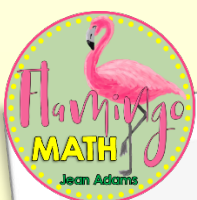


AP[®] Precalculus

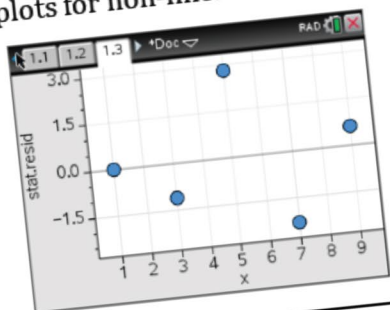
Competing Function

Model Validation

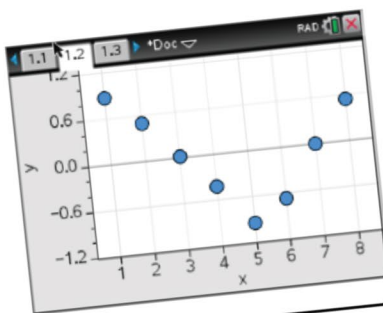


Residual Plots

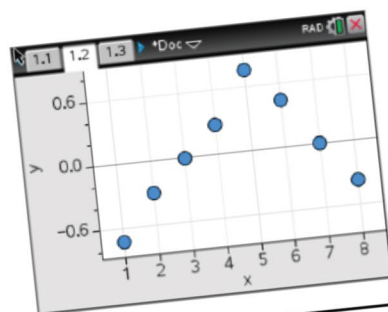
A **residual plot** is a scatter plot that shows the residuals on the vertical axis and the independent variable on the horizontal axis. We will use residual plots to help us decide whether the model chosen is an appropriate model. Notice the residual plot (below) for our example where $\hat{y} = 2.05x + 4.95$ that the residuals are randomly placed above and below the x-axis. You will see either random residual plots, appropriate for linear regressions, or patterned residual plots for non-linear regressions.



Random



U-Shaped Pattern
Non-Linear Model



Inverted U-Shaped Pattern
Non-Linear Model

Completely Editable

2.6

Lesson 6-A: Competing Function Model Validation

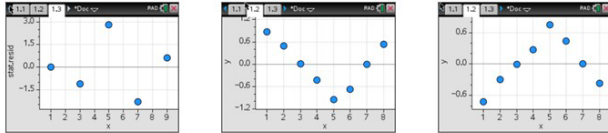
In this lesson we will explore how to construct linear, quadratic and exponential models based on a data set. We want to focus on whether our model can be justified as *appropriate* for a data set by using the graph of the residuals of a regression.

Revisiting Linear Regression Models

EX #1: Darrin injured his hamstring in a football game. His sports trainer sent him to physical therapy to build up his leg strength by using a regulated weight machine. The table and graph show the number of repetitions he has been able to do on each day he was in PT. The best-fitting linear function $\hat{y} = 2.05x + 4.95$ (read as "y hat") is also shown. We can

Residual Plots

A **residual plot** is a scatter plot that shows the residuals on the vertical axis and the independent variable on the horizontal axis. We will use residual plots to help us decide whether the model chosen is an appropriate model. Notice the residual plot (below) for our example where $\hat{y} = 2.05x + 4.95$ that the residuals are randomly placed above and below the x-axis. You will see either random residual plots, appropriate for linear regressions, or patterned residual plots for non-linear regressions.



Random U-Shaped Pattern Inverted U-Shaped Pattern

EX #6: The total number of views for a popular podcast is increasing. The table shows the estimates of the total views each month since January 2010.

Month (Jan '10 = 1)	1	6	14	25	27	31	33	38	42
Total Views (in millions)	0.2	0.55	1.1	2.9	3.8	8.2	9.3	10.6	15.7

A. Use a graphing calculator to find a quadratic, exponential, and power regression model.

Quadratic Model: _____

Exponential Model: _____

Power Model: _____

EX #8: While driving down a dirt road you punctured the tire of the pressure, y , measured in pounds per square inch (psi), measured in seconds after the tire was punctured.

x (sec)	5	10	15	20	25	30
y (psi)	28	22	17	14	10	8

A. Find an exponential function in the form of $y = ab^x$ that has the best correlation coefficient?

B. Plot the equation and scatter plot on the same screen. Does the equation fit the data points? Explain.

C. Using the model, what was the pressure of the tire before the puncture?

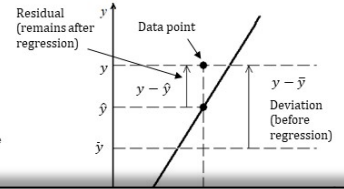
D. The residual graph is shown at right. Does this match your expectations? What does it tell you about the validity of your regression model?

EX #9: A set of data has an upward trend, shown in the plot below. Is the best-fitting function linear or exponential? Use the residual plot to help you decide.

Vocabulary and Definitions

Given a data point (x, y) , the **deviation of a data point** is $y - \bar{y}$, the directed distance of its y -value from $\bar{y}$, where $\bar{y}$ is the average of the y -values.

