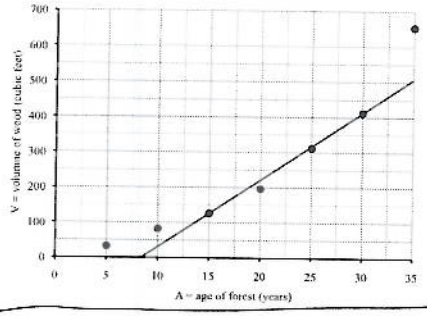
SOLUTIONS

4.5 Fitting lines to data – Practice exercises

1. The scatter plot shows the total volume of wood, V cubic feet, in managed forests of different ages, A years.

   (a) For each line, state some reason why the fit is not good. (We know the line will not go through all, or even most, of the points, so that is not the problem. Instead look at slope/steepness, intercept/height, etc.)

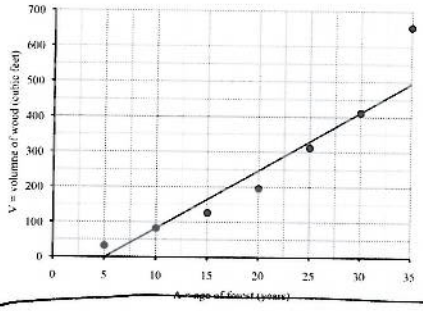
LINE A



Line A seems too low.

There are 3 points above the line and only 1 point below. Would prefer balanced.

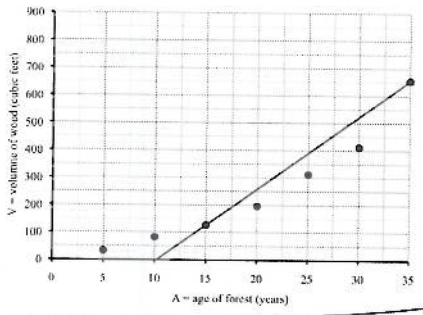
LINE B



Line B might not be steep enough.

The last point is way above the line.

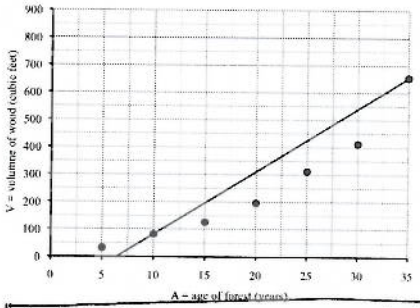
LINE C



Line C seems too steep.

The line is below the first two point and the above 3 points. Would prefer intermixed.

LINE D



Line D seems too high.

Almost all of the points fall below the line.

   (b) Which of these four lines do you think fits best, and why?

Line B seems best. It goes through or very close to almost all of the points.

2. Noel is considering investing in a company's stock so he looked up a few values.

Day	0	300	500
Value ($)	23.19	37.00	48.10

(a) Calculate the daily rate stock prices changed during the first 300 days.

$ROC = \frac{\$37.00 - \$23.19}{300 - 0 \text{ days}} = (37.00 - 23.19) \div 300 = .04603\ldots$ ≈ $.046/day

Prices dropped just under 5¢/day.

(b) Calculate the daily rate stock prices changed from Day 300 to Day 500.

