UNDERSTANDING SCATTER PLOTS

1. Eliseo performed a study to explore how TV viewing habits affect student performance. He collected data by surveying several students in his math class. He asked each student to provide their current math mark as well as the number of hours spent watching TV each day. The data are summarized in the following table:

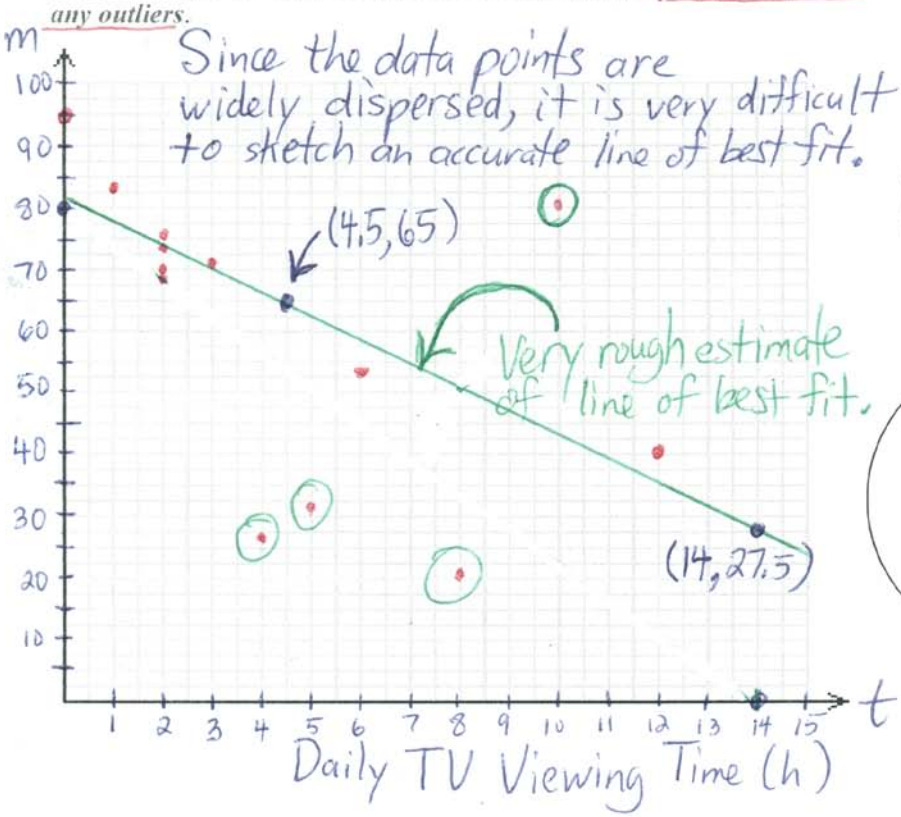
Daily TV Viewing Time (h)	1	6	3	2	10	0	2	5	2	12	8	5	2	4
Mark (%)	83	53	71	73	81	95	68	51	70	40	21	32	75	27

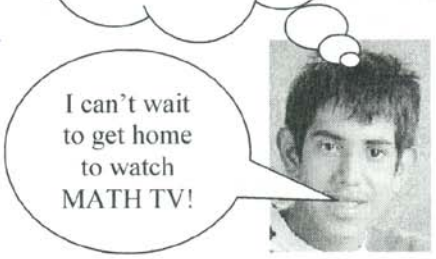
(a) State the independent and dependent variable.

SaherSami

Independent: TV Viewing Time **Dependent:** Mark

(b) Create a *scatter plot* of the data (plot the data points). ***Do not connect the dots!*** Label the axes and include a title for your graph. In addition, *circle any outliers*.





(c) Describe the relationship between the students' daily TV viewing time and their mathematics marks.

The math mark tends to decrease with increased TV viewing time.

(d) Draw a line of best fit. Then write the slope-intercept equation for the line of best fit. Show your work!

$\text{slope} = \frac{\Delta m}{\Delta t} = \frac{m_2 - m_1}{t_2 - t_1} = \frac{27.5 - 65}{14 - 4.5} \doteq -3.9$

(calculated using points identified on graph)

→ estimated from graph

Equation of Line of Best Fit: $m = -3.9t + 81$

(e) Use the equation of your line of best fit to estimate the math mark of a student who watches four hours of TV per day. $t = 4,\ m = ?$

$m = -3.9(4) + 81 \doteq 65$

The estimated mark for a student who watches four hours of TV daily is about 65%

(f) Again using your equation, estimate the number of hours of TV watched by a student with a mark of 45%.

$t = ?,\ m = 45$

$45 = -3.9t + 81$

$\therefore -3.9t = 45 - 81$

$\therefore t = \frac{-36}{-3.9} \doteq 9$

The line of best fit predicts about 9 hours of TV viewing per day.

(g) How certain are you that your estimates are accurate?

The estimates are not very accurate because there is a large degree of variation in the data.

(h) Now check your answers to (e) and (f) by using your graph.

Equation Answer	*Graph Answer*	*Do the answers agree?*
(e) 65%	(e) 66%	Yes, there is close agreement.
(f) 9 hours	(f) 9.5 hours	Yes, " " " "

2. Now use TI-Interactive to create a scatter plot and to determine the line of best fit for the same data given in question 1. Print out the TI-Interactive document that you create and staple it to this sheet. In addition, summarize your results below.

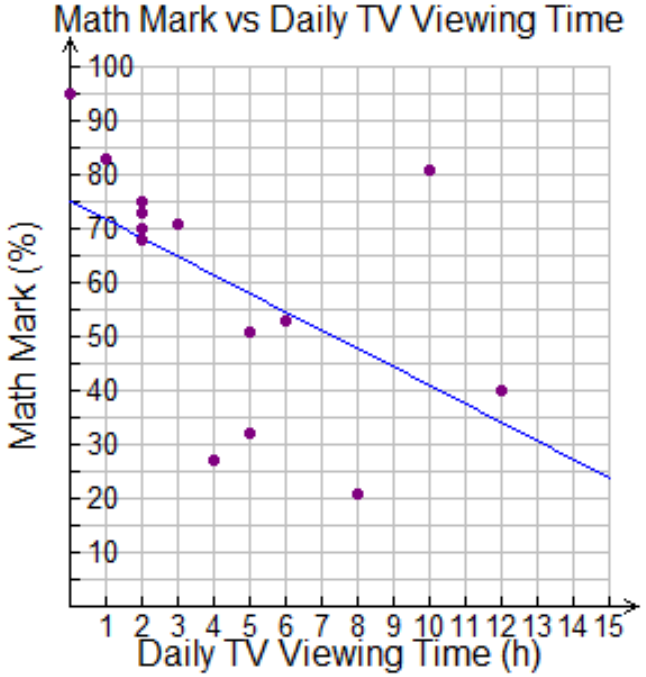
Equation obtained using your line of best fit: $m = -3.9t + 81$

Equation obtained using TI-Interactive: $m = -3.4t + 75$

3. Complete the following table. Use point form.

Similarities between Unit Four and Unit Five	*Differences between Unit Four and Unit Five*
• both units involve graphing on a Cartesian plane • both units involve linear relations • both units involve independent and dependent variables • both units involve slope and intercepts	• In unit 5, the relations were EXACTLY described by an equation. Once one of the variables is given a value, the value of the other can be calculated EXACTLY using the equation • In unit 6, data are collected for the purpose of finding out the strength of the relationship between two variables. The equations of the lines of best fit only allow us to ESTIMATE!

