

Department of Computer Science, University of Karachi  
**BSCS: 306: Probability and Statistical Methods**  
 Course Tutor: Dr Tahseen A. Jilani  
**Assignment: 11: Correlation and Linear Regression Models**

---

- Q 1.** An article in the *Journal of Monetary Economics* assesses the relationship between percentage growth in wealth over a decade and a half of savings for baby boomers of age 40 to 55 with these people's income quartiles. The article presents a table showing five income quartiles, and for each quartile there is a reported percentage growth in wealth. The data are as follows.

Income quartile	1	2	3	4	5
Wealth growth (%)	17.3	23.6	40.2	45.8	56.8

Run a simple linear regression of these five pairs of numbers and estimate a linear relationship between income and percentage growth in wealth.

- Q 2.** A financial analyst at Goldman Sachs ran a regression analysis of monthly returns on a certain investment ( $Y$ ) versus returns for the same month on the Standard & Poor's index ( $X$ ). The regression results included  $SS_X=765.98$  and  $SS_{XY}=934.49$ . Give the least-squares estimate of the regression slope parameter.

- Q 3.** Recently, research efforts have focused on the problem of predicting a manufacturer's market share by using information on the quality of its product. Suppose that the following data are available on market share, in percentage ( $Y$ ), and product quality, on a scale of 0 -100, determined by an objective evaluation procedure ( $X$ ):

<b>X</b>	27	39	73	66	33	43	47	55	60	68	70	75
<b>Y</b>	2	3	10	9	4	6	5	7	9	10	13	12

Estimate the simple linear regression relationship between market share and product quality rating.

- Q 4.** The following are data on annual inflation and stock returns. Run a regression analysis of the data and determine whether there is a linear relationship between inflation and total return on stocks for the periods under study.

<b>Inflation (%)</b>	1	2	13	-10.3	0.5	2	-1.8	5.8	5.9
<b>Total Return on Stocks (%)</b>	-3	36	12	-8	53	-2	18	32	24

- Q 5.** For the data given below, regress one variable on the other. Is there an implication of causality, or are both variables affected by a third?

**Sample of Annual Transactions (\$ millions)**

<b>Year</b>	2002	2003	2004	2005	2006	2007
<b>Credit Card</b>	156	204	279	472	822	1,213
<b>Online Debit Card</b>	211	280	386	551	684	905

- Q 6.** Use the template to regress McDonald's international sales, then answer the following questions:

**10 Years Sales for McDonald's at Year End (in billions)**

<b>U.S. Sales</b>	7.6	7.9	8.3	8.6	8.8	9	9.4	10.2	11.4	12.1
<b>International Sales</b>	2.3	2.6	2.9	3.2	3.7	4.1	4.8	5.7	7	8.9

(a). What is the regression equation? (b). What is the 95% confidence interval for the slope? (c). What is the standard error of estimate?

- Q 7.** A study was carried out to determine whether there is a linear relationship between the time spent in negotiating a sale and the resulting profits. A random sample of 27 market transactions was collected, and the time taken to conclude the sale as well as the resulting profit were recorded for each transaction. The sample correlation coefficient was computed:  $r= 0.424$ . Is there a linear relationship between the length of negotiations and transaction profits?