6.2 Lines of Best Fit

YOU WILL NEED

- ruler
- grid paper

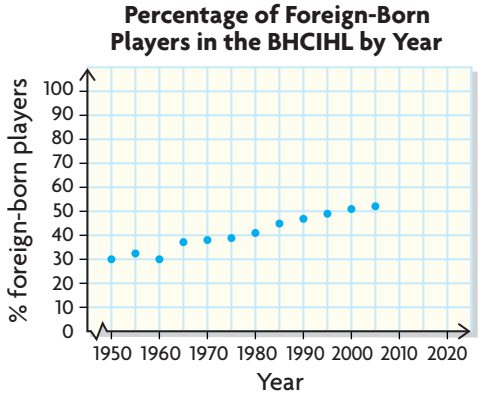
GOAL

Sketch a line of best fit for a given set of data and determine the equation of the line.

LEARN ABOUT the Math

Over the past half century, the percentage of foreign-born players in the Black Horse Corners International Hockey League (BHCIHL) has increased. The data are summarized in the table and on the scatter plot.

Year	Percentage of Foreign-Born Players in the BHCIHL
1950	30
1955	33
1960	30
1965	37
1970	38
1975	39
1980	41
1985	45
1990	47
1995	49
2000	51
2005	52

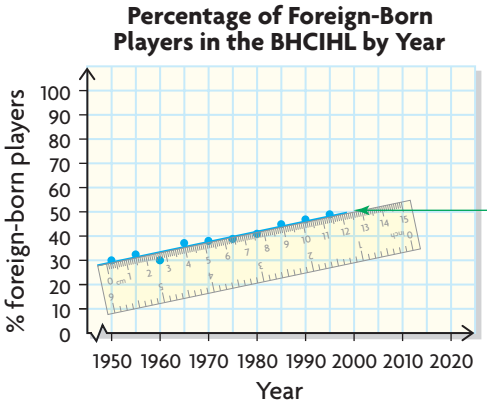


? How might the percentage of foreign-born players in BHCIHL in the year 2020 compare with the percentage in 1993?

EXAMPLE 1 Representing a situation using a line of best fit

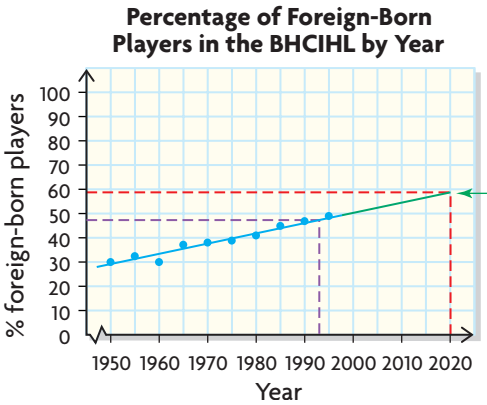
Use the scatter plot to estimate, and then, compare the number of foreign-born players in 2020 and in 1993.

Ryan's Solution: Using a graphing strategy



I drew a **line of best fit** to help me see the **trend**. I placed my transparent ruler over the plotted points so that most of them were close to the edge. I tried to "balance" the points on either side of the line.

I can see that the percentage of foreign-born players in the league is increasing as time passes.



I assumed that this increasing trend would continue and I extended the line of best fit until it went to 2020.

If the trend continues, the percentage of foreign-born players in the BHCIHL in 2020 will probably be about 59%.

I used the extended line of best fit to **extrapolate** the percentage for 2020.

The percentage of foreign-born players in the BHCIHL in 1993 was probably about 48%.

I used the line of best fit to **interpolate** the percentage for 1993.

Based on the trend I see, the number of foreign-born players in the BHCIHL will likely be quite a bit higher in 2020 than it was in 1993.

line of best fit

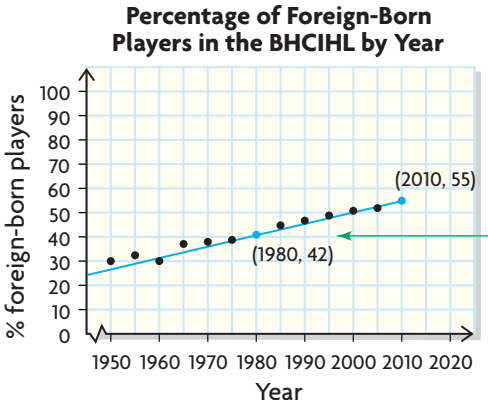
a line that best describes the relationship between two variables in a scatter plot

trend

a relationship between two variables for which the independent variable is time

Omar determined an equation to describe the relationship between the percentage of foreign-born players and time. He then used his equation to make a prediction for the year 2020.