L1	L2
3	71
2	73
10	81
0	95
2	68
5	51
2	70
12	40
8	21
5	32
2	75
4	27



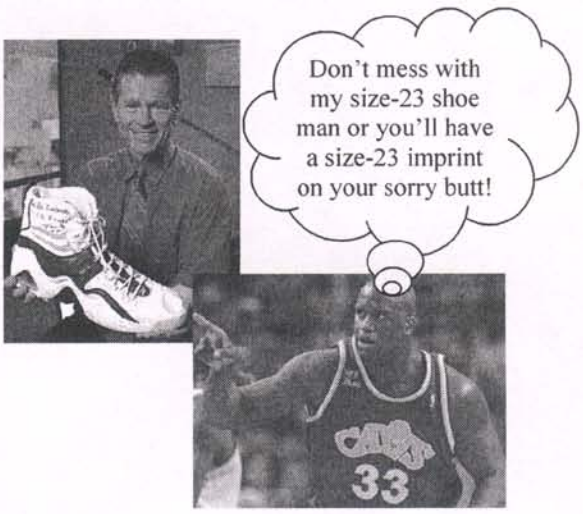
Linear Regression (ax+b)

regEQ(x) = -3.42566x + 75.1708

UNDERSTANDING SCATTER PLOTS #2

Predicting Shaquille O'Neal's Hand Span

In this activity you will collect data by measuring foot lengths and hand spans. You will then use your data to predict Shaquille O'Neal's hand span.



Step One – Calculating Shaq's Foot Length in Centimetres

It is well known that Shaquille O'Neal (also known as "Shaq") wears a size-23 shoe. What is not well known is his foot length in centimetres. Luckily, there are formulas that relate shoe size, as measured with a Brannock device (see diagram below and to the right), to foot length, in inches.

$m \rightarrow$ represents men's shoe size as measured by a Brannock device

$w \rightarrow$ represents women's shoe size as measured by a Brannock device

$f \rightarrow$ represents foot length ***in inches***

$$\boxed{m = 3f - 22} \qquad \boxed{w = 3f - 21}$$

(a) Use the appropriate formula above to calculate Shaq's foot length ***in inches***. Show all work!

$m = 23,\ f = ?$

$\therefore\ 23 = 3f - 22$

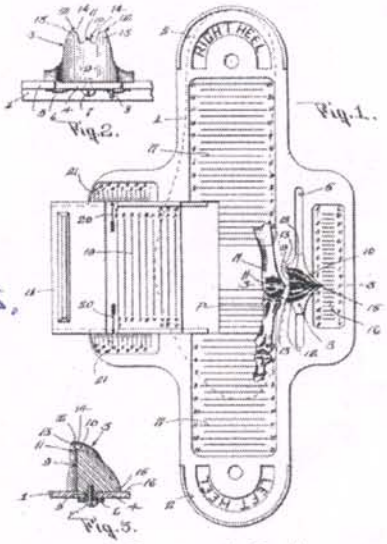
$\therefore\ 23 + 22 = 3f - 22 + 22$

$\therefore\ 45 = 3f$

$\therefore\ \frac{45}{3} = \frac{3f}{3}$

$\therefore\ 15 = f$

Shaquille O'Neal's foot length is about 15 inches.



A Brannock Device

(b) Now convert Shaq's foot length ***to centimetres*** by using the equation $C = 2.54I$, where C represents the length in centimetres and I represents the length in inches.

$I = 15,\ C = ?$

$\therefore\ C = 2.54(15)$

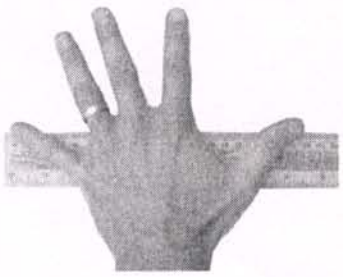
$= 38.1$

Conclusion: Shaq's foot length in cm is about 38.1 cm.

Step Two – Collecting the Data by Measuring Hand Span and Foot Length

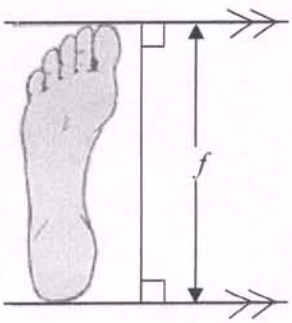
How to Measure Hand Span

- The hand is placed ***palm down*** on a flat surface.
- The fingers are outstretched as far as possible.
- Measure the distance between the ***outside of the thumb*** to the ***outside of the little finger***.



How to Measure Foot Length

- Shoes and socks must be removed.
- Place the most prominent toe and the most prominent part of the heel between two parallel lines that are perpendicular to the foot.
- Measure the distance between the two parallel lines.



Use the measuring procedures described above to complete the following table.

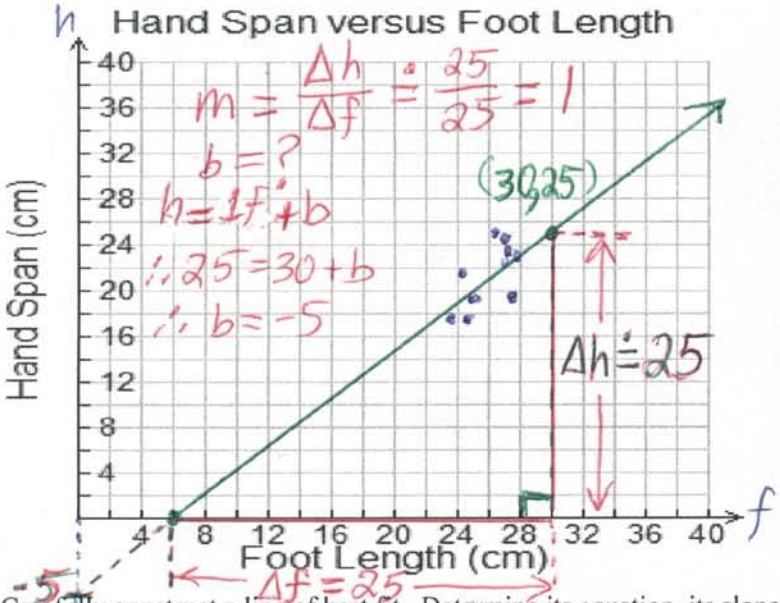
Measure to the nearest millimetre, that is, to one decimal place.

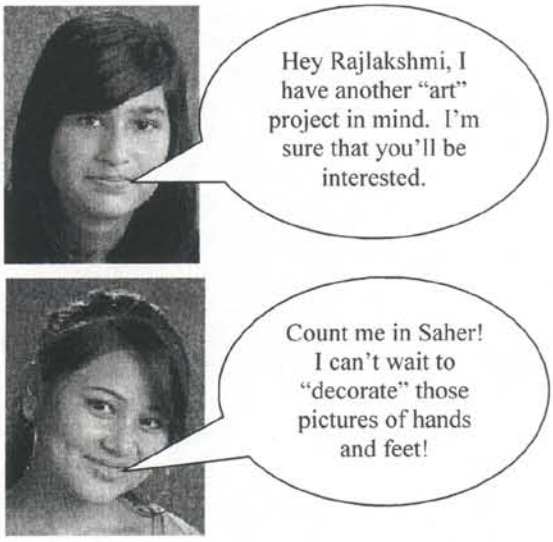
	Student Name	*Foot Length in cm (f)*	*Hand Span in cm (h)*	*Ratio of f : h*
1.	Pablo	27.3	23.9	1.14
2.	Hannah	24.2	20.8	1.16
3.	Kamaljeet	27.9	23.4	1.19
4.	Saher	24.4	17.8	1.37
5.	Meghan	23.9	19.9	1.20
6.	Aditya	26.4	21.1	1.25
7.	Naghan	27.6	21.5	1.28
8.	Gohar	26.8	24.1	1.11
9.	Wafa	23	19.5	1.18
10.	Justin	27.2	22.8	1.19

1.21 Average ratio of f:h

Step Three – Analyzing the Data

(a) Use the data in the table to create a scatter plot.
Do not connect the dots!





