Correlation Coefficient

Agenda

Warm-Up

HW Check

Notes p.64

HW: #1-6

Reminders

Unit 5 TEST next Tuesday!

Essential Question

How do I find the strength and direction of a correlation in a data set?

Warm-Up Wednesday

1. As a science experiment, Keith recorded the percent humidity and the number of stars he could see at 10:00 P.M. each evening.

Star Counting Experiment										
Humidity (%)	84	76	79	88	95	82	87	88	75	82
Number of Visible Stars	12	22	25	15	11	19	13	18	20	22

X
Y

a. Write the equation of the line of best fit $y = -.59x + 67.16$

b. Using the equation above, estimate the number of stars visible at 100% humidity.

$x = 100$ about 8 stars

c. Using the equation above, estimate the humidity when Keith sees 40 stars.

$y = 40$ 45% humidity

LinReg
y=ax+b
a=-.591607565
b=67.15839243

Algebra I Unit 5- Using Linear Regression to Estimate Solutions and Make Predictions

Student Practice–Using Linear Regression to Estimate Solutions and Make Predictions

Name _____ Date_____ Period _____

Read and answer the following questions. Round your answer to the nearest hundredth.

1. As a science experiment, Keith recorded the percent humidity and the number of stars he could see at 10:00 P.M. each evening.

Star Counting Experiment										
Humidity (%)	84	76	79	88	95	82	87	88	75	82
Number of Visible Stars	12	22	25	15	11	19	13	18	20	22

a. Write the equation of the line of best fit_____

b. Using the equation above, estimate the number of stars visible at 100% humidity.

c. Using the equation above, estimate the humidity when Keith sees 40 stars.

2. Hummingbird wing beat rates are much higher than those in other birds. Estimates for various species are given in the table.

Hummingbird Wing Beats

Mass (g)	3.1	2.0	3.2	4.0	3.7	1.9	4.5
Wing Beats	60	85	50	45	55	90	40

a. Write the equation of the line of best fit_____

b. Using the equation above, estimate the wing beat rate of a 6.5 gram hummingbird.

c. Predict the wing beat rate for a Giant Hummingbird with a mass of 9 grams. Does your answer make sense?

3. The table below represents the age of a person, x, and their normal systolic blood pressure, y.

Age	Systolic Blood Pressure
10	115
30	125
50	135
70	145

a. What equation could be used to determine a person's normal systolic blood pressure?

b. What is the age of a person when his Systolic Blood Pressure is 161?

Algebra I Unit 5- Using Linear Regression to Estimate Solutions and Make Predictions

4. As scuba divers descend, the pressure of the water increases. Scuba divers can determine their depth by the pressure. Pressure can be expressed in atmospheres. An atmosphere is equivalent to 14.7psi (pounds per square inch) of pressure. The table below shows the relationship between atmospheres of pressure and ocean depth.

Depth of Ocean (feet)	0	33	66	99	132
Pressure (atmosphere)	1	2	3	4	5

a. What equation could represent this situation?

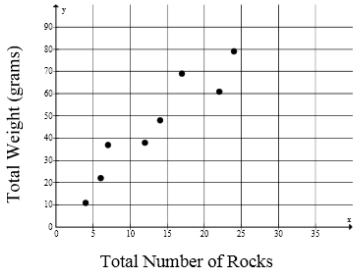
b. What is the atmospheric pressure when the depth of the ocean is 500 feet?

5. The table below lists corresponding *x*-and *y*-values of a linear function.
What is the value of *y* when *x* = 5?

x	***y***
0	3
1	12
2	21
3	30

A 39
B 40
C 48
D 50

6. A science class recorded the weight of different bags of rocks. Their results are displayed in the scatterplot.



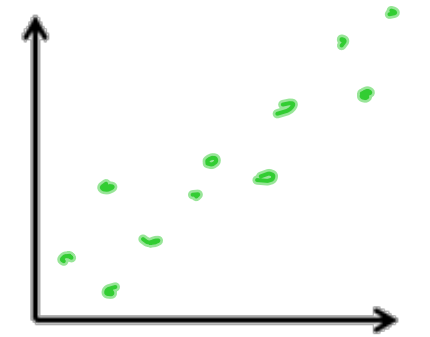
a. Write the equation of the line of best fit for this data and sketch it on the graph above.

b. If the bag had 45 rocks, what would be a reasonable estimate of its weight?

