A study was done to see if the body mass of carnivores was related to the bite force the animal produces. Twenty-five species of carnivore were randomly selected and the body mass (kg) and bite force (N) was determined for each species. The data are available in a JMP data table on the course web page:

http://www.public.iastate.edu/~wrstephe/stat401.html

Download the data table and use JMP to produce output that can be used to answer the following questions. It is not sufficient to simply hand in JMP output. You must answer the questions completely. Turn in the JMP output attached at the end of your written (word processed) solutions.

1. Use JMP's Fit Y by X to obtain a plot of Bite Force (Y) versus Body Mass (X). Describe the general relationship between the two variables.

   There appears to be a positive linear relationship between Body Mass and Bite Force. Above average values of Bite Force are associated with above average values of Bite Force. Below average values of Bite Force are associated with below average values of Bite Force.

2. Give the least squares equation of the line that relates Body Mass to Bite Force.

   Predicted Bite Force = 47.67 + 10.75*Body Mass

3. Give an interpretation of the estimated slope within the context of the problem.

   For each additional 1 kg of body mass a carnivore has, the bite force increases, on average, 10.75 N.

4. Why is there no reasonable interpretation of the estimated intercept within the context of the problem?

   There is no reasonable interpretation of the estimated intercept because a carnivore with no (zero) body mass makes no sense.

5. Give the value of $R^2$ and an interpretation of this value within the context of the problem.

   $R^2 = 0.952$. This indicates that 95.2% of the variation in bite force can be explained by the linear relationship with body mass.

6. Give the value of $s_{Y|X}$, the estimate of the error standard deviation, $\sigma$.

   $s_{Y|X} = 67.83$ N (The value of the Root Mean Square Error).
7. Test the hypothesis that the slope parameter, \( \beta_1 \), is zero against the alternative that it is not zero.

\[
H_0 : \beta_1 = 0 \\
H_A : \beta_1 \neq 0
\]

\[ t = 21.28 \]

\[ P-value < 0.0001 \]

Reject the null hypothesis because the P-value is so small.

There is a statistically significant linear relationship between body mass and bite force.

8. Construct a 95% confidence interval for the slope parameter, \( \beta_1 \). Does this agree with the results of your test in above? Explain briefly.

The 95% confidence interval for the slope parameter goes from 9.70 to 11.79 N/kg. This agrees with the results of the test of hypothesis because zero is not in the interval therefore it must be rejected as a possible value for the slope.

9. Obtain a plot of residuals versus Body Mass. Are there any unusual patterns or points? If so, comment on the pattern and/or indicate the unusual points.

There may be a sharp upward linear pattern for body masses between 0 and 25 kg. This is offset by the two large negative residuals. The carnivore with a body mass of 120 kg (\textit{Ursidae Ailuropoda melanoleuca}) may be highly influential.

10. Use Confid Curves Indiv to put prediction bands on your plot of the data. Do any of the points fall outside the prediction bands? If so, what species, body mass, bite force and predicted bite force are associated with these points?

11. Save Predicteds and Save Residuals. What is the predicted Bite Force for the species \textit{Lynx rufus}? What is the residual associated with this prediction?

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Body Mass (kg)</th>
<th>Bite Force (N)</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ursidae</td>
<td>Ursus malayanus</td>
<td>60.1</td>
<td>883.2</td>
<td>693.733632</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Body Mass (kg)</th>
<th>Bite Force (N)</th>
<th>Predicted</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Felidae</td>
<td>Lynx rufus</td>
<td>10.5</td>
<td>184.5</td>
<td>160.546599</td>
<td>23.9534011</td>
</tr>
</tbody>
</table>
12. Use Analyze + Distribution to analyze the residuals. Based on this analysis, and any plots from the previous analysis, comment on each of the conditions necessary for regression analysis.

The histogram is skewed toward larger values (skewed right). The box plot shows four species that have residuals that are potential outliers. These species, and related values, are:

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Body Mass (kg)</th>
<th>Bite Force (N)</th>
<th>Predicted Bite Force (N)</th>
<th>Residuals Bite Force (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canidae</td>
<td>Lycaon pictus</td>
<td>22.4</td>
<td>374.6</td>
<td>288.468488</td>
<td>86.1315121</td>
</tr>
<tr>
<td>Hyaenidae</td>
<td>Crocuta crocuta</td>
<td>63.1</td>
<td>565.7</td>
<td>725.982848</td>
<td>-160.28285</td>
</tr>
<tr>
<td>Mustelidae</td>
<td>Enhydra lutris</td>
<td>34.7</td>
<td>281.0</td>
<td>420.690272</td>
<td>-139.69027</td>
</tr>
<tr>
<td>Ursidae</td>
<td>Ursus malayanus</td>
<td>60.1</td>
<td>883.2</td>
<td>693.733632</td>
<td>189.466368</td>
</tr>
</tbody>
</table>

The normal quantile plot shows points that do not follow the diagonal line representing a normal model.

All of the above observations point to problems with the distribution of the random errors. The random errors are probably not normally distributed (due to the skewness in the sample). The random errors are probably not identically distributed (due to the potential outliers).

The problems with the distribution of the random errors do not affect the fit of the regression line or the fact that $R^2$ is very high. They do affect the true confidence associated with confidence intervals (and curves) and the true P-value associated with tests of hypothesis. In this situation, the test statistic was so extreme that we would still be safe in saying that there is a statistically significant linear relationship between body mass and bite force in carnivores.

Remember to turn in the output of Fit Y by X and the analysis of the residuals with the answers to the questions.