(b) Carefully construct a line of best fit. Determine its equation, its slope and its intercepts.

Equation: $h = f - 5$ Slope: 1 Vertical Intercept: -5 Horizontal Intercept: 6

(c) Now use TI-Interactive, a graphing calculator or spreadsheet software to determine the equation of the line of best fit as well as its intercepts. Summarize your results below.

Equation: $h = 0.92f - 2.24$ Slope: 0.92 Vertical Intercept: -2.24 Horizontal Intercept: 2.45

(d) Explain why it is better to use the equation obtained in (c) than it is to use the equation obtained in (b).

It is extremely difficult to sketch a line of best fit with pencil and paper. At best, we can only expect a rough estimate.

(e) Explain the *meaning*, in the context of this problem, of each of the following.

Slope = Constant of Variation = Rate of Change: 0.92 cm of hand span per centimetre of foot length

Vertical Intercept = Initial Value: No meaning in this context since a negative hand span is not possible

Horizontal Intercept: No meaning in this context because it is not possible to have a hand span of zero

(f) Does the data that you collected show a positive correlation, a negative correlation or no correlation? Explain.

There is definitely a positive correlation in the data because hand span increases with foot length.

(g) If you did everything correctly, your line of best fit should have a positive slope. Explain why you would expect this.

Hand span should increase with foot length. This can only happen if the slope (rate of change) is positive.

Step Four – Predicting Shaq's Hand Span

(a) You will use two different methods to predict Shaq's hand span.

Method 1	*Method 2*
Use the equation from (c) in step 3. $h = 0.92f - 2.24$ Shaq: $f \doteq 38$ cm $\therefore h = 0.92(38) - 2.24$ $= 32.72$	Calculate the average of the $f:h$ ratios from the table on the previous page. Then use this average to predict Shaq's hand span. average ratio $= \frac{1.14+1.16+1.19+1.37+1.2+1.25+1.28+1.11}{10}$ $\doteq 1.21$ This means that on average, the foot length is about 1.21 times the hand span, i.e. $f \doteq 1.21h$ $\therefore h \doteq \frac{f}{1.21} = \frac{38}{1.21} = 31.4$
Using method 1, I predict Shaq's hand span to be: about 32.7 cm	Using method 2, I predict Shaq's hand span to be: about 31.4 cm

(b) Predicting Shaq's hand span is an example of ***interpolation*** / ***extrapolation*** **(circle the correct answer)** because

we had to estimate a value BEYOND the range of our data set.

(c) The correct answer to the previous question is "extrapolation." In the space provided below, show an example of ***interpolation*** that involves the data you collected in step 3.

The smallest foot length in the data set is 23 cm and the largest foot length is 27.9 cm. There is no foot length of 26 cm in the data set, so we can interpolate the hand span for a foot length of 26 cm.

Method 1: Equation

$h = 0.92f - 2.24 = 0.92(26) - 2.24$
$\doteq 21.7$

Estimate: Hand span is 21.7 cm

Method 2: Averages

f	h

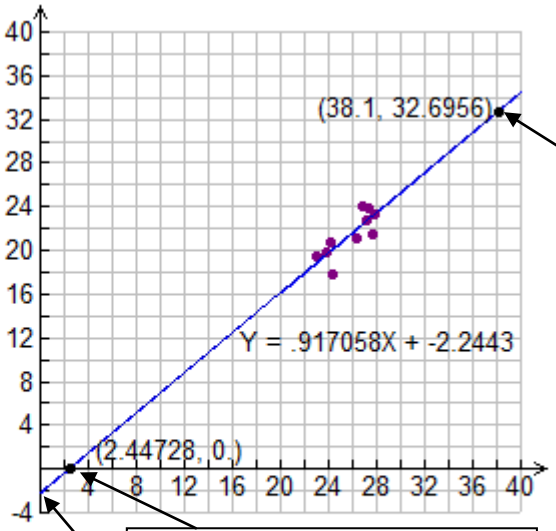
Closest points: (24.4, 17.8), (26.4, 21.1)

Average $= \frac{17.8 + 21.1}{2} \doteq 19.4$

Estimate: 19.4 cm

Line of Best Fit Obtained with TI-Interactive

L1	L2
27.3	23.9
24.2	20.8
27.9	23.4
24.4	17.8
23.9	19.9
26.4	21.1
27.6	21.5
26.8	24.1
23	19.5
27.2	22.8



Linear Regression (ax+b)
regEQ(x) = .917058x + -2.2443

Prediction for Shaquille O'Neal's Hand Span

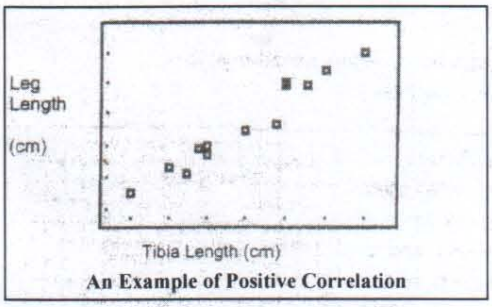
Foot length was calculated based on Shaq's shoe size: size-23 shoe size → 38.1 cm foot length

Line of best fit predicts a hand span of about 32.7 cm. This is an ***extrapolation***.

Intercepts have no meaning in the context of this scenario. (See step 3, part (e))

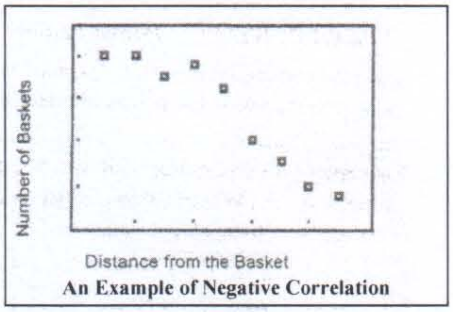
SCATTER PLOTS AND WHAT THEY TELL US

Investigating Correlation



An Example of Positive Correlation

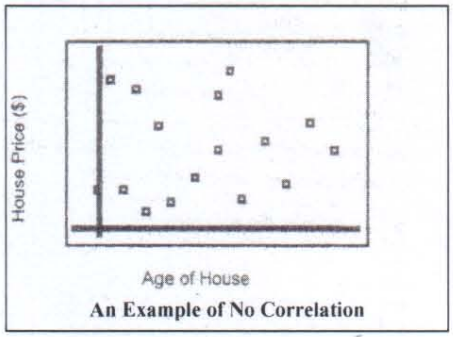
1. The graph shows the plotted points rising upwards to the right.
   - Agree
   - Disagree
   - Pass
2. As the length of the tibia increases the length of the leg increases
   - Agree
   - Disagree
   - Pass
3. The graph can be used to determine the length of a person's leg if you know the length of the tibia bone
   - Agree
   - Disagree
   - Pass



An Example of Negative Correlation

1. The graph shows the plotted points falling to the right
   - Agree
   - Disagree
   - Pass
2. As the distance from the net increases the number of baskets made decreases.
   - Agree
   - Disagree
   - Pass
3. The graph can be used to determine the number of baskets you will make if you know the distance from the basket.
   - Agree
   - Disagree
   - Pass

1. The graph shows the plotted points scattered.
   - Agree
   - Disagree
   - Pass
2. As the age of the house increases the price of the house is either large or small.
   - Agree
   - Disagree
   - Pass
3. The graph can't be used to determine the price of the house if you know how old it is.
   - Agree
   - Disagree
   - Pass



An Example of No Correlation

What is Correlation?