Omar's Solution: Using an algebraic strategy



I used a transparent ruler to draw a line of best fit.

$(x_1, y_1) = (1980, 42)$,
$(x_2, y_2) = (2010, 55)$

I wrote down the coordinates of two points that lie on the line of best fit.

$$\text{m} = \frac{y_2 - y_1}{x_2 - x_1}$$
$$= \frac{55 - 42}{2010 - 1980}$$
$$= \frac{13}{30}$$

I used the coordinates to calculate the slope of the line.

$$y = \text{m}x + \text{b}$$
$$y = \left(\frac{13}{30}\right)x + \text{b}$$
$$55 = \left(\frac{13}{30}\right)(2010) + \text{b}$$
$$55 = 871 + \text{b}$$
$$\text{b} = 55 - 871$$
$$\text{b} = -816$$

I wrote the equation for the line of best fit. I used the coordinates of a known point on the line, (2010, 55), to determine the value of b.

$$y = \frac{13}{30}x - 816$$

I wrote the final equation for the line of best fit by substituting the slope m and the y-intercept b.

$$y = \left(\frac{13}{30}\right)(2020) - 816$$
$$\doteq 59.3$$

I substituted the value $x = 2020$ to estimate the percentage of foreign-born players that will be in the BHCIHL by the year 2020.

I estimate that in 2020, about 59.3% of players in the BHCIHL will have been born outside of Canada.

$$y = \left(\frac{13}{30}\right)(1993) - 816$$
$$\doteq 47.6$$

I substituted the value $x = 1993$ to estimate the percentage of foreign-born players that were in the BHCIHL in 1993.

I estimate that in 1993, about 47.6% of players in the BHCIHL had been born outside of Canada.

These values are reasonable, since they are very close to the values that I could read off the graph of the line of best fit.

If the trend continues, in 2020 there will be almost 12% more foreign-born players in the league than there were in 1993.

Reflecting

A. Different students may use their rulers to draw different lines of best fit. How might that affect their estimates?

B. Do Ryan's and Omar's solutions seem to give similar results? Explain.

C. Which boy's strategy would you choose? Why?

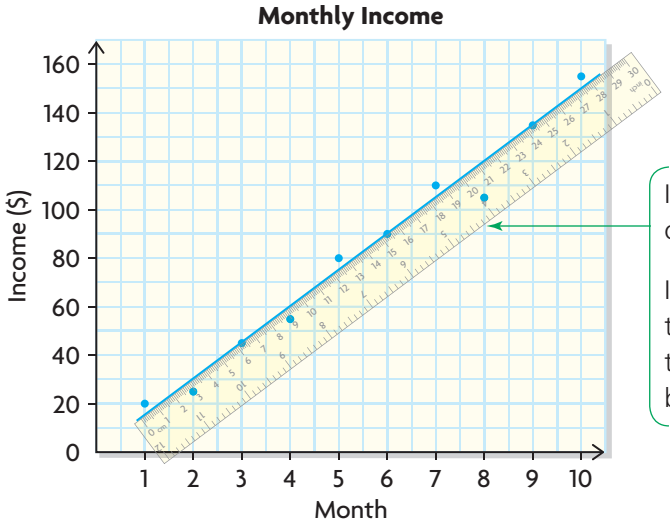
APPLY the Math

EXAMPLE 2 Using a line of best fit to describe a trend

Kajsa and Erika make bead jewellery in their spare time. Their monthly income for 10 consecutive months is shown in the table. Describe the trend in their income.

Month	1	2	3	4	5	6	7	8	9	10
Income ($)	20	25	45	55	80	90	110	105	135	155

Kylie's Solution: Using a graphical representation

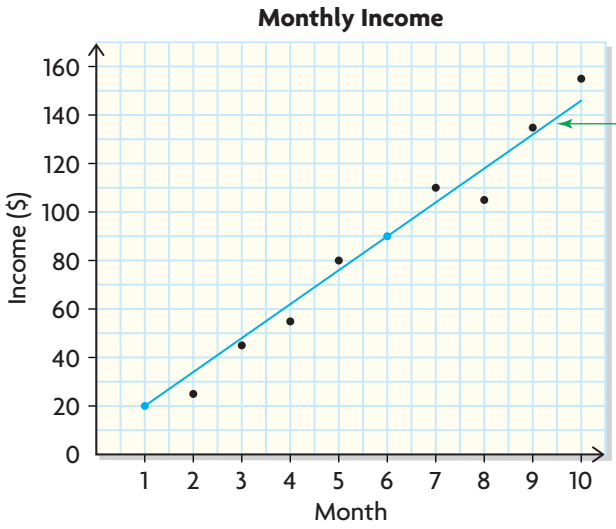


I sketched the data on a scatter plot.