The **residual of a data point** from the line $\hat{y} = mx + b$ is $y - \hat{y}$. This is the vertical directed distance of its y -value from the regression line. Residuals help to determine if a curve is **appropriate** for the data set.



EX #4: Suppose that these data have been measured for two related variables x and y .

- Enter the data in two lists on your grapher. Show by linear regression that the best-fitting linear function is $\hat{y} = 1.52x + 2.35$.
- Record the correlation coefficient, r .
- Complete the table by finding the regression points $(x, \hat{y})$.
- Find the residual points using L3 to complete the final column. (Round to 3 decimal places.)
- Make a residual plot of the data on your graphing calculator. How well does the linear function fit the data?

Data Points (x, y)	Regression Points ($x, \hat{y}$)	Residual Points
(4, 9)		
(7, 14)		
(10, 18)		
(13, 25)		
(16, 23)		
(19, 28)		

Lesson 6-B: Competing Function Model Validation

We want to observe both the residual plot and the endpoint behavior when choosing an appropriate model.

- ☐ If the residual plot follows a regular pattern, then there is a behavior that is not accounted for by the kind of function chosen.
- ☐ If the residual plot has no identifiable pattern, then the regression equation most likely will model all but random fluctuations in the data.

EX #7: Your biology teacher has bacterial cultures for a class of 24 students. Each student is measuring the number of bacteria, y , in his or her culture at intervals of time, x hours during the experiment. The data are shown in the table below.

x (hours)	y (bacteria)	x (hours)	y (bacteria)
0.5	200	3.8	1412

Competing Function Model Validation Practice Set 1

Name _____

Date _____ Period _____

Problems 1-3, for each scenario given, find the residuals.

1. Sarah is selling bottled water on a hot summer day at a tennis tournament. She records each day's high temperature and the number of bottled waters sold for the 7-day tournament. She found her data to be linear in nature, so she found a least squares regression equation to predict the sales of water, from the daily high temperature, in degrees Fahrenheit: $\hat{y} = 0.75x - 32$. What is the residual if a day had a high temperature of 93 degrees and Sarah sold 35 bottles of water?

7. The US bottled water producer's revenue, in billions of US dollars, from 2005 to 2012 is shown in the table at right. Use a graphing utility to write a regression equation. The time, in years, is given by t .

- A. Create a scatter plot to determine the type of regression model that best fits the data set. Find the equation and the correlation coefficient, r .

- B. Explain the meaning of the y -intercept in context.

- C. Using the residual plot for the regression model, do the residuals follow a definite pattern or are they randomly scattered? Explain.

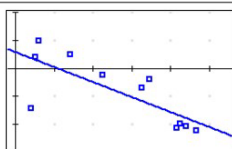
Year	Revenue, in billions USD
2005	\$10.01
2006	\$10.86
2007	\$11.55
2008	\$11.18
2009	\$10.60
2010	\$10.69
2011	\$11.07
2012	\$11.82

Source: Beverage Marketing Corp.

Problems 3-4, Answer the following by analyzing the data given.

3. The scatterplot displays a set of data along with its least-squares regression line. Consider removing the outlier $(2, -7)$ and calculating a new regression line. Which statement would be true? Explain.

- A. The slope of the regression line would increase.
B. The coefficient of determination (r^2) would decrease.
C. The y -intercept of the regression line would increase.



7. A jet-ski's average speed, y , (in knots) for several engine speeds, x , in hundreds of revolutions per minute (RPM's) are shown below.

Engine speed, x	8	11	14	17	19	22.5
Jet ski speed, y	8.97	14.18	19.62	26.05	31.29	43.13

- A. Use the data to determine a polynomial model of best fit. How did you determine your model?

- B. Use your model to estimate the average speed for the jet ski when the engine speed is 2500 RPM's.

- C. What engine speed will produce a jet-ski speed of 20 knots?

8. Vehicles are tracked at a busy intersection for a 24-hour period and reported in the table below, where $t = 0$ represents midnight.

Hours	1	3	6	10	14	18	21
Vehicles	17	4	22	45	90	78	42

- A. Find a regression equation model that fits the data.

5. A local print shop sells yard signs for realtors in bundles of 50 signs. The profit margins for each bundle are shown below. Assume the behavior is a quadratic regression for a up to 800 signs.

- A. Find a quadratic regression model for data in the table below. Use the model to complete the table.

Quadratic Regression Model: _____

- B. Record the correlation coefficient, r .

x (Signs)	y (Profit, \$)	$\hat{y}$	$y - \hat{y}$ (residuals)
50	15		
100	36		

- C. Find the residual plot for the data. Based on the residual plot, could the

2.6 Competing Function Model Validation Practice Set 2

Name _____

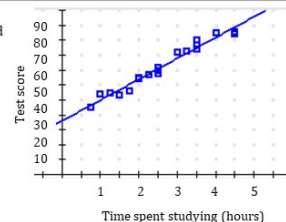
Date _____ Period _____

Problems 1-2, Interpret the slope and y -intercept of the regression line.

