S.ID.B.6a: Linear, Quadratic, and Exponential Regression

GRAPHS AND STATISTICS

S.ID.B.6a: Linear, Quadratic and Exponential Regression

B. Summarize, represent, and interpret data on two categorical and quantitative variables

6. Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

a. Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models. Includes the use of the regression capabilities of the calculator.

Overview of Lesson

- activate prior knowledge and review learning objectives (see above)
- explain vocabulary and/or big ideas associated with the lesson
- connect assessment practices with curriculum
- model an assessment problem and solution strategy
- facilitate guided discussion of student activity
- facilitate guided practice of student activity
Selected problem set(s)
- facilitate a summary and share out of student work
Homework – Write the Math Assignment

Vocabulary

Regression Model: A function (e.g., linear, exponential, power, logarithmic) that fits a set of paired data. T*he model may enable other values of the dependent variable to be predicted.*

Big Ideas

The **individual data points** in a **scatterplot** form **data clouds** with shapes that suggest relationships between dependent and independent variables.

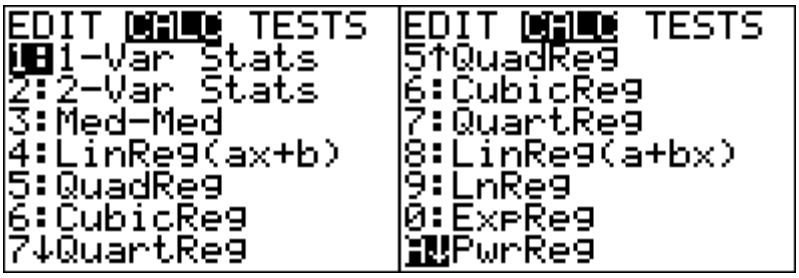
A **line of best fit** divides the data cloud into two equal parts with about the same number of data points on each side of the line. A line of best fit can be a straight line or a curved line, depending on the shape of the data cloud.

Calculating Regression Equations. Technology is almost always used to calcualte regression equations. .

- **STEP 1**. Use STATS EDIT to Input the data into a graphing calculator.
- **STEP 2**. Use 2nd STAT PLOT to turn on a data set, then ZOOM 9 to inspect the graph of the data and determine which regression strategy will best fit the data.
- **STEP 3**. Use STAT CALC and the appropriate regression type to obtain the regression equation.
- **STEP 4**. Ask the question, "Does it Make Sense (DIMS)?"

DIFFERENT TYPES OF REGRESSION

The graphing calculator can calculate numerous types of regression equations, but it must be told which type to calculate. All of the calculator procedures described above can be used with various types of regression. The following screenshots show some of the many regressions that can be calculated on the TI-83/84 family of graphing calculators.



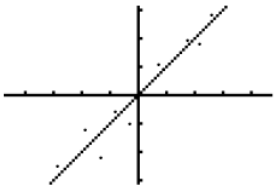
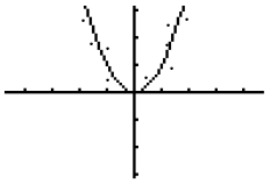
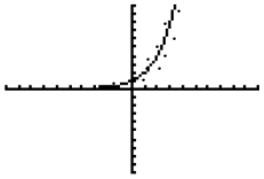
The general purpose of linear regression is to make predictions based on a line of best fit.

Choosing the Correct Type of Regression to Calculate

There are two general approaches to determining the type of regression to calculate:

- The decision of which type of regression to calculate can be made based on visual examination of the data cloud, or.
- On Regents examinations, the wording of the problem often specifies a particular type of regression to be used.

