Name _____ Section _____

Turn in the Results and Discussion Question sections.

Results

Table 4.1: *Guessed and True Distances to Objects*

Object	Object Description	Guessed Distance (ft)	Teams to Measure Distance*	Your Team's Measured Distance (ft)	True Distance (Median Measured Distance) (ft)
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					

*instructor determines several teams to measure each distance

Table 4.2: *Quality of fit for the linear and quadratic models*

Fitted Model	Do the scatter plot points seem to follow the fitted line or curve? (yes or no)	Error standard deviation s	95% Prediction Error $2s$
Linear			
Quadratic			

Table 4.3: *Predicting the 11th distance (show details of calculations)*

1	Chosen Model (Linear or Quadratic)	
2	Basis for model choice (check all that apply)	• Better visual quality of fit • Smaller value for prediction error 2s • Simpler model
3	Chosen Model's prediction formula ($\hat{y}$ as a function of x = guessed distance)	
4	Chosen Model's prediction $\hat{y}$ for 11th distance using guessed distance x =_____	
5	True distance y for the 11th object (supplied by instructor)	
6	Prediction error for 11th object (true y) - $\hat{y}$	
7	Does your prediction error lie within the bounds -$2s$ to $2s$? (yes or no)	

Discussion Questions

1. Why might the measurements of "true distances" be different for different groups even though they are measuring the same distance?

2. Describe any patterns you see in your raw scatter plot. Is the relationship linear? Or, does it seem to describe a curve? Are there any points that seem to depart from the pattern of the others (outliers)?

Name _____ Section _____

3. Look at how the points on your raw scatter plot lie in relation to the $y = x$ line. Are they scattered evenly above and below the line? If yes, you are approximately an unbiased guesser. (This is rare.) If most of the points lie below the $y = x$ line, you tend to overestimate the distances: you are positively biased. If most of the points lie above the $y = x$ line then you underestimate distances and are negatively biased. What type of guesser are you?

4. Which model (linear or quadratic) seemed to be the best model for your data? Explain why you chose this model.

5. Using your chosen model, look at its prediction error for the 11th object, $(y - \hat{y})$. Did it satisfy the empirical rule? Why?

6. **Reality check.** For predicting the health of one's kidneys using only a blood sample, what is the response variable (spell it out), and what is the predictor variable? Give at least one reason why the predictor variable is easier to measure than the response variable.

REFERENCES

Nih.gov. The kidneys and how they work. (2009, February). *NIH Publication*, 09-3195. Retrieved from http://kidney.niddk.nih.gov/kudiseases/pubs/yourkidneys/#tests