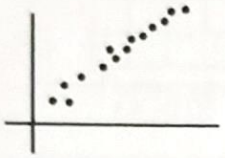
Name:_____ Date:_____

Review: Statistics and Probability

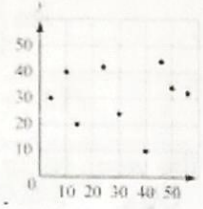
INTERPRETING CORRELATION FROM GRAPHS

For each of the following decide if it would be:

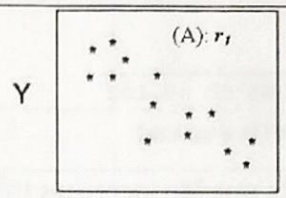
Strong Positive, Weak Positive, Strong Negative, Weak Negative or No Correlation



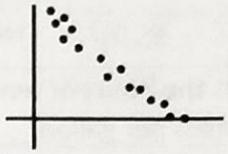
Strong positive



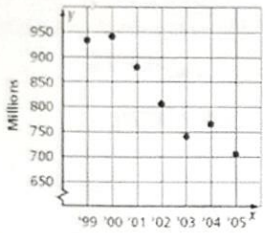
no correlation



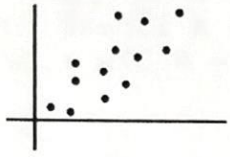
weak negative



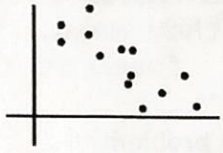
strong negative



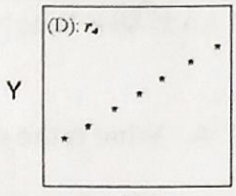
strong negative



weak positive



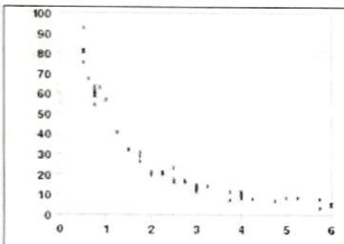
weak negative



perfect positive

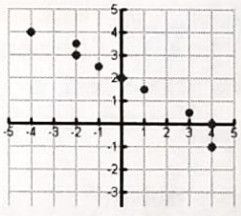
For each of the following decide if the graph would be:

Linear, Exponential or No correlation

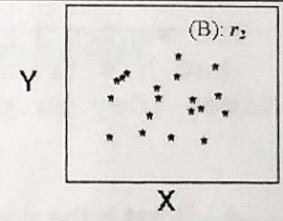


Exponential

exponential

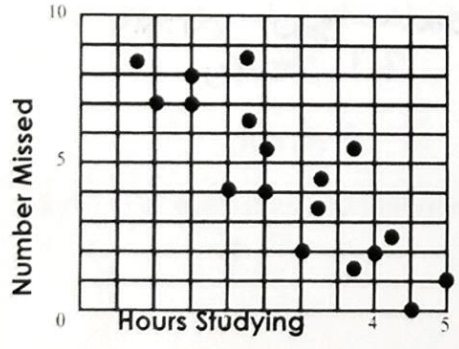


linear

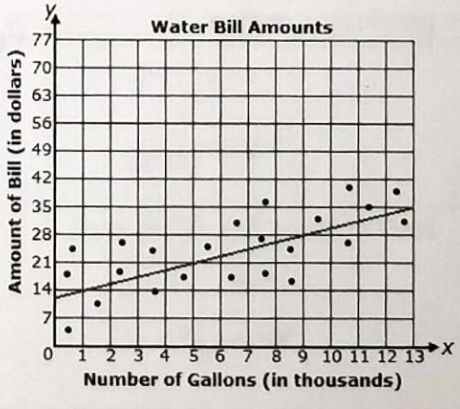


none

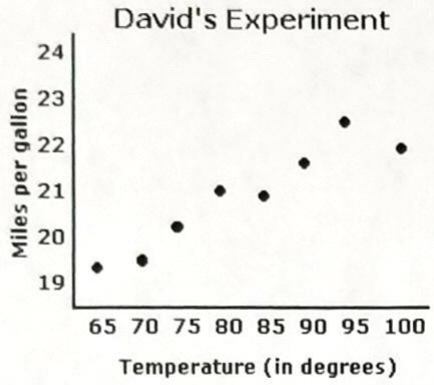
For each of the following fill in the blanks so the sentence makes sense



