Lab Task: Linear Regression: To find relationship between two variables

Task:

Build an excel worksheet calculates a linear regression equation. First two columns will contain two integers on each row and multiple rows in worksheet.

Regression Definition:
A regression is a statistical analysis assessing the association between two variables. It is used to find the relationship between two variables.

Regression Formula:
Regression Equation$(y) = w_0 + w_1x$
Slope$(w_1) = (N\Sigma XY - (\Sigma X)(\Sigma Y)) / (N\Sigma X^2 - (\Sigma X)^2)$
Intercept$(w_0) = (\Sigma Y - b(\Sigma X)) / N$

where
       x and y are the variables.
       w_1 = The slope of the regression line
       w_0 = The intercept point of the regression line and the y axis.
       N = Number of values or elements
       X = First Score
       Y = Second Score
       ΣXY = Sum of the product of first and Second Scores
       ΣX = Sum of First Scores
       ΣY = Sum of Second Scores
       ΣX^2 = Sum of square First Scores

Regression Example: To find the Simple/Linear Regression of

X Values	Y Values
60	3.1
61	3.6
62	3.8
63	4
65	4.1

To find regression equation, we will first find slope, intercept and use it to form regression equation.

Step 1: Count the number of values.   N = 5

Step 2: Find XY, X^2 See the below table

X Value	Y Value	X*Y	X*X
60	3.1	60 * 3.1 = 186	60 * 60 = 3600
61	3.6	61 * 3.6 = 219.6	61 * 61 = 3721
62	3.8	62 * 3.8 = 235.6	62 * 62 = 3844
63	4	63 * 4 = 252	63 * 63 = 3969
65	4.1	65 * 4.1 = 266.5	65 * 65 = 4225

Step 3: Find ΣX, ΣY, ΣXY, ΣX^2.

$\Sigma X = 311$
$\Sigma Y = 18.6$
$\Sigma XY = 1159.7$
$\Sigma X^2 = 19359$

Step 4: Substitute in the above slope formula given.

$\text{Slope}(w_1) = (N\Sigma XY - (\Sigma X)(\Sigma Y)) / (N\Sigma X^2 - (\Sigma X)^2)$
$= ((5)*(1159.7)-(311)*(18.6))/((5)*(19359)-(311)^2)$
$= (5798.5 - 5784.6)/(96795 - 96721)$
$= 13.9/74$
$= 0.19$

Step 5: Now, again substitute in the above intercept formula given.

$\text{Intercept}(w_1) = (\Sigma Y - w_1 (\Sigma X)) / N$
$= (18.6 - 0.19(311))/5$
$= (18.6 - 59.09)/5$
$= -40.49/5$
$= -8.098$

Step 6: Then substitute these values in regression equation formula

$\text{Regression Equation (y)} = w_0 + w_1x$
$= -8.098 + 0.19x.$

Suppose if we want to know the approximate y value for the variable x = 64. Then we can substitute the value in the above equation.

$\text{Regression Equation(y)} = w_0 + w_1x$
$= -8.098 + 0.19(64).$
$= -8.098 + 12.16$
$= 4.06$

This example will guide you to find the relationship between two variables by calculating the Regression from the above steps.

Data Set:

Year end foreign exchange rate Pak Rupees per US $ taken from State Bank of Pakistan [1].

YEAR	YEAR #	PKRS PRICE PER USD
1981	1	13
1982	2	14
1983	3	14
1984	4	17
1985	5	17

1986	6	18
1987	7	19
1988	8	22
1989	9	22
1990	10	25
1991	11	26
1992	12	28
1993	13	31
1994	14	32
1995	15	36
1996	16	41
1997	17	47
1998	18	52
1999	19	53
2000	20	65
2001	21	61
2002	22	58
2003	23	59
2004	24	60
2005	25	61
2006	26	61
2007	27	69
2008	28	82
2009	29	86
2010	30	86
2011	31	95
2012	32	99

Submission Procedure

Submit your excel sheet file like [NAME].[xlsx] for example ALI.xlsx etc. in an email with subject title "[ALI]-LABWORK-01" at swjaffry@pucit.edu.pk.

[1]   http://www.sbp.org.pk/ecodata/HER-USDollar.xls