Using the Data Cloud to Select the Correct Regression Calculation Program

$y = x$ 	$y = x^2$ 	$y = 2^x$ 
If the data cloud takes the general form of a straight line, use **linear regression**.	If the data cloud takes the general form of a parabola, use **quadratic regression**.	If the data cloud takes the general form of an exponential curve, use **exponential regression**.
Note: The general forms of some data clouds are difficult to interpret. In difficult to interpret cases, the strength of the correlation coefficient can be used to determine which type of regression best fits the data. See lesson for standard S.ID.C.8,		

Drawing a Line of Best Fit on a Scatterplot

A line of best fit may be drawn on a scatterplot of data by using values from the regression equation.
STEP 1. Input the regression equation in the y-editor of a graphing calculator
STEP2. Use ordered pairs of coordinates from the table of values to plot the line of best fit.

- In linear regression, the line of best fit will always go through the point $(\bar{x}, \bar{y})$, where $\bar{x}$ is the mean of all values of *x*, and $\bar{y}$ is the mean of all values of *y*. For example, the line of best fit for a scatterplot with points (2,5), (4,7) and (8,11) must include the point $\left(x = \frac{14}{3}, y = \frac{23}{3}\right)$, because these *x* and *y* values are the averages of all the x-values and all the y-values.
- If the regression equation is linear and in $y = mx + b$ form, the y-intercept and slope can be used to plot the line of best fit.

Making Predictions Based on a Line of Best Fit

Predictions may be made based on a line of best fit.

STEP 1. Input the regression equation in the y-editor of a graphing calculator

STEP2. Use ordered pairs of coordinates from the table of values to identify expected values of the dependent (y) variable for any desired value of the independent (x) variable.

REGENTS PROBLEMS TYPICAL OF THIS STANDARD

1. An application developer released a new app to be downloaded. The table below gives the number of downloads for the first four weeks after the launch of the app.

Number of Weeks	1	2	3	4
Number of Downloads	120	180	270	405

Write an exponential equation that models these data. Use this model to predict how many downloads the developer would expect in the 26th week if this trend continues. Round your answer to the nearest download. Would it be reasonable to use this model to predict the number of downloads past one year? Explain your reasoning.

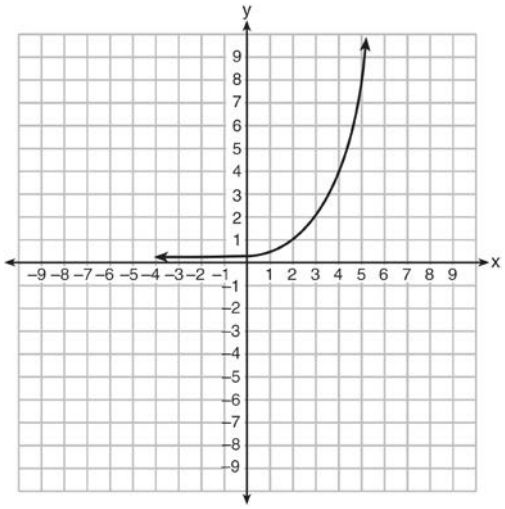
2. The table below shows the number of grams of carbohydrates, *x*, and the number of Calories, *y*, of six different foods.

Carbohydrates (*x*)	**Calories** (*y*)
8	120
9.5	138
10	147
6	88
7	108
4	62

Which equation best represents the line of best fit for this set of data?

a. $y = 15x$
b. $y = 0.07x$
c. $y = 0.1x - 0.4$
d. $y = 14.1x + 5.8$

3. Write an exponential equation for the graph shown below.



Explain how you determined the equation.

4.The table below shows the attendance at a museum in select years from 2007 to 2013.

Attendance at Museum

Year	2007	2008	2009	2011	2013
Attendance (millions)	8.3	8.5	8.5	8.8	9.3

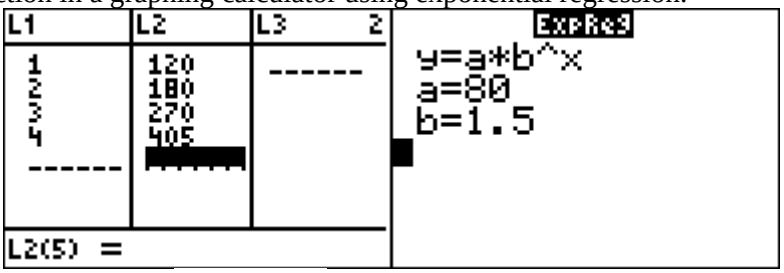
State the linear regression equation represented by the data table when $x = 0$ is used to represent the year 2007 and *y* is used to represent the attendance. Round all values to the *nearest hundredth*. State the correlation coefficient to the *nearest hundredth* and determine whether the data suggest a strong or weak association.

