Statistical Reasoning   **Name:** _____

Is there a relationship between the ticket price and distance from the field?

Is there any relationship between the price of the ticket and the distance from the soccer field? Assume I am going to get a seat right at the half-way line (centered) and that row 1 is close to the field and row 50 is further away up in the stands.

Row #	Price
2	$205
8	$175
10	$170
15	$105
20	$185
25	$100
30	$65
32	$78
40	$50
48	$20

1. Would it be appropriate to assign one variable as explanatory and the other as response, or is this just a relationship between two variables? How do you know?

2. Make a scatterplot of the data.
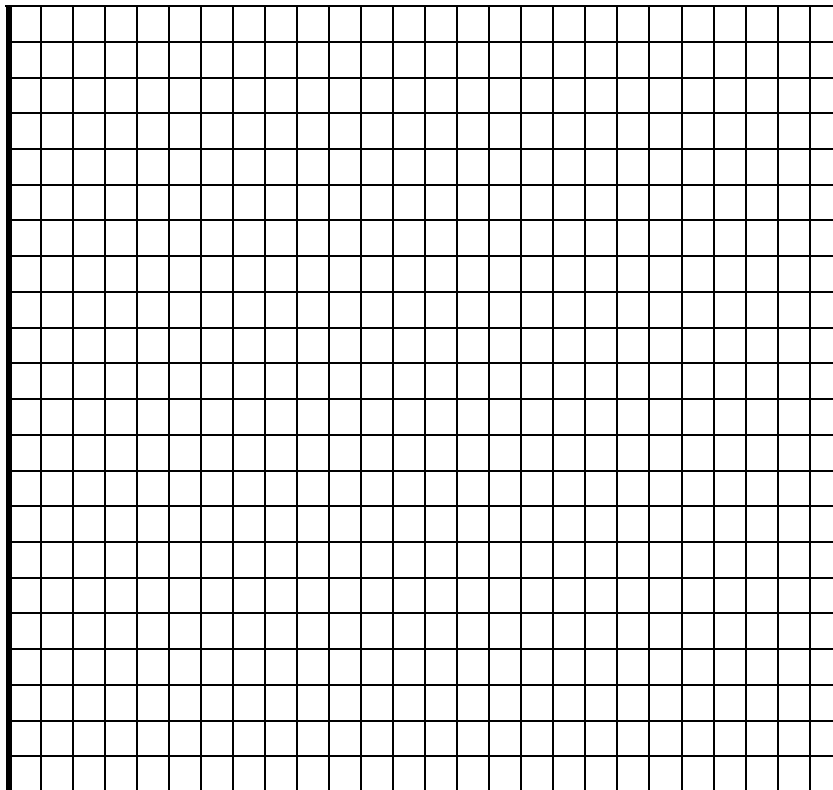
3. The correlation coefficient is:_____. Describe what this means in this situation.

4. The line of best fit is:

5. What does the slope (a) mean in this situation?

6. What does the y-intercept (b) mean in this situation?

7. What would you predict the cost of a seat on the 21st row? 37th row?

8. If your seat cost $127, what row would you expect to sit in?

9. What would happen to the correlation if you measured cost in euros instead of dollars?

10. Use residuals to determine if the equation you wrote in Question #3 is a good fit for the data.

x	Actual Y	Equation	Predicted Y	Residual
2				
8				
10				
15				
20				
25				
30				
32				
40				
48				

11. Using the residual data, do you think that the line is a good fit or not. Explain.

12. A residual is the difference between a piece of data (observed value) and the regression equation (predicted value). What do you think a positive residual means in terms of price vs seat at the stadium? A negative residual?

13. What is the value for R^2? Explain what this number tells us about the regression line.

14. What percent is due to lurking variables? Name at least 2 lurking variables.