2013

Modeling the World

Jeremy Martin
Trinity University

UNDERSTANDING BY DESIGN

Unit Cover Page

Unit Title: Modeling the World

Grade Level: 11th – 12th

Subject/Topic Area(s): Statistics / Pre-Cal

Designed By: Jeremy Martin

Time Frame: 2 weeks

School District: KIPP San Antonio

School: KIPP UPrep

School Address and Phone: KIPP UPrep
319 E Mulberry Ave
San Antonio, TX 78212
(210) 290-8720

Brief Summary of Unit (Including curricular context and unit goals):

Pre-Requisites for the Unit: Students should have studied exponential and polynomial functions and linear regression before starting this unit.

Context: This unit is envisioned as an end of course unit for Pre-Calculus. The unit is designed specifically to incorporate elements of Calculus and Statistics for students interested in taking either AP course.

Unit Goals: Students should come out of this unit with the ability to perform polynomial and exponential regression and interpret the residuals and coefficient of determination in context of the model's goodness of fit. Students will gain proficiency in using computers and calculators to display data and create regression models.

UbD Template 2.0

Stage 1 – Desired Results

Precalculus standards:

P1B: TSW use a problem-solving model that incorporates analyzing given information, formulating a plan or strategy, determining a solution, justifying the solution, and evaluating the problem-solving process and the reasonableness of the solution;

P1D: TSW communicate mathematical ideas, reasoning, and their implications using multiple representations, including symbols, diagrams, graphs, and language as appropriate

P2I: TSW determine and analyze the key features of exponential and polynomial functions including zeros, and intervals over which the function is increasing or decreasing;

Transfer

Students will independently use their learning to…

Analyze polynomial and exponential functions/growth models and determine which best applies in a given situation.

Meaning

Understandings

- Polynomial and exponential functions can be used to model real-life situations.
- Exponential models can be linearized by graphing on a logarithmic scale.
- Exponential growth will eventually outstrip polynomial growth, given enough time.

Essential Questions

- What is a *good* model? Does such a thing exist? How do we distinguish between good and better? Useful and not useful?
- Is a simpler model preferable to a more complex model that fits the data better?

Acquisition

Knowledge

- Students analyze patterns in the scatterplot to determine correct function type for regression.
- Students will be able to transform data and perform regression to generate models from existing data.
- Students evaluate transformed data for linearity.
- Students will justify conclusions using statistically sound reasoning.
- Students will use R^2 values and residual plots to evaluate a model's strength.

Skills

- Students will be able to use stat plot to generate a scatterplot
- Students will analyze the shape of a scatterplot to determine what type of functional regression is appropriate.
- Student will transform response data into log(response) and analyze scatterplot to determine linearity
- Students will perform polynomial and exponential regressions on the TI 84.
- Students will use the TI 84 to estimate the point of intersection between two functions.

Stage 2 – Evidence		
CODE (A, M or T)	**Evaluative Criteria** (for rubric)	
T		Performance Task(s) *Students will demonstrate meaning-making and transfer by…* Completing the Modeling performance task --- Other Evidence (e.g., formative)
A	Grade	Calculator skills quiz: Students use stat plot, quadratic, cubic, quartic, exponential regression
M	Grade	Evaluation of models: Students use metadata from regression models to determine the most useful model
T	Grade	Logarithm scale: Students identify data and determine if the standard xy plane, $x \log y$, or $\log x \log y$ is the best transformation.
A	Grade / Informal	Homework, Classwork and Warm-Up questions as needed.