S.ID.B.6a: Linear, Quadratic, and Exponential Regression
Answer Section

1. ANS:
   a) $y = 80(1.5)^x$
   b) $80(1.5)^{26} \approx 3,030,140.$
   c) No, because the prediction at $x = 52$ is already too large.

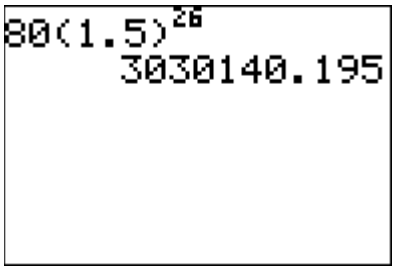
   Strategy: Use data from the table and exponential regression in a graphing calculator.

   STEP 1: Model the function in a graphing calculator using exponential regression.



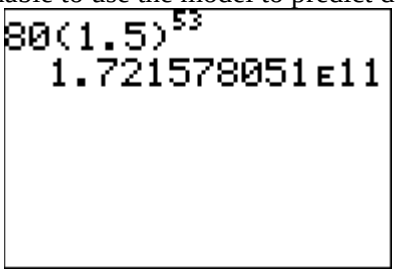
   The exponential regression equation is $y = 80(1.5)^x$

   STEP 2. Use the equation to predict the number of downloads when $x = 26$.



   Rounded to the nearest download, the answer is 3,030,140.

   STEP 3. Determine if it would be reasonable to use the model to predict downloads past one year.



   It would not be reasonable to use this model to make predictions past one year. The number of predicted downloads is more 170 billion downloads, which is more than 20 downloads in one week for every person in the world.

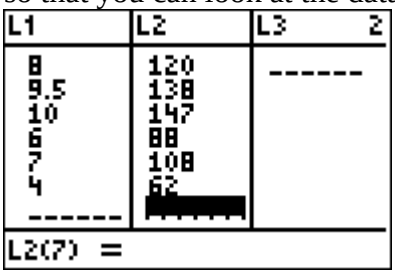
   DIMS? Does It Make Sense? For near term predicitons, yes. For long term predictions, no.

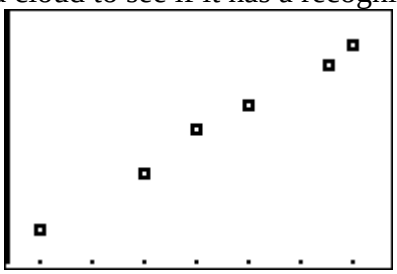
   PTS: 4 REF: 061536ai NAT: S.ID.B.6a TOP: Regression
   NOT: NYSED classifies this as A.CED.A.2

2. ANS: D

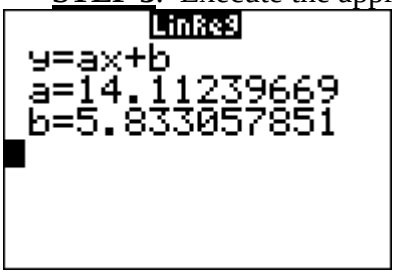
Strategy: Input the data into a graphing calculator, inspect the data cloud, and find a regression equation to model the data table, input the regression equation into the y-editor, predict the missing value.

- **STEP 1**. Input the data into a graphing calculator or plot the data cloud on a graph, if necessary, so that you can look at the data cloud to see if it has a recognizable shape.





- **STEP 2**. Determine which regression strategy will best fit the data. The graph looks like the graph of an linear function, so choose linear regression.
- **STEP 3**. Execute the appropriate regression strategy in the graphing calculator.



