University of Guelph Numeracy Project

About Bivariate Correlations and Linear Regression: Examples

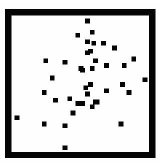
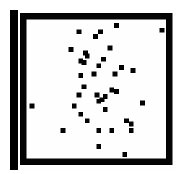
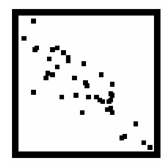
UNIVERSITY of GUELPH

TABLE OF CONTENTS

About Bivariate Correlations and Linear Regression: Examples

Types of Bivariate Correlations

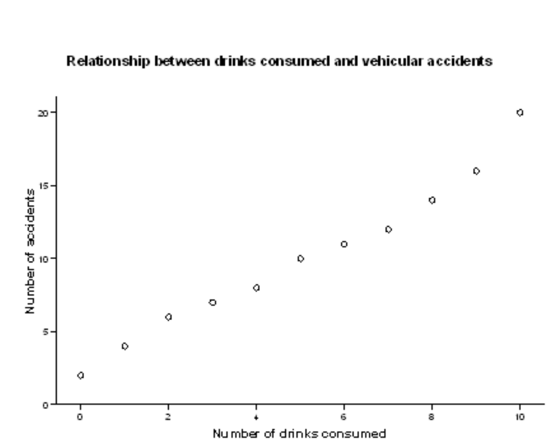
► Given the following three graphs, select the correct description for each relationship

weak positive   zero correlation   strong negative

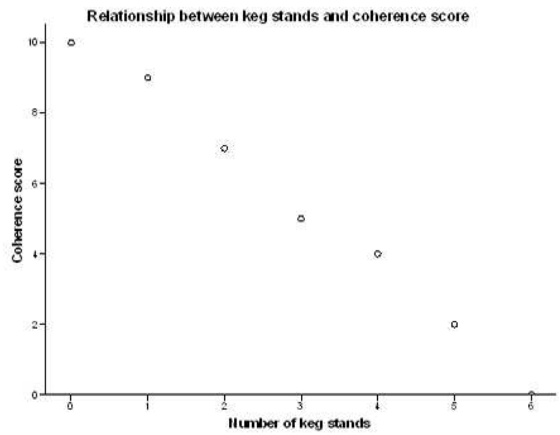
Positive Correlation

► A researcher examined the relationship between the number of alcoholic beverages consumed prior to driving and the number of vehicular accidents. A plotted graph indicated a strong positive relationship between the number of drinks consumed and the number of of accidents.



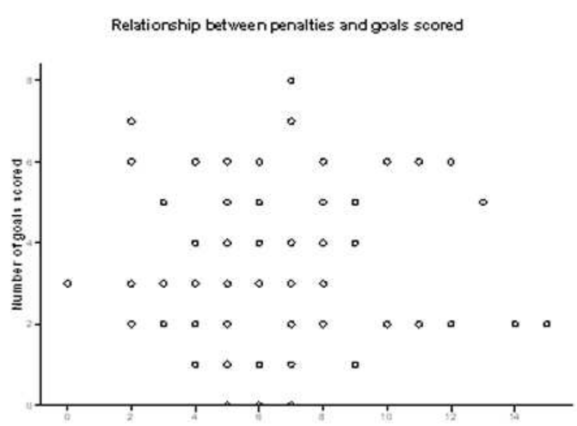
Negative Correlation

► A researcher attended a keg party and examined the relationship between coherence and the number of keg stands performed by an individual. Coherence was measured as a score from 0 to 10 (0 indicated no coherence and 10 indicated high coherence). A plotted graph indicated a negative relationship between coherence and the number of keg stands performed.

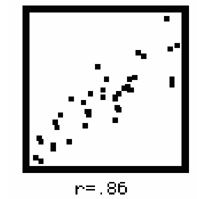


Zero Correlation

► A researcher examined the relationship between the number of goals that the Ottawa Senators scored and the number of penalties in each game of the 2006-2007 season. A plotted graph indicated a zero relationship between the number of penalties committed and the number of goals scored.



