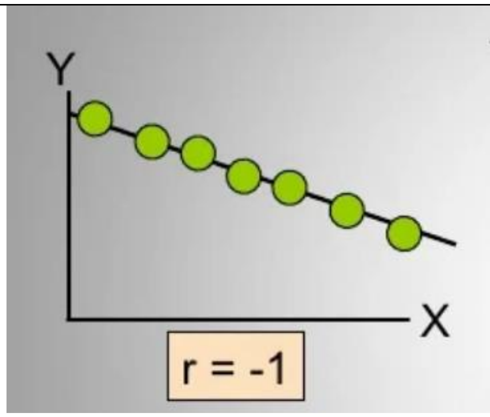
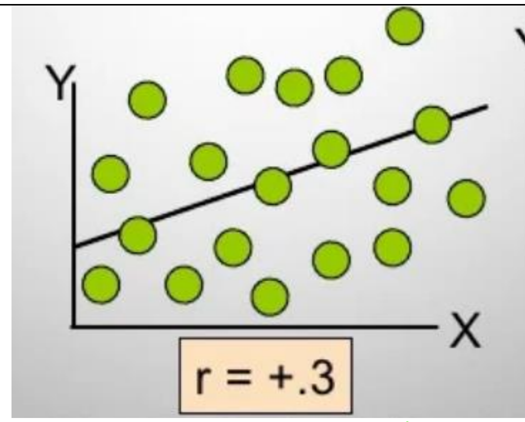
Chapter 11 Practice

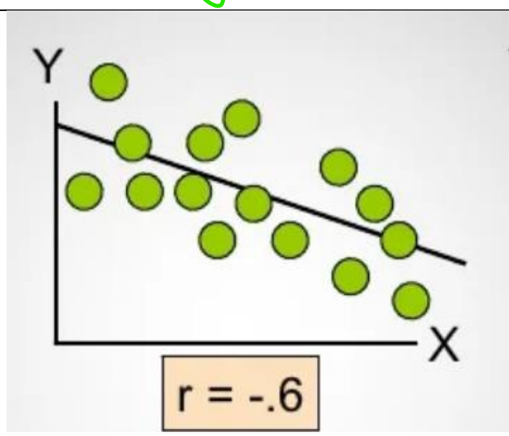
1. Look at the graphs below and interpret them.



