Section 1.6 – Linear Regression by Hand

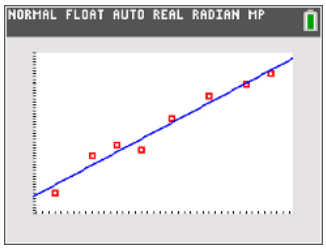
MDM4U
Jensen

Part 1: Linear Regression Using Technology Practice

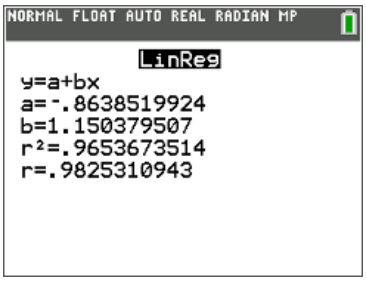
This table shows data for the full-time employees of a small company.

Age (years)	Annual Income ($000)
33	33
25	31
19	18
44	52
50	56
54	60
38	44
29	35

a) Generate a scatterplot of the data.



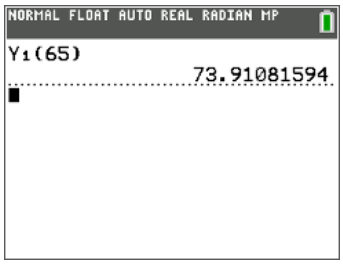
b) Perform a linear regression and state the equation of the line of best fit. Explain what the slope and y-intercept mean in context.



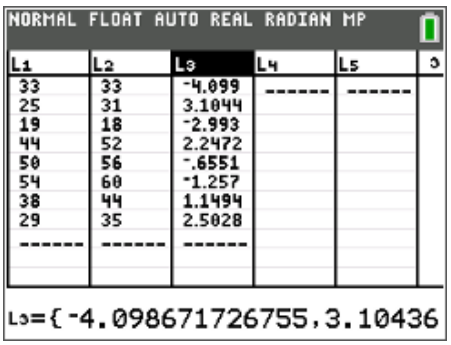
c) What is the correlation coefficient? What does this tell you about the relationship between age and annual income?

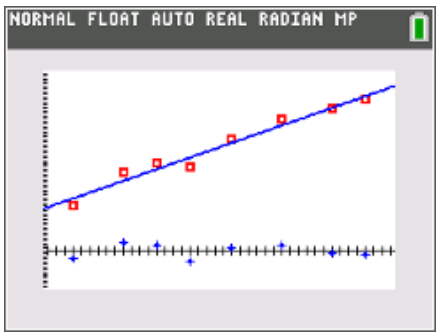
d) What is the coefficient of determination? What does it mean?

e) Use the line of best fit to predict the income for a 65 year old employee.



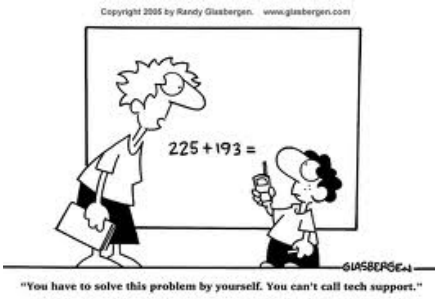
f) Find the residual values. What do they tell you about the correlation between the two variables?





Part 2: Linear Regression by Hand

Example: The following table lists the mathematics of data management marks and grade 12 averages for a small group of students. Start by completing filling in the missing cells. You will need these values to calculate the correlation coefficient and equation of the line of best fit.



MDM4U Mark (x)	Grade 12 Average (y)	x^2	y^2	xy
74	77	5476		
81	87	6561	7569	7047
66	68		4624	4488
53	67			3551
92	85	8464		7820
45	55	2025	3025	
80	76		5776	
$\sum x =$	$\sum y =$	$\sum x^2 =$	$\sum y^2 =$	$\sum xy =$

a) Determine the equation of the least squares regression line (line of best fit)

b) Calculate the correlation coefficient by hand