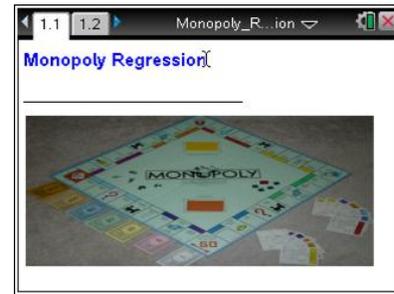## Math Objectives

- Students will analyze the linear association between two variables and interpret the association in the context of a given scenario.
- Students will calculate a least-squares regression line and interpret the coefficients.
- Students will determine the appropriateness of a least-squares regression plot by analyzing residuals and the effects of unusual points on the line.

Vocabulary

- coefficient of determination
- correlation coefficient
- explanatory variable
- influential point
- least-squares regression line
- outlier
- response variable
- scatterplot
- slope
- *y*-intercept

About the Lesson

- This lesson involves analyzing the association between the number of spaces from Go and the cost of the property on a standard Monopoly board.
- As a result, students will:
  - Calculate a least-squares regression line for data and interpret the coefficients.
  - Discuss unusual points and their effects on the line.
  - Interpret a residual plot.

TI-Nspire™ Technology Skills:
- Download a TI-Nspire document
- Open a document
- Move between pages
- Grab and drag a point

Tech Tips:
- Make sure the font size on your TI-Nspire handhelds is set to Medium.
- You can hide the function entry line by pressing [ctrl] [G].

Lesson Files:
Student Activity
Monopoly_Regression_Student. pdf
Monopoly_Regression_Student. doc

TI-Nspire document
Monopoly_Regression.tns

Visit www.mathnspired.com for lesson updates and tech tip videos.

Discussion Points and Possible Answers

The following data are the number of spaces from GO and the cost of the property for each property on a standard Monopoly board.
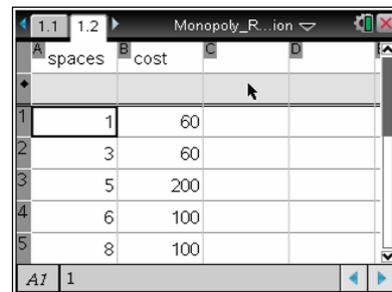
Property	Spaces from GO	Cost
Mediterranean Avenue	1	60
Baltic Avenue	3	60
Reading Railroad	5	200
Oriental Avenue	6	100
Vermont Avenue	8	100
Connecticut Avenue	9	120
St. Charles Place	11	140
Electric Company	12	150
States Avenue	13	140
Virginia Avenue	14	160
Penn Railroad	15	200
St. James Place	16	180
Tennessee Avenue	18	180
New York Avenue	19	200
Kentucky Avenue	21	220
Indiana Avenue	23	220
Illinois Avenue	24	240
B & O Railroad	25	200
Atlantic Avenue	26	260
Ventnor Avenue	27	260
Water Works	28	150
Marvin Gardens	29	280
Pacific Avenue	31	300
North Carolina Avenue	32	300
Pennsylvania Avenue	34	320
Short Line Railroad	35	200
Park Place	37	350
Boardwalk	39	400

Move to page 1.2.

Page 1.2 of the .tns file on the handheld shows a table of the data displayed above.

1. Use the scroll bar at the right to look through the data in the table.

    a. Do you notice any trends or any noteworthy data values?

    Sample Answers: Students might notice that in general as the spaces from GO increase so does the cost of the properties. Students might also notice that a few of the properties do not follow this trend.
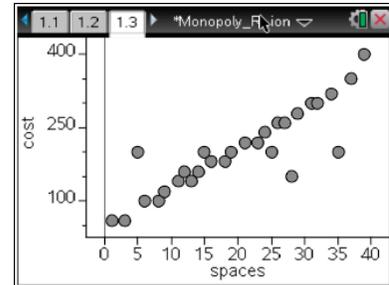
b. Which variable would make sense to be the explanatory variable? The response variable? Explain your reasoning.

Answer: The explanatory variable is the number of spaces from GO, and the response variable is the cost of the property because it would make sense to predict the cost by knowing the number of spaces.

2. Insert a new page (1.3) by selecting [doc▾] > **Insert > Data & Statistics**.

    - Create a scatterplot for the data by moving the cursor to the bottom center where it says **Click or Enter to add variable**.
    - Click on the text and select the variable **spaces**.
    - Move the cursor to the center of the left axis, and select **cost** as the variable0.