In ***statistics***, the ***correlation coefficient*** is used to measure the ***"strength" of the relationship between two variables***. Researchers collect data (make measurements of some kind, usually involving two variables) and then try to determine whether the variables are related to each other. The purpose of this process is to help us make predictions about one variable based on what we know about another variable.

For example, there is a correlation between income and education. We find that people with higher income usually have more years of education. When we know there is a correlation between two variables, we can make a prediction. If we know a group's income, we can predict their years of education.

Direction

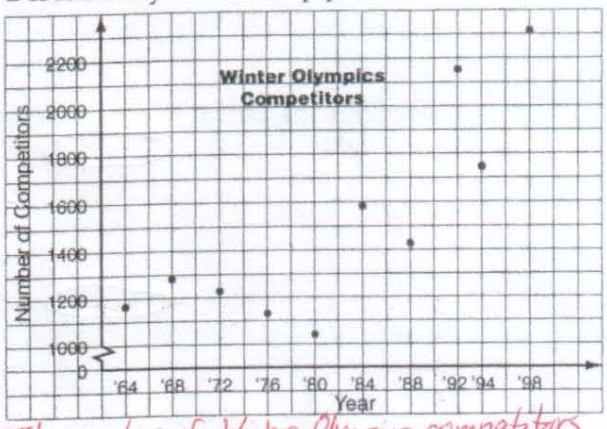
There are two ***types*** or ***directions*** of ***correlation***. In other words, there are two patterns that correlations can follow. These are called ***positive correlation*** and ***negative correlation***. (Keep in mind that in a correlational study, the researcher is measuring conditions that already exist.)

4.10 Scatter Plots

MATHPOWER™ 9, Ontario Edition, pp. 204–208

You can find out if a relationship exists between two variables by drawing a **scatter plot**. Plot points to represent the pieces of data, and then look for a pattern.

1. The scatter plot shows the total number of competitors at each Winter Olympics since 1964. Describe any relationship you see.

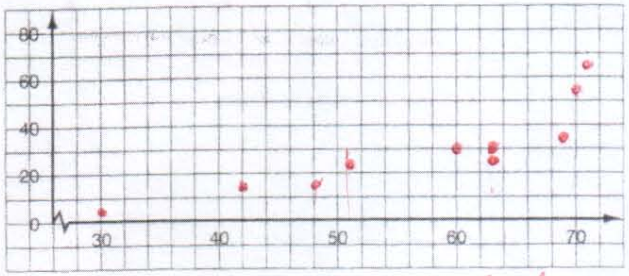


The number of Winter Olympics competitors tends to increase with time.

2. The table gives average heights and average masses of different types of dogs.

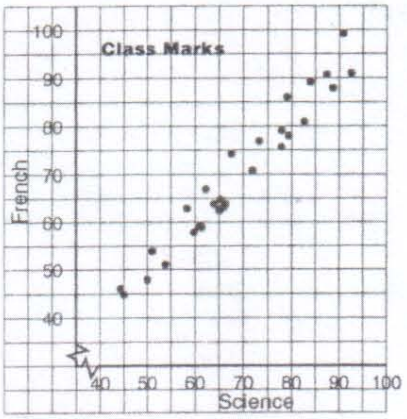
Type of Dog	Height (cm)	Mass (kg)
Belgian Sheepdog	63	30
Cocker Spaniel	42	14
Collie	63	27
English Springer Spaniel	51	23
Irish Setter	69	32
Irish Terrier	48	12
Japanese Chin	30	4
Labrador Retriever	60	30
Newfoundland	71	65
Saint Bernard	70	57

Draw a scatter plot of mass versus length. Describe any relationship you see.



The mass of a dog tends to increase with height.

3. A class recorded their marks in science and in French, and drew a scatter plot of the data.



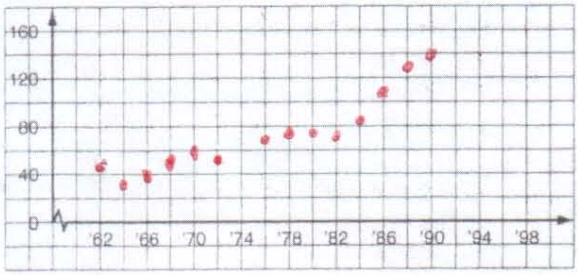
From this data, does it appear that a student's science mark is a good indication of the French mark? Explain.

For each student, the science and French marks are roughly the same. Thus, the science mark is a good indication of the French mark.

4. The table gives the total population of whooping cranes in North America over several years.

Year	Population	Year	Population
1962	45	1978	75
1964	35	1980	77
1966	40	1982	73
1968	50	1984	82
1970	60	1986	110
1972	52	1988	130
1976	68	1990	140

a) Draw a scatter plot of population versus year.



b) Use your scatter plot to estimate the total whooping crane population in 1994 and 1998.

extrapolation

c) Use your research skills to find the actual numbers in 1994 and 1998. Compare with your estimates.

estimates 1994: 150, 1998: 160

4.11 Lines of Best Fit

MATHPOWER™ 9, Ontario Edition, pp. 209–211

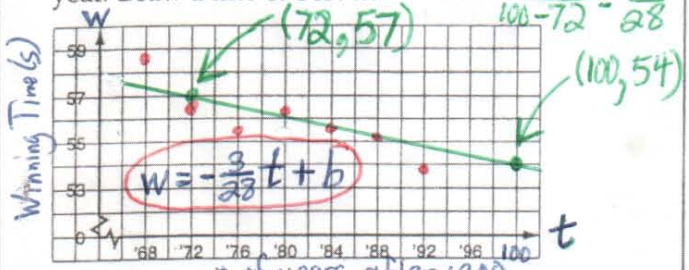
The **line of best fit** on a scatter plot is a line drawn as close as possible to all the data points. There are about as many points above the line as there are below the line.

1. The table gives the winning times in the men's 100-m backstroke swimming event at several Summer Olympics.

Year	Winner	Winning Time(s)
1968	Matthes (E. Germany)	58.7
1972	Matthes (E. Germany)	56.58
1976	Naber (U.S.)	55.49
1980	Baron (Sweden)	56.33
1984	Carey (U.S.)	55.79
1988	Suzuki (Japan)	55.05
1992	Tewksbury (Canada)	53.98

a) Draw a scatter plot of winning time versus year. Draw a line of best fit.

$m = \frac{54-57}{100-72} = -\frac{3}{28}$



$54 = -\frac{3}{28}\left(\frac{100}{1}\right) + b \Rightarrow b = \frac{453}{7}$

b) In 1960, David Thiele of Australia won the event. Extrapolate to estimate his winning time.

$W = -\frac{3}{28}t + \frac{453}{7} = -\frac{3}{28}\left(\frac{60}{1}\right) + \frac{453}{7} = 62.6\text{ s}$

c) David Thiele's actual winning time was 61.0 s. Compare this time with your estimate. How close were you?

$62.6 - 61 = 1.6$ 1.6s too high

d) Estimate the winning time in 2000.

