

Scatter It! (Predict Billy's Height)

Susan Haller
St. Cloud State University
skhaller@stcloudstate.edu



Published: March 2012

Overview of Lesson

In this lesson, students explore the relationship between age and height in order to help a hypothetical student predict his height in two years. Students will examine data that will enable them to create a scatterplot and approximate a line of best fit. The scatterplot and line of best fit will be used to predict height. The slope of the line of best fit will be interpreted in context.

GAISE Components

This investigation follows the four components of statistical problem solving put forth in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report*. The four components are: formulate a question, design and implement a plan to collect data, analyze the data by measures and graphs, and interpret the results in the context of the original question. This is a GAISE Level A activity.

Common Core State Standards for Mathematical Practice

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.

Common Core State Standards Grade Level Content (Grade 8)

8. SP. 1. Construct and interpret scatterplots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.
8. SP. 2. Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association, informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line.
8. SP. 3. Use the equation of the linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept.

NCTM Principles and Standards for School Mathematics

Data Analysis and Probability Standards for Grades 6-8

Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them:

- formulate questions, design studies, and collect data about a characteristic shared by two populations or different characteristics within one population;

- select, create, and use appropriate graphical representations of data, including histograms, box plots, and scatterplots.

Select and use appropriate statistical methods to analyze data:

- use observations about differences between two or more samples to make conjectures about the populations from which the samples were taken;
- discuss and understand the correspondence between data sets and their graphical representations, especially histograms, stem-and-leaf plots, box plots, and scatterplots.

Develop and evaluate inferences and predictions that are based on data:

- make conjectures about possible relationships between two characteristics of a sample on the basis of scatterplots of the data and approximate lines of fit;
- use conjectures to formulate new questions and plan new studies to answer them.

Prerequisites

Students will have knowledge of and experience with graphing points on a coordinate system. Students will have knowledge of and experience with constructing and evaluating properties of lines.

Students will have experience using technology to create graphs of data.

Learning Targets

Students will correctly construct a scatterplot, both manually and with the use of technology. Students will analyze the scatterplot visually, looking for trends and patterns. In addition, students will approximate a line of best fit manually, and then create one using technology.

Time Required

Part of one class period

Materials Required

- Graphing calculators or computers with software capable of creating scatterplots.
- Spaghetti noodles.
- Graph paper.

Instructional Lesson Plan

The GAISE Statistical Problem-Solving Procedure

I. Formulate a Question

Begin the lesson by formulating a hypothetical situation:

Billy is an eighth-grade student who loves to play basketball. Because of this, he wants to know how tall he will be when he is in tenth-grade – the first year he will be eligible to play on his high school basketball team. Billy was looking through his school pictures and noticed that he seemed to grow in height about the same amount each year and wants to know if he can use his height from each year to predict how tall he will be in two years.

Ask students to write some questions that they would be interested in investigating about students’ age and height. Some possible questions might include:

1. What are some typical heights of students?

2. Is there a relationship between age and height? If so, what kind of relationship?
3. Is the relationship between age and height the same for male and female students? If so, how are they similar? If not, how are they different?
4. Can we use the relationship between age and height to predict the height of a student in the future? If so, for how many years is this relationship valid?
5. What kind of graph can we use to illustrate the relationship between age and height? How can the graph help us determine someone's future height?

II. Design and Implement a Plan to Collect the Data

Explain to the class that Billy's parents measured each of their children's heights on the first day of school every year. The table below shows Billy's age (in years) and height (rounded to the nearest quarter inch) at the start of every school year since kindergarten.

Table 1. Billy's age and height.

Person	Grade	Age	Height (in inches)
Billy	K	5	42.00
Billy	1	6	44.50
Billy	2	7	46.25
Billy	3	8	50.00
Billy	4	9	53.50
Billy	5	10	56.25
Billy	6	11	59.25
Billy	7	12	63.25
Billy	8	13	68.00

III. Analyze the Data

Explain to the class that they will be graphing Billy's data on graph paper. Ask whether age or height should be the independent variable, and discuss a reasonable scale for the graph. If needed, discuss the rationale for each of these. If students have not used the terminology for 'scatterplot', be sure to discuss this. The scatterplot for Billy's data is shown in Figure 1.

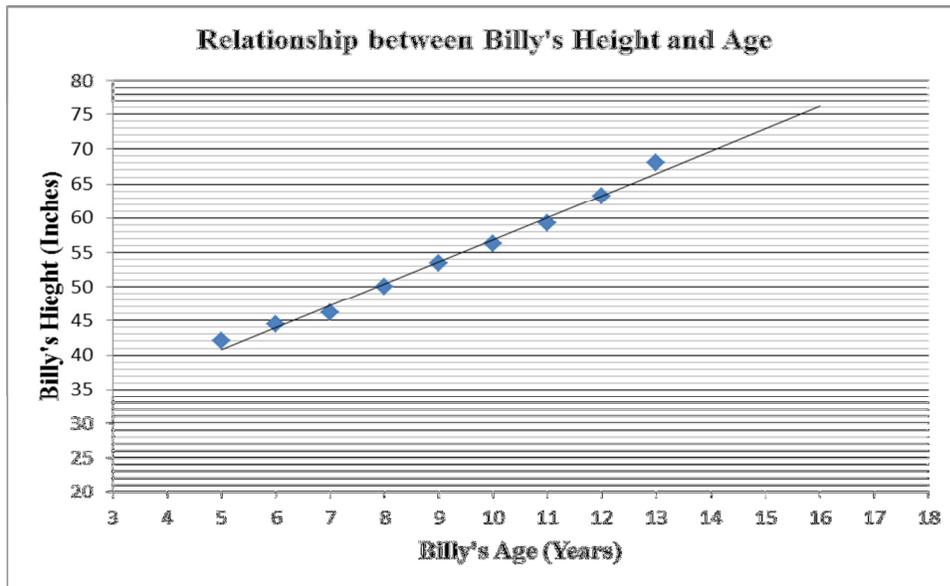


Figure 1. Scatterplot for Billy's age and height.

