Hwk: Pg. 145: # 13.

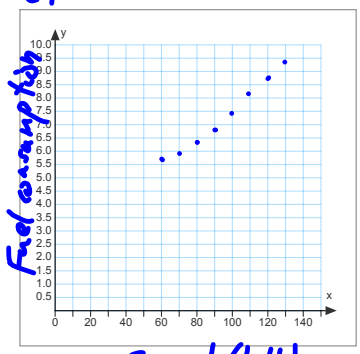
Speed (km/h)	60	70	80	90	100	110	120	130
Fuel consumption (L/100km)	5.6	5.9	6.2	6.7	7.4	8.1	8.7	9.4

a) Speed is the independent variable.

b) The fuel consumption increases as speed increases.

d) There is a positive correlation between the two.

c)



Mar 9-11:26 AM

Linear Regression

Regression is an analytic technique for determining the relationship between a dependent variable and an independent variable. When the two variables have a linear correlation, you can develop a simple mathematical model of the relationship between the two variables by finding a line of best fit. To do this we draw the line through as many points as possible, follow the general trend of the data, and have a roughly equal number of points above and below the line. You can then use the equation for this line to make predictions by interpolating (estimating between data points) and extrapolating (estimating beyond the range of the data).

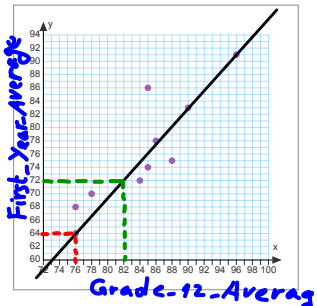
Other names for a line of best fit are regression line or trend line.

A point that is substantially above or below all other data points is called an outlier.

Example:

Grade 12 Average vs First Year University Average

	Grade_12_Avera...	First_Year_Average
1	85	74
2	90	83
3	76	68
4	78	70
5	88	75
6	84	72
7	76	64
8	96	91
9	86	78
10	85	86
11		



a) Graph the data and classify the linear correlation.

There seems to be a strong positive correlation between the grade 12 average and the first year average (both increase).

b) Draw a line of best fit.

Use this linear model to predict:

i. the first-year average for a student who had an 82% average in grade 12.

∴ We would expect such a student to have a first year average of 72%.

ii. the grade-12 average for a student with a first-year average of 64%.

∴ We would expect such a student to have had a grade 12 average of 76%.

Oct 3-10:23 AM

c) Find the equation of the line of best fit using software technology and by hand.
Note that when the linear correlation is not strong, a line of best fit is not very useful. For more accurate results we use the **least-squares formula** to find the equation of the line of best fit.

By hand: $y = mx + b$

1. Pick any 2 points that lie on the line of best
2. Find the slope between the 2 points
   $m = \frac{y_2 - y_1}{x_2 - x_1}$
3. Use the slope, m and one of the points to find the y-intercept, b of the line
4. Substitute m and b into $y = mx + b$

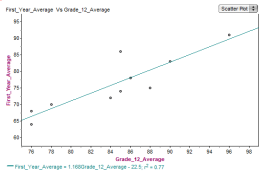
Using Technology: Fathom (can use Excel)

First_Year_Average = 1.168 × Grade_12_Average - 22.5

Equation of the line of best fit

$r^2 = 0.77$ ← coefficient of determination

$r = \sqrt{0.77} \doteq 0.88$ r-correlation coefficient



Since r = 0.88 is between 0.67 and 1, then there a strong positive correlation between the two.

Homework: Pg. 153: #2, 4i,6,8,10,12,14a

Oct 3-10:32 AM