3. a. Describe the association between the two variables.

Sample Answers: There appears to be a positive linear association between the number of spaces from GO and the cost of a property. Most of the data values follow the same pattern, but four points do not.

b. Describe any unusual points in your scatterplot.

Sample Answers: Students should notice that several points that do not seem to follow the same pattern as the rest of the data. Four properties have the same *y*-value (cost) regardless of the number of spaces from GO. These points correspond to the four railroad properties. In addition, the properties 12 and 28 spaces from GO, Electric Company and Water Works, cost $150 and while the point for Electric Company is close to the pattern of the other properties, Water Works at (28, 150) does not follow the pattern of the rest of the data.

> **Teacher Tip:** A discussion on outliers and influential points might be beneficial here. Before students calculate a least-squares regression line, ask them to predict how the unusual points might affect the least-squares regression line. Discuss whether these points would be classified as influential points or outliers. You might refer to the Statistics Nspired Activities Influential Points and Influencing Regression as background or follow up activities.
>
> For beginning algebra students, you might want to skip question 3b below; and for questions 4-6, have them choose a moveable line instead of the

linear regression line. They can grab the line in the center to shift it vertically and at the ends to rotate the line to approximate the trend in the data. Note that this line will not reflect the "pull" of the points representing the railroads and utilities. If you choose to use a moveable line, you would most likely omit the parts of this activity discussing the strength of the linear relationship and the outliers.

4. Insert a new page (1.4) by selecting [doc▾] > **Insert > Lists & Spreadsheet**.
    - Move the cursor to the shaded area in the second row of Column A.
    - Calculate the least-squares regression line for the data by selecting **MENU > Statistics > Stat Calculations > Linear Regression (a+bx)**.
    - Press the right arrow for **X List:**, and choose **Spaces**.
    - Tab to **Y List:**, and choose **Cost**.
    - Press [enter] or select **OK**.

5. a. Write the equation that describes the relationship between the number of spaces from GO and the cost of the properties.

    Answer: Predicted Cost = 67.28 + 6.78 (spaces from GO).

    b. Interpret the slope in terms of the context.
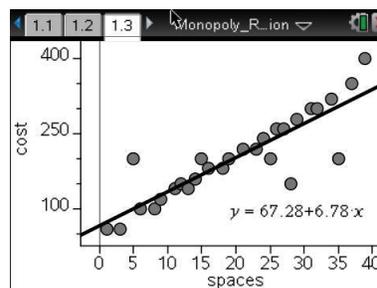
    Answer: On average, the cost of the property increases by $6.78 each space from GO.

    c. Interpret the *y*-intercept. Does it have a meaning in this context? Explain your reasoning.

    Answer: To start at GO costs $67.28. This could mean that it costs $67.28 to play the game.

Move back to page 1.3.

6. On the scatterplot of the data, plot the least-squares regression line on your scatterplot by selecting **MENU > Analyze > Regression > Show Linear (a+bx)**.

7. How well does the line fit the data?

Sample Answers: The line is drawn towards the outliers, so it does not capture the trend for the properties close to GO or the farthest from GO.

> **Teacher Tip:** Although the plot might appear linear, students still need to determine the strength and validity of their regression model by analyzing a residual plot and correlation.

8. Interpret the correlation coefficient.

Sample Answers: The correlation is approximately 0.88; there is a strong positive linear association between number of spaces from GO and the cost of the properties.
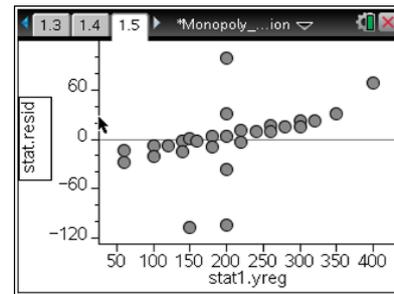
9. What is the meaning of the coefficient of determination, r^2?

Sample Answers: 77% of the variation in cost of the properties can be accounted for by the number of spaces from GO.

10. Insert a new page (1.5) by selecting [doc▾] > **Insert**
    > **Data and Statistics**.
    - Create a residual plot by moving the cursor to the lower part of the screen until you see **Click or Enter to add variable**.
    - Select the variable **stat.yreg** (these are the predicted cost values).
    - Move the cursor to the left of the screen until you see **Click or Enter to add variable**, and select **stat.resid**.

