Enter the data into your graphing calculator.

Hours Spent Studying (x)	Test Score
2	85
3	81
4	88
5	91
6	98

LinReg
y=ax+b
a=3.6
b=74.2
r²=.784503632
r=.8857220963

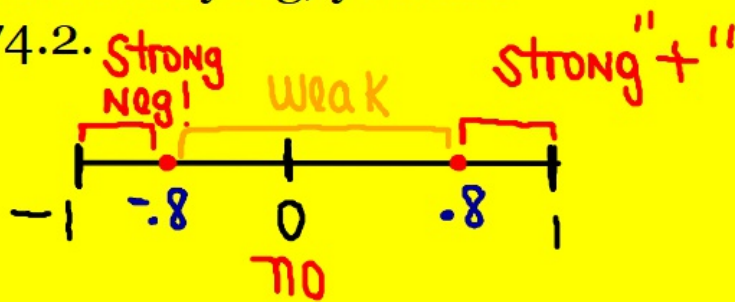Line of Best Fit
y = 3.6x + 74.2

Slope = 3.6
Y-Intercept = 74.2

For every hours spent studying, your grade is predicted to increase by 3.6 points.

If you spend zero hours studying, you are predicted to earn a 74.2.

Strong Neg!   Weak   Strong "+"

±0.8
±0.81
±0.82

-1   -.8   0   .8   1

no

Plotting Data on Scatterplot

1.) Press 2nd Y=

```
STAT PLOTS
1:Plot1…Off
    L1    L2
2:Plot2…Off
    L1    L2
3:Plot3…Off
    L1    L2
4↓PlotsOff
```
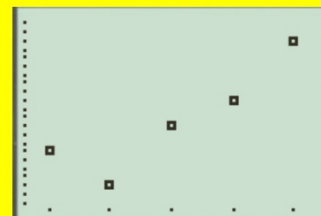
2.) Press Enter

```
Plot1  Plot2  Plot3
On  Off
Type:
Xlist:L1
Ylist:L2
Mark: ▫  +  ·
```

3.) Press Enter to turn on Plot 1

```
Plot1  Plot2  Plot3
On  Off
Type:
Xlist:L1
Ylist:L2
Mark: ▫  +  ·
```
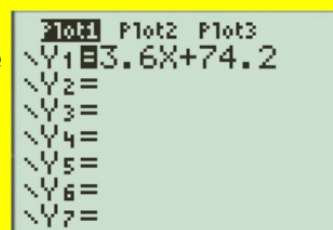
4.) Press Zoom 9: ZoomStat

Displaying the Line of Best Fit

5.) Press Y=
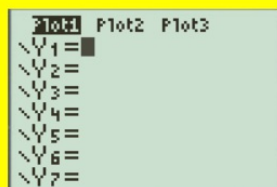
6.) Press VARS--> 5: Statistics--> EQ--> 1:RegEq--> Ente

$Y_1 = 3.6X + 74.2$

7.) Press Graph

pos. [   ] neg.
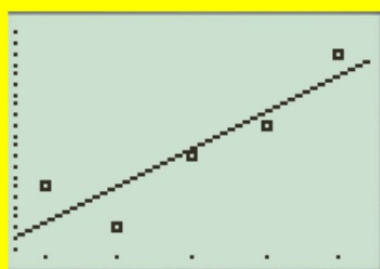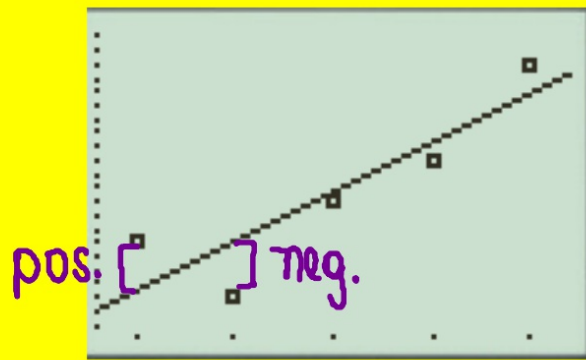
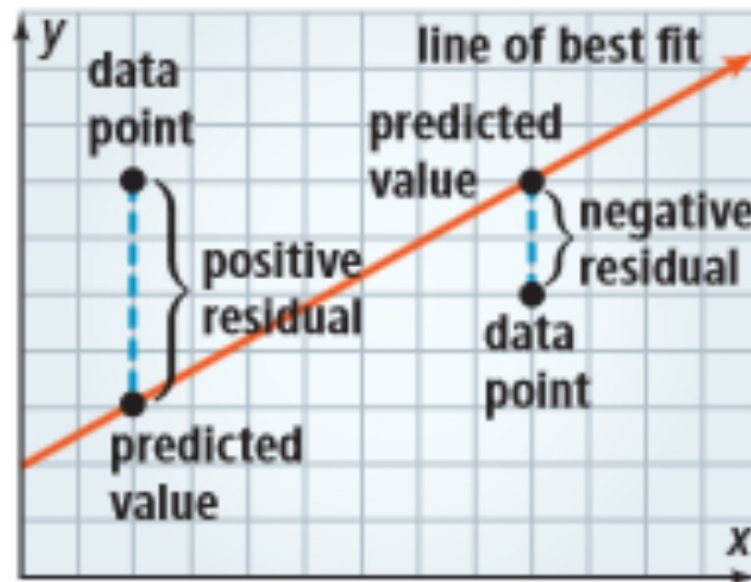Do the actual values measure perfectly to the predicated values (line of best fit)?

