

MATHEMATICS

Algebra I

A Fossil Puzzle

Table of Contents

Rationale for Lesson and Associated Tasks	Page 1
Lesson and Associated Tasks Overview	Pages 1 - 2
West Virginia College-and Career-Readiness Standards	Page 2
Mathematical Habits of Mind (MHM)	Page 2
Mathematics Teaching Practices to Support Student Growth	Pages 2-3
Essential Understandings	Page 3
Set-up Phase	Page 3
Establish Small Groups	Page 3
Develop Open-Ended Questions	Pages 3-4
Gather Materials	Page 4
Anticipated Common Student Misconceptions	Page 4
Explore Phase	Pages 4-5
Prior Instruction/Knowledge	Pages 5-6
Implementation Phase	Pages 6-12
Share, Discuss and Analyze Phase	Pages 12-13



Task Title: A Fossil Puzzle

Grade or Content Area: Grade 9

Toolkit Author: Natalie Dillinger, Diane Munza, and Adam Riazi

Original Task Creator: Illustrative Mathematics

Quarter: 2

Rationale for Lesson and Associated Tasks

Students in grade 9 are asked to predict the height of an ancient human based on the find of a fossilized humerus bone. Based on the length of the humerus bone, students are tasked with using statistics to predict the height of the ancient human. Students collect data by measuring the length of the humerus and height for each student in the class. Data will be graphed on a scatter plot and a linear model generated to best fit the data. From the model, students make a prediction for the height of the human and determine the strength of their prediction.

This task invites students to collect, summarize, interpret, and draw conclusions from bivariate data using scatter plots, best fit lines, residuals, and correlation coefficients. According to NCTM*, statistical association between two variables is a fundamental statistical idea. Students are introduced to statistical association through the line of best fit as an extension of linear equations. As students determine the line of best fit, they also determine the closeness of the line to data points. For this task, the focus is determining a linear model for the data and the strength of fit for the model.

**Finding What Fits* by Stephanie A. Casey, NCTM.org/Publications, April 2016, Vol 21, Issue 8.

Lesson and Associated Tasks Overview

Task: Illustrative Mathematics Algebra I Unit 3 Two-Variable Statistics Lesson 10 *A Fossil Puzzle*

Preparation: ([click here](#))

Lesson: ([click here](#))

Practice: ([click here](#))

Student Lesson: ([click here](#))

For grade 9 students, this lesson and associated tasks may serve as a culmination for summarizing, representing, and interpreting data on two categorical and quantitative variables. Students create a linear model for the collected data. By using residuals and the correlation coefficient, students analyze the fit of their model. Students predict an outcome using the linear model. Suggestions for developing content for this lesson and associated tasks can be found in the explore section of the task guide. This is intended as a three- day lesson. The following provides an outline.

Day 1

- Introduce *A Fossil Puzzle*. Provide students with a picture of a human arm showing humerus bone to launch a discussion of humerus length and height.
- Collect data for humerus length and height of students
- Generate scatter plot graph for collected data
- Analyze the shape of scatter plot

Day 2

- Determine the line of best fit for the data using string
- Determine a linear function to fit data using graphing technology
- Analyze the linear function and predict the height of the ancient human.

Day 3:

- Calculate and graph residual values
- Interpret the fit of your linear function based on the residual graph
- Determine and analyze the correlation coefficient
- Compare linear model and height predictions found by other groups

West Virginia College- and Career-Readiness State Standards

M.A1HS.37 or M.1HS.35

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 given functions or choose a function suggested by the context. Emphasize linear and exponential models.
- b. Informally assess the fit of a function by plotting and analyzing residual Instructional Note: Focus should be on situations for which linear models are appropriate.
- c. Fit a linear function for scatter plots that suggest a linear association.

Instructional Note: Students take a more sophisticated look at using a linear function to model the relationship between two numerical variables. In addition to fitting a line to data, students assess how well the model fits by analyzing residuals.

M.A1HS.38 or M.1HS.36

Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data. Instructional Note: Build on students' work with linear relationships in eighth grade and introduce the correlation coefficient. The focus here is on the computation and interpretation of the correlation coefficient as a measure of how well the data fit the relationship.

M.A1HS.39 or M.1HS.37

Compute (using technology) and interpret the correlation coefficient of a linear fit. Instructional Note: Build on students' work with linear relationships in eighth grade and introduce the correlation coefficient. The focus here is on the computation and interpretation of the correlation coefficient as a measure of how well the data fit the relationship.

