Stat 104 – Homework 3
Due Thursday September 23, 2010

Reading:
September 7 – September 16 Chapter 3
September 21 – September 28 Chapter 4

Assignment:

1. Complete the following problems from the text: 3.1, 3.5, 3.11, 3.16, 3.27, 3.28, 3.33, and 3.47.
2. One definition of binge drinking is consuming 4 or more alcoholic drinks in a 2-hour period for females or 5 or more alcoholic drinks in a 2-hour period for males. Of the 257 Stat 104 students who completed the questionnaire at the beginning of the semester, 160 students indicated that they drank alcohol the previous weekend. Using the definition above, 93 participated in binge drinking. Below is a summary of the drinking status and gender of the 257 students who completed the questionnaire.

<table>
<thead>
<tr>
<th>Drinking status</th>
<th>Did not drink</th>
<th>Drank</th>
<th>Binged</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>59</td>
<td>43</td>
<td>21</td>
<td>123</td>
</tr>
<tr>
<td>Male</td>
<td>38</td>
<td>24</td>
<td>72</td>
<td>134</td>
</tr>
<tr>
<td>Total</td>
<td>97</td>
<td>67</td>
<td>93</td>
<td>257</td>
</tr>
</tbody>
</table>

a) Identify the explanatory and response variables.
b) Compute row percentages, conditional proportions for the various drinking status categories for females and males.
c) Below is a mosaic plot depicting the percentages calculated in b). Use the percentages from b) and the mosaic plot to compare females to males in terms of their drinking status.

Mosaic Plot
3. The December 2003 issue of Kiplinger’s Personal Finance published data on the 2004 model year cars and light trucks (vehicles). A random sample of 10 vehicles was taken from the data published. Below are the names of the vehicles, the EPA city mileage (mpg) ratings and the weights (tons).

<table>
<thead>
<tr>
<th>Vehicle name</th>
<th>Ford Focus</th>
<th>Chevy Malibu</th>
<th>Ford Taurus</th>
<th>Kia Optima</th>
<th>Nissan Altima</th>
<th>Toyota Camry</th>
<th>Mitsubishi Diamante</th>
<th>VW Jetta</th>
<th>Lexus ES</th>
<th>Cadillac Deville</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight, x</td>
<td>1.30</td>
<td>1.59</td>
<td>1.65</td>
<td>1.64</td>
<td>1.52</td>
<td>1.54</td>
<td>1.77</td>
<td>1.59</td>
<td>1.73</td>
<td>1.99</td>
</tr>
<tr>
<td>City mpg, y</td>
<td>26</td>
<td>24</td>
<td>20</td>
<td>20</td>
<td>21</td>
<td>24</td>
<td>18</td>
<td>21</td>
<td>20</td>
<td>18</td>
</tr>
</tbody>
</table>

a) Plot the data. Use Weight as the explanatory variable, x, and City mpg as the response, y. Describe the relationship between Weight and City mpg.
b) Compute the mean and standard deviation for the Weight. Round final answers to 3 decimal places.
c) Compute the mean and standard deviation for the City mpg. Round final answers to 3 decimal places.
d) Using \( \sum (x - \bar{x})(y - \bar{y}) = -3.674 \), show that the correlation between Weight and City mpg is \( r = -0.853 \). Explain in words what this correlation means.
e) Compute the estimate of the slope for the least squares regression line. Round final answer to 2 decimal places.
f) Give an interpretation of the estimated slope within the context of the problem.
g) Compute the estimate of the intercept for the least squares regression line. Round final answer to 2 decimal places.
h) Give the equation of the least squares regression line. Use this equation to predict the City mpg for the Taurus that weighs 1.65 tons. Give the residual for the Taurus.

4. We often read or hear reports about the relationship between diet and health. Data were collected on the fat intake (grams per day per capita) and the death rate (number of deaths per 100,000 people) from colon cancer for 30 nations. The data are available by following the link to (all) sections of Stat 104 from www.stat.iastate.edu/courses/. Follow the instructions in the JMP Guide to download/open the data set.

a) Use Analyze – Distribution to create graphical displays and numerical summaries of the death rate data. The histogram should be oriented so that death rate is along the horizontal axis. The intervals for death rate should go from 0 to 28 with an increment of 4 and no minor ticks. Use the JMP output to fully describe the distribution of death rates for the 30 nations. Be sure to attach the JMP output to your assignment.
b) Use Analyze – Fit Y by X to construct a scatter plot of death rate versus fat intake. From the red triangle pull down menu next to Bivariate Fit select Fit Line to have JMP compute the least squares regression equation relating fat intake to death rate. From the red triangle pull down menu next to Linear Fit select Plot Residuals to have JMP construct a plot of residuals versus fat intake values. Using the JMP

- Describe the scatter plot of death rate versus fat intake.
- Give the regression equation for predicting death rate from fat intake.
- Give an interpretation, within the context of the data, of the estimated slope.
- Give the value of \( R^2 \) and an interpretation of this value.
- Describe the plot of residuals versus living area and indicate what it tells you about the fit of the least squares regression line.

Be sure to attach the JMP output to your assignment.