

Least squares fit line

Introduction

Scatter diagrams depict data. Often it is desirable to model the data with a linear equation. One way to find a linear equation that models the data is to find the data's least squares fit line. Once the least squares fit line is known substituting values into the equation allows you to make inferences.

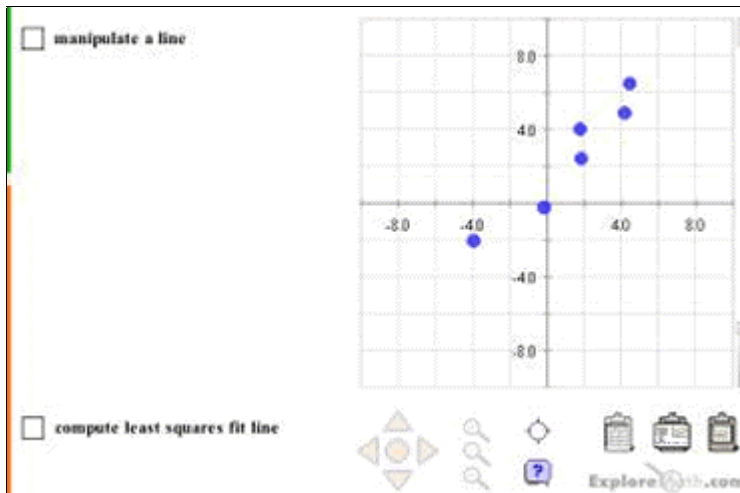
Let's explore least squares fit lines. Use your web browser to go to the "Least squares fit line" activity:

<http://www.explorelearning.com/index.cfm?method=cResource.dspView&ResourceID=68>

Below are several questions designed to get you thinking about the activity. Answer them on a separate piece of paper.

Activity

Drag the six blue points onto the plane similar to the graph below.



Question 1a. What do you observe about the general trend of the data points?

Question 1b. If you constructed a line on the plane that passed as close as possible to as many points as possible, would the line have a positive or negative

ExploreLearning

slope?

Question 1c. Would the line described in the last question have a positive or negative y-intercept?

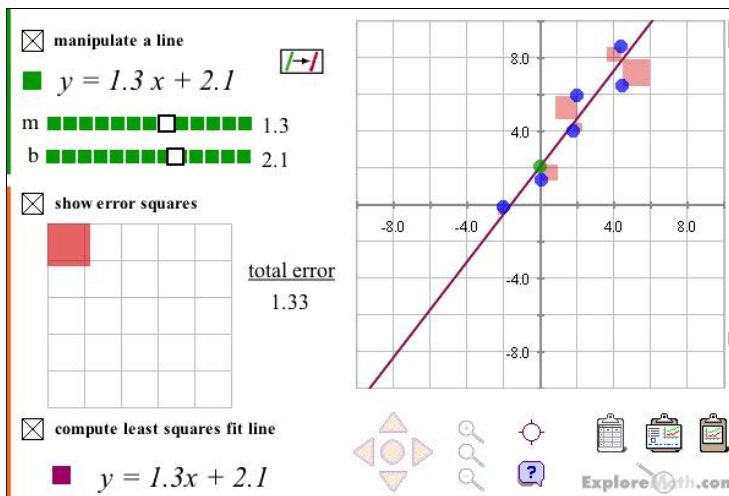
Select the “compute least squares fit line” to check the answers to the last two questions.

Question 1d. What would happen to the y-intercept if every data point were shifted upward 2 units?

Question 1e. What would happen to the slope if every data point was shifted upward 2 units?

Test your answers to the last two questions by moving every blue point upward two units.

Select the “manipulate a line” box, select the “green to red” box, and select the “show error squares” box.



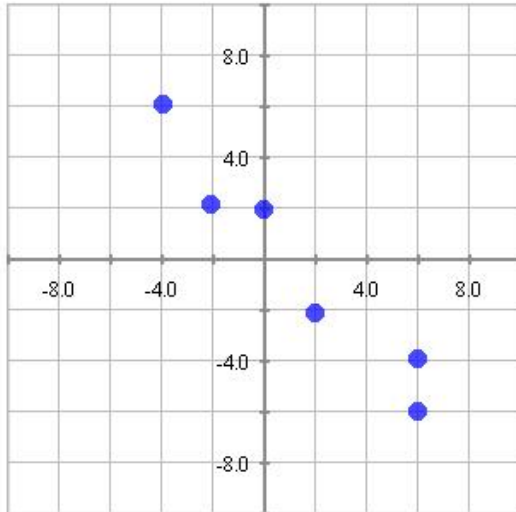
Question 1f. How do you think the total error is calculated?

Question 1g. How could the blue points be moved to reduce the total error?

Position the blue points close to the following coordinates:

(-4, 6) (-2, 2) (0, 2) (2, -2) (6, -4) (6, -6)

ExploreLearning



Question 2a. What is the least squares fit line for the set of data points?

Question 2b. Using the least squares fit line, what would be an estimate for y when $x = 10$?

Question 2c. What are some real world applications that could involve least squares fit lines?

Conclusion

The least squares fit line models the data that is basically linear in nature. The total error of squares is calculated by summing up the square of the vertical distances from each data point to the least squares fit line.