Mathematical Habits of Mind (MHM)

MHM1. Make sense of problems and persevere in solving them.

MHM2. Reason abstractly and quantitatively.

MHM4. **Model** with mathematics.

MHM5. Use **appropriate tools** strategically.

*While several MHMs are listed, the words in bold font are the focus of the lesson and associated tasks.

Mathematics Teaching Practices to Support Student Growth

- **Implement tasks that promote reasoning and problem solving.**
- **Use and connect mathematical representations.**
- Pose purposeful questions. Build procedural fluency from conceptual understanding.
- Support productive struggle in learning mathematics.

*While several Mathematics Teaching Practices are listed, those in bold font are the focus of the lesson and associated tasks.

Essential Understandings

- Data can be collected for two quantitative variables (bivariate) and represented on a scatter plot to be analyzed using both student-generated graphs and graphing technology.
- A function **can be created** to model the behavior of the bivariate variables displayed on the scatterplot and used to **interpret** the data and **solve problems**.
- The fit of the linear model can be assessed by **plotting and analyzing the residuals** and determining the correlation coefficient using graphing technology.

Set-up Phase

1. Become an Expert Regarding All Lesson and Associated Task Content

It is essential that you, the teacher, become familiar with all “*A Fossil Puzzle*” materials. This lesson comes with many associated tasks, the use of each may vary depending on the time frame available for the lesson and student levels of engagement and understanding. Reviewing, analyzing, and completing all tasks prior to the implementation with students is imperative for classroom success.

2. Establish Small Groups

“*A Fossil Puzzle*” tasks promote both individual and small group thinking. While working in small groups, students collect data by measuring the length of the humerus and height for each student in the small group which will be shared with the whole class. Students will continue to work in small groups to create a scatter plot and find a linear function that best models the data for the entire class. Students contribute to meaningful small group conversations as they solve problems using their linear model. Students explore the meaning of residuals and correlation coefficients to determine the fit of their linear model to the collected data.

Skill levels, leadership skills, and personalities are all considered when creating small groups. Small group collaboration works best when students have been provided previous opportunities to work together on a regular basis. Weeks of teacher observation of student behavior will be extremely helpful when creating small groups for this lesson and associated tasks. Taking notes regarding leadership skills, personalities, the ability to take criticism, to question, and to think deeply about a task or problem will enable teachers to make sound decisions regarding small group placement.

When creating the groups, all these factors help to eliminate the potential situation in which one student takes the lead and makes the decisions for the group. In this scenario, one student is gaining all the benefits of the task, while others do not. If a student is not engaged in conversation, this lesson and associated tasks will not be beneficial in helping all students to have meaningful discussions about the mathematics involved nor in analyzing the relationships inherent to the tasks. Students will be working both individually and together when placed in small groups to complete this lesson and associated tasks.

3. Develop Open-Ended Questions

- What are the variables given in the situation?
- Is there a way to collect information from your classmates to answer the question?
- How can you determine an appropriate scale for the graph?
- How does the method for measuring the humerus impact the approximation for the fossil?
- Can a model be generated that passes through all data points?
- How can you create a stronger model for the data?
- How do outlier points effect the model?

- How can residuals and the correlation coefficient be used to determine the strength of a linear model?

4. Gather Materials

- Computer and presentation device (for the teacher)
- Internet access
- Graphing calculators (or graphing software)
- Desmos App downloaded on teacher and student computers
- *A Fossil Puzzle* Handout to Record Data (1 copy provided for each student) Attached document
- Rulers or tape measure in centimeters
- Graph Paper
- String
- Chart paper
- Markers

5. Anticipated Common Student Misconceptions

- Student measurements for humerus may not be consistent among students.
- Students may not accurately label axis units.
- Students may calculate slope incorrectly when the x-axis and y-axis have different scales.
- Students may not relate the linear function to the regression equation.
- Students may interpret a negative residual as an error instead of comparing residual to zero.
- Students may misinterpret a pattern in residual values as the model for the function.
- Students may misunderstand how to interpret a negative correlation coefficient.

Explore Phase

Before the teaching of this lesson and associated activities, the following content should have been covered in previous lessons. Reviews may be necessary therefore some suggested activities/ lessons are included.