I used a transparent ruler to draw a line of best fit.

From the scatter plot and the line of best fit, the trend is that the monthly income increases steadily.

Jasper's Solution: Using an algebraic representation



I sketched a line of best fit.

$(x_1, y_1) = (1, 20)$,
$(x_2, y_2) = (6, 90)$

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$
$$= \frac{90 - 20}{6 - 1}$$
$$= 14$$

Their monthly income is increasing by $14 each month.

$$y = mx + b$$
$$y = 14x + b$$
$$20 = 14(1) + b$$
$$20 = 14 + b$$
$$20 - 14 = b$$
$$6 = b$$
$$y = 14x + 6$$

The equation of the line of best fit has a positive slope, so the trend for Kajsa and Erika's income is that their monthly income is increasing steadily.

I chose two points on the line and used their coordinates to calculate the slope of the line.

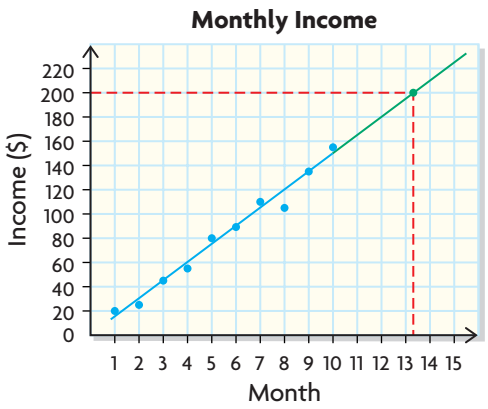
The slope is the rate of change in their income each month.

To determine b, I substituted 14 for m and the coordinates of the point (1, 20) into the equation for the line of best fit.

EXAMPLE 4 Using a line of best fit to solve a problem

Using the data in Example 3, estimate when Kajsa and Erika's income will reach $200 for a month.

Mathieu's Solution: Using a graphing strategy



I sketched a line of best fit.

I located the *x*-coordinate of the point on the line that represents an income of $200.

The x-value of this point is about $x = 13$.

Kajsa and Erika's income will reach \$200 after about 13 months.

I rounded the value of x because only whole numbers make sense for this situation.

Cassandra's Solution: Using an algebraic strategy

$$\text{Slope} = \frac{\text{rise}}{\text{run}}$$

$$= \frac{150}{10}$$

y-intercept $= 0$

$$y = 15x$$

$$200 = 15x$$

$$x = \frac{200}{15}$$

$$\doteq 13.3$$

I used Mathieu's scatter plot and line of best fit to estimate its slope and y-intercept.

I used this information to write the equation of this line.

I substituted 200 for y and solved for x.

Kajsa and Erika's income will reach \$200 after about 13 months.

I rounded the value of x because only whole numbers make sense for this situation.

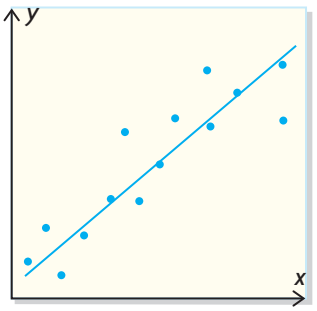
In Summary

Key Ideas

- You can use a line of best fit to make predictions for values not actually recorded or plotted. This is done by interpolating or extrapolating.
- Predictions can be made by reading values off a graph, or by using an equation of the line of best fit.

Need to Know

- If the pattern of points on a scatter plot looks like it follows a straight line, a line of best fit can be used to represent the relationship between the variables.
- When you draw a line of best fit, the points on the scatter plot should be "balanced" on each side of the line.
- You can use the coordinates of two points on the line of best fit to determine its slope and its equation.

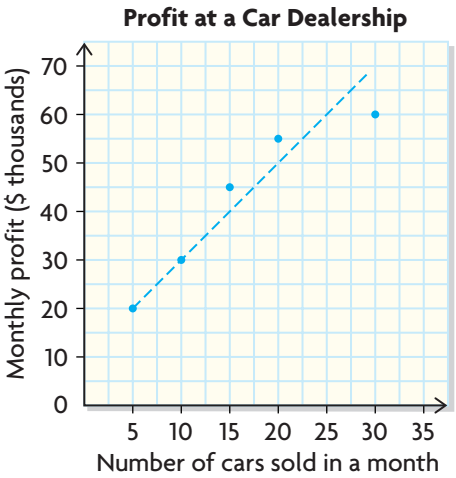


CHECK Your Understanding

1. This scatter plot shows the monthly profit for a car dealership when a certain number of cars are sold.

a) Use the graph to estimate the monthly profit in a month where 23 cars are sold.

b) Use the graph to estimate the number of cars sold in a month where the profit is $67 000.



