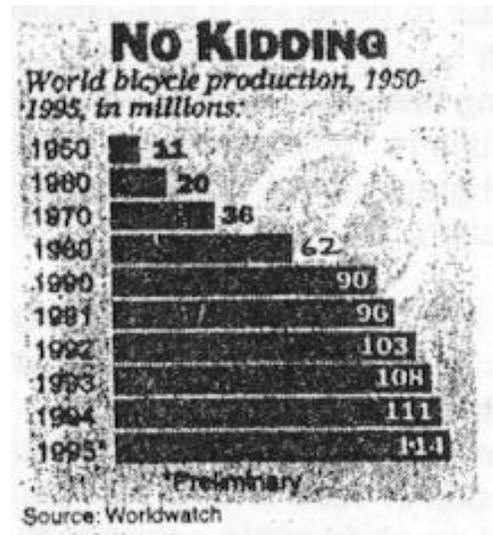


## DATA ANALYSIS: BICYCLE PRODUCTION

(Page 1 of 6)

The graph at the right appeared in the Youngstown Vindicator in September, 1996.



1. Look at the graph and understand what it says.

2. Notice that 1995's data is "preliminary", that is, it is estimated. We are going to explore how that could have been estimated.

3. Enter the data in to the STAT LISTs in your calculator. Put the years: 50, 60, 70, ..., 94 into the  $L_1$  list and the number of bicycles produced (in millions) 11, 20, 36, ..., 111 into the  $L_2$  list. NOTE: Enter the first 9 pairs only. Do not enter the data for 1995 as we will use this data as our "test" point.

4. Using an appropriate window graph these points on your graphing calculator. Record the dimensions of the window you used.

5. Using any 3 points create a system of equations to find a quadratic equation of the form  $y = Ax^2 + Bx + C$  that could model this data. Type this equation into  $y_1 =$  and graph it to see how closely it models this data.

6. Using this equation that you found and stored in  $y_1$  predict what the bicycle production in 1995 should be. Compare this answer to the "preliminary" value in the original graph from the newspaper. Compute the % error. Record all of this information.

7. a) Now use the STAT CALC Regression Equation feature on your graphing calculator to generate a quadratic regression using this data. Store the equation in  $y_2$ . Turn off  $y_1$  and graph  $y_2$  to see how closely it models this data.

b) Using this equation that you found and stored in  $y_2$  predict what the bicycle production in 1995 should be. Compare this answer to the "preliminary" value in the original graph from the newspaper. Compute the % error. Record all of this information.

8. a) Using any 4 points create a system of equations to find a cubic equation of the form  $y = Dx^3 + Ex^2 + Fx + G$  that could model this data. Turn off  $y_1$  and  $y_2$  and type this cubic equation into  $y_3$  and graph it to see how closely it models this data.

b) Using this equation that you found and stored in  $y_3$  predict what the bicycle production in 1995 should be. Compare this answer to the "preliminary" value in the original graph from the newspaper. Compute the % error. Record all of this information.

9. a) Now use the STAT CALC Regression Equation feature on your graphing calculator to generate a cubic regression using this data. Store the equation in  $y_4$ . Turn off  $y_1, y_2, y_3$  and graph  $y_4$  to see how closely it models this data.  
b) Using this equation that you found and stored in  $y_4$  predict what the bicycle production in 1995 should be. Compare this answer to the "preliminary" value in the original graph from the newspaper. Compute the % error. Record all of this information.
10. a) Using the STAT CALC Regression Equation feature find the Exponential Regression equation and put this into  $y_5$ . Graph only  $y_5$  to see how closely it models this data.  
b) Using this equation that you found and stored in  $y_5$  predict what the bicycle production in 1995 should be. Compare this answer to the "preliminary" value in the original graph from the newspaper. Compute the % error. Record all of this information.
11. a) Using the STAT CALC Regression Equation feature find the Power Regression equation and put this into  $y_6$ . Graph only  $y_6$  to see how closely it models this data.  
b) Using this equation that you found and stored in  $y_6$  predict what the bicycle production in 1995 should be. Compare this answer to the "preliminary" value in the original graph from the newspaper. Compute the % error. Record all of this information.
12. a) In your opinion, which of the 4 "calculator generated regression" equations best modeled this data? Explain why you picked this equation.  
b) Use this equation to predict the bicycle production in the year 2000. How far off would you expect this answer to be and in what direction?  
c) Use this equation to approximate the bicycle production in the years: 1955, 1972, 1986.  
d) Explain the difference between interpolation and extrapolation.
13. Write a paper summarizing what you just did and outline how you did it. Also include your impressions of what you learned.
14. *(optional)* Using the RCL feature on your calculator, create a program with this data and call it bikedata or some other appropriate name.
15. *(optional)* Using the "best" fitting equation, create and store a picture that better illustrates the graph of the data and the graph of the equation, showing appropriate units.

