You are wondering how in Australia in 1991 the average price of a Mazda car related to the age of the car. The data set that you have available to investigate this is at

http://www.statsci.org/data/oz/mazdas.html

Answer each of the following questions for your own edification, but do not include your answers to them in your write-up (so do not turn in your answers to these questions).

- What are the individuals being measured?
- What characteristic does each variable measure, and what variable type is each variable (numerical, categorical, logical, etc.)?
- How were individuals sampled for this data set? In particular, are the observations in this data set made independently?
- From what population are individuals randomly drawn in this data set? This population tells you the scope of inference, or how widely your statistical inferences will extend.
- Which variables are random variables? What is the random process behind each random variable?

Turn in your answers to the following questions related to the above data set. You should define a new variable \( \text{Age} = 91 - \text{Year} \).

1. Based on your understanding of what the variables represent, should any of them be log transformed?
2. Keeping in mind the question of interest, make a scatterplot of the observed values of your response variable versus those of the explanatory variable (including a regression line), and explain what this plot tells you.
3. What is the “true” model equation for a linear model that will allow you to conduct statistical inference to address your question of interest? Fit this model. What is the fitted model equation? Interpret the coefficients of the fitted model equation.
4. Check the sampling variability assumptions and comment on the severity of any departures from them.
5. What is the coefficient of determination \( R^2 \) for this fitted model, and what does this tell you about the goodness of fit of the fitted model?
6. Report the results of a hypothesis test that addresses your question of interest. In doing so, state what type of test you conducted, and for each test the null and alternative hypotheses, the distribution of the test statistic under the null hypothesis, the value of the test statistic, the significance level, the \( p \)-value, and what this means in terms of statistical significance.
7. Give both a point estimate and a 95% confidence interval for the coefficient of age in the “true” model.
8. Interpret your statistical inferences from the previous two problems in terms of what they tell you about Mazda prices.

9. Use the model to give both a point estimate and a 95% confidence interval for the median selling price of a 1985 Mazda advertised in this paper in Australia in 1991.

10. Use the model to give both a point estimate and a 95% prediction interval for the selling price of a 1985 Mazda advertised in this paper in Australia in 1991.