Stat 301 Lab 8
Group Activity on Multiple Regression

Names of Group Members: _______________, _______________, _______________
________________, _______________, _______________

The laboratory 5 assignment looked at data on a random sample of 100 new vehicles for the 2004 model year. We will revisit that data in this group activity. We will look specifically at predicting Highway MPG using the Horsepower and Cylinders along with new explanatory variables constructed from Horsepower and Cylinders.

1. Fit a simple linear regression with Highway MPG as the response and Cylinders as the explanatory variable.
   a) Describe the general relationship between Highway MPG and Cylinders.
   
   b) Give the value and an interpretation of the estimated slope coefficient for Cylinders.
   
   c) Why does the estimated intercept not have an interpretation within the context of this problem?
   
   d) Is Cylinders, by itself, a statistically significant variable for explaining variation in Highway MPG? Support your answer statistically.
2. Fit a no-interaction model for the relationship between Highway MPG and the two explanatory variables Cylinders and Horsepower.
   a) Give the prediction equation for Highway MPG regressed on Cylinders and Horsepower.

   b) How much has $R^2$ changed by adding Horsepower to the model with Cylinders?

   c) Is the change in $R^2$ in b) statistically significant? Support your answer by referring to the appropriate F- and P-values.

   d) How much variation in Highway MPG is explained by Horsepower alone? Note: You should figure this out, show your work, from the output for the no-interaction model.
3. Fit an interaction model for the relationship between Highway MPG and the two explanatory variables Cylinders and Horsepower. Note: You should leave the center polynomials option turned on in Model Specification of Fit Model.
   a) Give the prediction equation for Highway MPG regressed on Cylinders, Horsepower and \((\text{Cylinders} - \text{Mean Cylinders}) \times (\text{Horsepower} - \text{Mean Horsepower})\).
   
   b) Is there significant interaction between Cylinders and Horsepower? Support your answer statistically.
   
   c) A significant interaction between Cylinders and Horsepower would tell you what about the relationship between Highway MPG and Cylinders?

4. Fit a model that has Cylinders, Horsepower, \(\text{Cylinders} \times \text{Horsepower}\) and \(\text{Cylinders}^2\). Again, be sure the center polynomials option is turned on.
   a) Is this model useful? Support your answer using the appropriate F- and P-values.
b) Does Cylinders\(^2\) add significantly to the model? Support your answer using the appropriate F- and P-values.

c) Using this model give the predicted value and residual for the Chevrolet Corvette convertible 2dr that has 350 horsepower and 8 cylinders.

d) Give a 95% prediction interval for the Chevrolet Corvette convertible 2 dr.

e) For the distribution of residuals for the model with Cylinders, Horsepower, Cylinders*Horsepower and Cylinders\(^2\) there are two potential outliers displayed on the box plot. What vehicles correspond to these two potential outliers? Give the Vehicle Name, Horsepower, Cylinders, Highway MPG, Predicted Highway MPG and Residual.