**BICYCLE PRODUCTION TEACHER'S SOLUTIONS AND IDEAS**

3.

L1	L2	L3	9
50	11	-----	
60	20		
70	36		
80	62		
90	90		
91	96		
92	103		

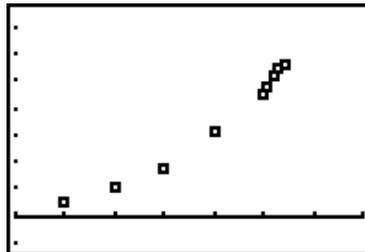
L1 = {50, 60, 70, 80...}

This is the screen you should get when entering the data.

4.

```

WINDOW
Xmin=40
Xmax=110
Xscl=10
Ymin=-20
Ymax=150
Yscl=20
Xres=█
    
```



This is the window I suggest, although you should emphasize that many windows are correct. I like to leave space below the x-axis for tracing.

$$\begin{aligned}
 5. \quad & 2500A + 50B + C = 11 \\
 & 6400A + 80B + C = 62 \\
 & 8100A + 90B + C = 90
 \end{aligned}$$

I used points 1, 4, and 5 -- probably not the best choice. These are the 3 equations in 3 variables that I generated. Next we will solve this system using "inverse of a matrix."

$$\begin{pmatrix} 2500 & 50 & 1 \\ 6400 & 80 & 1 \\ 8100 & 90 & 1 \end{pmatrix} \cdot \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 11 \\ 62 \\ 90 \end{pmatrix} \Rightarrow \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 2500 & 50 & 1 \\ 6400 & 80 & 1 \\ 8100 & 90 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 11 \\ 62 \\ 90 \end{pmatrix}$$

$$[A] \cdot [C] = [B] \Rightarrow [C] = [A]^{-1} \cdot [B] \text{ (store in [C])}$$

Now store the values for A, B, C for the equation  $y_1 = Ax^2 + Bx + C$  by either :

One way:  $.0275 \rightarrow A$   
 $-1.875 \rightarrow B$   
 $36 \rightarrow C$

**OR** More clever:  $[C](1,1) \rightarrow A$  **NOTE:** [C] is matrix C  
 $[C](2,1) \rightarrow B$   
 $[C](3,1) \rightarrow C$

$$6. \quad y_1(95) \approx 106.0625 \quad \% \text{ Error} = \frac{100(y_1(95) - 114)}{114} \approx 6.96\%$$

**NOTE:** Here are some options for this part of the exercise:

**OPTION1:** Let students see which 3 points generate the "best" prediction for 1995. You could even make it a competition.

**OPTION2:** Assign several different combinations of 3 points to different students to see which is the best predictor.

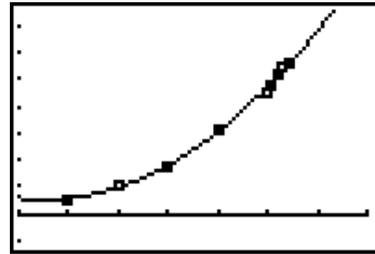
7a. STAT CALC 5  
QuadReg L1, L2, y2:

The regression equation  
was automatically stored:

This is the graph of the  
data and Reg Eq:

```
QuadReg
y=ax2+bx+c
a=.0420331637
b=-3.816904589
c=97.17004837
R2=.9975607285
```

```
Plot2 Plot3
\Y1=AX2+BX+C
\Y2=.04203316372
031X2+ -3.816904
5894063X+97.1700
4837312
\Y3=
\Y4=
```



7b.

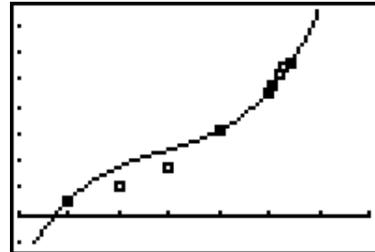
```
Y2(95)
113.913415
100(Y2(95)-114)/
114
-.0759517936
```

The percent error is -.076% -- WOW!!  
Notice that if we round to the nearest unit,  
the result we obtain is exactly what the  
newspaper's preliminary estimate was for  
1995.

8a. I used the values from 1950, 1980, 1991, 1994 and solve system of 4 equations  
in 4 unknowns and got the following. I then stored the resulting matrix into matrix C.

```
[A]-1[B]
[[.0023281596 ]
[-.4805986696]
[34.14456763 ]
[-785.751663 ]]
```

```
[C](1,1)→D
.0023281596
[C](2,1)→E
-.4805986696
[C](3,1)→F
34.14456763
[C](4,1)→G
```



As you can see, the graph doesn't look very good for interpolation, but looks very good  
for extrapolation, or data in the 1990's.