Correlation Coefficient p.64

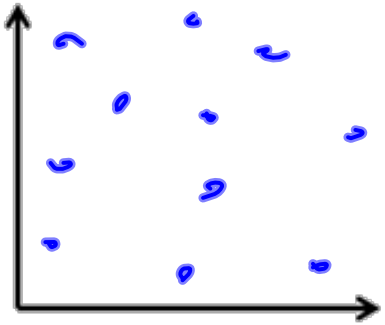
Essential Question

How do I find the strength and direction of a correlation in a data set?



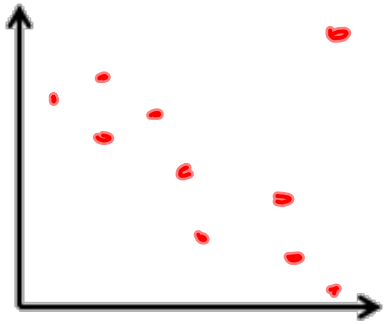
Positive Correlation

When one quantity goes UP, the other goes UP



No Correlation

There is no relationship between the two quantities



Negative Correlation

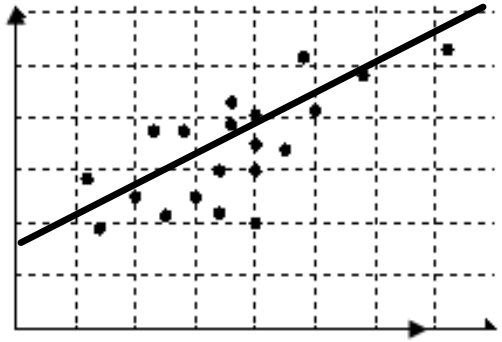
When one quantity goes UP, the other goes DOWN

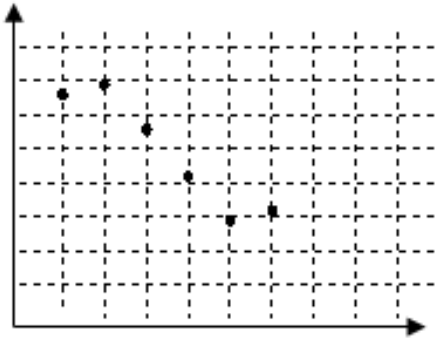
Correlation Coefficient p.64

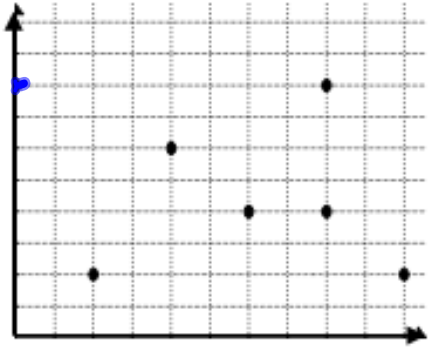
Essential Question

How do I find the strength and direction of a correlation in a data set?

EXAMPLES







1. Correlation: POSITIVE	2. Correlation: NEG.	3. Correlation: NO
(Linear) Nonlinear	(Linear) Nonlinear	Linear (Nonlinear)
Describe relationship between x and y: As $x\uparrow$, $y\uparrow$	Describe relationship between x and y: $x\uparrow$ $y\downarrow$	Describe relationship between x and y: no relationship

Correlation Coefficient p.64

Essential Question

How do I find the strength and direction of a correlation in a data set?

4. Determine if the following relationships represent positive, negative, or no correlation.

Hours of studying and grades	study ↑ grades ↑ Positive
Test scores and shoe size	test ↑ shoe? NO correlation
Number of pets a person has and number of books the person reads	pets ↑ books? NO correlation
How tall a person is and how fast they drive	tall ↑ drive? No correlation
Number of days absent from school and grades	abs ↑ grades ↓ negative
Temperature and the number of people wearing jackets	Negative

