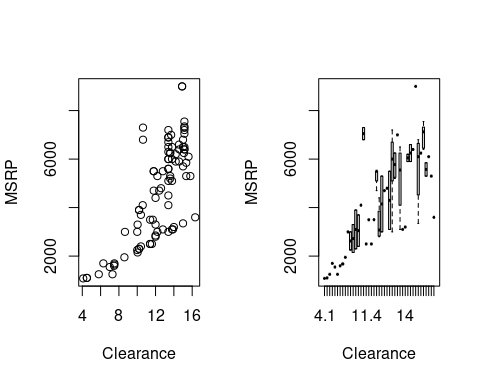
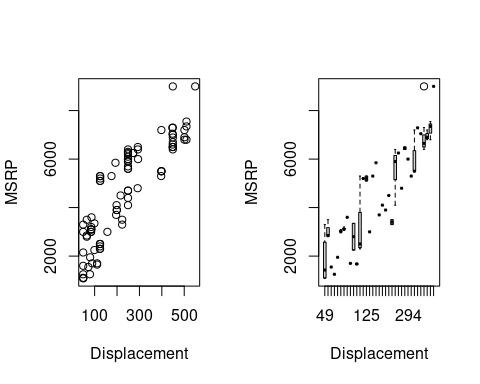
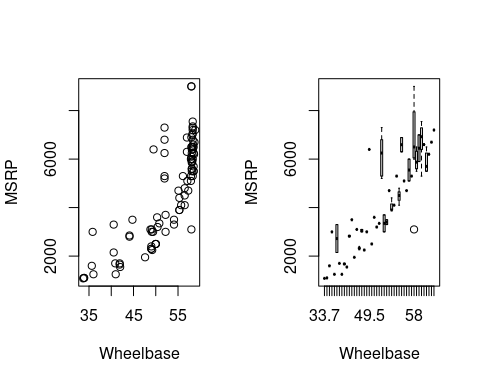
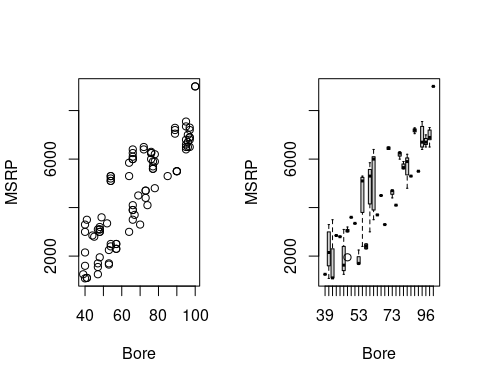
Question 1

Bore and Displacement can be used appropriately without transformation in the linear regression model, as there is a clear linear positive relationship between them and MSRP. For Clearance and Wheelbase, we see that there is a nonlinear positive relationship, as such, we have to transform them before using them in linear regression. All of them are potential predictors that can be used in the model, however, with some transformation of Clearance and Wheelbase, we can use these as predictors more accurately in the linear regression.



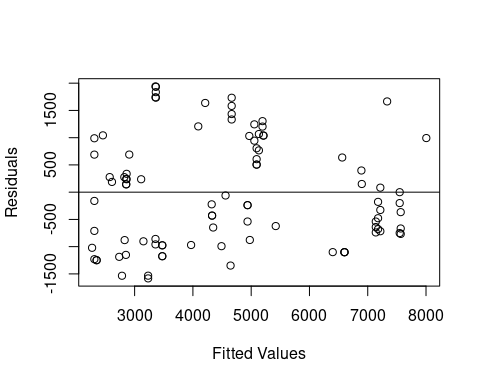
Question 2

The Fitted Model is MSRP = 423.025 + 6.722(Displacement) + 38.915(Bore). The R^2 for the model is 75.66% and the adjusted R^2 is 75.12%. This means that 75.66% of the variation in the model can be accounted for by the two variables. The Displacement coefficient means that holding all other values the same if Displacement rises by one cubic inch then the Price increases by $6.772. The Bore coefficient means that holding all other values the same if the Bore in the motorcycle rises by one inch then the Price increases by $38.915.