54s (see graph for $t = 100$)

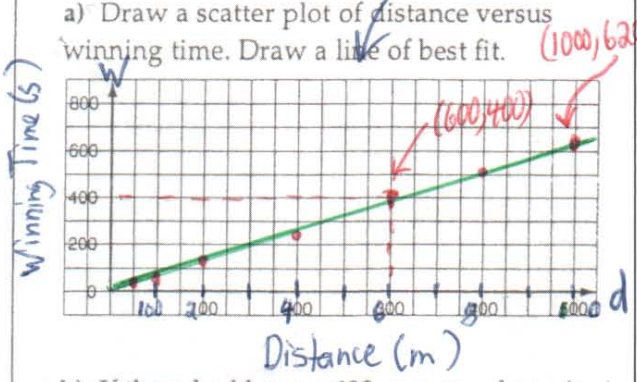
2. The table gives the winning times for five women's freestyle swimming events at the 1992 Summer Olympics.

Distance	Winning Time(s)
50 m	24.79
100 m	54.64
200 m	117.90
400 m	247.18
800 m	505.52

Independent Dependent

Axes reversed in original grid.

a) Draw a scatter plot of distance versus winning time. Draw a line of best fit.



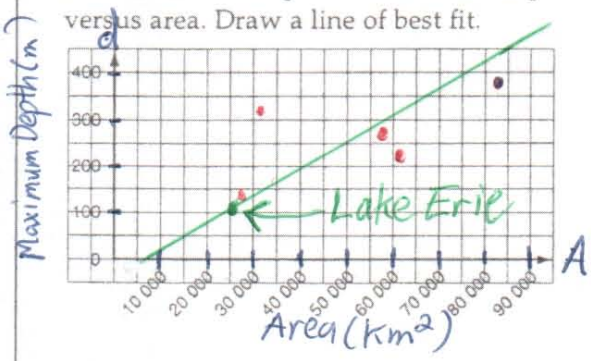
b) If there had been a 600-m race, what winning time would you estimate for it? 400s (see graph)

c) If there had been a 1000-m race, what winning time would you estimate for it? 620s (see graph)

3. The table gives the areas and maximum depths of several Canadian lakes.

Lake	Area (km²)	Maximum Depth (m)
Superior	83 270	393
Huron	60 700	229
Michigan	58 020	281
Great Bear	31 790	319
Great Slave	28 440	140

a) Draw a scatter plot of maximum depth versus area. Draw a line of best fit.



b) The area of Lake Erie is 25 680 km². Use your scatter plot to predict the maximum depth of Lake Erie. 100 m (see graph)

c) The actual maximum depth of Lake Erie is 64 m. Compare this with your estimated depth. Does your line of best fit give reasonable estimates of maximum lake depths? Explain.

The estimates are NOT reasonable because there are too few data points and a large degree of variation in the data.

Linear and Non-Linear Relations Revisited

1. Given the table of values, identify whether each relation is linear or non-linear. **Explain** your reasoning.

x	y	first differences
1	−1	–
2	1	1−(−1)=2
3	3	3−1=2
4	5	5−3=2
5	7	7−5=2

Δy

First differences are constant

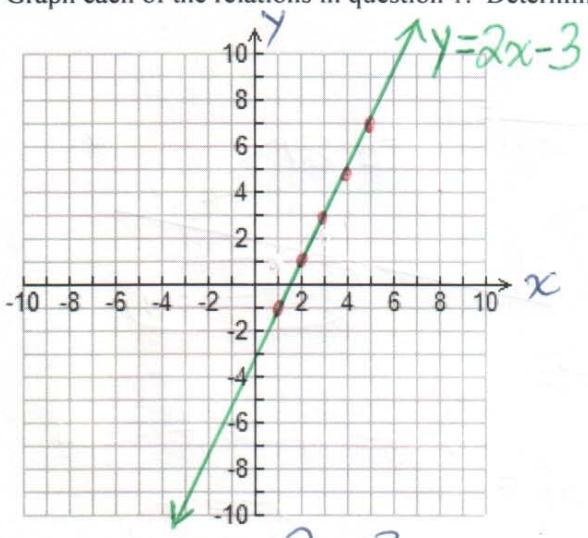
Linear because first differences are constant.

x	y	first differences
−2	1	–
−1	−2	−2−1=−3
0	−3	−3−(−2)=−1
1	−2	−2−(−3)=1
2	1	1−(−2)=3

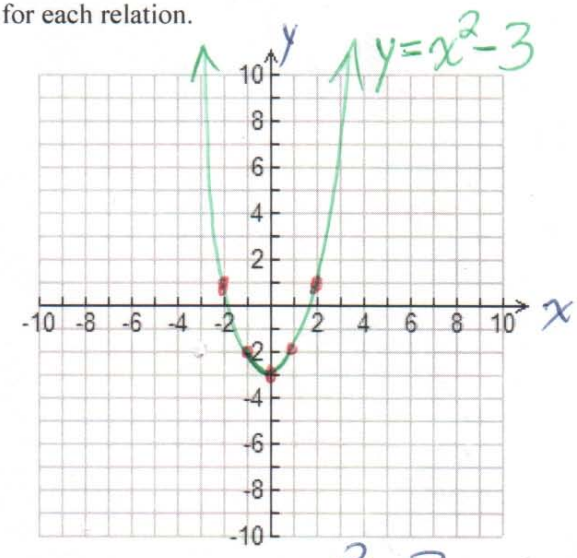
not constant

Non-linear because the first differences are NOT constant.

2. Graph each of the relations in question 1. Determine an equation for each relation.



Equation: $y=2x-3$



Equation: $y=x^2-3$

3. For each graph shown on the grid, state whether the graph represents a linear or non-linear relation.

Linear or non-linear?

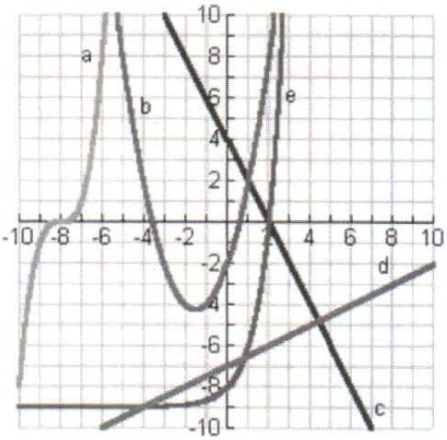
(a) non-linear

(b) non-linear

(c) linear

(d) linear

(e) non-linear



4. **Dependent and Independent Variables**

A change in the independent variable *causes* a change in the dependent variable. For example, a change in the number of bus passengers (independent variable) affects the weight of the bus (dependent variable).

1. For each pair of quantities, decide which is the independent variable and which is the dependent variable. Draw an arrow from the independent variable to the dependent variable. For example,

bus passengers → weight of bus.

a) number of customers → total sales

b) body temperature ← time spent in cold shower

c) average traffic speed ← number of cars on the highway

d) number of schools in a city ← total population of city

e) number of cigarettes smoked per day → money saved

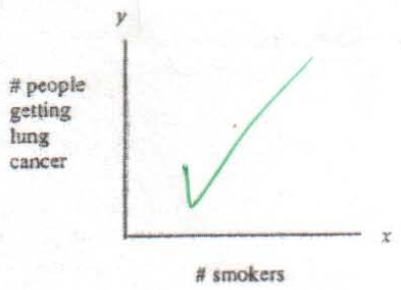
f) number of traffic accidents ← number of drunk drivers

g) humidity level → sales of air conditioners

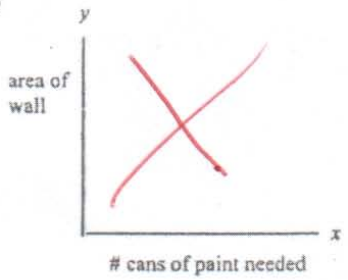
h) number of homes flooded ← amount of rainfall

2. The independent variable should always be on the x-axis. Using this rule, label each set of axes as "correct" or "incorrect".