1. Create a scatter plot to represent bivariate data.
2. Determine if data is linear based on the shape of the graph.
3. Determine and interpret the slope and y-intercept for a linear function.
4. Find an equation in point-slope form using a point and slope.
5. Simplify an equation in point-slope form to y-intercept form.
6. Use Desmos or a graphing calculator to generate lists (or tables), graphing lines, and finding regression equations. Desmos is a free app that can be downloaded by both teachers and students. The following is a link for the Desmos app: www.desmos.com. Directions for using Desmos can be found later in this section and in the teacher notes.
Students can also use graphing calculators to generate lists, StatPlots, linear models, and regression equations.
7. Create a scatter plot and determine a model with linear data using graphing technology. Illustrative Mathematics provides a lesson, *Linear Models*, for which students create and analyze scatterplots. In section 4.2 of the lesson, *Orange you Glad We're Boxing Fruit*, students collect data while watching the video *Oranges in a Box*. Students create a scatterplot using Desmos technology found in the student lesson. Directions for using Desmos are found on the graph. Questions in the student lesson can be used to determine understanding of the linear model generated from the data. Below are some links for this lesson:

Student lesson: ([click here](#))

Lesson: ([click here](#))

Oranges in a Box video: ([click here](#))

Direct link to Desmos application for the problem: ([click here](#))

Helpful resources from the Kendall Hunt website:

Resource 1 – ([click here](#))

Resource 2 – ([click here](#))

At the bottom of the page under lesson preparation are links to print formatted materials.

8. Analyze the fit of a linear model for a scatterplot using residuals.

Illustrative Math provides the lesson *Residuals* that can be used to develop or review the concept. In lesson 6.2, *Oranges Return*, questions 1 – 9 can be used to determine student understanding of residuals. The following link may be helpful for developing and reviewing residuals. ([click here](#))

It is important for students to understand the concept of determining and graphing residuals. The steps found in the attached document, ***Using Desmos***, may be helpful for using the Desmos app for questions 1,2, and 7. If students are using a graphing calculator, a table can be generated after the best-fit equation is determined.

9. Find a regression equation using graphing technology

Steps for finding a regression line using Desmos can be found in the attached handout, ***Using Desmos***.

10. Use technology to determine and analyze the Correlation Coefficient.

Illustrative Mathematics lesson *Matching Correlation Coefficients* provides 8 graphs to match with correlation coefficients. ([click here](#))

In addition, the Illustrative Mathematics lesson *Using the Correlation Coefficient* provides information regarding the correlation coefficient, how to graph a regression equation using Desmos, and what constitutes a weak or strong correlation. The student lesson provides practice for students. ([click here](#))

The following links from Kendall Hunt may be helpful in providing additional materials and student responses.

Link One ([click here](#))

Link Three ([click here](#))

Link Two ([click here](#))

Link Four ([click here](#))

Prior Instruction/Knowledge:

In grade 8, students create scatter plots. In grade 9, students extend their knowledge to determining mathematical models that capture key elements of the relationship between two variables. In grade 8, students work with linear relationships that will be built upon in grade 9 followed by the introduction of the correlation coefficient to measure the data fit of the relationship.

Please review the following:

Educators Guide for Mathematics: Algebra 1 (pages 38 – 41, pdf pages 40-43) ([click here](#))

Educators Guide for Mathematics: Math 1 (pages 32 – 36, pdf pages 34-38) ([click here](#))

Prerequisite Skills

- Organize, display, and interpret information in scatter plots. Approximate a trend line and identify the relationship as positive, negative, or no correlation.
- Make predictions based on results from surveys and samples.
- Derive a linear equation that models a set of data (line of best fit) using calculators. Use the model to make predictions.
- Use and interpret function notation in number and word problems: determine a value of the function given an element in the domain.
- Identify outliers and clusters in bivariate data in tables and scatter plots.

Supporting Skills

- Given a list of ordered pairs in a table or graph, identify either verbally or algebraically the rule used to generate and record results.

Impending Skills

- Derive a linear equation that models a set of data (line of best fit) using calculators. Use the model to make predictions.
- Derive a quadratic or exponential function that models a set of data (curve of best fit) using calculators. Use the model to make predictions.

Source: *The Quantile Framework for Mathematics*

<https://metametricsinc.com/educators/quantile-for-educators/>

2020 MetaMetrics Inc.

Implementation Phase

Day 1

- Introduce *A Fossil Puzzle*. Provide students with a picture of a human arm showing humerus bone to launch a discussion of humerus length and height.
- Collect data for humerus length and height of students
- Generate scatter plot graph for collected data
- Analyze shape of scatter plot

Teacher Notes Day 1:

1. Introduce the activity to students using the following:

An anthropologist finds a fossilized humerus bone of an ancient human ancestor. The humerus is an arm bone running from the shoulder to the elbow. It is 24 centimeters in length. Working in small groups, discuss how could you determine the height of the ancient human?

