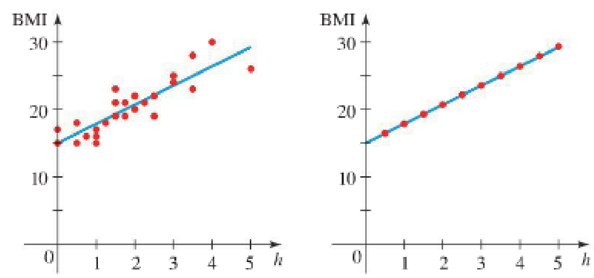
FOCUS ON MODELING Fitting Lines to Data



A model is a representation of an object or process. For example, a toy Ferrari is a model of the actual car; a road map is a model of the streets in a city. A **mathematical model** is a mathematical representation (usually an equation) of an object or process. Once a mathematical model has been made, it can be used to obtain useful information or make predictions about the thing being modeled. The process is described in the diagram in the margin. In these *Focus on Modeling* sections we explore different ways in which mathematics is used to model real-world phenomena.

The Line That Best Fits the Data

In Section 1.10 we used linear equations to model relationships between varying quantities. In practice, such relationships are discovered by collecting data. But real-world data seldom fall into a precise line. The **scatter plot** in Figure 1(a) shows the result of a study on childhood obesity. The graph plots the body mass index (BMI) versus the number of hours of television watched per day for 25 adolescent subjects. Of course, we would not expect the data to be exactly linear as in Figure 1(b). But there is a linear *trend* indicated by the blue line in Figure 1(a): The more hours a subject watches TV, the higher the BMI. In this section we learn how to find the line that best fits the data.



(a) Line of best fit (b) Line fits data exactly

FIGURE 1

Table 1 gives the nationwide infant mortality rate for the period from 1950 to 2000. The *rate* is the number of infants who die before reaching their first birthday, out of every 1000 live births.

TABLE 1
U.S. Infant Mortality

Year	Rate
1950	29.2
1960	26.0
1970	20.0
1980	12.6
1990	9.2
2000	6.9

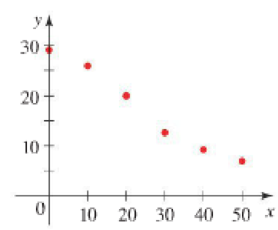


FIGURE 2 U.S. infant mortality rate

The scatter plot in Figure 2 shows that the data lie roughly on a straight line. We can try to fit a line visually to approximate the data points, but since the data aren't *exactly* linear, there are many lines that might seem to work. Figure 3 shows two attempts at "eyeballing" a line to fit the data.

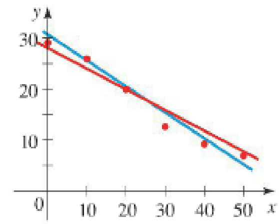


FIGURE 3 Visual attempts to fit line to data

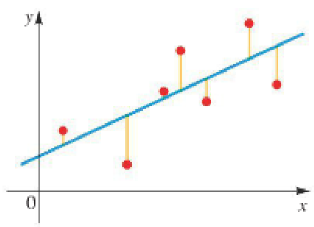


FIGURE 4 Distance from the data points to the line

Of all the lines that run through these data points, there is one that "best" fits the data, in the sense that it provides the most accurate linear model for the data. We now describe how to find this line.

It seems reasonable that the line of best fit is the line that is as close as possible to all the data points. This is the line for which the sum of the vertical distances from the data points to the line is as small as possible (see Figure 4). For technical reasons it is better to use the line where the sum of the squares of these distances is smallest. This is called the **regression line**. The formula for the regression line is found by using calculus, but fortunately, the formula is programmed into most graphing calculators. In Example 1 we see how to use a TI-83 calculator to find the regression line for the infant mortality data described above. (The process for other calculators is similar.)

EXAMPLE 1 ■ Regression Line for U.S. Infant Mortality Rates

(a) Find the regression line for the infant mortality data in Table 1.