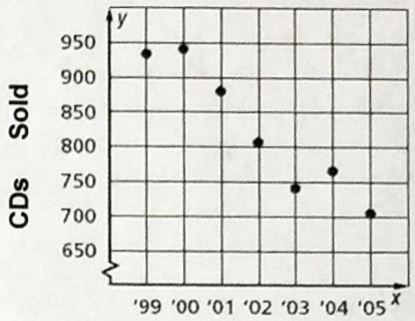
As the hours of studying increases, The number of problems missed on the test goes down



As the number of gallons of water used increases the Amount of bill increases



As the hours of temperature outside increases the Miles per gallon increases.



As the years go up, the CDs sold decreases

Harley interviewed her friends to find out how many hours they work a week and how much money they make. She did a linear regression equation and found it to be y = 8.32x + 9.

A. What is the slope in this problem?

8.32

B. Using the slope, fill in the blanks of the following:

For every hour the students work they would get an extra $8.32

Gus wanted to see the connection between drinking soda and burping. He had his friends drink different amount of sodas and measured how many times they burped. He created a line of best fit equation for the data and found it to be y = 3.1x + 0.45

A. What is the slope in this problem?

3.1

B. What does the slope tell you in this problem? (You should answer it similar to how you answered B in the previous question

For every soda you drink you burp 3.1 times

INTERPRETING CORRELATION COEFFICIENTS

For each of the following decide if it would be:

Strong Positive, Weak Positive, Strong Negative, Weak Negative or No Correlation

r = 0.86 SP	r = -0.40 None	r = 0.62 none	r = 0.17 none	r = -0.71 weak negative
r = -0.05 none	r = -0.75 weak negative	r = -1 perfect strong negative	r = 0.9 strong positive	r = 0.29 none
r = 0.89 strong positive	r = -0.75 weak	r = 0.09	r = -0.27	r = 0.55
r = 0.89	r = -0.75	r = 0.09	r = -0.27	r = 0.55

For each of the following decide if it would be:

Positive, Negative or No Correlation

Hours of Sleep and Energy ↑ ↑ positive	Age and video games played ↑ ↓ negative
Orange juice drunk and length of flu ↑ ↓ negative	Number of shoes owned and number of sodas drunk in a day ↑ ? none
Days until the end of school and happiness ↑ ↑ positive	Temperature and sunscreen used ↑ ↑ positive
Number of eyelashes and text messages sent ↑ ? none	Days in ISS and grade in class ↑ ↓ negative

DETERMINING LINEAR REGRESSION AND CORRELATION COEFFICIENT

1. Find the correlation coefficient and the equation of the line of best fit for the following data set.

Year	1984	1988	1992	1996	2000
Gallons Available	150	125	104	76	50

Correlation Coefficient: $-.999$ Line of best fit equation: $y = -6.23x + 12501.2$

2. Determine the correlation coefficient and the linear regression model for the following data set.

Drop Height (cm)	Bounce Height (cm)
100	26
90	23
80	21
70	18
60	16

Correlation Coefficient: $r = .997$ Linear Regression Model: $y = .25x + .9$

3. What is the equation of the best fitting line and the correlation coefficient for the following data?

Team	1	2	3	4
Wins	52	49	47	39

Correlation Coefficient: $r = -.952$ Best fitting line: $y = -4.1x + 57$

4. What is the correlation coefficient and the best fitting linear line for the following ordered pairs?

(1, 860), (2, 930) , (3, 1000), (4, 1150), (5, 1200), (6, 1360)

Correlation Coeffecient: $r = .989$ Best fitting line: $y = 98.86x + 737.37$

5. What is the linear regression model and the correlation coefficient of the following points?

(0, 10), (1, 7)

