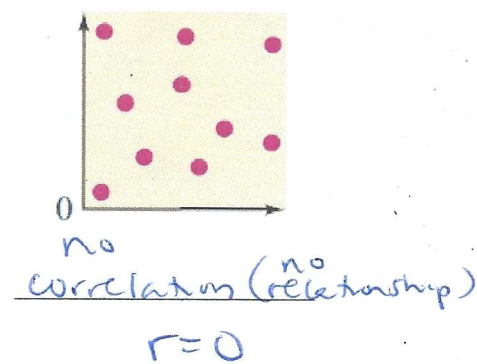
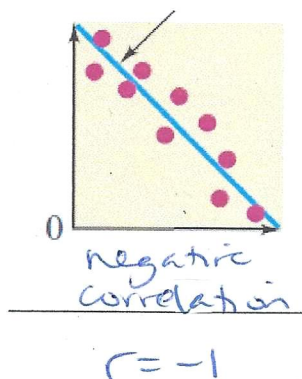
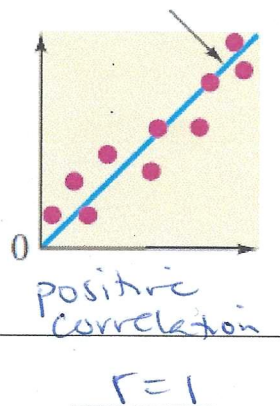


## 5.1 – Linear Regression

- **scatter plot** → a graph that relates two groups of data using ordered pairs <sup>(x,y)</sup> (or points)
  - **linear regression** → a trend line that shows the relationship between two sets of data that can be used to make predictions shown in a scatter plot
  - may also ask to find “line of best fit” or data has a “linear” relationship
  - Must put data in calculator in list 1 ( L1 ) and list 2 ( L2 ) and create a scatter plot
  - Equation for linear regression (using the calculator) →  $y = ax + b$  where you fill “a” part and “b” part with numbers that the calculator produces
  - After finding linear regression equation → will also state the equation’s correlation coefficient
- The closer “r” is to -1 or 1, then the stronger the correlation (points will cluster together)



**Examples:** Complete each problem.

L1 L2

1.)

x	y
1	21
2	15
3	12
4	9
5	7

- State the correlation coefficient:  $r = -0.9782$
- State the line of best fit:  $y = -3.4x + 23$
- Is the model a good fit? Yes NO Explain: r is close to -1
- Using the regression equation, find y when x = 18  $y = -38.2$  (table)
- Using the regression equation, find x when y = 46.8  $x = -7$  (intersection)

L1 Recreation Expenditures L2

2.)

Year	Dollars (billions)
1993	340
1994	369
1995	402
1996	430
1997	457
1998	489
1999	527
2000	574

- Circle the correlation coefficient that represents the given data:  
A.)  $r = -0.99645$  B.)  $r = 0.99745$  C.)  $r = 0.99645$  D.)  $r = 0.98645$
- Circle the equation that BEST represents the given data:  
A.)  $y = \frac{96}{3}x - 64105$  B.)  $y = \frac{97}{3}x - 64105$  C.)  $y = \frac{97}{3}x + 64105$   
→ 32 → 32.3 → 32.3
- What amount of money will recreation earn in the year 2011? about \$9176.11 billion  
 $x =$   $y = ?$  (table)

3.) **Olympic 500-Meter Men's Gold Medal Speed Skating Times**

Year	1980	1984	1988	1992	1994	1998
Time (seconds)	422	432	404	420	395	382

a.) Find the linear regression model that fits this data where let  $x$  (L1) = numbers of years since 1980

$$y = -2.29217x + 430.56024$$

b.) Predict the amount of time in the year 2013.  $x = 33$

about 355 seconds  $y = ?$   
(table)

4.) Which of the following tables represents a direct variation? If it is direct variation, state the equation.

Note: Direct Variation equation is represented by the power function  $y = k \cdot x^p$  where  $p$  is a positive number

a.)

x	y
-2	3.2
1	2.4
4	1.6

No

$$y = -2.67x + 2.67$$

b.)

x	y
-2	1
3	6
8	11

No

$$y = x + 3$$

c.)

x	y
4	6
8	12
10	15

yes

$$y = 1.5x$$

5.) **Average Temperatures in Northern Latitudes**

Latitude ( $^{\circ}$ N)	0	10	20	30	40	50	60	70	80
Temp. ( $^{\circ}$ F)	79.2	80.1	77.5	68.7	57.4	42.4	30.0	12.7	1.0

a.) What is the average temperature for New Bern, NC (Latitude line is  $35.1083^{\circ}$  N)?  $55.1^{\circ}$  F

b.) What is the Latitude line for a city that has an average temperature of  $33.5^{\circ}$  F?  $55.45^{\circ}$  N

6.) Your height and arm span have a linear relationship.

I've selected 6 students from the class and measured each person's height and arm span in inches.

Find the linear regression equation to answer each question.

Student's Name	Height (inches)	Arm Span (inches)
Celeste Blocker	70	72
Tay Quan Bryan	70	76
Colby Parker	70	69
Jahyra Dancy	65	64
Moe Eh	69	64
Eden Tellis	62	60

$$y = 1.412790698x - 28.09883721$$

$(r = .8075002161)$

a.) Using the regression equation, predict the height of a person with 5 ft arm span.  $62.4$  in

b.) Using the regression equation, predict the arm span of a person who is 7'4" tall.  $96.2$  in

(table)  $x = 88$   $y = ?$