

## AP Statistics – Chapter 4 Practice Free Response

- The productivity of American agriculture has grown rapidly due to improved technology (crop varieties, fertilizers, mechanization). Here are data on the output per hour of labor on American farms. The variable is an “index number” that gives productivity as a percent of the 1967 level.

<u>Year</u>	<u>Productivity</u>	<u>Year</u>	<u>Productivity</u>
1940	21	1965	91
1945	27	1970	113
1950	35	1975	137
1955	47	1980	166
1960	67	1985	217

You will now examine whether an exponential model is appropriate for the growth of productivity over time. Be sure to let  $1940 = 0$ , and all other years are set as years since 1940.

- Perform an appropriate logarithmic transformation on this data. Then do a linear regression on this transformed data and give the values of “a” and “b” for the regression.
  - Use the results of part (a) to determine the coefficients of the exponential model (A and B). Also write the exponential function that results from these coefficients.
  - Using a residual plot, determine whether an exponential model is appropriate for this data. Explain your reasoning.
- Over the past 30 years in the United States there has been a strong positive correlation between cigarette sales and the number of high school graduates.
    - Draw a diagram of the relationship and identify all variables.
    - The statement in this problem is an example (circle the correct answer):
 

causation                      common response                      confounding

- A 1969 study among the Pima Indians of Arizona investigated the relationship between a mother’s diabetic status and the appearance of birth defects in her children. The results appear in the two-way table below.

<u>Birth Defects</u>	<u>Diabetic Status</u>		
	<u>Nondiabetic</u>	<u>Prediabetic</u>	<u>Diabetic</u>
None	754	362	38
One or more	31	13	9

- Fill in the row and column totals in the margins of the table.
- Compute (in percents) the conditional distributions of birth defects for each diabetic status.
- Comment on any clear associations you see.

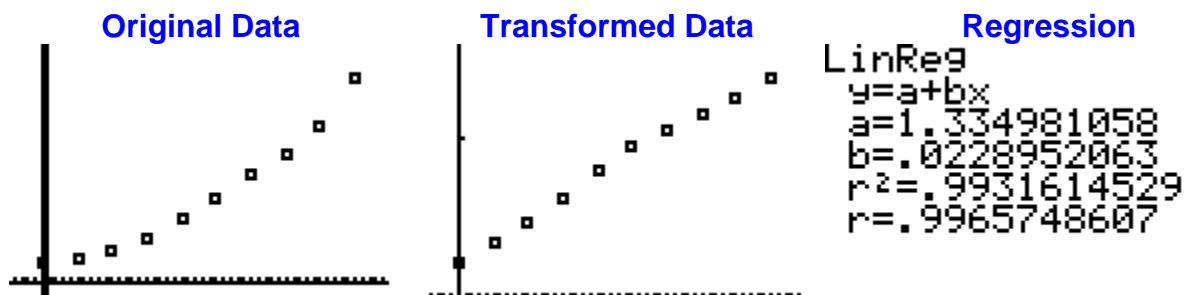
## Chapter 4 Practice Free Response - Answers

1. The productivity of American agriculture has grown rapidly due to improved technology (crop varieties, fertilizers, mechanization). Here are data on the output per hour of labor on American farms. The variable is an “index number” that gives productivity as a percent of the 1967 level.

<u>Year</u>	<u>Productivity</u>	<u>Year</u>	<u>Productivity</u>
1940	21	1965	91
1945	27	1970	113
1950	35	1975	137
1955	47	1980	166
1960	67	1985	217

You will now examine whether an exponential model is appropriate for the growth of productivity over time. Be sure to let 1940 = 0, and all other years are set as years since 1940.

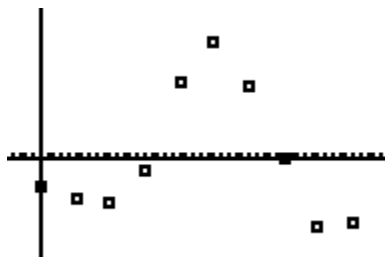
- a. Perform an appropriate logarithmic transformation on this data. Then do a linear regression on this transformed data and give the values of “a” and “b” for the regression.



- b. Use the results of part (a) to determine the coefficients of the exponential model (A and B). Also write the exponential function that results from these coefficients.

$$A = 10^a = 10^{1.335} = 21.63 \text{ and } B = 10^b = 10^{0.023} = 1.054, \text{ so } y = 21.63(1.054)^x$$

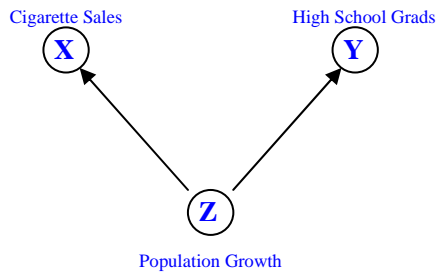
- c. Using a residual plot, determine whether an exponential model is appropriate for this data. Explain your reasoning.



**This residual plot seems to show a pattern; therefore the data does not accurately follow an exponential growth pattern.**

2. Over the past 30 years in the United States there has been a strong positive correlation between cigarette sales and the number of high school graduates.

a. Draw a diagram of the relationship and identify all variables.



b. The statement prior to #10 represents (circle the correct answer):

causation                      **common response**                      confounding

3. A 1969 study among the Pima Indians of Arizona investigated the relationship between a mother's diabetic status and the appearance of birth defects in her children. The results appear in the two-way table below.

<b>Birth Defects</b>	<b>Diabetic Status</b>		
	Nondiabetic	Prediabetic	Diabetic
None	754	362	38
One or more	31	13	9

a. Fill in the row and column totals in the margins of the table.

<b>Birth Defects</b>	<b>Diabetic Status</b>			
	Nondiabetic	Prediabetic	Diabetic	
None	754	362	38	<b>1154</b>
One or more	31	13	9	<b>53</b>
	<b>785</b>	<b>375</b>	<b>47</b>	

b. Compute (in percents) the conditional distributions of birth defects for each diabetic status.

<b>Nondiabetic</b>	<b>Prediabetic</b>	<b>Diabetic</b>
<b>31 / 785 = 4%</b>	<b>13 / 375 = 3%</b>	<b>9 / 47 = 19%</b>

c. Comment on any clear associations you see.

**Mothers who are diabetic are much more likely to have children with birth defects.**