Write the regression equation in a format that can be compared to the answer choices: $y = 14.11x + 5.83$

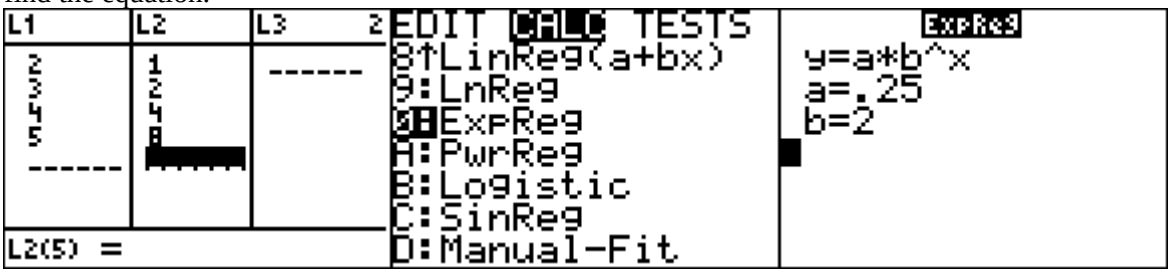
- **STEP 4**. Compare the answer choices to the regression equation and select choice d.

PTS: 2 REF: 081421ai NAT: S.ID.B.6a TOP: Regression

3. ANS:

$y = 0.25(2)^x$.

Strategy: Input the four integral values from the graph into a graphing calculator and use exponental regression to find the equation.



Alternative Strategy: Use the standard form of an exponential equation, which is $y = ab^x$.

Substitute the integral pairs of (2,1) and (3,2) from the graph into the standard form of an exponential equation and obtain the following: $1 = ab^2$ and $2 = ab^3$.

Therefore, $2ab^2 = ab^3$

$$2 = \frac{ab^3}{ab^2}$$

$$2 = b$$

Accordingly, the equation for the graph can now be written as $y = a \cdot 2^x$.
Substitute the integral pair (4,4) from the graph into the new equation and solve for *a*, as follows:

$$y = a \cdot 2^x$$

$$4 = a \cdot 2^4$$

$$4 = a \cdot 16$$

$$\frac{4}{16} = a$$

$$\frac{1}{4} = a$$

The graph of the equation can now be written as $y = \frac{1}{4}(2)^x$

PTS: 2 REF: 011532ai NAT: F.LE.A.2 TOP: Modeling Exponential Equations

4. ANS:
$y = 0.16x + 8.27$ $r = 0.97$, which suggests a strong association.

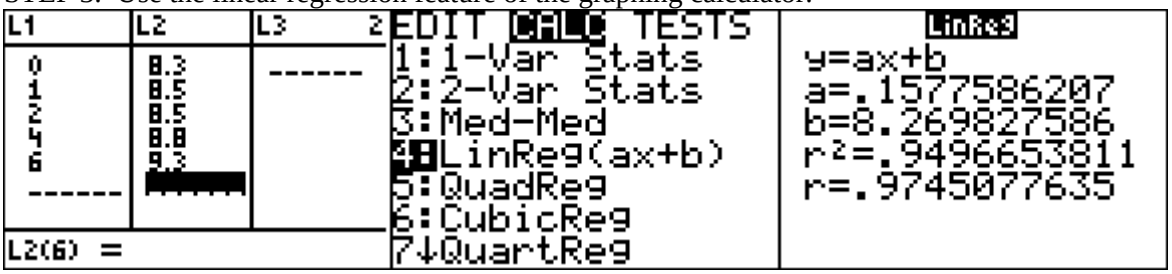
Strategy: Convert the table to data that can be input into a graphing calculator, then use the linear regression feature of the graphing calculator to respond to the question.

STEP 1. Convert the table for input into the calculator.

Attendance at Museum					
Year (L1)	0	1	2	4	6
Attendance (L2)	8.3	8.5	8.5	8.8	9.3

STEP 2. Make sure that STAT DIAGNOSTICS is set to "On" in the mode feature of the graphing calculator. Setting STAT DIAGNOSTICS to on causes the correlation coefficient (*r*) to appear with the linear regression output.