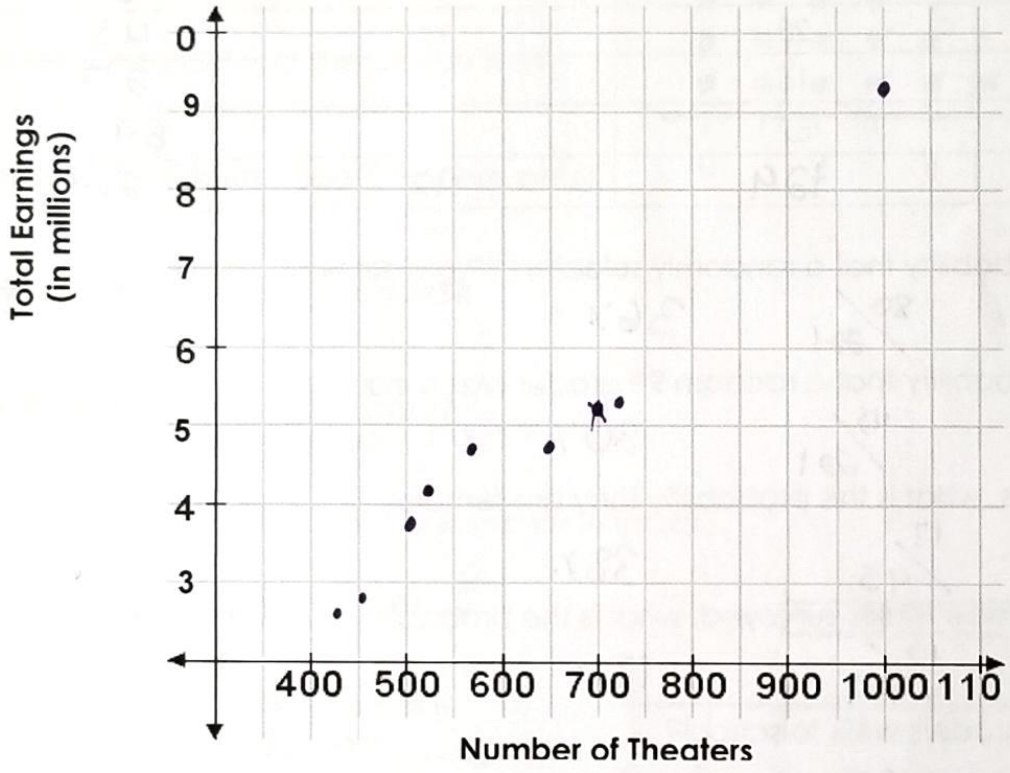
Correlation Coeffecient: $r = -1$ Best fitting line: $y = -3x + 10$

PUTTING ALL THE PIECES TOGETHER

1. The table below shows the number of movie theaters showing Skyfall and the film's weekly total earnings (in millions).

a) Construct a scatter plot for this set of data:

Number of Theaters (x)	443	455	503	530	560	657	723	1064
Total Earnings (y) (in millions)	2.57	2.65	3.73	4.05	4.76	4.76	5.15	9.35



b) What do you predict the correlation coefficient will be? (Strong/weak/none, positive/negative) Why do you predict this?

Strong positive. It is increasing and close together

c) What is the correlation coefficient?

r = .980

d) Is is a correlation or causation and why?

CAUSATION. The more theaters that show infinity war, the more $ they will make

e) Write the line of best fit model for this data.

y = 0.01x + -1.68

f) What is the projected earnings when 2000 theaters show the film?

0.01(2000) + -1.68 $18.32 million

TWO-WAY FREQUENCY

3. Below is a table the shows the results of a survey about how the freshman class at Sprayberry High School gets to school. Complete the table and use the information to answer the questions.

	Male	Female	Total
Walk	34	46	80
Car	28	17	45
Bus	15	12	27
Bike	52	17	69
Total	129	92	221

a. What is the probability that a randomly selected 9th grader walks to school?

80/221 36%

b. What is the probability that a random 9th grader was a car rider?

45/221 20%

c. Of the car riders, what is the probability they are female?

17/45 38%

d. Given the 9th grade males surveyed, what is the probability they ride the bus?

15/129 12%

e. How many 9th graders walk to school?