11. Describe the residual plot. Based only on the residual plot, would you consider your original data to be approximately linear? Explain why or why not.

Sample Answers: The residual plot appears fairly linear. The points that were unusual on the original scatterplot have much larger residuals than the rest of the data. Based on the residual plot alone, I would not consider the data to be approximately linear since there appears to be a predictable pattern in the plot: the line seems to over-predict prices for costs lower than about $200 and under predict prices for costs higher than $200.
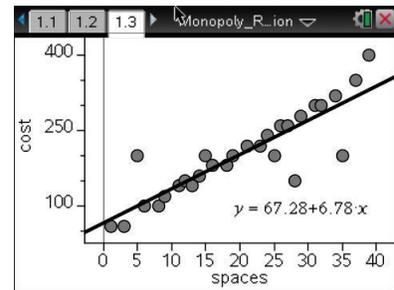
> **Teacher Tip:** Be sure students understand how to interpret positive and negative residuals. You might need to refer them back to the plot with the regression line on Page 1.3.

Move back to page 1.3.

12. In question 6, you displayed the least-squares regression line on the scatterplot. If you removed the outliers from the scatterplot, predict how the regression line would change.

     Sample Answers: The regression line will have a steeper slope and more closely fit the rest of the data. The data overall would appear more linear than previously.
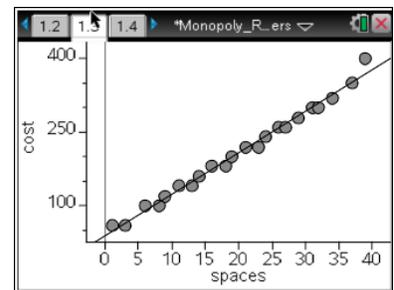
> **Teacher Tip:** One of the primary reasons for fitting a line to data is for prediction purposes. Have students consider whether you would need a model to predict the prices of the railroads and utilities.

Move back to page 1.2.

13. On Page 1.2, delete the following data points for the railroad and utility properties: (5, 200), (12, 150), (15, 200), (25, 200), (28, 150), and (35, 200).

14. Return to Page 1.3, and answer these questions.
    a. Was your prediction in question 12 correct? Explain your reasoning.

    Sample Answers: Yes, the line fits the data very closely.
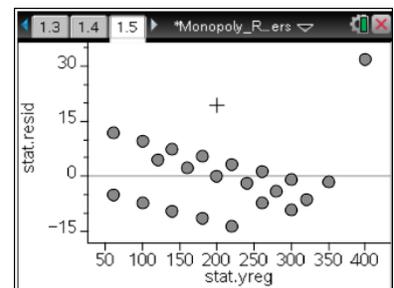
    b.   Describe the slope of the new regression line.

    Sample Answers: The new slope is 8.33, which is an increase from the previous slope. The cost of the properties increases an average of $8.33 for each additional space from GO.

Move to page 1.5.

15. Select **MENU > Window/Zoom > Zoom-Data**, and examine the new residual plot. Does it support a conclusion that the data are more linear? Explain your reasoning.

Sample Answers: Yes, the residuals seem to be fairly random on the left side and close to 0, but there is one unusual point in the upper right hand corner. This means that overall the predicted costs are very close to the actual costs of the properties and supports the conclusion that there is an approximately linear association between the number of spaces from GO and property cost.

16. There is one unusual point in the upper right-hand corner. This is the residual for Boardwalk. Explain why it is so large in context of the problem.

    Answer: Boardwalk's actual cost is much higher than the expected cost based on the linear model. It does not follow the general pattern of the rest of the data.

Extension

1. How can you determine whether a linear model is appropriate for a set of data? Discuss at least two ways.

   Answer: Students should discuss the appearance of the scatterplot and the appearance of the residual plot. They might also discuss the strength of the correlation but should be careful not to associate a strong correlation coefficient as implying a model is a good fit; the square of the correlation coefficient can also give a good sense of the relationship in terms of predictability.

Wrap Up

Upon completion of the lesson, the teacher should ensure that students are able to understand:
- when a least-squares regression line is an appropriate model for a data set.
- the meaning of the coefficients of a least-squares regression line in a given context.
- the effects of unusual points on a regression line.