Stage 3 – Learning Plan		
CODE (A, M, T)	Pre-Assessment *How will you check students' prior knowledge, skill levels, and potential misconceptions?*	
	Learning Activities	Progress Monitoring (e.g., formative data)
A	Day 1 Pre-test: Review of Functions, logarithms	Formative
	Pre-work: Students read (Modeling the World page 1-2 and complete the answers.	Formative
A	Day 2 Direct teach: Linear models and linear regression review. Concept of fit, R^2. Focus: Calculator skills (DiagnosticOn, STAT PLOT, Regression), interpreting data, identify the domain of a model. Potential problems with forecasting.	Formative: Check the answers
M	Classwork: Sketch line of fit. Interpretation questions. (Modeling the World Pages 3-4): What is a good model?	
	Day 3 Graphing on a log-transformed axis.	
A	Direct Teach: Examples include graphing the function $y = 2^x$ on [0,10] in standard x-y and in x-logy form. Students will follow along on paper and calculator. Key understanding: the log transformation will straighten out the exponential function into a linear function. Other examples could include $y = k^{-x}$	Formative: Graphing on personal white boards

M A T	to demonstrate that the line could have a negative slope. Classwork: Students will graph a log-transformed function ($y = 5^x$) and estimate its slope. Day 4: Guided practice: Students graph a polynomial function of high degree (>5) on interval [0,20], then determine that the log transformation and graph it. Students will be prompted to answer questions and they will notice that the log transformation does not straighten the equation out. Quiz: The students will use a log transformation to determine when an exponential function and a high-degree polynomial intersect. This is direct preparation for performance task 1.	Summative: Quiz (Performance Task 1 could be moved here to take the place of the quiz)
A	Day 5: Exploration 1: Use a time series to demonstrate quadratic model (ball falling from a height). Discuss when the model falls apart (hitting the ground). Exploration 2: Use time series to demonstrate exponential growth (bacteria). Discuss transformations. Students will do the log transformation and look for patterns there. We will discuss performing a linear regression on the transformed data. This will introduce an exponential model.	Informal assessment
M	Day 6: Validating polynomial and exponential models. Focus: Deepen the exploration of the bacteria. Transfer residual analysis to the exponential model. Guided practice: Exponential regression on the calculator. Exploration 1: Residual plot for the exponential regression. The students will learn that the residuals are very large. Exploration 2: Students transform the response into log(response). They will perform a linear regression and confirm that the R^2 values are the same as with exponential regression. They will also examine the residuals for the log-transformed data and transfer what they learned in the linear residual analysis to this. Is a simpler model preferable to a more complex model?	Formative assessment: Check answers
M	Day 7: Transferring to calculator and computer. Physical skills (collecting and entering data) and interpretive skills in evidence. Students will use Excel to perform regression analysis. Guided practice will involve recycling previous explorations and transferring the skill set from calculator to the computer.	Formative: Students email completed documents to me.
T	Day 8 Quiz to be completed on computer or paper / calculator. Students self-differentiate.	Summative/Formative

M	Day 9 Breakdown of the model. World example: Stock market crash of 1929 and its effect on the Dow Jones. Conditions for regression. Exploration: Students will be given the Dow Jones or the American GDP values from Dec 1920 to Dec 1928. They will not know what they are tracking, but they will be told to choose and defend a regression model. The students will then put their estimates for the next year on the board. The real value will be revealed once everyone's best guess is displayed and it will be shockingly lower. Class discussion: Assumptions for a predictive model. The idea is for the class to arrive at a set of conditions that must be met if we are to use a predictive model with any certainty. Discussion of quiz results at end of class. Day 10: Performance Tasks 1 & 2 to be completed.	Informal Summative

Note: Pages 6-9 is a 4 page intro to the linear model that I wrote and am including here. The first two pages are designed to be completed by the student prior to beginning the unit while the last two pages are designed to be completed after a refresher lesson on linear modeling.
Page 10 Includes a description of the two performance tasks that will follow.

The Linear Model

The first functional relationship that an Algebra student learns is the linear function. Linear functions are terrific for a lot of reasons. They are easy to understand. They have real-world applications and they are a starting point for abstractly representing a real-world scenario into a mathematical equation. Consider a customer buying gasoline: If one gallon of gas costs \$3.29, then two gallons costs \$6.58. We can "Algebra-tize" this by stating that amount that the customer paid, y, for x gallons of gasoline is modeled by the equation (Fill In the Blank): ____________________________

The above scenario is an example an introduction to linear functions. Many other factors can be added to complicate the problem. The customer could choose to purchase a hotdog and soda from the convenience store for \$2.94. The new linear function is (FIB): ____________________________. The amount \$2.94 is called the ____________________________.