Working in small groups, students discuss the following questions as they determine how to estimate the ancient human's height. The following are examples of questions to ask students as they are working:

- What are the variables for the situation?
Answer: the length of the humerus bone and the height of the human
- Is there a way to collect information from your classmates to answer the question?
Answer: Measure the length of the humerus bone and height of the students in the class
- How can the data collected be used to estimate the height of the ancient human?
Answer: Graph the data collection (scatter plot) and model the data with a function. What are some factors to consider when measuring students?

After 2 -3 minutes, use whole-group discussion regarding student ideas to determine student understanding for the task and given directions as to data collection.

2. Students collect data for humerus length and height for each student in their small group. Both measurements are in centimeters. Record measurements on handout *A Fossil Puzzle* (attached document Handout for a Fossil Puzzle). A student representative from each group records measurements on the whiteboard to generate data for the whole class.

3. After data collection, students discuss in small groups appropriate labels for the x and y-axis as well as the scale to use for each.
The following are examples of questions to ask students as they are working:
 - Which variable is the independent variable?
 - Which variable is the dependent variable?
 - Based on the independent and dependent variables, what are the labels for the x and y-axis?
 - Will all data points be visible with the chosen scale?
 - Is the scale for the x-axis the same as the y-axis?

4. Students individually create a scatter plot for the class data.
Students graph their data points on the graph provided, *A Fossil Puzzle*, using the scale determined by the small group.

5. Working in small groups, students analyze the shape of the graph by answering the following questions:
 - Does the data appear to be linear?
 - Is the slope positive or negative?
 - What is your best estimate for the slope?
 - If the scales for the x and y-axis are different, how does this impact your slope value?
 - In terms of the data, what does the slope represent?
 - What is your best estimate for the y-intercept?
 - In terms of the data, what does the y-intercept represent?

6. As a teacher-led whole-class discussion, answer the questions from #5.

Question: Does the data appear to be linear?
Answer: Yes

Question: Is the slope positive or negative?
Answer: Positive

Question: What is your best estimate for the slope?
Answer: Will vary

Question: If the scales for the x and y-axis are different, how does this impact your slope value?
Answer: The scale for each axis must be considered when discussing the change in value

Question: In terms of the data, what does the slope represent?
Answer: Change in height compared with the change in humerus length

Question: What is your best estimate for the y-intercept?
Answer: Will vary

Question: In terms of the data, what does the y-intercept represent?
Answer: The height of the human when the humerus bone is 0 cm.

Day 2

- Determine the line of best fit for the data using string
- Create a scatter plot using graphing technology (graphing calculator, Desmos, or other graphing applications)

- Determine a linear function to fit data using graphing technology
- Analyze the linear function and predict the height of the ancient human

Teacher Notes Day 2:

1. Working in small groups, students discuss how to use the string to determine the best fit line for their data points:

- Does the string need to pass through all data points?
- If the string cannot pass through all data points, how do you decide where to place the string?
- What would it mean if your string is below most of the data points? Or above most of the data points?

After a 2 – 3-minute small group discussion, answer the above questions in a whole group discussion.

Question: Does the string need to pass through all data points?

Answer: Answers will vary, but to create a model, it does not need to pass through every point. Discuss what it means if the string would pass through every point and if a point is completely separate from most of the points.

Question: If the string cannot pass through all data points, how do you decide where to place the string?

Answer: Answers will vary, but the string should be “in the middle” of the data points.

Question: What would it mean if your string is below most of the data points? Or above most of the data points?

Answer: String above- data points less than the estimate. String below – data points more than the estimate. Later in the lesson, this effects the residual graph.

2. Working in small groups, students individually find the line of best fit for their data points. Students place string on the scatter plot to represent a line that best fits the data. Continue to move the string until it moves through the middle of the data points. Students in small groups determine if individual placements are the “best” placement and make changes as needed. Draw a line where the string is the best fit.

The following represent questions to ask students as you circulate throughout the room:

- Can the string be moved to pass through more data points?
- Are there more data points either above or below the line?
- Are there outliers – points that do not fit any line?

