
Peter Regan teaching MBA Math at Tuck

MBA Math Sample Exercise

Statistics: Linear Regression

Moving well beyond generic GMAT aptitude questions, the MBA Math sample exercises allow prospective MBA students to self assess their proficiency with the quantitative building blocks of the MBA first year curriculum.

MBA focused, time efficient, and convenient, the MBA Math online course builds your quantitative skills. Furthermore, the MBA Math transcript provides a concise summary for admissions officers that you are up to speed.

Visit **www.mbamath.com** to learn more!

Exercise

Consider the following sample data for the relationship between advertising budget and sales for Product A:

Observation	1	2	3	4	5	6	7	8	9	10
Advertising ($)	60,000	70,000	70,000	80,000	80,000	90,000	100,000	100,000	100,000	110,000
Sales ($)	362,000	416,000	417,000	499,000	485,000	536,000	602,000	623,000	616,000	663,000

What is the slope of the "least-squares" best-fit regression line?

Solution

Solution Commentary

(5:15)

Manual Solution

The equation for the best-fit regression line through a set of points is given by y = mx + b where m is the slope and b is the y-intercept.
The slope m is given by the covariance of x and y divided by the variance of x.
The y-intercept b is given by y_{avg} - m * x_{avg} where y_{avg} is the mean of y, m is the slope, and x_{avg} is the mean of x.

We take these definitions as given without explanation and focus on computation.

We now compute the various intermediate values needed to define the equation of the regression line.

Let's compute first the means of x and y, which in our case are the advertising and sales values, respectively.

Mean of x = (60,000 + 70,000 + 70,000 + 80,000 + 80,000 + 90,000 + 100,000 + 100,000 + 100,000 + 110,000)/10
Mean of x = 86,000

Mean of y = (362,000 + 416,000 + 417,000 + 499,000 + 485,000 + 536,000 + 602,000 + 623,000 + 616,000 + 663,000)/10
Mean of y = 521,900

Let's now compute the variance of x:

Variance of x = $((60{,}000 - 86{,}000)^2 + (70{,}000 - 86{,}000)^2 + (70{,}000 - 86{,}000)^2 + (80{,}000 - 86{,}000)^2 + (80{,}000 - 86{,}000)^2 + (90{,}000 - 86{,}000)^2 + (100{,}000 - 86{,}000)^2 + (100{,}000 - 86{,}000)^2 + (110{,}000 - 86{,}000)^2)/10$

Variance of x = 244,000,000

Next we compute the covariance of x and y:

Covariance of x and y = ((60,000 - 86,000)*(362,000 - 521,900) + (70,000 - 86,000)*(416,000 - 521,900) + (70,000 - 86,000)*(417,000 - 521,900) + (80,000 - 86,000)*(499,000 - 521,900) + (80,000 - 86,000)*(485,000 - 521,900) + (90,000 - 86,000)*(536,000 - 521,900) + (100,000 - 86,000)*(602,000 - 521,900))/10

Covariance of x and y = 1,518,600,000

We can now put together the pieces of the regression line equation:

The slope m = 1,518,600,000/244,000,000 = 6.22
The y-intercept b = 521,900 - 6.22 * 86,000 = -13,344

The value requested in this exercise is the slope, which is 6.22

Excel Solution

The Excel solution image omits six of the data points for space reasons.

	A	B	C	D	E
1	Observation	1	2	3	4
2	Advertising (X)	60,000	70,000	70,000	80,000
3	Sales (Y)	362,000	416,000	417,000	499,000
4					
5	Square Diff. from Mean X	676,000,000	256,000,000	256,000,000	36,000,000
6	Square Diff. from Mean Y	25,568,010,000	11,214,810,000	11,004,010,000	524,410,000
7	Diff. from Mean XY	4,157,400,000	1,694,400,000	1,678,400,000	137,400,000
8					
9		Basic X	Basic Y	Basic XY	
10	Mean	86,000	521,900		
11	Variance (Population)	244,000,000	9,527,290,000		
12	Std Devn (Population)	15,620	97,608		
13					
14	Covariance (Population)			1,518,600,000	
15	Correlation			1.00	
16					
17	Slope	6.22			
18	Y-Intercept	-13,344			
19					

Sheet1

Peter Regan teaches decision science courses at Dartmouth's Tuck School and Duke's Fuqua School. He also teaches pre-term quantitative skills courses at Tuck and Cornell's Johnson School. He created the MBA Math self-paced, online pre-MBA quantitative skills course covering finance, accounting, economics, statistics, and spreadsheets.