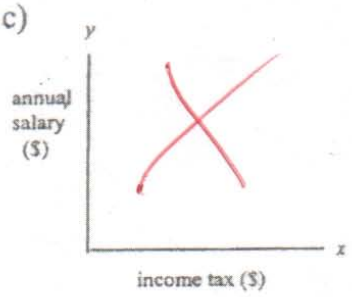
a)



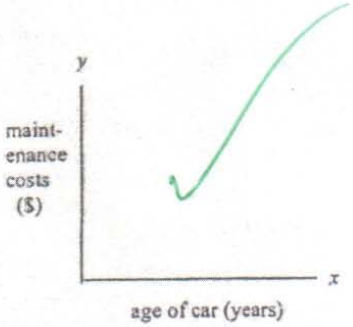
b)



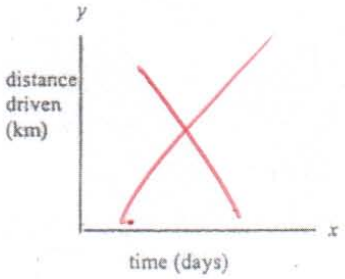
c)



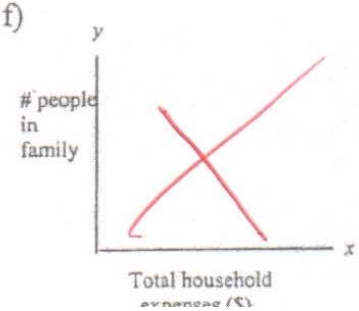
d)



e)



f)



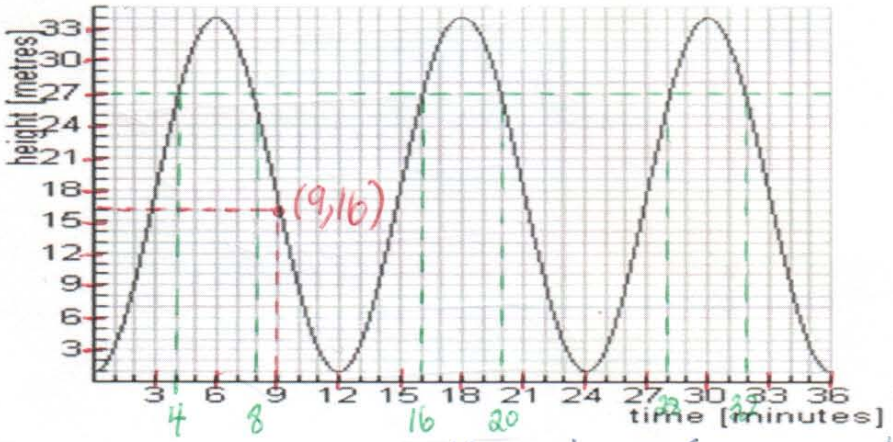
5. Consider the graph of the relation shown below.

(a) Identify the independent and dependent variables.

independent: Time dependent: Height

(b) Describe how the dependent variable changes as the independent variable increases.

The height varies periodically between a minimum of 1 m and a maximum of 34 m. This cycle repeats every 12 minutes.



(c) Use the graph to estimate the height at 9 minutes. 16 metres (see graph)

(d) Use the graph to estimate the times at which the height is 27 metres. 4s, 8s, 16s, 20s, 28s, 32s

(e) This relation is said to have a ***periodic*** behaviour. Give at least one real-life example of what this relation could model.

Height above the ground of a passenger on a Ferris wheel.

6. High Energy Gas Company charges its customers \$12 per month plus ten cents per cubic metre of gas used. The New Gas Company charges \$20 a month plus five cents per cubic metre.

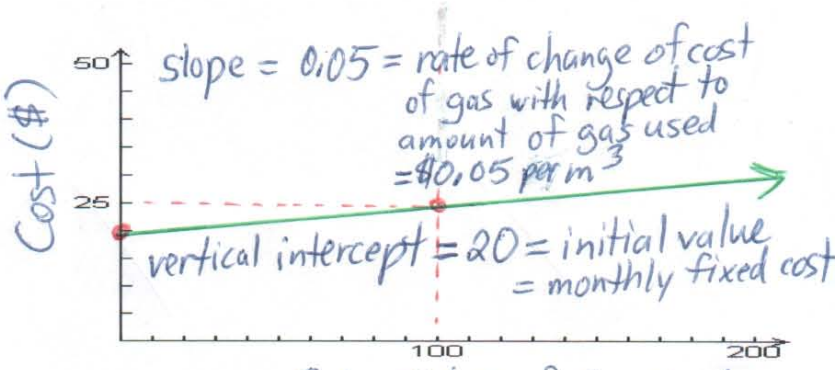
(a) For each company, write an equation to represent the total cost per month (C) in terms of the number of cubic metres of gas used (n).

$C = 0.05n + 20$

(b) Identify the independent and dependent variables.

Independent: n
Dependent: C

(c) Graph the relationship for each company.



(d) Use both the graphs and the equations to determine the following:

(i) the cost for a usage of 80 cubic metres of gas

(ii) the usage of gas for a cost of \$50.00

(i) $C = 0.05(80) + 20 = 4 + 20 = 24$ \$24.00 if 80 m³ of gas used

(ii) $50 = 0.05n + 20 \longrightarrow 0.05n = 50 - 20 \rightarrow 0.05n = 30 \rightarrow n = \frac{30}{0.05} = 600$

For a cost of \$50, gas usage is 600 m³

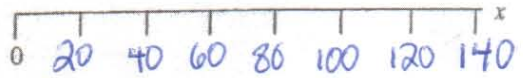
7.

Name: _____

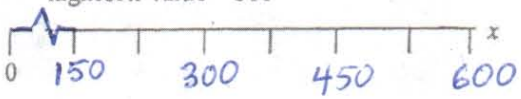
Choosing Scales for a Scatter Plot

1. Create a scale on the x-axis for each of the following situations, so that the distance from the lowest x-value to the highest x-value covers at least half the length of the axis. Use a break if necessary.

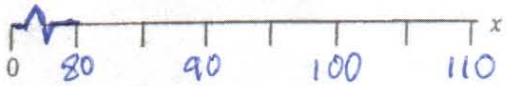
a) lowest x-value = 22
highest x-value = 130



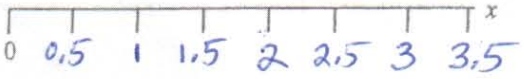
b) lowest x-value = 165
highest x-value = 511



c) lowest x-value = 82
highest x-value = 103

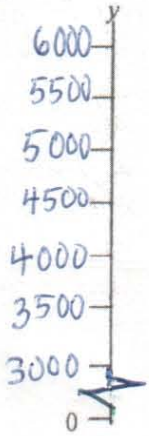


d) lowest x-value = 0.7
highest x-value = 2.8

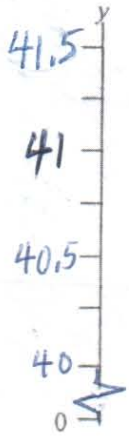


2. Create a scale on the y-axis for each of the following situations, so that the distance from the lowest y-value to the highest y-value covers at least half the length of the axis. Use a break if necessary.