80

f. How many total students were surveyed?

221

g. What is the probability of walking to school?

80/221 36%

h. Given the freshmen boys, what is the probability they walk to school?

34/129 26%

i. What is the probability that a randomly selected student was a female?

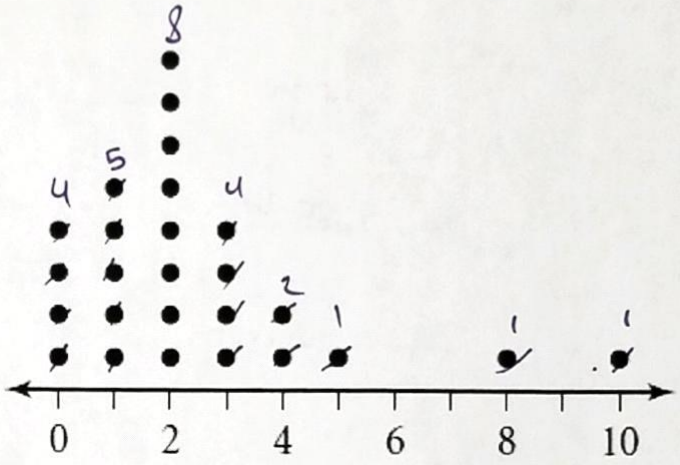
92/221 42%

j. If a random student is selected what is the percent that they are a biker?

69/221 31%

To the right is a dot plot represent the number of pets in Dr. Oldham's 3rd period class

17. How many students were surveyed?



18. How many students have 7 pets?

26

19. How many students have 0 pets?

4

20. What is the center of the distribution?

8

21. How many students had 5 or more pets?

3

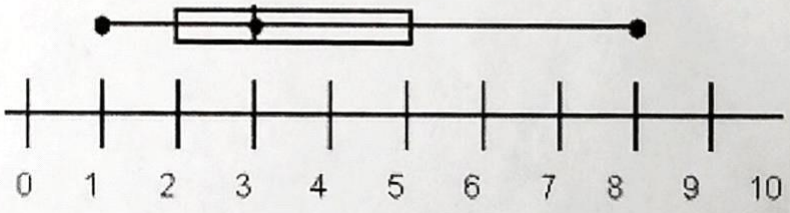
22. What is the spread of the data?

10

23. What is the shape of the data?

skew right

1. Use the box and whisker plot to answer the following



a. 25% of the data is less than 2

b. The middle 50% of the data is Between 2 and 6

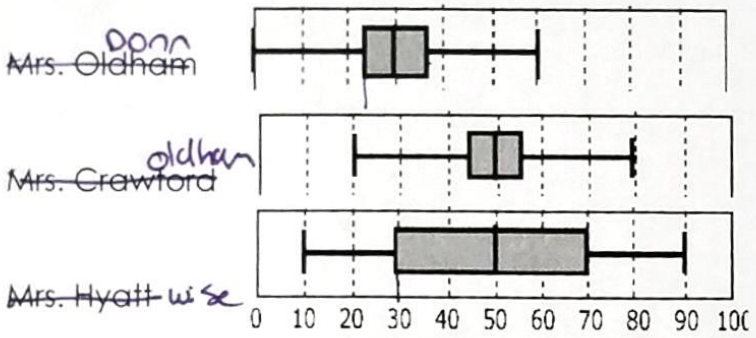
c. 25% of the data is greater than 5

d. All of the data is lower than 8.

2. Define the following in your own words

Mean:	Mode:	Range:	Interquartile Range:
Average	most frequent	HIGHEST MINUS LOWEST	Q_3 MINUS Q1
Median:	**Q1:**	**Q3:**	**Minimum:**
middle number	middle of lower half of data	middle of upper half of data	LOWEST VALUE

Below are 3 box and whisker plots for grades that students received in 3 different classes.



3. Which class had the highest Q3?

wise

4. Which class had the lowest Q1?

Donn

5. Which class had the lowest grade?

Donn

6. Which class has the highest grade?

wise

7. What class had lowest IQR?

oldham

8. What class had highest IQR

wise