Before you begin this activity, you'll need to enable a feature on your graphing calculator. I believe this is only necessary for students with TI-82 or TI-83 calculators.

1. Enter the CATALOG (it's located above the zero button)
2. Move the cursor down to DIAGNOSTICON
3. Hit ENTER twice (you should see DONE printed on the screen).

Situation: The following table reports the distance (in miles) from Baltimore to each of 12 cities along with the corresponding airfare

Destination	Distance	Airfare
Atlanta	576	178
Boston	370	138
Chicago	612	94
Dallas	1216	278
Detroit	409	158
Denver	1502	258
Miami	946	198
New Orleans	998	188
New York	189	98
Orlando	787	179
Pittsburgh	210	138
St. Louis	737	98

1. Is *airfare* a function of the *distance* traveled? Describe the relationship between these variables.

2. Create a scatterplot for this data. It is common practice to place the independent variable on the horizontal axis and the dependent variable on the vertical axis. Would distance be a good predictor or airfare? Would it enable you to make perfect predictions?
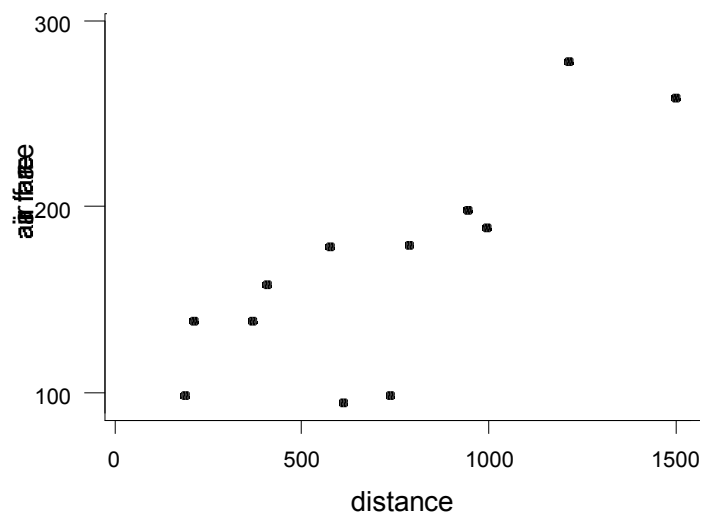
3. Draw a line on the scatterplot that seems to summarize well the relationship between airfare and distance. Estimate the slope and y-intercept for the line you just sketched and write out its equation.

4. Use your equation to predict the airfare for Atlanta, which is 576 miles away. Does this match the actual airfare for Atlanta (which is $178)? The difference between the actual airfare and the predicted airfare is called a *residual*. Complete the following table.

City	Distance	Predicted Airfare	Actual Airfare	Residual
Atlanta	576		$178	
Chicago	612		$94	

5. I imagine you all came up with slightly different prediction lines. How can we determine which prediction line is the best? Suppose two students came up with the following prediction lines. How do we determine which line is better?
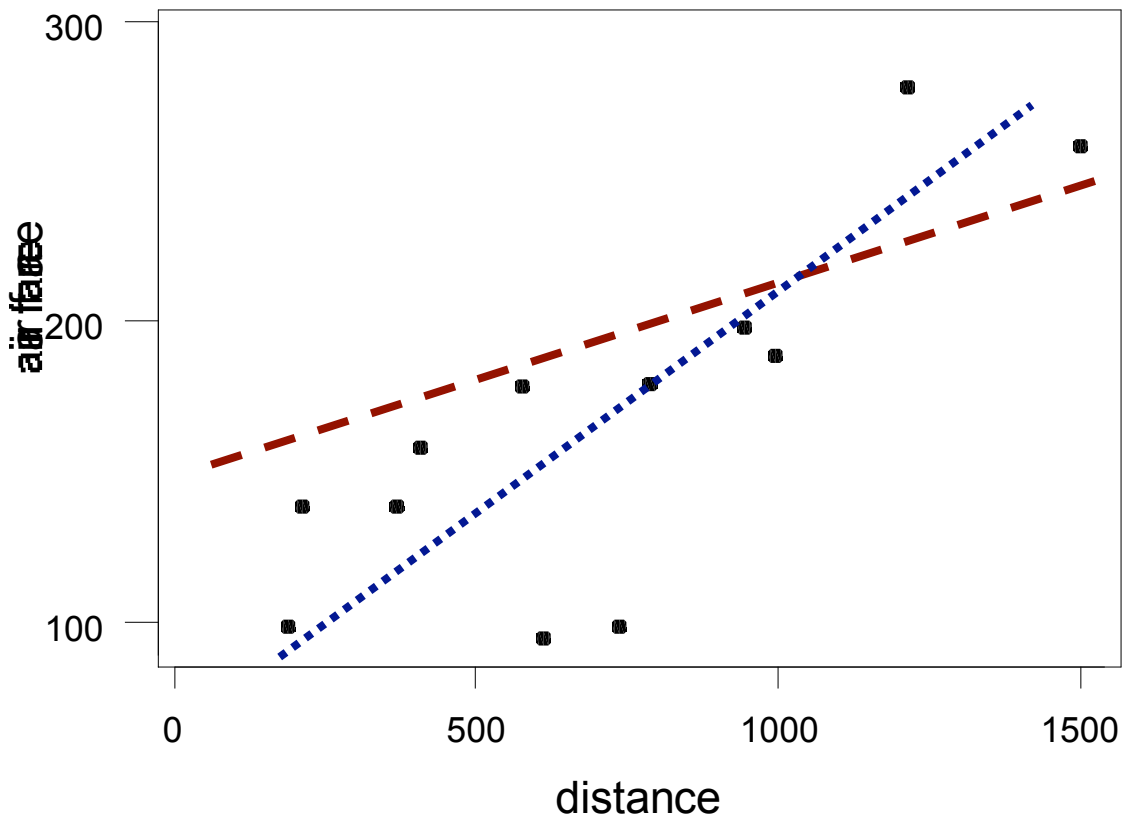
A common approach of determining which line is best is by examining the residuals (the vertical distance between the actual value and the value predicted by the line). Some of these residuals will be negative and some will be positive (which means they will cancel each other out). To take care of this problem, we commonly examine the <u>squares</u> of these residuals. Whichever line produces the smallest *sum of squared residuals* is the best-fitting line.