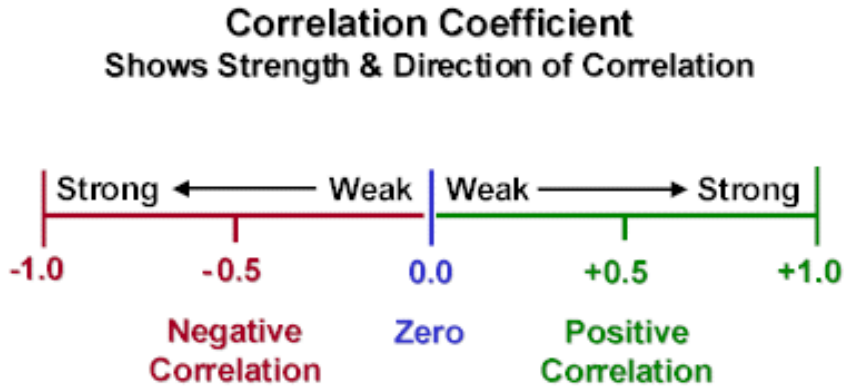
temp ↑ ↓ jackets ↓ ↑

Correlation Coefficient p.64

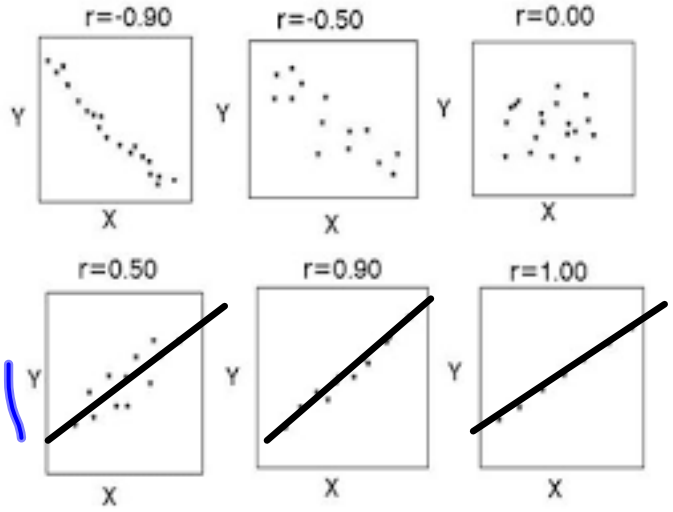
Essential Question How do I find the strength and direction of a correlation in a data set?

Correlation Coefficient: r

a NUMBER that determines if the correlation is positive or negative, as well as how closely the equation models the data



strong → r close to 1



Correlation Coefficient p.64

Essential Question How do I find the strength and direction of a correlation in a data set?

Correlation Coefficient:

a NUMBER that determines if the correlation is positive or negative, as well as how closely the equation models the data

Finding the Correlation Coefficient, *r*, on the TI-84

1. Turn on Diagnostics
   - This only has to be done once (or if you reset the calculator)
   - [2nd] and then [0] to enter the catalog.
   - Scroll to "diagnosticsOn" and press Enter until the screen says "Done"
2. Enter data into the calculator selecting 1:Edit.
3. Go to [STAT] and then then CALC menu up top. Finally, select 4:LinReg and press Enter.

```
LinReg
 y=ax+b
 a=1.690909091
 b=.2727272727
 r²=.9701626472
 r=.9849683483
```

Correlation Coefficient p.64

Essential Question

How do I find the strength and direction of a correlation in a data set?

Write the equation for each set of data. Determine the correlation coefficient.

5.

x	y
5	-7
10	-10
15	-13

LinReg
y=ax+b
a=-.6
b=-4
r²=1
r=-1

$y = -.6x - 4$

c.c. = $r = -1$

neg. corr.

strong

6. {(-3, 4),(0, 10),(3,16)}

Correlation Coefficient p.64

Essential Question How do I find the strength and direction of a correlation in a data set?

7.

Average Maximum Daily Temperature in January for Northern Latitudes							
Latitude (°N)	35	33	30	25	43	40	39
Temperature (°F)	46	52	67	76	32	37	44

X

Y

A. Write the equation of the trend line. $y = -2.49x + 137.62$

B. Calculate the correlation coefficient and describe the measure of association between this data.

$r = -.98$ Negative, strong association

C. The average maximum daily temperature in January for a particular city is 20°F. Estimate the latitude of the city.

$y = 20$

about 47° N

LinReg
y=ax+b
a=-2.487179487
b=137.6227106
r²=.9575476499
r=-.9785436372

Correlation Coefficient p.64

Essential Question

How do I find the strength and direction of a correlation in a data set?

8. Seven adults were surveyed about their education and earnings. The table shows the survey results.

Years of Education	14	18	12	16	14	14	16
Earnings Last Year (Thousand $)	63	73	42	68	48	52	84

Calculate, using your calculator, an equation that best fits this data. Which of the following accurately describes the correlation and causation for the data set?

A. Weak correlation; no causation
B. Weak correlation; likely causation
C. Strong correlation; no causation
D. Strong correlation; likely causation

Algebra I Unit 5 – Correlation Coefficient, *r*

Student Practice – Correlation Coefficient, *r*

Name _____ Date_____ Period _____

Go to this link to discover more about correlation coefficients. **http://tinyurl.com/qjrqsvs**

Read and answer the following questions. Round answers to the nearest hundredth.

