Name:_____

AP Statistics ASSignment 4.1 Regression Models

1. Definitions: Define each of the following using your own words:

Transforming/re expressing data:

Linear Regression: .

Exponential Regression:

2. *Basic Algebraic properties of logarithms:* $log_b x = y$ if and only if $b^y = x$
   Write the rules for each logarithmic expression:

Rules:

- $log_b(MN) =$
- $log_b(M/N) =$
- $log_b\, x^p =$

3. Given each set of data, indicate which regression is best for modelling it. Perform the regression on your Ti-83 and write the equation:

i)

x	5	10	15	20	25	30	35	40	42
y	1.1	3.1	8.7	26.5	84.1	252.77	1350.3	2368.22	3000

ii)

x	4	6	9	16	18	21	23	25	26
y	33.5	112.5	252.2	1003.2	1403	1950.2	2395	2796	3300

iii)

x	1	2	3	4	5	6	9	12	14
y	24	31.6	36	39	41.3	43.5	48.5	51.5	55.3

4. Which regression would be best when comparing each pair of variables: Quadratic, Exponential, Logarithmic, Power, Logistic, or Linear. Explain your

i) The size of a salmon vs the amount of lice infection

ii) The size of a lake vs the amount of aquatic life

iii) Height of a child vs the age of the child

5. A scatterplot of a company's revenue versus time indicates a possible exponential relationship. A linear regression of $y = \log(\textit{revenue in } \$10,000)$ against $x = \textit{years}(\text{since } 2005)$ gives $\hat{y} = 0.875 + 0.59x$ with $r = 0.72$.

   a) What does the slope represent?

   b) What does the Y-intercept represent?

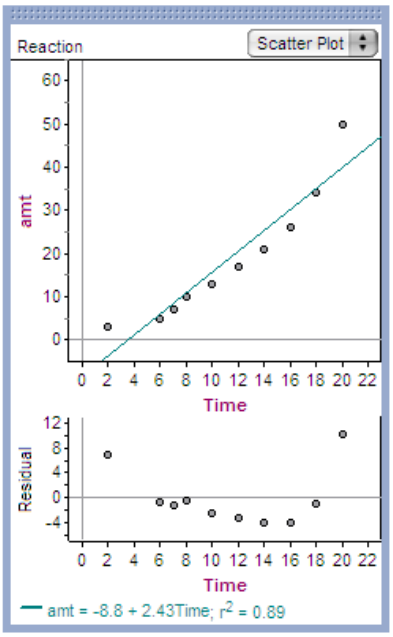
   c) What is the equation in exponential form?

   d) What is the predicted revenue for the year 2009?

   e) What percent in the variation in revenue can be explained by the variation in time with the LSRL?

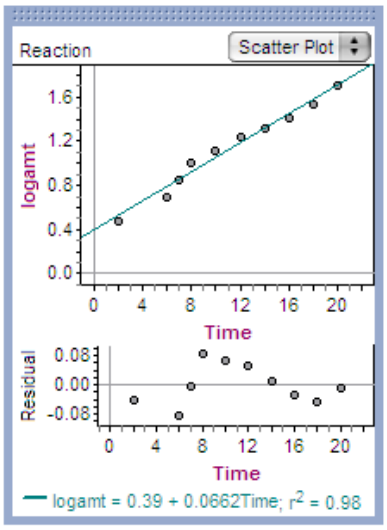
   f) With this equation, can we use the revenue to predict the year? Explain

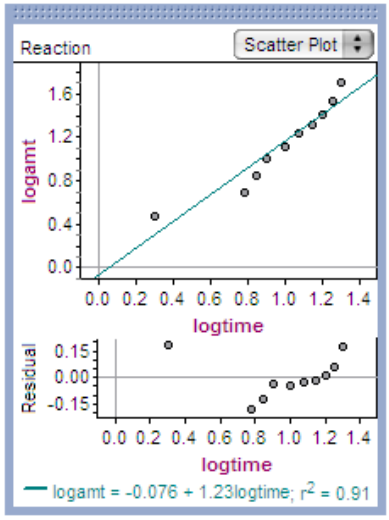
6. Some data were collected during a chemical reaction in which reactants A and B were reacting to form products C and D. As the products were formed, scientists measured their masses (in grams) at several times (in minutes) during the experiment. The Fathom screen shot displays a scatterplot with the least-squares regression line superimposed, and a residual plot.



a. How well does the linear model fit these data? Justify your answer.

b. We used Fathom to transform the data and to produce the two screen shots below. Would an exponential model or a power model provide a better description of the relationship between reaction time and amount of product? Justify your answer.

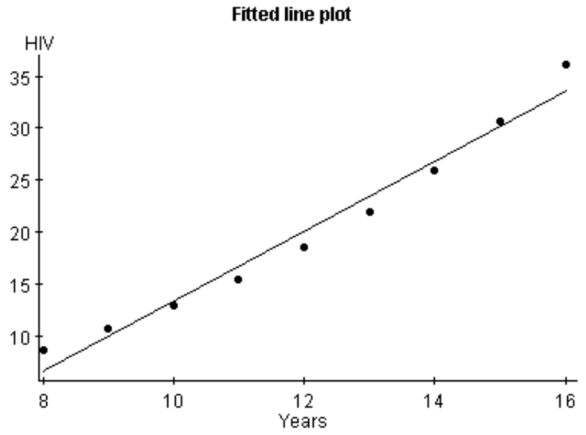


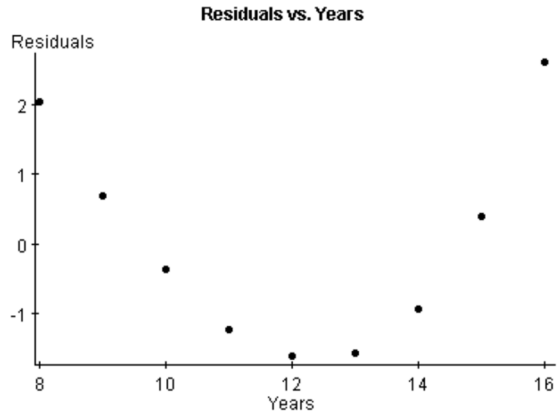


c. Use the model you chose in Question 6b to predict the amount of product that had been produced after 5 minutes. Show your method.

7. This dataset contains the global estimate of cumulative HIV cases worldwide. A scatterplot with the least-squares regression line superimposed and a residual plot from CrunchIt! are displayed.

Years since 1980	HIV infection (millions)
8	8.7
9	10.7
10	13.0
11	15.5
12	18.5
13	21.9
14	25.9
15	30.6
16	36.2





a. How well does the linear model fit these data? Justify your answer.

b. Some experts warned that the number of cases of HIV infection was growing exponentially. Use your calculator to perform an appropriate transformation of the data to achieve linearity if the underlying relationship between the variables is exponential. Perform least-squares regression on the transformed data. Record your regression equation below. Define any variables you use.

c. Use your model from Question 7c to predict the number of HIV infections in the year 2000.