Gestation and Longevity

(Based on an activity in <u>Workshop Statistics</u>, Rossman and Oehsen, 1997)

The following table lists the average longevity (in years) and gestation period (in days) for a sample of animals as reported in *The 1993 World Almanac and Books of Facts*.

Animal	Gestation	Longevity	Animal	Gestation	Longevity
Ass	365	12	Guinea pig	68	4
Baboon	187	20	Hippopotamus	238	25
Bear, black	219	18	Horse	330	20
Bear, grizzly	225	25	Kangaroo	42	7
Bear, polar	240	20	Leopard	98	12
Beaver	122	5	Lion	100	15
Buffalo	278	15	Monkey	164	15
Camel	406	12	Moose	240	12
Cat	63	12	Mouse	21	3
Chimpanzee	241	20	Opossum	15	1
Chipmunk	31	6	Pig	112	10
Cow	287	15	Puma	90	12
Deer	201	8	Rabbit	31	5
Dog	61	12	Rhinoceros	450	15
Elephant	645	40	Sea lion	350	12
Elk	250	15	Sheep	154	12
Fox	52	7	Squirrel	44	10
Giraffe	425	10	Tiger	105	16
Goat	151	8	Wolf	63	5
Gorilla	257	20	Zebra	365	15

- 1. Create a scatterplot with gestation as the response variable.
- 2. Determine the regression line for predicting an animal's gestation period from its longevity. Draw (and label) this line on your scatterplot.
- 3. Carefully describe the slope and what it means in this situation. Be specific.
- 4. What proportion of the variability in animal's gestation periods is explained by longevity?
- 5. Create a scatterplot of the animal's residual values vs. their longevities. Don't forget to label the axes.
- 6. Suppose 2 new species are discovered. They're longevities are observed to be 6 years and 15 years. What does your model predict for gestation? Which do you think is more accurate? Explain.
- 7. There is a pattern in the residual plot. Explain in a sentence or two what the pattern is and what it signifies about the accuracy of predictions for animals with long vs. short lifetimes.
- 8. The elephant is clearly an outlier. Calculate its residual value (<u>show your work</u>). How does this residual seem to compare in size to the residuals of the other animals?
- 9. The giraffe appears to have the largest residual (in absolute value). Calculate its residual (<u>show your</u> <u>work</u>). Is its gestation period longer or shorter than expected for an animal with its longevity?

- 10. Eliminate for the moment the giraffe's data and determine the regression line for predicting gestation period from longevity. Record the equation and also the r-squared value.
- 11. Is this new regression line substantially different from the original? Add this line to your original scatterplot (use a different color or a dashed line and label each line with the appropriate regression equation).
- 12. Return the giraffe's data but now eliminate the elephant's data. Determine the regression equation and the r-squared value. Add this new regression line to the original scatterplot. Is there a substantial difference from the original?
- 13. Which animal (giraffe or elephant) represents an influential observation? Explain.
- 14. What is the approximate longevity for humans? Use this value (and your original regression equation) to predict the gestation period for humans. Do you think your estimate is reasonable? Explain.