1. [30 pts] About 120 km south of Sydney, Australia the ocean creates eruptions as water passes through a hole in the cliff at Kiama. Refer to the JMP analysis of the times between eruptions of the Kiama blowhole. Times are in seconds.

(a) [4] What is the mean time between eruptions? What is the median time between eruptions? What does the comparison of these two values indicate about the shape of the distribution of times? Explain briefly.

(b) [7] Test the hypothesis that the population mean time between eruptions is 60 seconds against the alternative that it is less than 60 seconds.

(c) [4] According to the 95% confidence interval on the population mean time could the population mean be 30 minutes? Explain briefly.
(d) [8] Comment on the distribution of residuals. Be sure to mention what you see in each plot.
   - Histogram

   - Box plot

   - Normal quantile plot

(e) [3] What do your comments in (d) indicate about the conditions necessary for the statistical analysis you performed in (b) and (c)?

(f) [4] What impact will your answer in (e) have on the statistical analysis performed in (b) and (c)?
2. [20 pts] Rugby football is a popular sport played in the United Kingdom, New Zealand, Australia, South Africa, France as well as other countries. A rugby game consists of bursts of action which end when points are scored, the ball goes out of the field of play or there is an infraction of the rules. In 1992, new rules were adopted for international matches. Below is a summary of data on the length of the bursts of action (in seconds) for games played under the old rules and those under the new rules.

<table>
<thead>
<tr>
<th></th>
<th>(1) Old rules</th>
<th>(2) New rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_i )</td>
<td>97</td>
<td>104</td>
</tr>
<tr>
<td>( \bar{Y}_i )</td>
<td>21.24</td>
<td>15.09</td>
</tr>
<tr>
<td>( s_i )</td>
<td>14.946</td>
<td>11.810</td>
</tr>
<tr>
<td>( s_p = 13.415 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) [9] Test the hypothesis that the population mean time of a burst of action under the old rules is the same as that under the new rules against the alternative that the population mean times are different.

(b) [7] What are plausible values for the difference in population mean times? Use 98% confidence.
(c) [4] Looking at the sample means and sample standard deviations and noting that times must be greater than zero, what can you tell me about the skew or symmetry of the distribution of times. Explain briefly.

3. [40 pts] Data on the price, Y (in Singapore dollars) of diamond rings and the carat size, X of the diamond in the ring are collected for 48 rings. All rings are made with 20 carat gold but have diamonds of various sizes. Refer to the JMP analysis on diamond ring prices.

(a) [3] Describe the general trend between diamond size and ring price.

(b) [3] Give the least squares regression equation for predicting ring price from diamond size.

(c) [5] Give an interpretation, within the context of the problem, for the estimated slope coefficient.
(d) [6] Construct a 95% confidence interval for the population slope coefficient, $\beta_1$.

(e) [4] **Based on your confidence interval in (d),** is there a statistically significant linear relationship between diamond size and ring cost? Explain briefly.

(f) [3] How much of the variation in ring cost is **not** explained by the linear relationship with diamond size?

(g) [4] Give an interpretation, within the context of the problem, for the estimated Y-intercept.

(h) [3] How much would you expect to pay, using the least squares regression equation, for a diamond ring with a 0.25 carat diamond in it?

(j) [5] Based on the plots of residuals, is the condition of normally distributed errors satisfied? Explain briefly.

4. [10 pts] Politicians are always talking about crime and how to lower the crime rate. Some politicians say that the way to reduce crime is to spend more money on police. Data are collected on the crime rate, \( Y \) (number of crimes per 100,000 people) and the amount of money spent on police in the previous year, \( X \) (in millions of dollars). The least squares regression equation is: \( \hat{Y} = 144.5 + 8.95X \). There is a statistically significant linear relationship between the amount of money spent on police and the crime rate.

(a) [5] Give an interpretation of the slope coefficient within the context of the problem.

(b) [5] Given your interpretation in (a) and the fact that the linear relationship is statistically significant, does this contradict the belief that crime can be reduced by increasing the amount of money spent on police? Explain briefly.