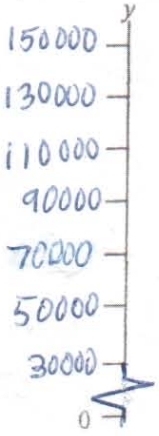
a) lowest y-value = 3,200
highest y-value = 5,700



b) lowest y-value = 40.25
highest y-value = 41.3

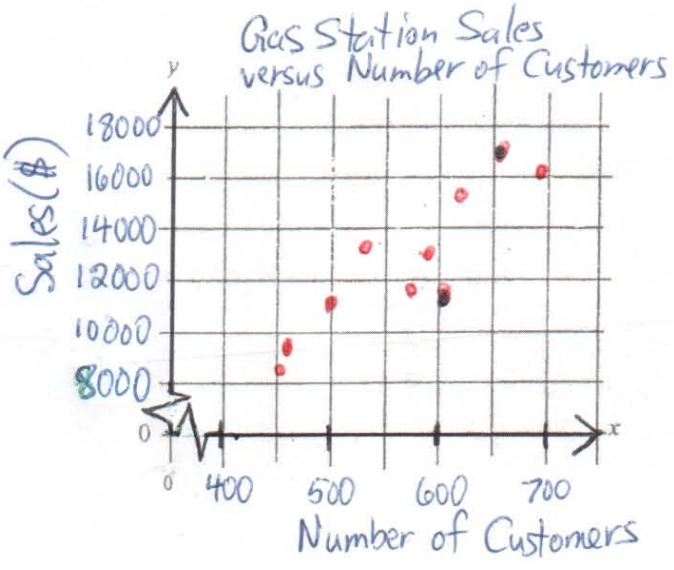


c) lowest y-value = 30,000
highest y-value = 125,000



3. Create a scatter plot of the following data, showing the number of customers and total sales for several gas stations on the same day. Put the number of customers on the x-axis and sales on the y-axis and choose appropriate scales. Label the axes and put a title at the top.

station	#customers	Sales ($)
Gerrard St.	460	9,605
Woodbine Ave.	501	11,022
Main St.	455	8,645
Kingston Rd.	658	17,108
Jones Ave.	524	13,672
Pape Ave.	620	15,511
Mortimer Ave.	589	12,908
Sammon Ave.	607	11,836
O'Connor Ave.	570	11,970
Leslie St.	695	16,070

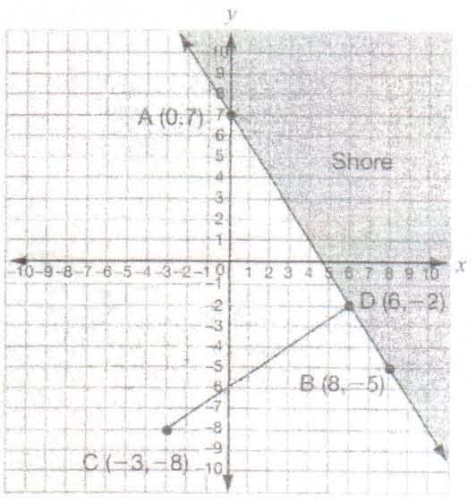


EQAO Practice – Linear Relations and Data Analysis

1. ***Analytic Geometry and Linear Relations***

Washed Up on the Shore

A boat is travelling from Point C toward Point D, which is on the shoreline.
The shoreline is represented by the line through points A and B.



slope of AB $= \frac{y_2-y_1}{x_2-x_1} = \frac{-5-7}{8-0} = \frac{-12 \div 4}{8 \div 4} = -\frac{3}{2}$

slope of CD $= \frac{y_2-y_1}{x_2-x_1} = \frac{-2-(-8)}{6-(-3)} = \frac{6}{9} = \frac{2}{3}$

Since $-\frac{3}{2}$ and $\frac{2}{3}$ are negative reciprocals of each other,

$AB \perp CD$

Therefore, the path from C to D must be perpendicular to the shoreline.

Determine whether the path from C to D is perpendicular to the shoreline.

Justify your answer.

Excellent Equations

A line is perpendicular to the line $y = 2x + 3$ and has the same *x*-intercept as $x + 3y + 10 = 0$.

Slope $= -\frac{1}{2}$ → passes through $(-10, 0)$

Find the equation of this line. Express your answer in the form $y = mx + b$.

Justify your answer.

Intercepts of $x + 3y + 10 = 0$

x-int: $y = 0$

$\therefore x + 3(0) + 10 = 0$

$\therefore x + 10 = 0$

$\therefore x = -10$

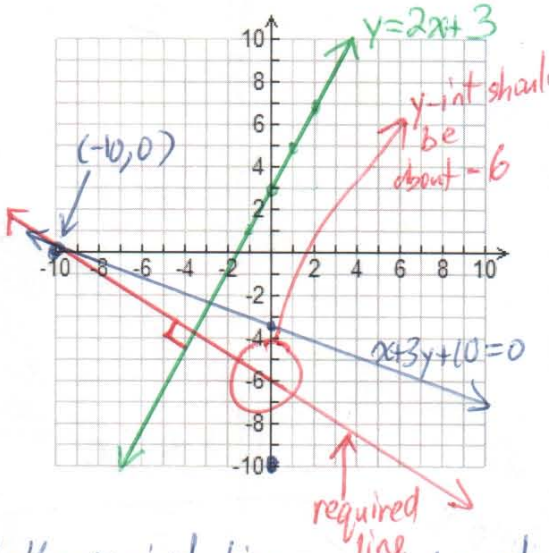
y-int: $x = 0$

$\therefore 0 + 3y + 10 = 0$

$\therefore 3y + 10 = 0$

$\therefore 3y = -10$

$\therefore y = -\frac{10}{3}$



Slope of $y = 2x + 3$ is 2. Therefore, the slope of the required line must be $-\frac{1}{2}$ (perpendicular lines).

∴ equation of required line is of form $y = -\frac{1}{2}x + b$

Since the *x*-intercept is $(-10, 0)$, the required line passes through $(-10, 0)$

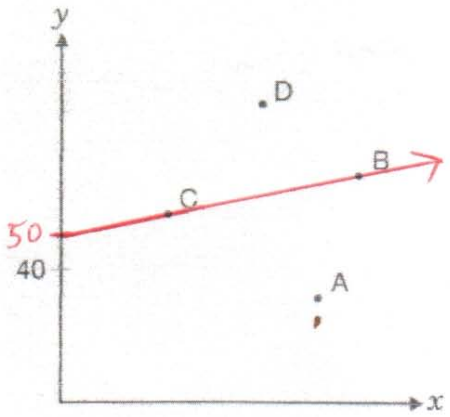
$\therefore 0 = -\frac{1}{2}(-10) + b \rightarrow 0 = 5 + b \rightarrow b = -5$

∴ equation is $\boxed{y = -\frac{1}{2}x - 5}$

Lineup

slope positive and small → line leans right and has a shallow slope.

The line $y = \frac{1}{5}x + 50$ passes through only one pair of points below. (→ y-intercept)



The line could only pass through the points C and B because this is the only pair of points that produces a y-intercept of about 50 and has a positive slope.