Notice that there is an artificiality imposed on our problem. The convenience store has set the cost of the gasoline and the cost of the hot dog and soda. How could we create a linear model if we do not know what the rate of change is? What if there is a process that appears to be linear but the data set does not exactly fit? To determine a model that best fits the data, we can use a process called *linear regression*.

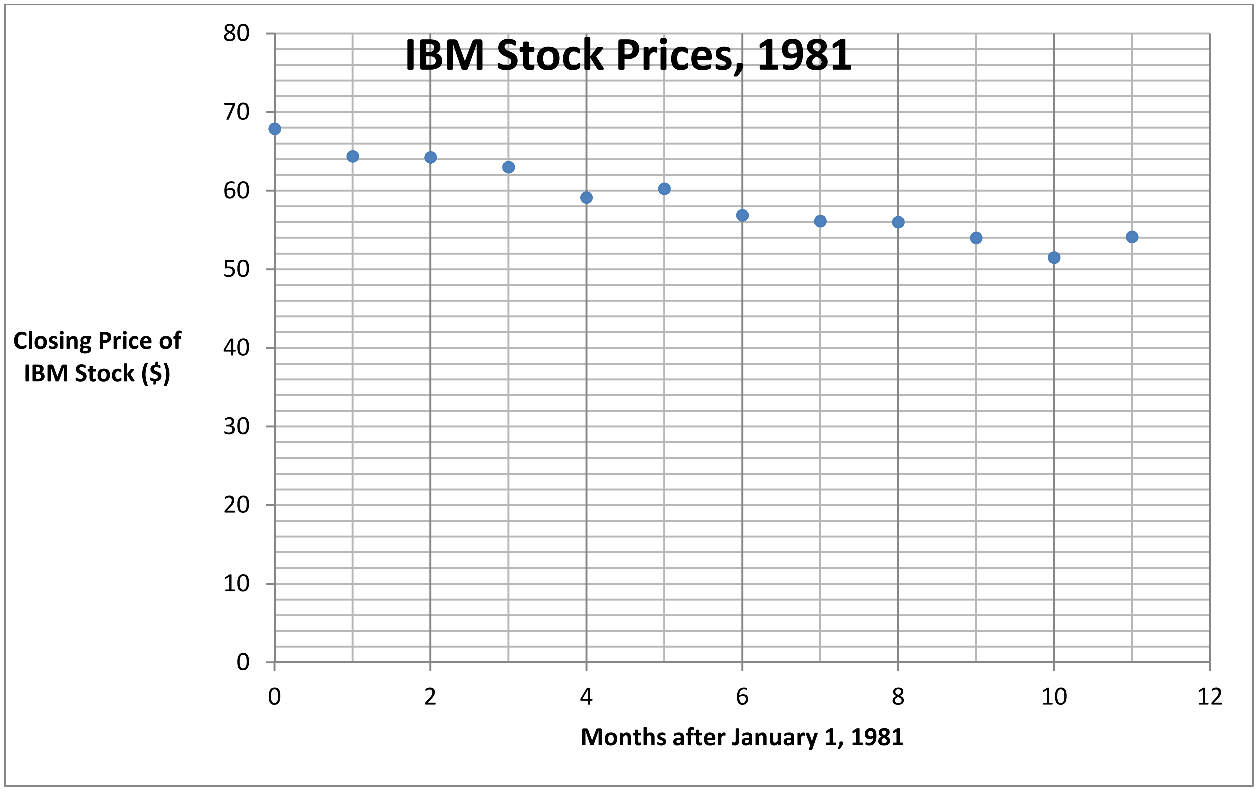
> Note: A linear regression model assumes that there is error in a linear process. The error is represented by the lower-case Greek letter ϵ (epsilon). The general form of the linear equation is changed to: $y = ax + b + \epsilon$, where a is the rate of change, b is the intercept and ϵ is the error. The error is also called the *residual*. A *best-fit* model minimizes the residual error.

