Algebra 2
Introduction to Data Analysis

Goals:

- Describe graphically, algebraically and verbally real-world phenomena as functions; identify the independent and the dependent variable (3.01)
- Translate among graphic, algebraic, and verbal representations of relations (3.02)
- Graph relations and functions and find the zeros of functions (3.03)
- Write and interpret an equation of a curve (linear) which models a set of data (4.01)
- Find the equation of best fit (linear) for a set of data. Interpret the constants, coefficients, and based in the context of the data. Check the equation for goodness of fit and use equation for prediction (4.02)

Materials and Equipment Needed:

- Copy of handout for each student
- Graphing calculator
- Graph paper for each student
- Paper for note taking during class
- Spring stand, spring, weight hanger, and weights (20 to 100 gm weights).

Activity One: Discussion of when it is appropriate to fit data with a linear function (using LinReg on calculator)
Using a group of prepared slides with data sets, discuss with students what data are appropriate to fit with a line. Refer to the PowerPoint file of **DataSets.ppt**. The list of data sets are:

- Land Speed Records over time (linear)
- Joseph Forte data 1 (game number, number of points scored) (cloud, not linear)
- Joseph Forte data 2 (game number, accumulated points scored) (linear)
- Distance an object has fallen versus time (quadratic)
- Average airline flight length versus number of seats on the plane (linear)
- Temperature of cooling water versus time (exponential)
- Average maximum temperature versus month in Henderson (not linear, trig)
- Day in April of first cherry blossom versus mean temp in March (linear, decreasing)

Activity Two: Fit linear data with linear regression model, interpret meaning of slope and y-intercept, and use linear function to predict outside given data. (This data can either be collected using spring stand, spring and weights or you can use the existing data set.)

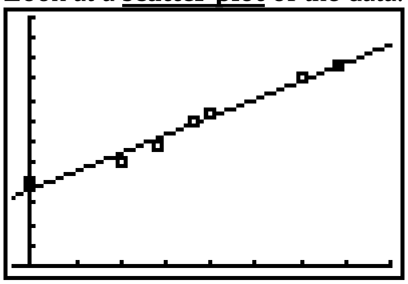
When weights are added to a spring, the length of the spring increases. Describe the relationship using data analysis.
Using a spring suspended from a spring stand, add weights to the spring. Measure the length of the spring (in centimeters) and record the total weight (in grams) on the spring.

weight (gms)							
spring length (cm)							

1. Using a spring stand with a spring, add weights to the spring and measure the length of the spring. Be sure to determine a method for length measurement.
2. As teacher takes data, students record data. A sample data set is shown.

weight (gms)	0	50	70	90	100	150	170
spring length (cm)	4	5.1	5.8	7	7.4	9	9.6

3. Based on observations in the data, students will discuss what the relationship might be. Discuss **independent** and **dependent** variables.
4. Sketch a plot of the data on the coordinate grid on the handout. Talk about scale and necessary labels. Discuss the shape of the data and the information we might get.
5. Students should **enter the data into the lists** of the calculator. Determine which variable is <u>independent</u> and which is <u>dependent</u>. Do a scatter plot on the graphing calculator. The window for the graph should parallel what was used on the scatter plot done on the grid. Talk about reasonable values for <u>domain</u> and <u>range.</u>
6. Look at a **<u>scatter plot</u>** of the data. Discuss **window** settings.



7. Use **LinReg** to determine the least squares line that fits this data. Use <u>best-fit line</u>. $y = 0.03475x + 3.71513$ where *x*=weight in grams and *y*=length of spring in centimeters. **Write the equation in the function list.**
8. **Superimpose the line over the data**. Discuss whether the model fits the data.
9. Discuss **the <u>slope</u> of the line** and its meaning. Be sure to use units. In this case the slope is 0.03475 cms/gram or for every gram added to the spring the length increases by 0.03475 centimeters.
10. Discuss the meaning of any intercepts. In this case the $y-$**intercept** should have a direct connection to the data collection. Technically, the $y-$ intercept should be (0,4). In this model the intercept of (0,3.71513) is showing some error. This intercept represents the length of the spring with no weights added.
11. Is this model a good model for the data? Determine how far each point is from the best-fit line. We will calculate the *y*-values of the line that go with the data points. This will allow us to determine the error in the model.
12. Forecast the weight that will stretch the spring to 15 centimeters by using the <u>prediction equation.</u> For this example, we will let *y*=15 and find the value of *x*-value. In this case the line passes through approximately (324.7, 15).

Follow-Up Activity: Return to How Old are They data set. Since each student will have a slightly different data set, as a class agree upon one data set to use for this activity.

How Old are They?

- Students should **enter the data into the lists** of the calculator. Ordered pairs are of the form $(age, guess)$.
- Look at a **<u>scatter plot</u>** of the data. Discuss **window** settings. Talk about <u>domain</u> and <u>range.</u>
- Use **LinReg** to determine the least squares line that fits this data. Use <u>best-fit line</u>. Results will vary based on the data set you select. In any case, it should be near $y = x$. **Write the equation in the function list.**
- **Superimpose the line over the data**. Discuss whether the model fits the data.
- Discuss the **meaning of the $y-$intercept and the <u>slope</u> of the line**.
- Determine how far each point is from the best-fit line.
- **Forecast** the actual age of a person that you guessed to be 37 years old.

Student Handout
Introduction to Data Analysis
Algebra II

1. When weights are added to a spring, the length of the spring increases. Describe the relationship using data analysis.

Using a spring suspended from a spring stand, add weights to the spring. Measure the length of the spring (in centimeters) and record the total weight (in grams) on the spring.

weight (gms)							
spring length (cm)							
length given by equation							

1. Show a scatter plot of the data above.

2. What is the Linear Regression model (best fit line) for the data? What does each variable represent?
3. What is the meaning of the slope? What is the meaning of the y-intercept?
4. How far is each point from the best-fit line? Find the $y-$value of the line that pairs with each $x-$value of the data. Put in chart above.
5. Do you think the linear regression model is a good fit for this data set? Why or why not?
6. Using this linear function as an equation of prediction, estimate the weight needed to stretch the spring to 15 centimeters.

Follow-Up Activity
Introduction to Data Analysis
Algebra II

Estimate the Ages of Famous People

The following list contains the names of famous people. Without talking to anyone, write down your estimate of the age of each person. If you do not know the person, make a guess.

	Actual Age ------ Estimated Age
Nancy Reagan	____________________________
Tiger Woods	____________________________
Mister Rogers	____________________________
Chelsea Clinton	____________________________
Eddie Murphy	____________________________
Tom Brokaw	____________________________
Oprah Winfrey	____________________________
Mick Jagger	____________________________
Heather Locklear	____________________________
Elizabeth Taylor	____________________________
Garth Brooks	____________________________
Jennifer Lopez	____________________________
Ringo Starr	____________________________

1. Enter data into lists of calculator $(age, guess)$.
2. Create a scatter plot. What should be the window settings?
3. Find the line of best-fit using LinReg on the calculator. Write the equation into Y1.
4. Superimpose the line over the data.
5. What are the meaning of the $y-$intercept and the **slope** of the line? Write answers in full sentences.
6. Determine how far each point is from the best-fit line. Write answers in the table above.
7. Forecast the actual age of a person you guessed to be 37 years old.