After the graph of Billy's data has been constructed, ask students to share anything they notice within the pattern of Billy's growth – does there seem to be a relationship between age and height? Does this relationship appear to be linear? Discuss the properties of a line of best fit, and have students place the spaghetti in their scatterplot to approximate the line of best fit for Billy's data. Next, have students locate Billy's age in grade 10 (he would be 15 at the start of that school year), and use the line of best fit to predict Billy's height. Students should compare their results with the rest of the class (predictions will vary because the lines will be slightly different).

Students should then enter the data into a graphing calculator or appropriate computer statistical software program and create a scatterplot of and line of best fit for Billy's data. Alternatively, the teacher can calculate the line of best fit. In this case, the line of best fit is $y = 3.21x + 24.8$.

IV. Interpret the Results

The students should next consider the line of best fit in context of the scenario presented. Students should recognize that the slope indicates that Billy's height increases 3.21 inches for every year of age. Students can then use the equation for the line of best fit and calculate that Billy should be approximately 73 inches tall when he is 15 years old.

Students should also assess the relationship between the graphed data and the line of best fit. For example, when Billy was 7 years old, his height was below the line of best fit, and it was above when he was 13 years old. This means that his height less than average when he was 7, but above average when he was 13. Students should consider this in context as well, because it may mean that his height might continue to be above average in his teenage years, thereby affecting the predicted height.

The teacher should engage the students in a discussion about the constraints for the line of best fit. For example, the equation would not be appropriate to use when Billy is an adult.

Assessment

1. The table below shows Jack's age in years and height, rounded to the nearest quarter-inch.

Age	4	5	6	7	8	9	10	11	12
Height	39.75	41.50	44.00	47.25	50.50	53.25	55.75	58.75	61.50

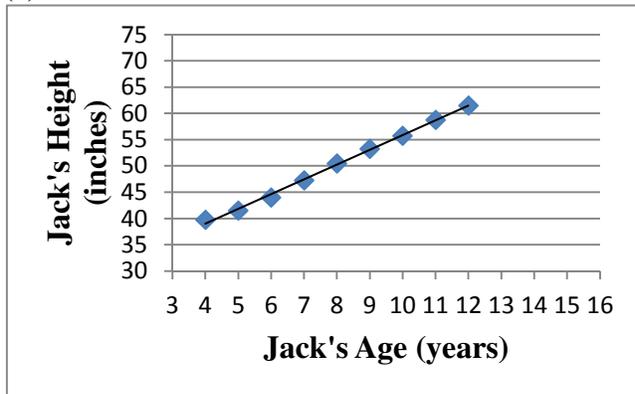
(a) Construct a graph of the data and describe the relationship between Jack's age and height.

(b) Find the equation for the line of best fit for the data.

(c) Predict Jack's height at age 16.

Answers

(a)



(b) $y = 2.8x + 27.8$.

(c) According to the line of best fit, Jack will be 72.6 inches tall at age 16.

Possible Extensions

Students can explore other relationships that may be based on age and/or height. For example, height and shoe size, or age and the number of text messages received each day.

References

This lesson was adapted from an investigation (Scatter It!) found at Census at School – New Zealand (<http://www.censusatschool.org.nz/classroom-activities/>).

Scatter It! Activity Sheet

Billy is an eighth-grade student who loves to play basketball. Because of this, he wants to know how tall he will be when he is in tenth-grade – the first year he will be eligible to play on his high school basketball team. Billy was looking through his school pictures and noticed that he seemed to grow in height about the same amount each year and wants to know if he can use his height from each year to predict how tall he will be in two years.

1. What are some questions you have about the relationship between a person's age and height?

2. The table below shows Billy's age (in years) and height (rounded to the nearest quarter inch) at the start of every school year since kindergarten.

Person	Grade	Age	Height (in inches)
Billy	K	5	42.00
Billy	1	6	44.50
Billy	2	7	46.25
Billy	3	8	50.00
Billy	4	9	53.50
Billy	5	10	56.25
Billy	6	11	59.25
Billy	7	12	63.25
Billy	8	13	68.00

Which variable, age or height, is the independent variable? Why?

3. Construct a graph showing the relationship between Billy's age and height. Be sure to use an appropriate scale.

4. Describe the relationship you notice about Billy's age and height.

5. Place a piece of spaghetti on your graph so that it passes through most of the points on your graph. The spaghetti represents an approximation for the line of best fit of the data. Estimate the slope of your line of best fit.

Use your line of best fit to determine Billy's height when he is 15 (his age in grade 10).

6. Enter the data into a spreadsheet or graphing calculator and calculate the line of best fit.

7. Explain what the slope of the line of best fit means in this scenario.

8. Use the line of best fit to predict Billy's height at age 15.

9. Would you feel confident using the line of best fit to predict Billy's height when he is 30? Why or why not?