STEP 3. Use the linear regression feature of the graphing calculator.



NOTE: Round the graphing calculator output to the *nearest hundredth* as required in the problem.

STEP 4. Record your solution.

PTS: 4 REF: 081536ai NAT: S.ID.C.8 TOP: Regression
KEY: linear NOT: NYSED classifies as S.ID.B.6a

Homework - Write the Math Assignment

START Write your name, date, topic of lesson, and class on your paper.

NAME: Mohammed Chen
DATE: December 18, 2015
LESSON: Missing Number in the Average
CLASS: Z

PART 1a. Copy **the problem** from the lesson and underline/highlight key words.
PART 1b. State your understanding of **what the problem is asking**.
PART 1c. **Answer** the problem.
PART 1d. Explanation of **strategy** with all work shown.

PART 2a. Create **a new problem** that addresses the same math idea.
PART 2b. State your understanding of **what the new problem is asking**.
PART 2c. **Answer** the new problem.
PART 2d. Explanation of **strategy** used in solving the new problem with all work shown.

Clearly label each of the eight parts.

Grading Rubric

Each homework writing assignment is graded using a four point rubric, as follows:

Part 1. The Original Problem	**Up to 2** points will be awarded for: a) correctly restating the original problem; b) explicitly stating what the original problem is asking; c) answering the original problem correctly; and d) explaining the math.
Part 2. My New Problem	**Up to 2** points will be awarded for: a) creating a new problem similar to the original problem; b) explicitly stating what the new problem is asking; c) answering the new problem correctly; and d) explaining the math.

This assignment/activity is designed to incorporate elements of Polya's four step universal algorithm for problem solving with the idea that writing is thinking. Polya's four steps for solving any problem are:

1. Read and understand the problem.
2. Develop a strategy for solving the problem.
3. Execute the strategy.
4. Check the answer for reasonableness.

EXEMPLAR OF A WRITING THE MATH ASSIGNMENT

Part 1a. The Problem

TOP Electronics is a small business with five employees. The mean (average) weekly salary for the five employees is $360. If the weekly salaries of four of the employees are $340, $340, $345, and $425, what is the salary of the fifth employee?

Part 1b. What is the problem asking?

Find the salary of the fifth employee.

Part 1c. Answer

The salary of the fifth employee is $350 per week.

Part 1d. Explanation of Strategy

The arithmetic mean or average can be represented algebraically as:

$$\overline{X} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

I put information from the problem into the formula. The problem says there are 5 employees, so $n = 5$. The problem also gives the mean (average) salary and the salaries of 4 of the employees. These numbers can be substituted into the formula as follows:

$$360 = \frac{340 + 340 + 345 + 425 + x_5}{5}$$

$$1800 = 340 + 340 + 345 + 425 + x_5$$

$$1800 = 1450 + x_5$$

$$1800 - 1450 = x_5$$

$$350 = x_5$$

Check: $360 = \frac{340 + 340 + 345 + 425 + 350}{5} = \frac{1800}{5} = 360$

Part 2a. A New Problem

Joseph took five math exams this grading period and his average score on all of the exams is 88. He remembers that he received test scores of 78, 87, 94, and 96 on four of the examinations, but he has lost one examination and cannot remember what he scored on it. What was Joseph's score on the missing exam?

Part 2b What is the new problem asking?

Find Joseph's score on the missing exam.

Part 2c. Answer to New Problem

Joseph received a score of 85 on the missing examination.

Part 2d. Explanation of Strategy

I substitute information from the problem into the formula for the arithmetic mean, as follows:

$$88 = \frac{78 + 87 + 94 + 96 + x_5}{5}$$

$$440 = 355 + x_5$$

$$85 = x_5$$

$$88 = \frac{78 + 87 + 94 + 96 + 85}{5} = \frac{440}{5} = 88$$

The answer makes sense.