In statistical modeling, the independent variable is usually called the *predictor* and the dependent variable is called the *response*. Data can be displayed on a *scatterplot* to determine if there appears to be a relationship between variables. The predictor is displayed on the $x-$ axis and the response is on the $y-$ axis. The shape of the scatterplot can be examined to determine if there is a correlation between the two.

The closing cost of IBM's stock was recorded on the first day of each month in the year 1981. The table below reflects the cost of the stock through the year.[1]

Jan 1	Feb 1	Mar 1	Apr 1	May 1	Jun 1	Jul 1	Aug 1	Sep 1	Oct 1	Nov 1	Dec 1
67.88	64.38	64.25	63.00	59.13	60.25	56.88	56.13	56.00	54.00	51.50	54.13

From the beginning of the year to the end of the year, the stock's price dropped almost 20%. Examine the scatterplot on the following page and look for a trend.



Your Turn:

A. Name the response and predictor variables.

B. In pencil with a straightedge, sketch a line that you believe fits the points.

C. On your line, what is the approximate value of y when $x = 12$? What does this point represent in the scenario?

Using the Linear Model - Classwork

Use the linear regression methods from class to create a linear model to fit the IBM Stock Problem. Record your equation and the R^2 value below. What does the R^2 value tell you about the goodness of fit?

Graph your linear model (in pen) on the previous page's scatterplot. How close was your original line to the line of best fit?

The table below will be used to calculate the residuals for the TI stock problem. Put the function in the heading next to y and evaluate at the x −value. Then enter the residual in the fourth column.

No. Months after 1/1/81	IBM Stock Price ($)	Line of Best Fit Value	Residual
x	y	$\hat{y} =$	$y - \hat{y}$
0	67.88		
1	64.38		
2	64.25		
3	63.00		
4	59.13		
5	60.25		
6	56.88		
7	56.13		
8	56.00		
9	54.00		
10	51.50		
11	54.13		

Classwork Problem Set:

A. In the space below, plot the residuals against the predictor variable. What does the residual plot tell you about the fit?

B. On June 15, 1981, the closing price of IBM stock was $59.88. According to the linear model, what is the best-fit value of the stock on that day?

C. On 1/1/1982, the closing price for IBM stock was $56.88. On 1/1/83, the closing price was $96.25. According to the model, what is the predicted closing price for those two dates?

D. Compare the predicted values in question C to the actual values. Do you believe that the model should be used to predict values outside of the time period that was sampled from?

Performance Task Descriptions

Both pieces of the performance task will be given to the students at the beginning of the unit. The students will have the option to work ahead or to complete the work at the end. Two class periods will be provided to complete the task.

Part 1: Functional analysis. Students are randomly assigned a polynomial and exponential function from the group below (subject to change) with domain restriction $x \geq 0$. Make sure the leading coefficient is > 0 and that constant term is > 1 for all polynomials used.

Polynomial	Exponential
$f(x) = 5x^2 + 3x + 6$	$g(x) = 2^x$
$h(x) = x^4 - x^2 + 13$	$k(x) = 1.5^x$
$m(x) = x^5 + 31$	$n(x) = 1.2^x$
$p(x) = \frac{1}{100}x^{10} + 3x^7 - x^5 + 32$	$q(x) = 1.1^x$
Etc.	

The students will do a complete graphical analysis on their two functions to include end behavior and intersection of the two graphs. There will be guided questions asking the student to relate the polynomial to the exponential and they will have to answer the final question: Over a long period of time, which type of growth is greater, exponential or polynomial.

Output for part 1: Students will submit the following:

1 graph with both functions sketched with appropriate scaling to include the point of intersection

1 completed handout answering the guided questions

Part 2: Data analysis. Students will be provided with data sets that indicate growth. Each set will have a predictor and a response. The students will perform a logarithmic transformation for the response variable to determine if the resultant graph is linear. Students will perform polynomial and exponential regressions (upwards differentiation: Power regression option can be included) for the data and make an informed choice about which approach would best model the relationship. The data sets will be generated so that each person gets his own data set. The creation of the data will be done using a statistical software program or Excel.

Modeling the World: Performance Task 1

You will receive two function rules from your teacher, one exponential and one polynomial. Write the equations in the spaces below.