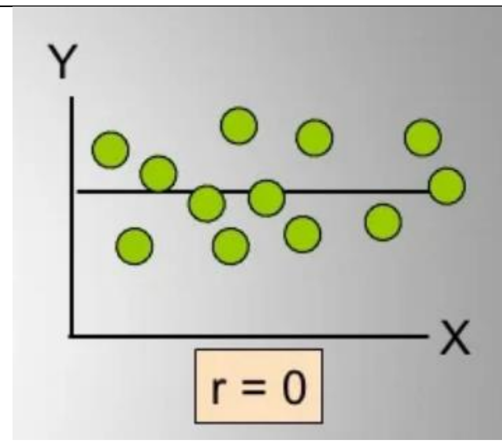
Perfect Negative Linear



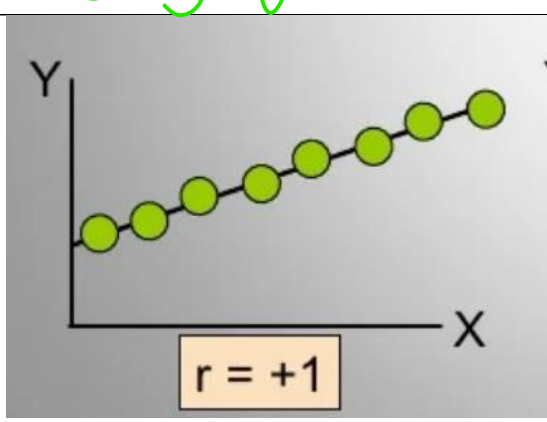
Weak Positive Linear



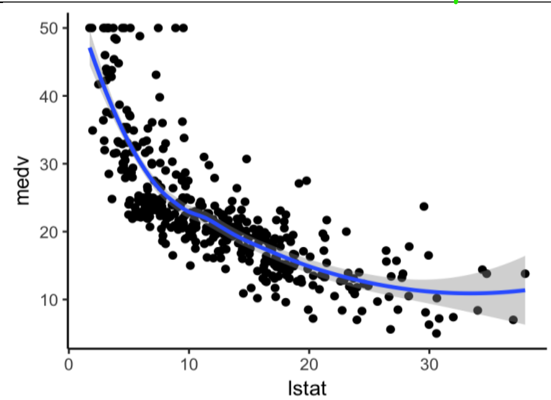
Strong Negative Linear



No Relationship



Perfect Positive Linear

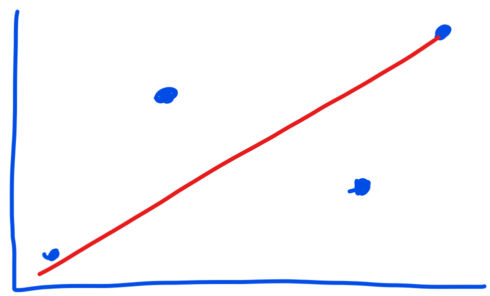


Negative Non-Linear

2. Trevor wants to examine the relationship between two field surveys on prey abundance. Answer the following questions using the data below.

	Sector 1	**Sector 2**	**Sector 3**	**Sector 4**
Trial 1 (X)	13	17	15	14
Trial 2 (Y)	10	19	12	17

a. Create a Scatter Plot on the calculator and draw the result.



b. Find the values below by hand.

$\sum x_i = 13+17+15+14 = 59$

$\sum y_i = 10+19+12+17 = 58$

$\sum x_i^2 = 13^2+17^2+15^2+14^2 = 879$

$\sum y_i^2 = 10^2+19^2+12^2+17^2 = 894$

$\sum((x_i)(y_i)) = (13\times10)+(17\times19)+(15\times12)+(14\times17) = 871$

c. Find the slope using the values from part b. $n=4$

$$\hat{\beta}_1 = \frac{(4)(871)-(59)\cdot(58)}{(4)(879)-(59)^2} = \frac{62}{35} = 1.77$$

d. Find the vertical intercept. $\bar{y} = \frac{10+19+12+17}{4} = 14.5$; $\bar{x} = \frac{13+17+15+14}{4} = 14.75$

$\hat{\beta}_0 = (14.5)-(1.77)(14.75) = -11.61$

e. Write the Linear Regression Model for this question.

$\hat{y} = -11.61 + 1.77x$ OK $\hat{y} = 1.77x - 11.61$

f. Using the values from part b, find the coefficient of correlation.

$$R = \frac{(4)(871) - (59)\cdot(58)}{\left(\sqrt{(4)(879) - (59)^2}\right)\cdot\left(\sqrt{(4)(894) - (58)^2}\right)} = 0.72$$

g. Find the coefficient of determination.

$$R^2 = (0.72)^2 = 0.52$$

h. Interpret what you got in parts f and g.

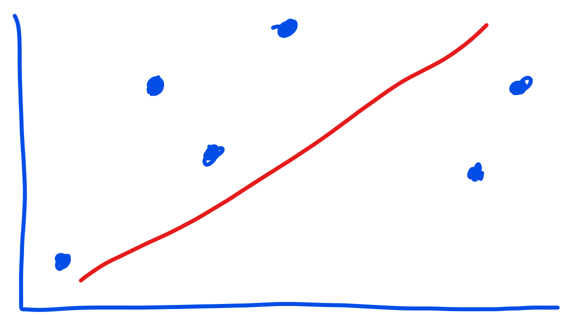
R: Strong Positive Linear Relationship

R^2: 52% of variation in y can be explained by X

3. Trevor saw that there was data on prey abundance from several decades ago. Analyze that data so Trevor can compare it to his.

	Sector 1	Sector 2	Sector 3	Sector 4	Sector 5	Sector 6
L_1 → **Trial 1** X	12	14	15	13	17	16
L_2 → **Trial 2** Y	9	13	17	16	14	12

a. Create a Scatter Plot on the calculator and draw the result.



b. Find the values below by calculator. (Via 2-Var Stats)

$\sum x_i$ = 87

$\sum y_i$ = 81

$\sum x_i^2$ = 1279

$\sum y_i^2$ = 1135

$\Sigma((x_i)(y_i))$ = 1183

Via LinRegTTest

c. Find the slope using the values from part b.

$\hat{\beta}_1 = b = 0.49$

d. Find the vertical intercept.

$\hat{\beta}_0 = a = 6.46$

e. Write the Linear Regression Model for this question.

$\hat{y} = 6.46 + 0.49x$

Still from LinRegTTest

f. Using the values from part b, find the coefficient of correlation.

$R = 0.32$

g. Find the coefficient of determination.

$R^2 = 0.10$

h. Interpret what you got in parts f and g.

