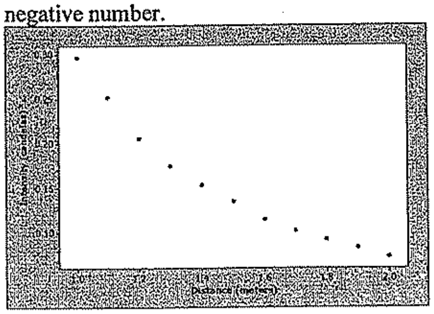
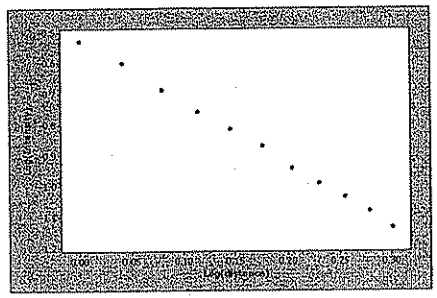
REVIEW ANSWERS CHAP 4 pp 316-319

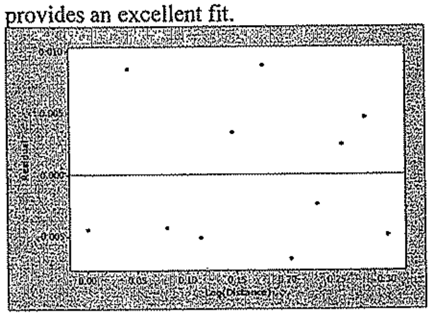
4.49 Spending more time watching TV means that less time is spent on other activities. Answers will vary, but some possible lurking variables are: the amount of time parents spend at home, the amount of exercise and the economy. For example, parents of heavy TV watchers may not spend as much time at home as other parents. Heavy TV watchers may not get as much exercise as other adolescents. As the economy has grown over the past 20 years, more families can afford TV sets (many homes now contain more than two TV sets), and as a result, TV viewing has increased and children have less physical work to do in order to make ends meet.

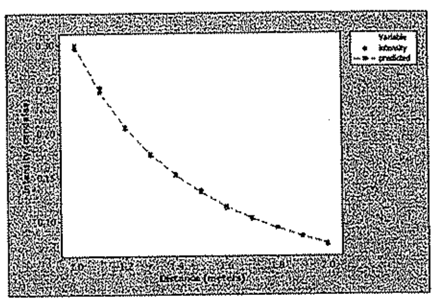
4.50 (a) Let y = intensity and x = distance. A scatterplot of the original data is shown below (left). The data appear to follow a power law model of the form $y = ax^b$ where b is some negative number.





(b) A scatterplot of the transformed data (above on the right), after taking the logarithms of both variables, shows a clear linear trend, so the power model is appropriate. The least-squares regression line for the transformed data is $\log \hat{y} = -0.5235 - 2.0126 \log x$. (c) The residual plot below shows no obvious patterns and $r^2 = 99.9\%$ so this linear model on the transformed data provides an excellent fit.





(d) Using the inverse transformation to find the predicted intensity gives $\hat{y} = 10^{-0.5235} x^{-2.0126} \doteq 0.2996x^{-2.0126}$. The plot of the original data with this model is shown above (right). (e) The predicted intensity of the 100-watt bulb at 2.1 meters is $\hat{y} = 0.2996 \times 2.1^{-2.0126} \doteq 0.0673$ candelas.

4.52 The explanatory variable is the amount of herbal tea and the response variable is a measure of health and attitude. The most important lurking variable is social interaction—many of the nursing home residents may have been lonely before the students started visiting.

4.53 (a) The column sums are shown below.

Single: 10,949 + 7,653 + 4,009 + 720 = 23,331
Married: 2,472 + 19,640 + 32,183 + 8,539 = 62,834
Widowed: 16 + 228 + 2,312 + 8,732 = 11,288
Divorced: 155 + 2,904 + 7,898 + 1,703 = 12,660

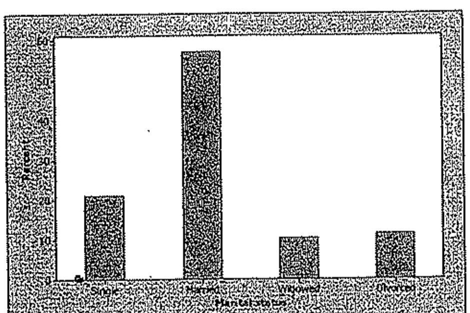
The sum of these column totals is 23,331 + 62,834 + 11,288 + 12,660 = 110,113, which is not equal to 110,115. The difference is due to rounding. (b) The marginal distributions, conditional distributions, and joint distribution are shown in the software output from Minitab below.

```
Rows: Age    Columns: Marital Status
              divorced  married  single  widowed      All

15-24              155     2472   10949       16    13592
                  1.14    18.19   80.55     0.12   100.00
                  1.22     3.93   46.93     0.14    12.34
                 0.141    2.245   9.943    0.015   12.344

25-39             2904    19640    7653      228    30425
                  9.54    64.55   25.15     0.75   100.00
                 22.94    31.26   32.80     2.02    27.63
                 2.637   17.836   6.950    0.207   27.631

40-64             7898    32183    4009     2312    46402
                 17.02    69.36    8.64     4.98   100.00
                 62.39    51.22   17.18    20.48    42.14
                 7.173   29.227   3.641    2.100   42.140

65+               1703     8539     720     8732    19694
                  8.65    43.36    3.66    44.34   100.00
                 13.45    13.59    3.09    77.36    17.89
                 1.547    7.755   0.654    7.930   17.885

All              12660    62834   23331    11288   110113
                 11.50    57.06   21.19    10.25   100.00
                100.00   100.00  100.00   100.00   100.00
                11.497   57.063  21.188   10.251  100.000

Cell Contents:          Count
                        % of Row
                        % of Column
                        % of Total
```

The table below provides just the marginal distribution for marital status.

