

# ALGEBRA - CALCULATING A LINEAR REGRESSION BY THE METHOD OF "LEAST SQUARES."

SEE  
PAGE 242  
EXAMPLE 2

IN CLASS, WE USED THE GRAPHING CALCULATOR TO FIND A "TREND LINE" OR "BEST FIT" LINE FOR THE DATA IN EXAMPLE 2 ON PAGE 242 OF OUR TEXT. NO MATTER HOW YOU DO A "LEAST SQUARES" FIT TO THE DATA, WHETHER BY USING A CALCULATOR, A SPREADSHEET, COMPUTER PROGRAM, OR BY "HAND", THE FINAL EQUATION OF THE BEST FIT LINE SHOULD BE THE SAME IN EACH CASE.

BELOW, I'LL SHOW YOU HOW TO DO A LEAST SQUARES FIT OF THE DATA BY HAND. YOU CAN USE THIS METHOD FOR ANY SCATTER PLOT OR TABLE OF X-Y PAIRS YOU THINK ARE RELATED BY A LINEAR FUNCTION. I THINK YOU WILL SEE WHY THIS METHOD IS RARELY DONE BY HAND! THE MATH IS EASY TO DO, IT IS JUST TEDIOUS - AN IDEAL APPLICATION FOR AUTOMATION!

---

STEP 1 MAKE A TABLE OF THE X-Y PAIRS, ADD A COLUMN FOR  $X^2$  AND X TIMES Y ( $XY$ ). FOR OUR DATA, THE  $X$ 'S ARE THE YEARS FROM 1980-1993.

STEP 2 FIND THE SUM OF EACH COLUMN AND FIND THE MEAN OF THE FIRST TWO COLUMNS. ALSO, COUNT THE TOTAL NUMBER OF X-Y PAIRS, FOR OUR PROBLEM THIS IS 14.

STEP 3 SUBSTITUTE THE COLUMN SUMS AND COLUMN MEANS INTO A FORMULA FOR THE SLOPE ( $a$ ) AND THE Y-INTERCEPT ( $b$ ) IN THE EQUATION OF THE LINE

$$y = ax + b$$

↑                      ↙  
SLOPE =  $a$           Y-INTERCEPT =  $b$

---

STEP 1

COLUMN	YEAR SALES			
	①	②	③	④
	X	Y	X <sup>2</sup>	X·Y
	80	2.05	6400	164.00
	81	2.3	6561	186.3
	82	2.45	6724	200.9
	83	2.7	6889	224.1
	84	3.2	7056	268.8
	85	3.45	7225	293.25
	86	3.65	7396	313.9
	87	3.75	7569	326.25
	88	3.9	7744	343.2
	89	4.2	7921	373.8
	90	4.6	8100	414
	91	5	8281	455
	92	5.35	8464	492.2
	93	5.6	8649	520.8

Where SALES (in billions of \$) of GREETING CARDS

N = 14  
N = NUMBER OF (X, Y) PAIRS OF DATA

STEP 2

COLUMN SUMS → 1211 52.2 104979 4576.5

← SUMS

COLUMN ① & ② MEANS → 86.5 3.7286

← MEANS (AVERAGES)

STEP 3

MOST OF THE WORK IS DONE, WE JUST NEED TO SUBSTITUTE THE CORRECT VALUES FROM OUR WORK INTO THIS FORMULA TO FIND THE SLOPE "a" IN OUR LINEAR EQUATION FOR THE BEST FIT LINE  $y = ax + b$

↑ SLOPE ↑ Y-INTERCEPT

$$* \text{ SLOPE} = a = \frac{\text{SUM OF COL. ④} - \left[ \frac{(\text{SUM OF COL. ①}) \cdot (\text{SUM OF COL. ②})}{N} \right]}{\text{SUM OF COL. ③} - \frac{(\text{SUM OF COL. ①})^2}{N}}$$

SUBSTITUTING:

$$\text{SLOPE} = a = \frac{4576.5 - \left[ \frac{(1211)(52.2)}{14} \right]}{104979 - \left[ \frac{(1211)^2}{14} \right]} = \frac{4576.5 - 4515.3}{104979 - 104751.5}$$

∴ **SLOPE = a = 0.269010989** RECOGNIZE THIS! See, you can do it!!

THE FORMULA FOR b, THE Y-INTERCEPT IS:  $b = \text{MEAN OF COL. ②} - [(\text{SLOPE}) \cdot (\text{MEAN OF COL. ①})]$

$$\therefore b = 3.7286 - [(0.269010989)(86.5)] = \boxed{-19.54085055 = b}$$

$$\therefore \boxed{Y = 0.269010989X - 19.54085055}$$

EQUATION OF LINE OF LEAST SQUARES FIT TO DATA.