► The following graph represents a:
A) Strong negative correlation
B) Weak positive correlation
C) Strong positive correlation
D) Zero correlation



Determining the Value of r

► A researcher examined the relationship between the number of alcoholic beverages consumed prior to driving and the number of vehicular accidents. The data is provided below:

X Drinks	X^2 Drinks^2	Y Accidents	Y^2 Accidents^2	XY (D)(A)
0	0	2	4	0
1	1	4	16	4
2	4	6	36	12
3	9	7	49	21
4	16	8	64	32
5	25	10	100	50
6	36	11	121	66
7	49	12	144	84
8	64	14	196	112
9	81	16	256	144
10	100	20	400	200
$\Sigma X=55$	$\Sigma X^2=385$	$\Sigma Y=110$	$\Sigma Y^2=1386$	$\Sigma XY=725$

To determine the value of r, simply plug the data values into the formula

$$r = \frac{N(\sum XY) - (\sum X)(\sum Y)}{\{ [N(\sum X^2)-(\sum X)^2] [N(\sum Y^2) - (\sum Y)^2] \}^{1/2}}$$

$$r = \frac{11(725) - (55)(110)}{\{ [11(385^2)-(55)^2] [11(1386^2)-(110)^2] \}^{1/2}}$$

$$r = \frac{7975 - 6050}{\{ [4235-3025] [15246-12100] \}^{1/2}}$$

$$r = \frac{1925}{\{ [1210] [3146] \}^{1/2}}$$

$$r = \frac{1925}{\{3806660\}^{1/2}}$$

$$r = \frac{1925}{1951.07}$$

$$r = 0.99$$

Therefore the relationship between the number of drinks consumed and the number of vehicular accidents is r = 0.99.

This is a strong, positive correlation.

Determining the Value of R^2

- In the previous example, the correlation coefficient was r = 0.99. The coefficient of determination is:
  $R^2 = r^2 = .99 = .98$

  Therefore, 98% of the variance in the number of accidents is accounted for by the number of drinks consumed, and 2% of the variance in the number of accidents is accounted by some other factor.

Determining the Value of r and R^2

- A researcher conducted a study to examine the relationship between the number of alcohol beverages consumed and the number of parties attended by adults in a month.

  The researcher obtained the following data:

  ΣX(alcohol beverages consumed) =120
  ΣX^2 =1240
  ΣY(parties attended) = 1536,
  ΣY^2 =157050
  ΣXY=13312
  N=16

What is the value of r? 0.99
What is the value of R? 0.98

What kind of relationship is this? Strong positive relationship

Determining the Value of a and b

► Using the data from the drinking and driving example, determine the value of a and b:

X	Y	$X-\bar{X}$	$(X-\bar{X})^2$	$Y-\bar{Y}$	$(X-\bar{X})(Y-\bar{Y})$
0	2	-5	25	-8	40
1	4	-4	16	-6	24
2	6	-3	9	-4	12
3	7	-2	4	-3	6
4	8	-1	1	-2	2
5	10	0	0	0	0
6	11	1	1	1	1
7	12	2	4	2	4
8	14	3	9	4	12
9	16	4	16	6	24
10	20	5	25	10	50
$\bar{X}=5$	$\bar{Y}=10$		$\Sigma=110$		$\Sigma=175$

To determine the value of b:

$$b = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\Sigma(X - X)^2}$$

$$b = \frac{175}{110}$$

$$b = 1.59$$

To determine the value of a :

$a = \overline{Y} – b\overline{X}$

a = 10 – 1.59(5)

a = 2.05

Therefore the regression equation for this example is:

Y' = 2.05 + 1.53X

► We want to determine how many accidents there would be if 5 drinks are consumed before driving:

Y' = 2.05 + 1.59X
Y' = 2.05 + 1.59(5)
Y' = 2.05 + 7.95
Y' = 10

Therefore, when 5 drinks are consumed before driving, there are approximately 10 accidents.