6. Enter the raw data into your calculator. The data and step-by-step procedures for the TI-83 are presented below.

Destination	Distance	Airfare
Atlanta	576	178
Boston	370	138
Chicago	612	94
Dallas	1216	278
Detroit	409	158
Denver	1502	258
Miami	946	198
New Orleans	998	188
New York	189	98
Orlando	787	179
Pittsburgh	210	138
St. Louis	737	98

1. Press the **STAT** button on your calculator.

2. Select **EDIT…** by pressing **ENTER**

3. Move your cursor under the **L1** column

4. Begin typing in the distance data
   576 ENTER 370 ENTER, etc.

5. When you have finished entering that data, move the cursor under the L2 column and enter in the airfare data

6. When finished, turn your calculator off.

7. We'll now find out how good your predicted line actually was by completing the following table. We could do this on our calculators, but we'll do this one by hand.

Destination	Distance	Actual Airfare	Predicted Airfare	Residual	Squared Residuals
Atlanta	576	178			
Boston	370	138			
Chicago	612	94			
Dallas	1216	278			
Detroit	409	158			
Denver	1502	258			
Miami	946	198			
New Orleans	998	188			
New York	189	98			
Orlando	787	179			
Pittsburgh	210	138			
St. Louis	737	98			
				Sum	

8. Who had the best-fitting line?

9. Now let's use our calculators to determine the actual best-fitting line. Here are the steps to follow.

    1. Press the **STAT** button
    2. Move the cursor to the right to select **CALC**
    3. Move the cursor down to **LinReg(ax+b)** and press **ENTER**
    4. Hit **ENTER** once again

You should get a screen that looks like this:

```
LinReg

y=ax+b

a=.1173750884

b=83.26735367

r² = .632001943

r = .7949854986
```

10. Write out the formula for the best-fitting regression line in slope-intercept form. Interpret the slope and y-intercept of the line.

11. The r statistic from your calculator is a *correlation coefficient*. It gives an index of how close the data points fit a straight line. Your textbook on page 44 gives a brief explanation of the correlation coefficient. We're more interested in the squared correlation coefficient, called the *coefficient of determination*.

R^2 measures the percent of variance in airfares that is explained by distance. In our situation, 63% of the reason why airfares differ can be accounted for by differences in distance.

What is the remaining 37% of variation in airfare due to?

12. Enter the following data into your calculator to find the best-fitting regression line. Write out the equation of that line along with the value of r-squared. Does it appear as though the line fits the data well?

Year	Population
1790	3.9
1810	7.2
1830	12.9
1850	23.2
1870	39.8
1890	58.6
1910	79.3
1930	122.8
1950	151.3
1970	203.3
1990	248.7

```
LinReg
y=ax+b
a = 1.210
b = -2200.703
r² = .9074
r = .9526
```

13. Create a scatterplot for this data and sketch the best-fitting line on the scatterplot. Comment on how well the line fits the data.