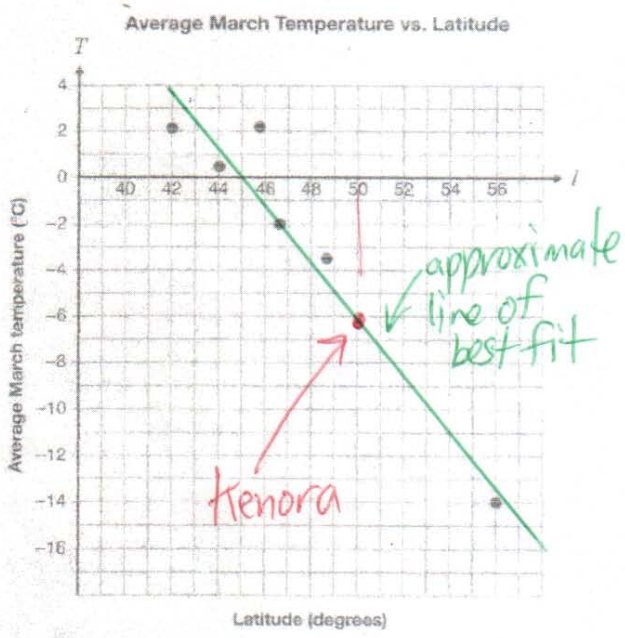
CD → y-intercept smaller than 40
CA → y-intercept larger than 50 and slope is negative
DA → y-intercept is much larger than 50 and the slope is negative

Which pair of points could the line pass through? Justify your response.

2. *Data Analysis – Lines of Best Fit*

March Temperatures

The average March temperatures for six Ontario communities are plotted according to their latitudes on the following scatter plot.



The general trend of the data is that average March temperature decreases as the latitude increases. The city of Kenora fits this trend quite nicely as is shown by its proximity to the line of best fit.

approximate.

The city of Kenora has a latitude of 50° and has an average March temperature of −6.3 °C. Does the community of Kenora follow the trend of the data?

Justify your answer.

Wing Length

Wing length is a reliable method for determining the age of young birds. Below is an example of data for a particular species.

Wing length (cm)	Age (days)
1.5	4
3.1	8
3.2	10
4.1	12
5.2	16

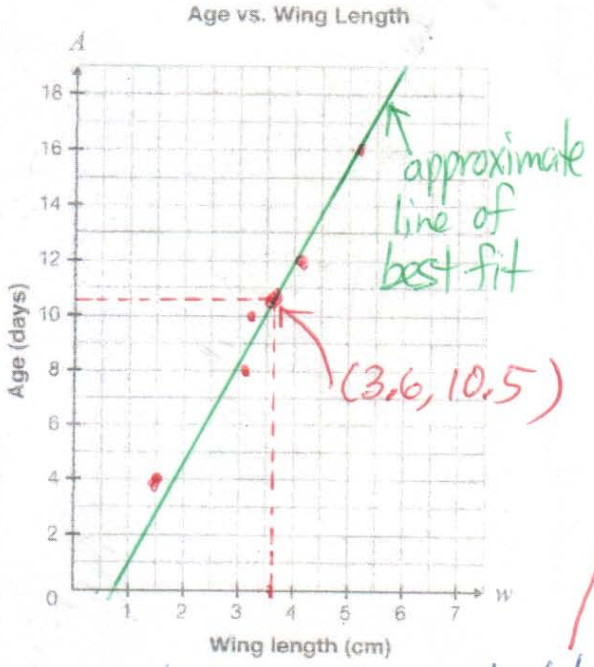
} 3.6 is between 3.2 and 4.1
∴ 11 days would be a good estimate since it lies between 10 and 12

Determine the age of a bird with a wing length of 3.6 cm.

You may use the grid if you wish.

Justify your answer.

An approximate line of best fit has been sketched. By using this line, we can see that for a wing length of 3.6 cm, the age is about 10.5 days.



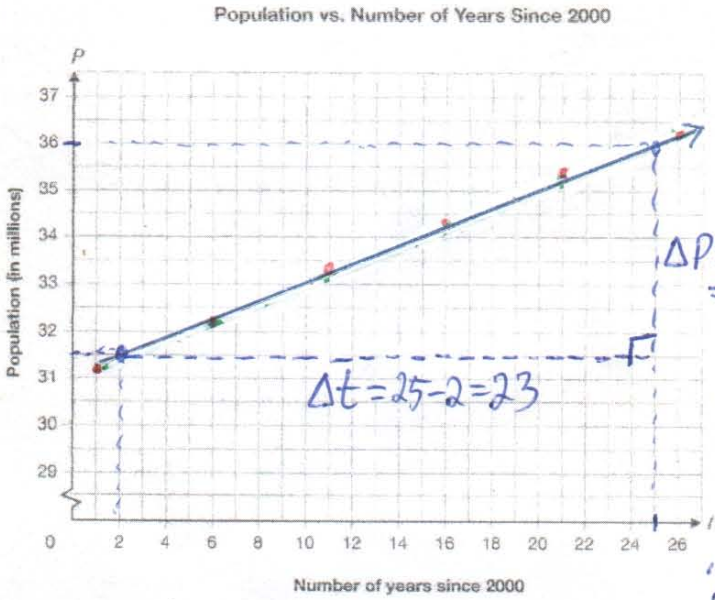
Both methods (i.e. interpolating from table and from graph) produce an estimate of an age of about 11 days.

Population Plans

Alvin is researching the population of Canada. He finds data for the year 2001 and predictions for every 5 years after that, as shown below.

P = population in millions
t = time since 2000, in years

Number of years since 2000, t	Population (in millions), P
1	31.1
6	32.2
11	33.4
16	34.4
21	35.4
26	36.2



$$\text{slope} = \frac{\Delta P}{\Delta t} = \frac{4.5}{23} \doteq 0.2$$

$\therefore P = 0.2t + b$

$\because$ (25, 36) lies on the line

$\therefore 36 = 0.2(25) + b$

$\therefore 36 = 5 + b$

$\therefore b = 31$

Determine an algebraic model for Alvin's data, and use it to make a reasonable prediction for the population of Canada in 2036.

Justify your answer.

Algebraic Model: $P = 0.2t + 31$

Prediction for 2036: $t = 36$

$\therefore P = 0.2(36) + 31 = 7.2 + 31 = 38.2$

A reasonable prediction for 2036 is a population of 38.2 million

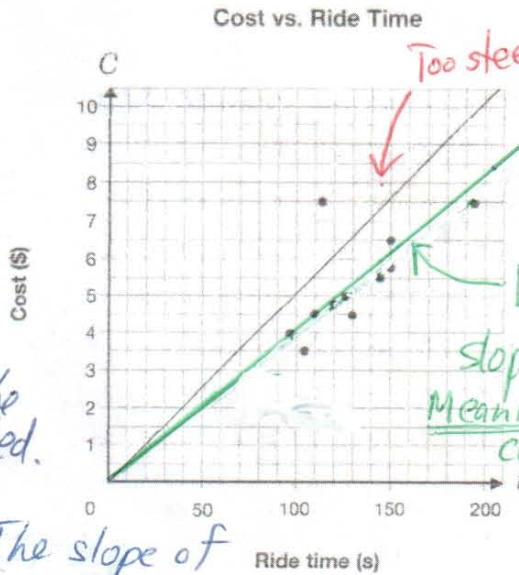
Amusement Park

Susanna collects data about the relationship between the **cost of each ride**, C, in dollars, and **the time the ride lasts**, t, in seconds. She plots the data on the graph below.

Susanna graphs the equation $C = 0.05t$. She notices that its line is **not** the line of best fit.

Describe how to change the equation so that it represents the equation of a line of best fit for her data.

Justify your answer.



slope $\doteq \frac{8}{200} = 0.04$

Meaning: average cost is $0.04 per second

The graph of $C = 0.05t$ does not pass through the region where the bulk of the data points are concentrated. The line labelled with an "*" provides a much better fit. The slope of the given line, 0.05, is too large. As shown in the diagram, a slope of about 0.04 gives a much better fit. (Cost of $0.04 per second provides a better fit.)