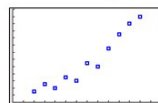
1. Your precalculus class completed a survey for your teacher at the end of Unit 1 Test. The scatter plot and trend line show the relationship between how many hours students spent studying, x , and their score on the test, y .

- A. The line fitted to the data has a slope of 14. What is the best interpretation of this slope?

- B. The y -intercept is 36. What is the best interpretation of this value?



5. Your agriculture club plants a different bean variety each week, for 11 weeks. Two weeks after the last been was planted, the plants have heights shown in the table and in the graph. The behavior of the data has an upward trend, can you determine the best-fitting function model?



Time (weeks)	2	3	4	5	6	7	8	9	10	11	12
Height (inches)	3	5	4	7	6	11	10	15	19	22	24

- A. Use the data set to find a linear regression and an exponential regression.

Linear model : _____ r : _____

9. The table below shows hours Shelby worked, x and tips received, y for a week.

Hours worked, x	4	5.5	4.75	3.5	7	6.5	3.5
Total Tips, y	\$38.50	\$54.30	\$148.75	\$35.90	\$115.65	\$87.40	\$34.85

- A. Use the data to determine a regression model of best fit and r -value.

- B. Find a linear regression model for the data, what do you notice about $\hat{y}(1)$?

- C. Shelby said that her father gave her a \$100 tip on the day she worked 4.75 hours. If you remove this \$100 tip, $(4.75, \$48.75)$ how does the regression equation change? What does the adjusted equation show you related to the data points?

- D. Can you calculate a better linear model by adjusting the y -intercept to $(0, 0)$ using $(6.5, \$87.40)$? Explain.

10. The table shows a lemonade stand data set based on temperature and revenue.

Temperature (degrees)	82	77	85	87	81	88
Revenue	\$50	\$32	\$58	\$72	\$47	\$92

- A. Find a linear regression model for the data. What is the correlation coefficient?

does your new regression model change the accuracy of the observed data?

- D. Use your model to find the number of vehicles that pass through the intersection at 5 pm each day.

Practice Sets

2.6 Competing Function Model Form A
Validation Set 1

Name _____

Date _____ Period _____

For problems 1 and 2, find the residual.

1. Taylor is selling lemonade. She found a least squares regression equation to predict sales of lemonade from the daily high temperature, $\hat{y} = 0.9x - 28$. What is the residual if a day had a high temperature of 95 degrees and she sold 61 glasses of lemonade?

2. The gas consumption rate in miles per gallon of 20 vehicles of various weights (pounds) showed data that fit an exponential regression model: $\hat{y} = 92.3(0.9995)^x$. Find the residual of a vehicle that weighs 2745 pound with an average of 21 miles per gallon.

3. Find a linear regression for the data in the table and look at the residual plot to determine if a linear model is appropriate. Tell why or why not.

x	1	2	3	4
$f(x)$	5	7	12	21

Grade:

2.6 Competing Function Model Form B
Validation Set 1

Name _____

Date _____ Period _____

For problems 1 and 2, find the residual.

1. Taylor is selling lemonade. She found a least squares regression equation to predict sales of lemonade from the daily high temperature, $\hat{y} = 0.85x - 24$. What is the residual if a day had a high temperature of 94 degrees and she sold 53 glasses of lemonade?

2. The gas consumption rate in miles per gallon of 18 vehicles of various weights (pounds) showed data that fit an exponential regression model: $\hat{y} = 92.6(0.9995)^x$. Find the residual of a vehicle that weighs 2810 pounds with an average of 24 miles per gallon.

3. Find a line
table and
determine
Tell why or
why not.

x	
$f(x)$	

Grade:

2.6 Competing Function Model Form A
Validation Set 2

Name _____

Date _____ Period _____

Use the following information for the problems. You collect data on how long you use your air pods and how much battery life is remaining (in hours). A trend line is created with x = hours using air pods, and y = hours of battery remaining.

1. The line fitted to the data has a slope of $m = -1.2$. What is the best interpretation of this slope?

2. The y-intercept is 8. What is the best interpretation of this value?

3. Write the equation of the trend line and use it to predict the amount of battery life left after using your air pods for 8 hours. Does the prediction seem reasonable in the context of the problem? Explain.

Grade:

2.6 Competing Function Model Form B
Validation Set 2

Name _____

Date _____ Period _____

Use the following information for the problems. You collect data on how long you use your air pods and how much battery life is remaining (in hours). A trend line is created with x = hours using air pods, and y = hours of battery remaining.

1. The line fitted to the data has a slope of $m = -1.1$. What is the best interpretation of this slope?

2. The y-intercept is 7. What is the best interpretation of this value?

3. Write the equation of the trend line and use it to predict the amount of battery life left after using your air pods for 7 hours. Does the prediction seem reasonable in the context of the problem? Explain.

Grade:

Daily Lesson Quiz