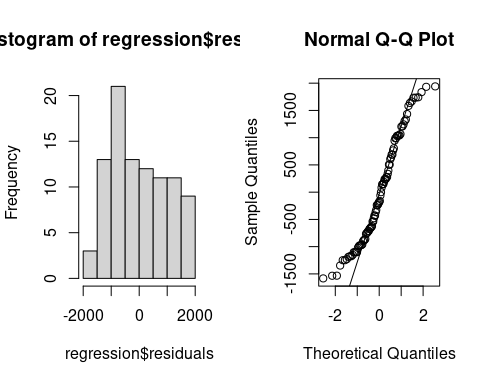
Question 3

## This model satisfies all the assumptions of a multiple linear regression model. The residual plot shows an equal spread across the line, which satisfies the Equal Variance Assumption. As there are no bends, it satisfies the Linearity Assumption. For normality Assumption as the histogram is immodial and mostly symmetrical. The QQ plot doesn’t have a lot of outliers or a lot of skewness. Also, there are 93 data points, which is a lot. All of this points to the fact that Normality Assumption is satisfied. For Independence Assumption, there is no way to be sure, but as each of the data points is independent of other data points, we can be mostly sure that Independence Assumption is satisfied.

Question 4.

This model is statistically significant. The F test is 139.9, while the p-value is <2.2e-16. Considering we use the standard a = 0.05, we can see that the F-critical value is 3.09. As the F test is greater than the F critical value, it means that the test is statistically significant. If we use the 5% significance level the Displacement is making a significant contribution to the MSRP given the other variables. This is because the p-value of Displacement is less than 0.05, while for Bore the p-value is larger than 0.05. As Displacement can reject the null theory, while Bore can’t, we can conclude that only Displacement makes a significant contribution to MSRP given other variables. 

##   
## Call:  
## lm(formula = MSRP ~ (Displacement) + (Bore), data = motor)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1582.7 -877.6 -178.2 805.6 1941.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 423.025 1036.588 0.408 0.6842   
## Displacement 6.722 3.324 2.022 0.0461 \*  
## Bore 38.915 26.221 1.484 0.1413   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 998.6 on 90 degrees of freedom  
## Multiple R-squared: 0.7566, Adjusted R-squared: 0.7512   
## F-statistic: 139.9 on 2 and 90 DF, p-value: < 2.2e-16



## Question 5

## We can add Wheelbase and Clearance as predictors in the new regression model to calculate MSRP. So the predictors in this model are Wheelbase, Clearance, Displacement, and Bore. The new Fitted model is MSRP = 7.786(Displacement) + 4.889(Bore) + 9.996(Wheelbase) + 299.343(Clearance). The new R^2 is 89.47% and the adjusted R^2 is 88.99%. The p-value is < 2.2e-16, while the F-test value is 187. All these points to the model being statistically significant and better at predicting y variables. The model satisfies all the assumptions as well. It the residual plot shows an equal spread across the line, which satisfies the Equal Variance Assumption. As there are no bends, that satisfied the Linearity Assumption. For normality Assumption as the histogram is immodial and mostly symmetrical. The QQ plot doesn’t have a lot of outliers nor a lot of skewness All of this points to the fact that the Normality Assumption is satisfied. Using a 5% significance level, we can see that Displacement and Bore make statistically significant contributions assuming all other predictors are the same because their p-value is less than 0.05. The Displacement coefficient means that, holding all other values the same, if Displacement rises by one cubic inch then the Price increases by $7.786. The Bore coefficient means that holding all other values the same if the Bore in the motorcycle rises by one inch then the Price increases by $4.889. The Wheelbase coefficient means that, holding all other values the same, if the Wheelbase rises by one inch then the Price increases by $9.996. The Clearance coefficient means that holding all other values the same, if the Clearance in the motorcycle rises by one inch then the Price increases by $299.343.

##   
## Call:  
## lm(formula = MSRP ~ (Displacement) + (Bore) + (Wheelbase) + (Clearance),   
## data = motor)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1191.70 -471.62 -68.88 514.01 1694.01   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1720.633 777.363 -2.213 0.02945 \*   
## Displacement 7.786 2.385 3.265 0.00156 \*\*   
## Bore 4.889 20.580 0.238 0.81277   
## Wheelbase 9.996 24.345 0.411 0.68236   
## Clearance 299.343 47.929 6.246 1.46e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 664.2 on 88 degrees of freedom  
## Multiple R-squared: 0.8947, Adjusted R-squared: 0.8899   
## F-statistic: 187 on 4 and 88 DF, p-value: < 2.2e-16