8b.

```
Y3(95)
116.6851441
100(Y3(95)-114)/
114
2.355389588
```

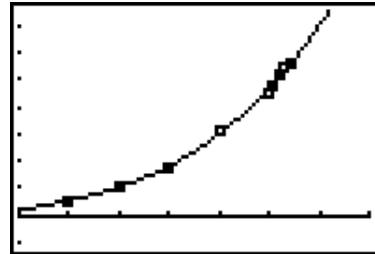
The percent error is 2.36%. Not bad.  
Point out that we should probably  
round this result to 117 because all the  
data in the original table is rounded to  
the nearest unit.

**NOTE:** Use the same options as listed for question 6 on the previous page.

9a. These are the 3 screens you should obtain on your calculator:

```
CubicReg
y=ax^3+bx^2+cx+d
a=3.84772E-4
b=-.0410146355
c=2.003649011
d=-35.01767393
R^2=.9978841326
```

```
Plot2 Plot3
\Y4=3.8477199543
091E-4X^3+-.0410
146354917X^2+2.0
03649011385X+ -35
.01767393456
\Y5=
```



9b.

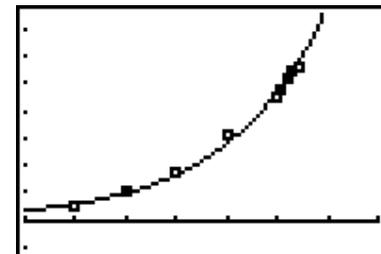
```
Y4(95)
115.0657864
100(Y4(95)-114)/
114
.9349003658
```

The percent error is .935%. Not bad. Again notice that if we round to the nearest unit, the result we obtain is exactly what the newspaper's preliminary estimate was for 1995.

10a.

```
ExpReg
y=a*b^x
a=.8981253206
b=1.052968867
r^2=.9953770826
r=.9976858637
```

```
Plot2 Plot3
146354917X^2+2.0
03649011385X+ -35
.01767393456
\Y5=.89812532058
842*1.0529688671
489^X
\Y6=
```



10b.

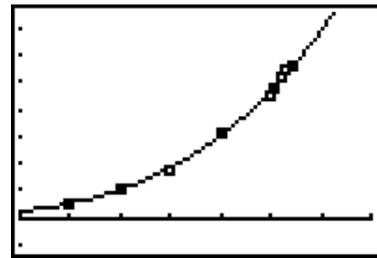
```
Y5(95)
121.0075192
100(Y5(95)-114)/
114
6.146946682
```

The percent error is 6.15% -- NOT ACCEPTABLE. Point out that some regression equations will generate good models of the data and some will not.

11a.

```
PwrReg
y=a*x^b
a=5.5036884E-6
b=3.699446159
r^2=.9988199068
r=.9994097792
```

```
Plot2 Plot3
\Y5=.89812532058
842*1.0529688671
489^X
\Y6=5.5036883716
996E-6X^3.699446
1594201
\Y7=
```



1997 REARDON DATA ANALYSIS GIFTS, INC. c:\winword\data analysis\biketec.doc 2/97

11b.

```
Y6(95)
114.0606472
100(Y6(95)-114)/
114
.0531992717
```

The percent error is .053% -- BIG WOW!!!  
Again notice that if we round to the nearest unit, the result we obtained is exactly the newspaper's preliminary estimate.

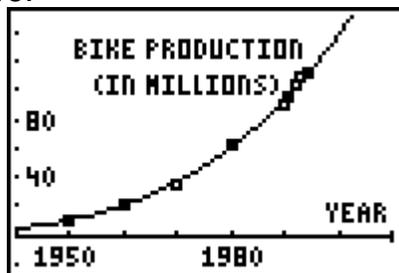
12a. Let each student pick what he/she thinks is best and argue *WHY*. I picked the Power Regression equation because it was closest, although the Quadratic Regression equation was outstanding also.

12b. Since 100 is the equivalent of the year 2000 in my data,  $y_6(100)$  was evaluated to predict the bicycle production in the year 2000 and it is: 137.89 million bikes. Since my prediction for 1995 was a little high, I would think that my prediction for the year 2000 might also be a little high. Point out that there are many other factors to consider, but this is a decent "educated guess."

13. This part of the assignment brings closure and it also has students summarize what they did. You will be surprised to hear some of their observations and interpretations of what they did.

14. See the Phone Bill Problem to see how to do store the data into a program.

15.



I think it is important to put "finishing touches" on the graph so that the units can easily be understood. It also looks much more impressive.

1997 REARDON DATA ANALYSIS GIFTS, INC. Bicycle Production 6 pages 97