NO! The distance between the line of best fit and the actual values is called a residual.

Residual = actual value - predicted value

Residual - Vertical distance from actual value to predicted value.

$(2, 85)$

$y = 3.6x + 74.2$

$y = 3.6(2) + 74.2$

$y = 81.4$

$85 - 81.4 = 3.6$

actual

predicted

Residual!

Residuals have numerical values!

STAT--> EDIT--> L3

L1	L2	▓	3
2	85	------	
3	81		
4	88		
5	91		
6	98		
------	------		

L3 =

2nd--> STAT--> 7: RESID

NAMES OPS MATH
1:L1
2:L2
3:L3
4:L4
5:L5
6:L6
7:RESID

L1	L2	▓	3
2	85	------	
3	81		
4	88		
5	91		
6	98		
------	------		

L3 = ∟RESID■

Press ENTER

L1	L2	L3	3
2	85	3.6	
3	81	-4	
4	88	-.6	
5	91	-1.2	
6	98	2.2	
------	------	------	

The numbers in L3 represent the residuals for the (x_1, y_1) points in L1 and L2.

Positive Residuals v. Negative Residuals

Positive Residuals = the predicted is value less than the actual value.

For example, using the equation y = 3.6x + 74.2, when x =2
y = __81.4__.

This was 3.6 units less than the value from the table.

Negative Residuals = the predicted value is greater than the actual value.
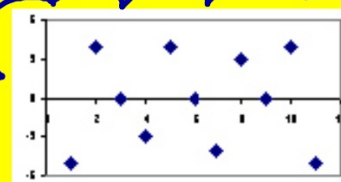
Residual Plots

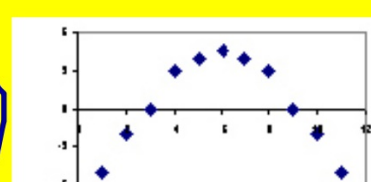A residual plot is a graph that shows the residuals on the vertical axis and the independent variable on the horizontal axis. If the points in a residual plot are randomly dispersed around the horizontal axis, a linear regression model is appropriate for the data; otherwise, a non-linear model is more appropriate.
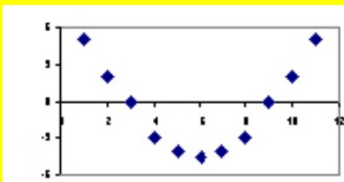
Non-random: U-shaped curve   Random pattern   Non-random: Inverted U

Ex. 1) The table shows the number of turtles that have hatched at a zoo since 2003.

Year	Turtles Hatched
0 ~~2003~~	21
1 ~~2004~~	17
2 ~~2005~~	16
3 ~~2006~~	16
4 ~~2007~~	14

A.) How many negative residuals are there? _____2_____

B.) What is the residual value for the year 2003? _____1.2_____

C.) Based on the line of best fit, the residual value for the year 2003 means that the predicted value is _____ the actual value.   (Less than or greater than)

Ex. 2) The table below shows the number of people (in millions) who voted in U.S. Presidential elections.

| Year	1980	1984	1988	1992	1996	2004
Voters	86.5	92.7	91.6	104.4	96.3	122.3

A.) How many positive residuals are there? _____

B.) What is the residual value for the year 1992? _____

C.) Based on the line of best fit, the residual value for the year 2004 means the predicted value is _____ the actual value.
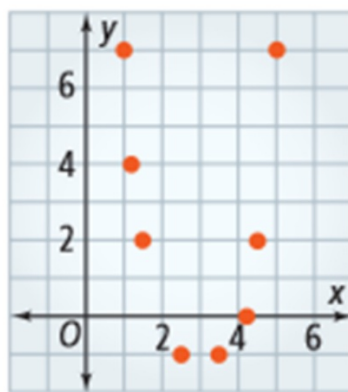
Ex. 3)

You model two data sets using linear models. The resulting residual plots are shown below. Which residual plot indicates that the linear model is a good fit for the data? Justify your answer.

Plot A
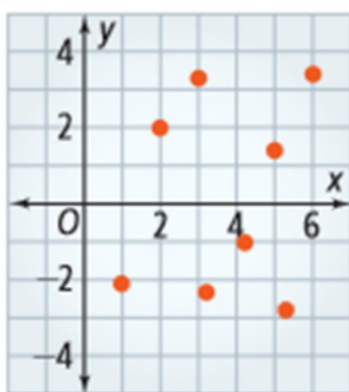
Plot B

Practice!

Complete #4, 6, and 12 on the Unit 4 Study Guide.