Polynomial Function $p(x)$: Exponential Function $q(x)$:

Using the methods we have discussed in class, determine the intersection of the graph of the two functions and then answer the following questions.

1. Write the ordered pair of the solution to the equation $p(x) = q(x)$. (10 points)

2. Let A represent the $x-$ coordinate from question 1. If $x < A$, which function appears to be greater, p or q? What about when $x > A$? Write both answers in conditional form ("If …, then …"). (10 points)

3. Using graph paper or technology, graph the functions $y = \log p(x)$ and $y = \log q(x)$. Label the functions and the axes. Choose appropriate boundaries to include the point of intersection. (30 points)

4. Divide the interval $[0, A]$ into 5 equal subintervals, $\left[0, \frac{A}{5}\right], \left(\frac{A}{5}, \frac{2A}{5}\right], \left(\frac{2A}{5}, \frac{3A}{5}\right], \left(\frac{3A}{5}, \frac{4A}{5}\right]$ and $\left(\frac{4A}{5}, A\right]$. Determine the average rate of change for $p(x)$ and $q(x)$ over each of the intervals. Compile the rate of change data into a table. (30 points)

5. Compare the rate of change for $p(x)$ to the rate of change for $q(x)$ over the last interval. Will there be another intersection point for these two functions? Justify your response with supporting evidence. (10 points)

6. Over long periods of time, which type of growth is greater, exponential or polynomial? Support your claim. (10 points)

Modeling the World: Performance Task 2

Objective: You have been given a data set. You will determine whether the data set is best modeled by a quadratic function or an exponential function. This can be turned in to me electronically or on paper.

Tasks: 1. You will display the data graphically.

2. You will run a quadratic regression on the data. The complete analysis will include:
- The best-fit quadratic function and its R^2 value.
- Residual plot analysis

3. You will run an exponential regression on the data. The complete analysis will include:
- The best-fit exponential function and its R^2 value.
- Residual plot analysis

4. You will determine which model is a better fit for the data. A complete answer will include:
- Comparison of the fit of the two models
- Comparison of the R^2 values
- Comparison of the residual plots

Key Points: In statistics, there is seldom a right or wrong answer. The most important thing is to perform the analysis deliberately and to support your answer with valid reasoning.

Follow the rubric to maximize your score

Rubric

Task	Mastery (4)	Proficient (3)	Emerging (2)	Developing (1)
Graphical Display		Appropriate representation; Axis labels; Units; Neat	Appropriate representation; Missing labels or units; Sloppy	Inappropriate representation
Quadratic Regression				
Equation	Accurate; Uses correct notation; R^2 value calculated	Accurate; R^2 value calculated; Uses incorrect notation	Inaccurate; Uses correct notation; R^2 value calculated	Inaccurate; Inaccurate notation; R^2 value missing
Residual Plot	States if using residual or log-residual; Axis labels; Units; Neat; Accurate	Axis labels; Units; Neat; Accurate	Missing labels or units or inaccurate.	Missing all elements
Residual analysis		Student describes the residual plot and discusses curve or pattern.		Student does not describe the residual plot
Exponential Regression				
Equation	Accurate; Uses correct notation; R^2 value calculated	Accurate; R^2 value calculated; Uses incorrect notation	Inaccurate; Uses correct notation; R^2 value calculated	Inaccurate; Inaccurate notation; R^2 value missing
Residual Plot	Student uses log-residual; Axis labels; Units; Neat; Accurate	Missing one of five elements	Missing 2 – 3 elements.	Missing 4 or more elements.
Residual analysis		Student describes the residual plot and discusses curve or pattern.		Student does not describe the residual plot
Comparison				
Model Chosen		Chooses		Does not choose
Argument $(\times 2)$	Student compares residual plot and R^2 to make a choice. Logical reasoning displayed	Student compares residual plot and R^2 to make a choice. Some flaw in reasoning	Student compares either residual plot or R^2	Student does neither.

Maximum Rubric Score: 36.

Your grade will be determined by the following ($G = \text{grade}$ and $S = \text{score}$) :

$$G = 2S + 30$$