3. Students determine a linear function to model the data using the best fit line.
 - Students choose 2 points on the best fit line and calculate the slope.
 - Students find the equation for the line in point-slope form.
 - Students determine the linear function for the line (slope-intercept form).
 - Students determine the y-intercept for the line

Students record work on the handout *A Fossil Puzzle*

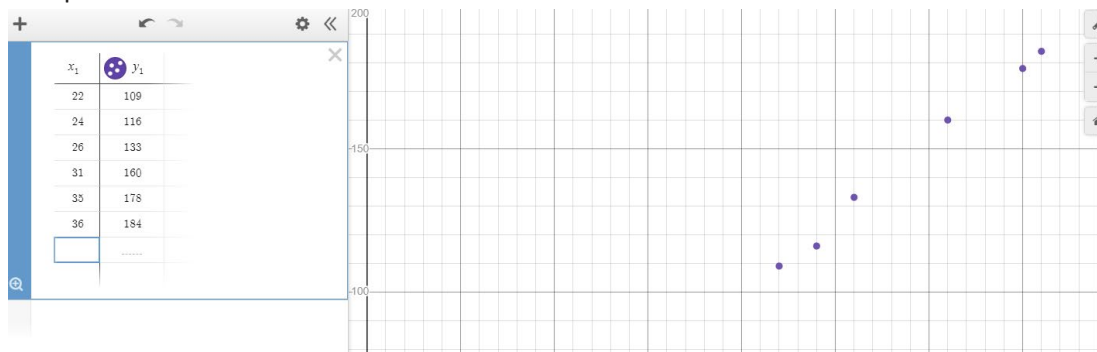
Students work individually to determine the equation but remain in small groups. Encourage students to collaborate as they find the linear model to fit their data.

As students individually determine the equation for their best fit line, below are some things to notice as you circulate throughout the room:

- Are students calculating the slope correctly?
- Does the sign for the slope fit the data points?
- Are students using point-slope form correctly?
- Do students have misconceptions when simplifying point-slope form to y-intercept form?
- Is the y-intercept in the equation the same (or close to) the y-intercept on the graph?

4. Students put data in Desmos application (www.desmos.com/calculator) or in the list menu on a graphing calculator.

Example:



5. Students put their equation in Desmos or a graphing calculator.

Working in small groups, students determine if their equation is the best fit using graphing technology.

- Does the line pass through or between all or most of the data points?
- If not, how can the slope or y-intercept be adjusted for a better fit?
- If most points are not close to the line, would other points create a better equation?

Example:



6. Analyze the linear function (best-fit line) which models the data. Determine and analyze the slope and y-intercept for a linear function
7. Determine the height of the fossil using the linear function for everyone in the small group. Based on all answers in the group, predict the height for your small group.
8. As a whole group discussion, discuss the height predictions for each small group.

Day 3

- Calculate and graph residual values
- Interpret the fit of your linear function based on the residual graph
- Using graphing technology, find the regression equation for your data
- Determine and analyze the correlation coefficient
- Compare linear model and height predictions found by other groups

Day 3 Teacher Notes

1. Review the definition of a residual: Predicted value (linear model) subtracted from the actual value (data point). Working in small groups, students discuss the following:
 - If the fit of the linear model is good, describe the residual values
Answer: the residual values will be close to zero
 - Can the residual values be positive or negative?
Answer: yes
 - What does a residual value of zero mean?
Answer: the predicted value is the same as the actual value
 - If the residual values form a pattern, what does that mean?
Answer: another model will provide a better fit
2. Individually students determine the residual value for each data point on their graph. Students using graphing calculators use the table menu to find the predicted value for each data point. Students using Desmos follow instructions in the explore section for residuals. Record residual values on the handout *A Fossil Puzzle*. In small groups, students share results of their individual graphs.
3. In small groups graph the residual values on the scatterplot graph (using technology). If colors are available, use a different color. Students using a graphing calculator can put values in a new list and create a new StatPlot. Desmos users can create a table in a new line.

Example:



Graphing
line after
choosing
wheel



x	$5.03x + 0.2$
-2	-9.86
-1	-4.83
0	0.2
1	5.23
2	10.26
22	110.86
24	120.92
26	130.98
31	156.13
35	176.25
36	181.28

You will need to type in x values following number 2.



Residual points are green

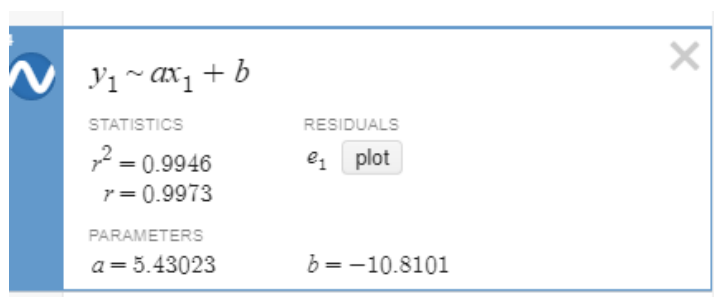
4. Working in small groups, students assess the fit of data points to the linear model by analyzing the residuals.

