Section 1.6 Worksheet – Linear Regression by Hand

MDM4U
Jensen

1) Sand driven by wind creates large dunes at the Great Sand Dunes National Monument in Colorado. Is there a linear relationship correlation between wind velocity and sand drift rate? A test site at the Great Sand Dunes National Monument gave the following information about x, wind velocity in 10cm/sec, and y, drift rate of sand in 100g/cm/sec.

a) Complete the chart

Wind Speed [x]	**Drift Rate [y]**	x^2	y^2	xy
70	3			
115	45			
105	21			
82	7			
93	16			
125	62			
88	12			
$\sum x =$	$\sum y =$	$\sum x^2 =$	$\sum y^2 =$	$\sum xy =$

b) Determine the equation of the least squares regression line ($\hat{y} = a + bx$). Interpret the slope and y-intercept in context.

Slope = $b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$

y-intercept = $a = \bar{y} - b\bar{x}$

c) Compute the correlation coefficient using the formula. Interpret r and r^2 in context.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

2) A study was conducted to determine if larger universities tend to have more property crime. Let x represent student enrollment (in thousands) and let y represent the number of burglaries in a year on the campus. A random sample of 8 universities in California gave the following information:

a) Complete the chart

Student Enrollment [x]	**Burglaries [y]**	x^2	y^2	xy
12.5	26			
30	73			
24.5	39			
14.3	23			
7.5	15			
27.7	30			
16.2	15			
20.1	25			
$\sum x =$	$\sum y =$	$\sum x^2 =$	$\sum y^2 =$	$\sum xy =$

b) Determine the equation of the least squares regression line ($\hat{y} = a + bx$) by hand. Interpret the slope and y-intercept in context.

Slope = $b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$

y-intercept = $a = \bar{y} - b\bar{x}$

c) Compute the correlation coefficient using the formula. Interpret r and r^2 in context.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$