R: Weak Positive Linear Relationship

R^2: 10% of variation in y can be explained by x

4. Use the R output below to answer the questions below.

```
Call:
lm(formula = myResponse ~ myFactor)

Residuals:
     1      2      3      4      5      6      7
 1.239 -9.249 -5.164  4.104 -1.030  5.617  4.483

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 102.4925     5.1381  19.948 5.85e-06 ***
myFactor     -3.6219     0.5649  -6.411  0.00137 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.055 on 5 degrees of freedom
Multiple R-squared:  0.8915,    Adjusted R-squared:  0.8699
F-statistic:  41.1 on 1 and 5 DF,  p-value: 0.00137
```

$\hat{\beta}_0$ $\hat{\beta}_1$ ϵ R^2 p-value

a. Write the Linear Regression Model for this question.

$$\hat{y} = 102.4925 - 3.6219x$$

b. What is the coefficient of determination?

$$R^2 = 0.8915$$

$\hat{\beta}_1$ is $-$ so R is too

c. What is the coefficient of correlation?

$$R = \sqrt{R^2} = \sqrt{0.8915} = -0.94$$

i. What is the strength?

Very Strong

ii. Positive or Negative?

Negative

d. Interpret the values found for part b and c.

R : Very Strong Negative Linear Relationship

R^2 : 89.15% of variation in y can be explained by x

5. For this last question, we are going to solve this example from the notes. We will be following along with putting the data in our calculators!

Example 11.20. consider the following situations.

Situation 3

Consider the data, which includes the point $(a, b) = (20, -20)$ at the end and figures shown below. Note that the least squares regression line is

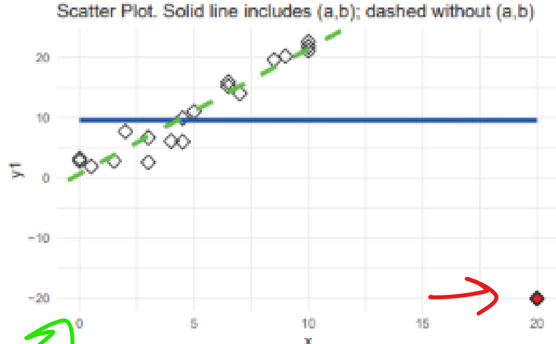
$$\hat{y} = 0x + 9.56.$$

		1	2	3	4	5	6	7	8	9	10
L1 →	x_i	1.50	3.00	0.00	0.50	5.00	6.50	8.50	9.00	0.00	10.00
L2 →	y_i	2.80	2.60	3.20	1.90	11.00	15.90	19.60	20.20	2.80	22.60
L3 →	$\hat{y}_i$	9.56	9.56	9.56	9.56	9.56	9.55	9.55	9.55	9.56	9.55
L4=L2-L3 →	Residual $e_i = y_i - \hat{y}_i$	-6.76	-6.96	-6.36	-7.66	1.44	6.35	10.05	10.65	-6.76	13.05
L5=L4/S_e →	Std Resid e_i/s_e	-0.66	-0.68	-0.62	-0.74	0.14	0.62	0.98	1.03	-0.66	1.27

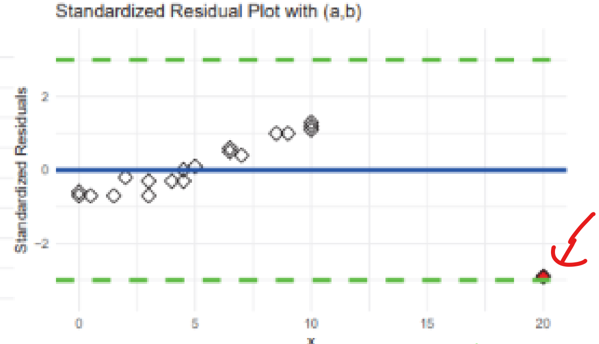
S_e — standard deviation for L4

	11	12	13	14	15	16	17	18	19	20
x_i	10.00	4.50	6.50	4.50	3.00	4.00	7.00	10.00	2.00	**20.00**
y_i	21.10	6.00	15.20	9.90	6.70	6.10	14.00	21.80	7.70	**-20.00**
$\hat{y}_i$	9.55	9.56	9.55	9.56	9.56	9.56	9.55	9.55	9.56	**9.53**
Residual $e_i = y_i - \hat{y}_i$	11.55	-3.56	5.65	0.34	-2.86	-3.46	4.45	12.25	-1.86	**-29.53**
Std Resid e_i/s_e	1.12	-0.35	0.55	0.03	-0.28	-0.34	0.43	1.19	-0.18	**-2.87**

Scatter Plot. Solid line includes (a,b); dashed without (a,b)



Standardized Residual Plot with (a,b)



A. Is the point $(20, -20)$ an outlier?

No b/c its just inside the boundry

B. Is the point $(20, -20)$ influential?

Yes because the dotted line & solid line are very different