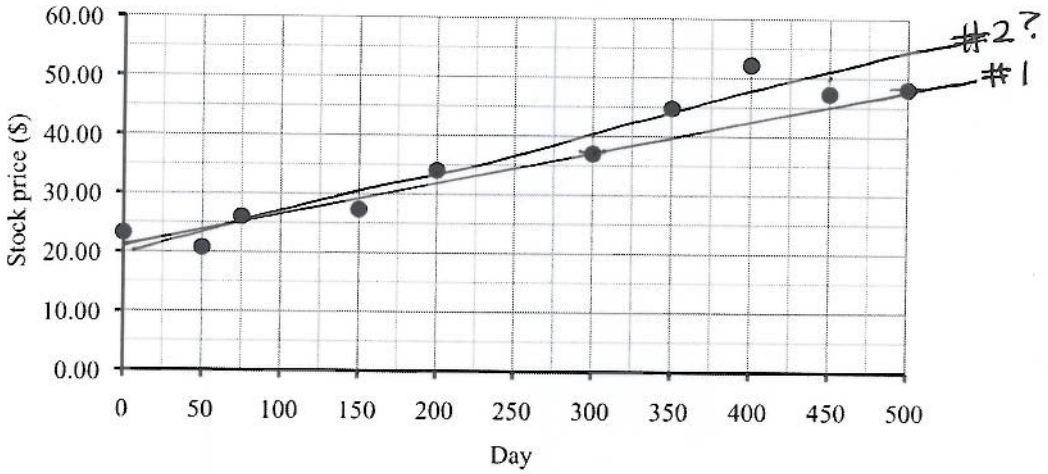
$ROC = \frac{\$48.10 - \$37.00}{500 - 300 \text{ days}} = (48.10 - 37.00) \div (500 - 300) =$ $.0555/day

Prices dropped just over 5¢/day

(c) Is this growth linear?

No. different rates of change.

(d) The scatter plot shows additional values of the stock Noel is considering buying.



Draw in a line that through the points for Day 300 and Day 500. Label this line #1. Explain why that line does not fit the data well.

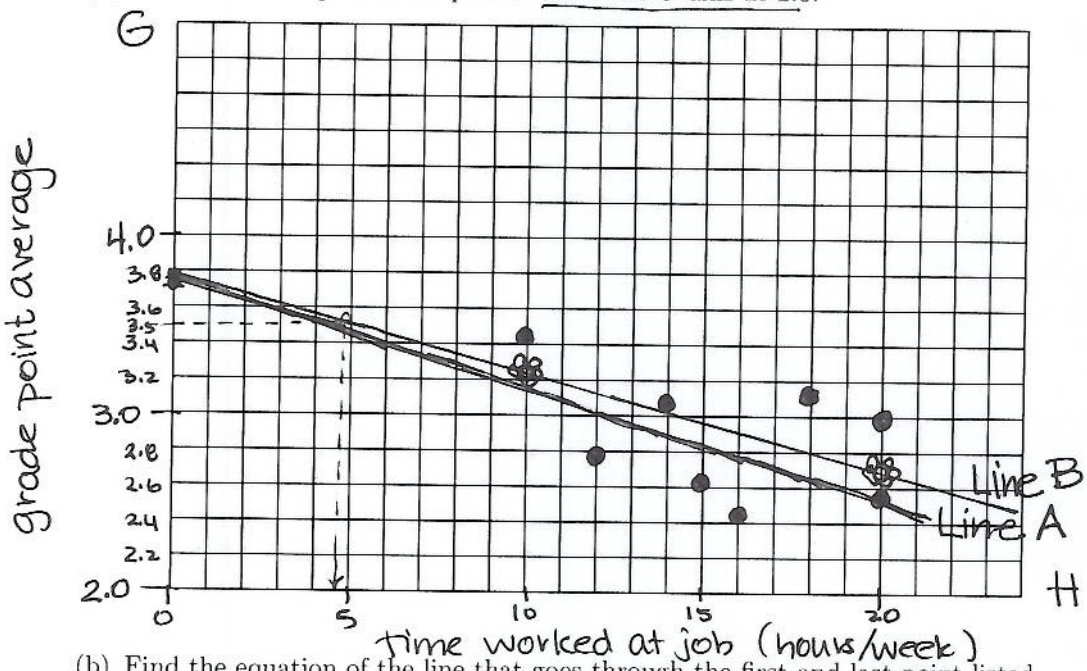
Line #1 does not seem steep enough.
At first the points are close or below the line, and later the points are above. Wanted intermixed.

(e) Draw in a line that fits the data better. It does not need to go through any of the points exactly. Label that line #2.