Single	Married	Widowed	Divorced
21.19%	57.06%	10.25%	11.50%

A bar chart of the marginal distribution is shown below.

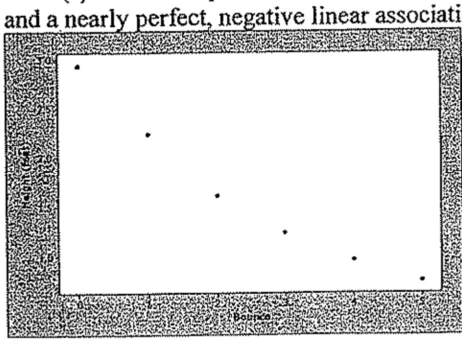


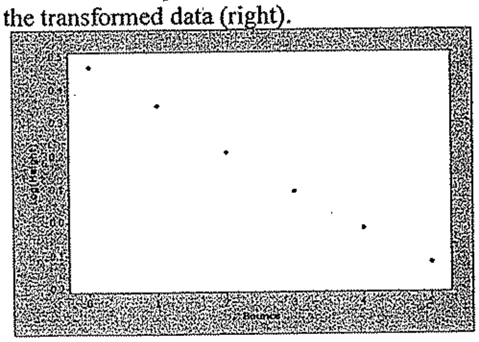
(c) The two conditional distributions are shown in the table below.

Age	Single	Married	Widowed	Divorced
15–24	80.55%	18.19%	0.12%	1.14%
40–64	8.64%	69.36%	4.98%	17.02%

Among the younger women, more than 4 out of 5 have not yet married, and those who are married have had little time to become widowed or divorced. Most of the older group is or has been married—only about 8.64% are still single. (d) Among single women, 46.93% are 15–24, 32.8% are 25–39, 17.18% are 40–64 and 3.09% are 65 or older.

4.54 (a) The scatterplots below show a strong nonlinear relationship for the original data (left) and a nearly perfect, negative linear association for the transformed data (right).

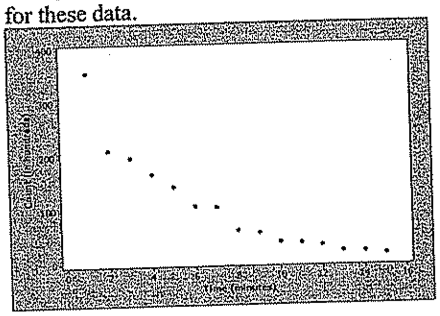


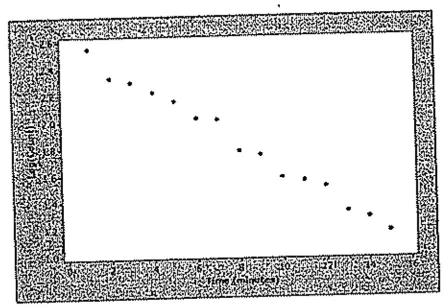


Not only is the linear association between the log(height) and bounce stronger than the linear association between the logarithms of both variables, but there is also a value of zero for the bounce number which means that the logarithm cannot be used for this point. The exponential model is more appropriate for predicting $y =$ height from $x =$ bounce number. (b) The least-squares regression line for the transformed data is $\log \hat{y} = 0.4610 - 0.1191x$. The residual plot below shows that the first two residuals are positive and the next three residuals are negative, but the residuals are all very small. The value of r^2 is 0.998, which indicates that 99.8% of the variability in log(height) is explained by linear relationship with bounce. This model provides an excellent fit.

4.57 *Who?* The individuals are cultures of marine bacteria. *What?* The two quantitative variables are $x =$ time (minutes) and $y =$ count (number of surviving bacteria in hundreds). *Why?* Researchers wanted to see if the bacteria would decay exponentially over time when exposed to X-rays. *When, where, how, and by whom?* It is not clear when or where the data were collected, but the counts were obtained after exposing cultures to X-rays for different lengths of time.

Graphs: Scatterplots below show the original data (left) and the transformed data (right) after taking the logarithm of count. Both plots suggest that the exponential decay model is appropriate for these data.





Numerical summaries: The least-squares regression line for the transformed data is $\log \hat{y} = 2.5941 - 0.0949x$. Using the inverse transformation, the predicted count is $\hat{y} = 10^{2.5941} 10^{-0.0949x} \doteq 392.7354 \times 10^{-0.0949x}$. *Interpretation:* The residual plot below shows no clear pattern and $r^2 = 98.8\%$, so the exponential decay model provides an excellent model for the number of surviving bacteria after exposure to X-rays.