2. The table shows temperatures at various times of the day.

Time (p.m.)	2	3	4	5	6	7
Temperature (°C)	−1	2	2.5	4	6.5	9

a) Construct a scatter plot for the data in the table.

b) Sketch a line of best fit.

c) Determine an equation for the line of best fit.

d) Predict the value of the temperature at 5:30 p.m.

e) Predict the time when the temperature is 8 °C.

PRACTISING

3. In this table, x represents the number of people enrolled in various classes at a health club, and y represents the number in each class that are male.

x	19	10	6	16	15	9	12	21
y	10	4	2	5	7	3	8	8

a) Construct a scatter plot for the data.

b) Sketch a line of best fit.

c) Use the line of best fit to estimate the value of y when $x = 14$.

d) Use the line of best fit to estimate the value of y when $x = 27$.

4. The following data show the final marks for 10 students in a math class and the average number of hours they studied math per week.

Final Mark	75	81	68	62	88	83	90	77	89	60
Average Number of Study Hours Per Week	3	3	5	1	5	3	6	3	5	2

a) Construct a scatter plot.
b) Sketch a line of best fit.
c) Determine an equation for the line of best fit.
d) Use the equation to estimate the mark for a student who studies an average of 4 h per week.
e) Use the graph to estimate the study time for a student whose final mark is 71.
f) Is there a relationship between final mark and number of hours of study per week? If so, describe it.

5. K A chair company has a contract to build all 1790 seats in a concert hall. The progress over the first week of work is shown in the table.

Number of Days	1	2	3	4	5	6	7
Total Number of Seats Completed	97	204	327	443	539	661	795

a) Estimate the number of seats built after 9 days. How many are built by the middle of day 5?
b) Estimate the number of days needed to build 1252 seats.
c) The company gets a bonus if it is able to finish all of the seats in two weeks or less. If the workers continue to make chairs at about the same rate in the second week, will the company be able to collect the bonus?

6. T Tomas is a member of the school track and field team. His times for running various distances are shown.

Distance (m)	50	100	150	200	250
Time (s)	6.1	12.0	18.3	25.2	31.7

a) Determine an equation for the line of best fit if the data were plotted on a scatter plot.
b) What is Tomas's time for a 175 m run likely to be?
c) Is it reasonable to use the same line of best fit to determine the time needed to run 3000 m? Explain.
d) What does the slope of the line of best fit tell you about how Tomas runs?

7. Kim is on her school basketball team. This table shows her statistics for the first 10 games of the season. (Each field goal made counts for two points, and each free throw made counts for one point.)

A

Game	Minutes Played	Field Goals		Free Throws		Points
		Made	Attempted	Made	Attempted	
1	32	5	13	4	6	14
2	30	4	10	3	3	11
3	24	2	6	1	1	5
4	29	1	3	2	4	4
5	36	3	6	0	1	6
6	19	5	11	2	2	12
7	12	0	3	0	4	0
8	21	1	5	1	2	3
9	18	3	5	1	5	7
10	19	3	7	2	6	8

a) Use a line of best fit to estimate the number of field goals Kim would make if nine were attempted.

b) Use a line of best fit to estimate the number of points Kim would score if she played for 40 min.

8. Suppose that you plotted some data on a scatter plot.

C

a) How would you draw a line of best fit?

b) How would you determine an equation for the line of best fit?

c) How would you use the line of best fit to interpolate or extrapolate?

Extending

9. The tables at the right show how the world record times for the 100 m sprint have changed over the years for both men and women.

a) Plot both sets of data on the same scatter plot.

b) Describe the trend in each set of data.

c) Sketch lines of best fit for each set of data.

d) Do the lines of best fit in part c) suggest that the women's world record time will someday be less than the men's? If so, predict when this might occur if current trends continue.

e) Do you expect current trends in each data set to continue forever? Explain.

Men:

Year	Time (s)
1960	10.0
1968	9.95
1983	9.93
1988	9.92
1991	9.86
1994	9.85
1996	9.84
1999	9.79
2005	9.77

Women:

Year	Time (s)
1960	11.3
1968	11.1
1976	11.01
1977	10.88
1983	10.81
1984	10.76
1988	10.49