Students answer the following:

- Based on the graph of the residuals, was your linear model a good fit? Explain your answer based on the following:
 - Closeness of points to the x-axis.
 - Residual values form a pattern or not
 - Based on the residuals, is measuring student humerus and height a strong tool for estimating the height of an ancient human? Record your answer on the *A Fossil Puzzle* handout.

5. Working in small groups, students find the regression equation for the class data. See notes in the explore section for directions to create regression equation graph.

Example:



It is important for students to grasp creating a linear model and analyzing fit based on residuals before using the graphing calculator to find the regression equation.

6. The correlation coefficient, r , quantifies the strength of a linear relationship. The values for r : $-1 \leq r \leq 1$. Based on the values, discuss in small groups then as a whole group, the following:
- What values for r would suggest a strong fit (data to model)?
Answer: $|r| \geq 0.8$
 - What values for r would suggest a weak fit (data to model)?
Answer: $|r| \leq 0.5$
 - What does the sign of r signify?
Answer: the slope
7. Determine the r -value from the regression equation. Based on this value, describe the strength of the linear model. Explain your reasoning. Record your answer on the handout *A Fossil Puzzle*.
8. Working in small groups using the regression equation, determine the fit of the graph based on the correlation coefficient.
- Review with students how to locate the correlation coefficient with the regression equation.
 - Review the values for r that refer to a weak and a strong relationship. Based on the r -value, describe the strength of the linear model.
 - Based on your r -value, is student measurement a good model for determining the height of an ancient human?
9. Small group presentation of findings:
In small groups present your findings to the class:
- Computer or calculator: Graph with data, linear model, residuals
 - Chart with Linear Model function, regression equation r value, and predicted height of the ancient human
 - Determination of strength of fit based on residuals and r value.
- Each small group share findings for the height of the ancient human and determination if the model based on student information can be used as a prediction for the height of an ancient human.

Share, Discuss, and Analyze Phase

Essential Question #1:

Data can be collected for two quantitative variables (bivariate) and represented on a scatter plot to be analyzed using both student-generated graphs and graphing technology.

Share – Students are given the following task:

An anthropologist finds a fossilized humerus bone of an ancient human ancestor. The humerus is an arm bone running from the shoulder to the elbow. It is 24 centimeters in length. Working in small groups, discuss how you could determine the height of the ancient human?

Discuss – Following the data collection, students explore the following questions before generating a scatter plot:

- Which variable is the independent variable?
- Which variable is the dependent variable?
- Based on the independent and dependent variables, what are the labels for the x and y-axis?
- Will all data points be visible with the chosen scale?
- Is the scale for the x-axis the same as the y-axis?

Analyze – Students individually create a scatter plot based on the collected data and answering questions in the share section.

Essential Question #2

*A function **can be created** to model the behavior of the bivariate variables displayed on the scatterplot and used to **interpret the data and solve problems**.*

Share - Working in small groups, students discuss how to use the string to determine the best fit line for their data points.

Discuss - Working in small groups, students individually draw the line of best fit for their data points. Then students find the equation that models the line of best fit. Students put both the data and the equation in their graphing application (Desmos, graphing calculator, other graphing software).

Analyze – Students working in small groups determine the following based on their linear model

- Determine and analyze the slope and y-intercept for a linear function
- Determine the height of the fossil using the linear function for everyone in the small group.

Essential Question #3

*The fit of the linear model can be assessed by **plotting and analyzing the residuals** and determining the correlation coefficient using graphing technology.*

Share Review the definition of a residual: Predicted value (linear model) subtracted from the actual value (data point).

Discuss - The correlation coefficient, r , quantifies the strength of a linear relationship. The values for r : $-1 \leq r \leq 1$. Working in small groups, students determine the r -value from the regression equation. After determining the r -value, describe the strength of the linear model.

Analyze – Working in small groups, students prepare short presentations for the class regarding their findings. Students share the height of the ancient human based on their model. Students share their small group determination of the strength of fit for their linear model based on both the residual graph and the r -value. Finally, the group determines if measuring student lengths can be used as a prediction of the height of an ancient human.