3. Is it true that students who work part-time have lower grades? Do the number of hours matter? The table shows the grade point average (GPA) of ten students compared to the number of hours per week each student works at a part time job. The variables we used are H for the time worked at job (hours/week) and G for the grades GPA, on the usual scale of 0.0 to 4.0.

H	0	0	10	12	14	15	16	18	20	20
G	3.72	3.91	3.43	2.79	3.08	2.62	2.44	3.17	3.00	2.55

(a) Make a scatter plot of the points. Start the G-axis at 2.0.



(b) Find the equation of the line that goes through the first and last point listed.
Hint: the first point tells you the intercept.

H	0	20
G	3.72	2.55

Intercept = 3.72

$$\text{slope} = \frac{2.55-3.72}{20-0 \text{ hours}} = (2.55-3.72)\div 20$$

$$= -.0585/\text{hour}$$

$$\boxed{G = 3.72 - .0585H}$$

check: $3.72 - .0585 \times 20 = 2.55$ ✓

(c) Draw this line on your graph and label it line A.

The problem continues ...

(d) Use your equation for line A to figure out what you would expect the GPA of a student working a 30 hour per week job to be.

← H = 30

$G = 3.72 - .0585 \times 30 = 1.965 \approx 1.97$ GPA

Working 30 hours per week corresponds to a GPA below passing (2.0). ☹

(e) It turns out, the best fitting line has equation $G = 3.7597 - .0551H$. Make a table of values for this equation using $H = 0, 10, 20$ hours.

H	0	10	20
G	3.76	3.21	2.66

$3.7597 - .0551 \times 20 = 2.6577 \approx 2.66$

$3.7597 \approx 3.76$

$3.7597 - .0551 \times 10 = 3.2087 \approx 3.21$

The additional points are labelled ✿ on my graph to distinguish them from actual data points.

(f) Use that table of values to graph this best fitting line on that same set of axes. Label it line B.

(g) According to line B, what's the most hours a student should work to be able to maintain a 3.5 GPA? Solve an equation, then check on your graph.

← just under 5 hours.

