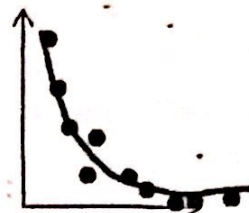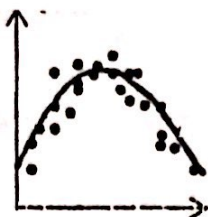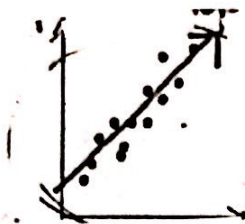## Line of Best Fit Notes

- When data is displayed on a scatter plot, it is often useful to represent the data with an equation of a line (Linear Regression) or a curve (Quadratic or Exponential Regression) for the purpose of

- A _____ is a straight line that best represents the data on a scatter plot. This is a line that goes through the _____ of all the points on a scatter plot with about half of the points _____ the line and half of the points _____ the line. (The line doesn't necessarily touch any of the points on the scatter plot).

- The formula for the line of best fit is similar to the equation $v = mx+b$, however your calculator gives you an a and b so we use the equation $y = $ _____
    - a= the slope: In context of the problem this will be how much the units for y increase (positive slope) or decrease (negative slope) for every one unit of x.
    - b= the y intercept: In context of the problem this will be how much you start with or if you had zero of your x variable, the amount of your y variable.
    - r= correlation coefficient: **Remember this tells you how strong or weak the correlation is between the variables [Also, tells you the direction]

Example 1: Is there a relationship between fat grams and the total calories for food?

Let's decide if we can predict the total number of calories based on the total grams of fat for a food item.

4x   y L₂

1. Use the following table to find the correlation coefficient and the line of best fit for the data.

Sandwich	Total Fat (g)	Total Calories
Hamburger	9	260
Cheeseburger	11	320
Quarter Pounder	21	420
Quarter Pounder with Cheese	30	530
Big Mac	31	560
Arch Sandwich Special	31	550
Arch Special with Bacon	10	320
Crispy Chicken	25	500
Fish Fillet	28	540
Grilled Chicken	20	440
Grilled Chicken Light	5	300

$r =$

$a =$

$b = .$

$y =$

2. Describe the strength and direction of the correlation of fat and total calories using your correlation coefficient.

3. What does the slope represent in context of the problem? $\frac{RISE}{RUN} = \frac{Y}{X} = \frac{calories}{Fat}$

4. What does the y-intercept represent in context of the problem?

5. Use your equation to determine how many calories are in a sandwich that contain <u>22 grams of</u> <u>fat,</u>

plug in for X

6. Using your equation, how many calories are in a sandwich that contain 50 grams of fat?

7. Using your equation, how many grams of fat do you expect to be in a sandwich with 500 calories?

Example #2: Can we predict the number of Incorrect answers you will give based on your hours of studying?

A student asks 15 classmates how many hours they spent studying for a test and how many incorrect answers they gave on that test.

X

y

# of Hours	1	5	6	9	10	8	5	4	10	4	10	7	9	2	8
Incorrect Answers	8	5	3	1	0	3	6	6	2	8	1	4	2	7	2

1. Use the table above and find the <u>correlation coefficient</u> and the <u>linear regression model</u> that best fits the <u>data above</u>

2. Describe the <u>strength and direction</u> of the correlation using your correlation coefficient.

3. What does the slope represent in context of the problem?

4. What does the y intercept represent in context of the problem?

5. If a classmate were to study for <u>3 hours</u>, use your equation to predict how many incorrect answers do you expect them to give?

6. If a classmate gave <u>9 incorrect answers,</u> use your equation to determine how many hours that person might have studied.