1. Below is a table showing the number of gold medals won by the United States at the Winter Olympics during various years.

Year	1992	1994	1998	2002	2006	2010
Gold Medals	5	6	6	10	9	9

a. Use a graphing calculator to write an equation for the best-fit line for that data._____

b. Calculate the correlation coefficient. *r*=_____. Describe the strength of its measure.

2. Below is a table showing the U.S. Federal Reserve's prime interest rate on January 1st of various years.

Year	2006	2007	2008	2009	2010
Prime Rate (percent)	7.25	8.25	7.25	3.25	3.25

a. Use a graphing calculator to write an equation for the best-fit line for that data._____

c. Calculate the correlation coefficient. *r*=_____. Describe the strength of its measure.

3. The table shows the price of a gallon of regular gasoline at a station in Los Angeles, California on January 1 of various years.

Year	2005	2006	2007	2008	2009	2010
Average Price	$1.47	$1.82	$2.15	$2.49	$2.83	$3.04

a. Use a graphing calculator to write an equation for the best-fit line for that data._____

b. Calculate the correlation coefficient. *r*=_____. Describe the strength of its measure.

c. What is the estimated price of gas in 2015?

Algebra I Unit 5 – Correlation Coefficient, *r*

4. A math club decided to buy T-shirts for its members. A clothing company quoted the following prices for the T-shirts. Which equations best describes the relationship between the total cost, *c*, and the number of T-shirts, *s*?

Math Club T-Shirts

Number of T-Shirts	Total Cost (dollars)
10	75
15	105
20	135

A $c = 6.75s$
B $c = 7.00s$
C $c = 2s - 20$
D $c = 15 + 6s$

5. The manager of a band has kept track of the price of tickets and the attendance at the band's recent concerts.

Concert Attendance by Ticket Price									
Price ($)	6	5	8.50	8	10	5.50	7	7.50	8
Attendance	213	256	155	194	160	267	258	210	235

a. Write the equation of the line of best fit that represents this data.

b. Predict the attendance at a concert where the price of tickets is $9.

6. A photographer hiked through the Grand Canyon. Each day she filled a photo memory card with images. When she returned from the trip, she deleted some photos, saving only the best. The table shows the number of photos she kept from all those taken for each memory card.

Grand Canyon Photos	
Photos Taken	Photos Kept
117	25
128	31
140	39
157	52
110	21
188	45
170	42

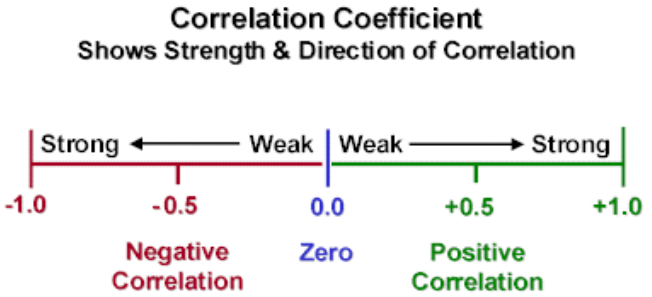
a. Write the equation of the line of best fit.

b. Predict the number of photos this photographer will keep if she takes 200 photos.

HW Help: Correlation Coefficient

Remember: You need to have DiagnosticsOn in your calculator to find the correlation coefficient, r!

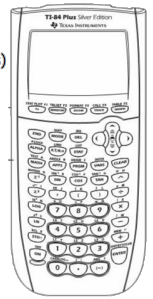
1. a. y = 0.25x - 493.95
   b. r = 0.84. What does this number say about the data?
2. a. y = -1.3x + 2616.25
   b. r = -0.85. What does this number say about the data?
3. You try!
4. You try! Your "y" is "c" and "x" is "s".
5. a. y = -20.95x + 368.89
   b. Use your model to find y when x =9
6. a. y = 0.33x - 11.33
   b. Use your model to find y when x =200.



Student Notes – Linear Regression

1. Press STAT, 1:Edit
2. X values (independent variables) go in L_1 and Y values (dependent variables) go in L_2.
3. Press STAT, go RIGHT to CALC, and select 4: LinReg (ax+b)
4. Make sure your screen looks like below. To store your equation in Y1, press ALPHA TRACE and select Y1.

```
LinReg(ax+b)
Xlist:L1
Ylist:L2
FreqList:
Store RegEQ:Y1
Calculate
```



5. Now Calculate! Make sure you replace a and b with the numbers the calculator gave you.