(b) Graph the regression line on a scatter plot of the data.

(c) Use the regression line to estimate the infant mortality rates in 1995 and 2006.

SOLUTION

(a) To find the regression line using a TI-83 calculator, we must first enter the data into the lists L_1 and L_2, which are accessed by pressing the STAT key and selecting `Edit`. Figure 5 shows the calculator screen after the data have been entered. (Note that we are letting $x = 0$ correspond to the year 1950 so that $x = 50$ corresponds to 2000. This makes the equations easier to work with.) We then press the STAT key again and select `Calc`, then `4:LinReg(ax+b)`, which provides the output shown in Figure 6(a). This tells us that the regression line is

$$y = -0.48x + 29.4$$

Here x represents the number of years since 1950, and y represents the corresponding infant mortality rate.

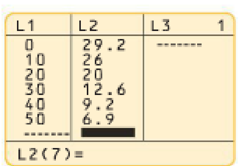
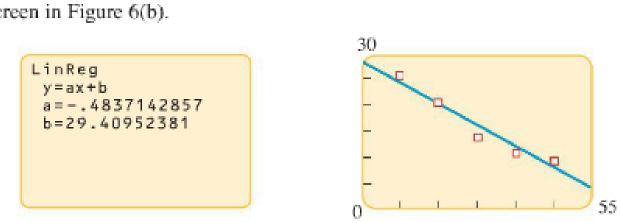


FIGURE 5 Entering the data

(b) The scatter plot and the regression line have been plotted on a graphing calculator screen in Figure 6(b).



(a) Output of the `LinReg` command

(b) Scatter plot and regression line

FIGURE 6

(c) The year 1995 is 45 years after 1950, so substituting 45 for x, we find that $y = -0.48(45) + 29.4 = 7.8$. So the infant mortality rate in 1995 was about 7.8. Similarly, substituting 56 for x, we find that the infant mortality rate predicted for 2006 was about $-0.48(56) + 29.4 \approx 2.5$. ■

Demand for Soft Drinks A convenience store manager notices that sales of soft drinks are higher on hotter days, so he assembles the data in the table.

(a) Make a scatter plot of the data.

(b) Find and graph a linear function that models the data.

(c) Use the model to predict soft drink sales if the temperature is 95°F.

High temperature (°F)	Number of cans sold
55	340
58	335
64	410
68	460
70	450
75	610
80	735
84	780

Extent of Arctic Sea Ice The National Snow and Ice Data Center monitors the amount of ice in the Arctic year round. The table below gives approximate values for the sea ice extent in millions of square kilometers from 1986 to 2012, in two-year intervals.

(a) Make a scatter plot of the data.

(b) Find and graph the regression line.

(c) Use the linear model in part (b) to estimate the ice extent in the year 2016.

Year	Ice extent (million km^2)	Year	Ice extent (million km^2)
1986	7.5	2000	6.3
1988	7.5	2002	6.0
1990	6.2	2004	6.0
1992	7.5	2006	5.9
1994	7.2	2008	4.7
1996	7.9	2010	4.9
1998	6.6	2012	3.6

Source: National Snow and Ice Data Center

Demand for Candy Bars In this problem you will determine a linear demand equation that describes the demand for candy bars in your class. Survey your classmates to determine what price they would be willing to pay for a candy bar. Your survey form might look like the sample to the left.

(a) Make a table of the number of respondents who answered "yes" at each price level.

(b) Make a scatter plot of your data.

(c) Find and graph the regression line $y = mp + b$, which gives the number of responents y who would buy a candy bar if the price were p cents. This is the *demand equation*. Why is the slope m negative?

(d) What is the p-intercept of the demand equation? What does this intercept tell you about pricing candy bars?

Would you buy a candy bar from the vending machine in the hallway if the price is as indicated?

Price	Yes or No
50¢	
75¢	
\$1.00	
\$1.25	
\$1.50	
\$1.75	
\$2.00	