$3.7597 - .0551H = 3.5$

$-3.7597 \qquad -3.7597$

$\frac{-.0551H}{-.0551} = \frac{-.2597}{-.0551}$

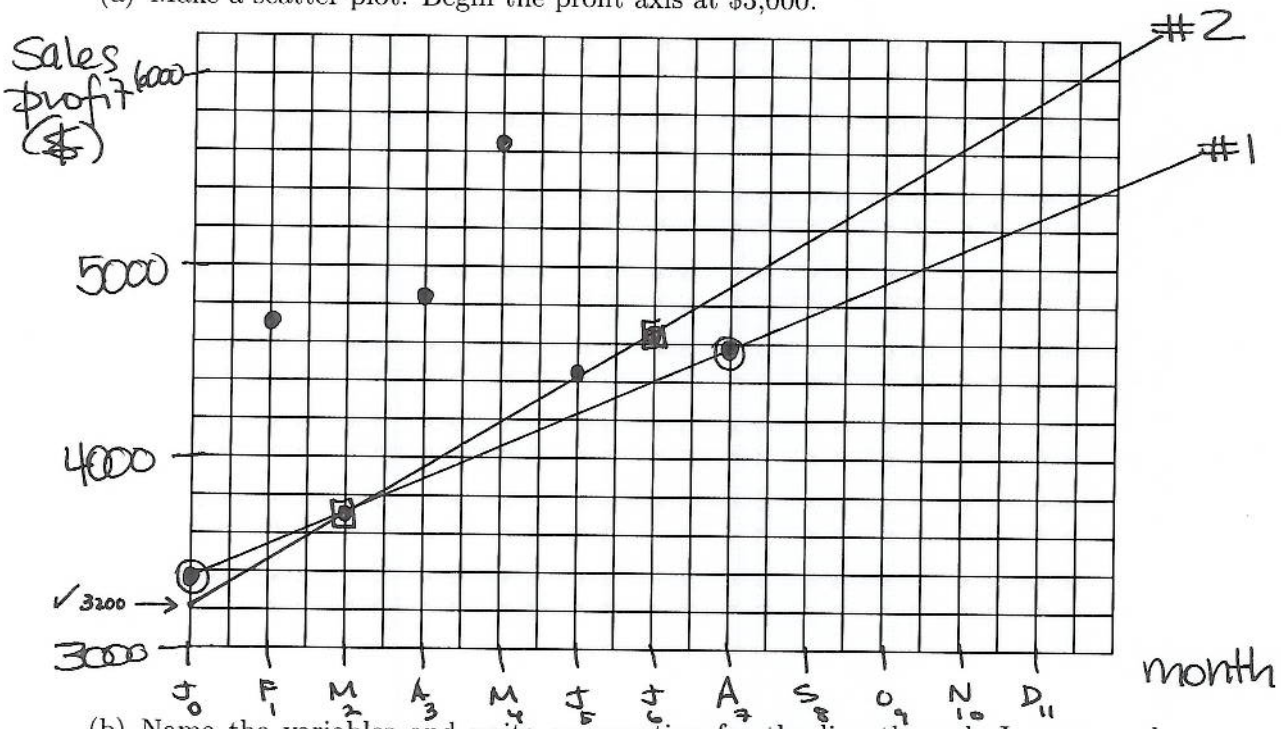
$H = 4.7132... \approx 5$ hours/week

According to this equation a student could work only 4-5 hours per week if s/he wants to maintain a 3.5 GPA.

4. Mia and Mandi and opened a candy shop this January. The table shows their monthly sales profit. Except for some seasonal fluctuation, Mia and Mandi generally expect your profits to rise steadily while their business is getting established.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Sales Profit ($)	3,394	4,702	3,683	4,840	5,632	4,432	4,649	4,590

(a) Make a scatter plot. Begin the profit axis at $3,000.



(b) Name the variables and write an equation for the line through January and August. Add this line (#1) to your graph. This line is too low.

M = time (months after Jan.) ~ indep

S = sales profit ($) ~ dep

$\text{slope} = \frac{\$4590 - \$3394}{7 - 0 \text{ mo.}} \approx \$170.86/\text{month}$

intercept = $3,394 ☺

$$S = 3{,}394 + 170.86\,M$$

	M	S
Jan	0	3394
Aug	7	4590

You would be correct to use Aug=8 and Jan=1 but then Jan is NOT where intercept is ☹

Check: $3394 + 170.86 \times 7 = 4590.02$ ✓

The problem continues ...

(c) Write an equation for the line through March and July. Notice that you need to find the intercept this time. Add this line (#2) to your graph. This line is too steep.

	M	S
March	2	3683
July	6	4649

slope $= \frac{\$4649 - \$3683}{6-2 \text{ months}} = \$241.50/\text{mo.}$

intercept $= 3683 - 241.50 \times 2$
$= 3200$ (yes! agrees with graph)

$$S = 3200 + 241.50M$$

check:
$3200 + 241.50 \times 3 = 3683$ ✓
$3200 + 241.50 \times 6 = 4649$ ✓

(d) Neither of these lines go anywhere near the data for February, April, and May, because those are outliers. Any idea why those months had much higher candy sales than the other months?

Feb = Valentine's Day
April = Easter
May = Mother's Day
} will sell more candy than usual

(e) What does each equation give as an estimate for September's sales?

#1) $3394 + 170.86 \times \underline{8} = \$4,760.08$

#2) $3200 + 241.50 \times \underline{8} = \$5,132.00$

$M = 8$
remember Jan = 0, Sept = 8
(instead of the usual Jan = 1, Sept = 9)

(f) Explain why Mia and Mandi should not use either of these lines to estimate October's sales.

Halloween will sell more candy again!