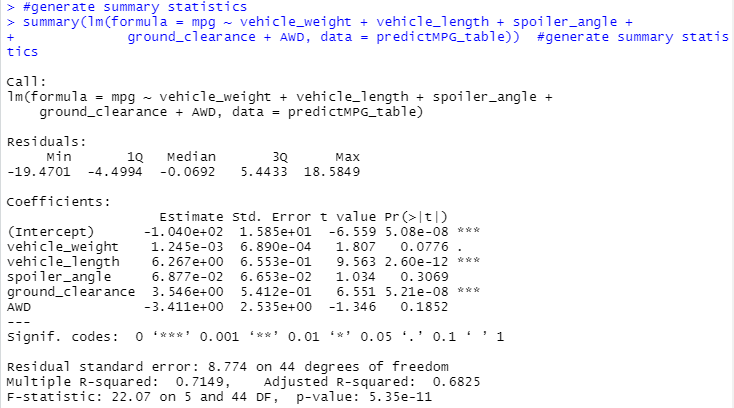
# MechaCar\_Statistical\_Analysis

## Linear Regression to Predict MPG



Which variables/coefficients provided a non-random amount of variance to the mpg values in the dataset?

Statistically, the intercept, vehicle\_length and ground\_clearance coefficients are unlikely to provide random amounts of variance to the linear model. The vehicle\_length and ground\_clearance coefficients have significant impact on mpg. The intercept is significant and there should be other variables and factors contributing to the mpg that are not included in this model.

Is the slope of the linear model considered to be zero? Why or why not?

mpg = .001 + -104 = -103.99

The slope of the linear model is not considered zero.

Does this linear model predict mpg of MechaCar prototypes effectively? Why or why not?

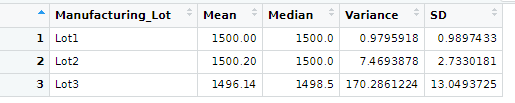
The r-squared is 0.6825 which means roughly 68% of all mpg predictions will be correct when using this multi-linear model. The p-value is 5.35e-11 which is smaller than our assumed significance level of 0.05%. Therefore, we can conclude that there is sufficient evidence to reject our null hypothesis which means the slope of linear model is not equal to zero.

## Summary Statistics on Suspension Coils

• Total Summary dataframe of the suspension coil’s PSI continuous variable across all manufacturing lots.



• Lot summary dataframe of the following PSI metrics for each lot: mean, median, variance, and standard deviation.

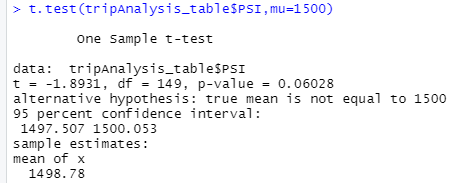


## Summary Statistics on Suspension Coils

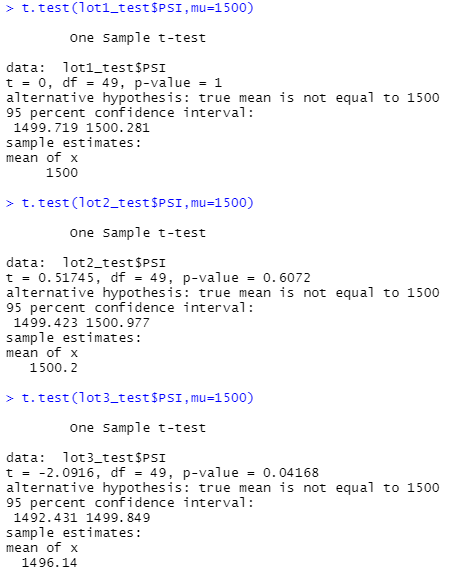
The total summary dataframe PSI metrics across all manufacturing lots show that the metrics are all within the limit and threshold. Even the variance of the suspension coils is below 100 pounds per square inch. Even the summary of each three lot in the second dataframe all metrics were within the limit except for Lot3 variance which is 170.29. Lot3 variance is above 100 pounds per square inch. The total summary dataframe show that the three lots show meeting design specification but drilling down to manufacturing data to see if each lot meet the design specification, it is clearly that Lot3 is not.

## T-Tests on Suspension Coils

An RScript using the t.test() function to determine if the PSI across all manufacturing lots is statistically different from the population mean of 1,500 pounds per square inch.



Three RScripts using the t.test() function and its subset() argument to determine if the PSI for each manufacturing lot is statistically different from the population mean of 1,500 pounds per square inch.



## T-Tests on Suspension Coils

All four t.test mean are close to 1500 except for Lot3. All the means were close in result but the p-values were all different. Only Lot3 p-value which is equal to 0.042 and less than the normal significance values of 0.05 percent.  There are sufficient evidence to reject the null hypotheses of the Lot1, Lot2 and the overall p-value across all manufacturing lots.

## Study Design: MechaCar vs Competition

This study will involve carrying out planning, designing, collecting data, analyzing, drawing meaningful interpretation and reporting of the research findings when comparing vehicles performance between MechaCar vs Competition. It will describe the relationship between variables in a sample or population and provide a summary of data or measures of central tendency in the form of mean, median and mode of cost, city or highway fuel efficiency, horse power, maintenance cost, or safety rating. The study will use random sampling t-test of data taken from a population to describe and inferences about the whole population and variable distribution. The null hypothesis will MechaCar can out perform the Competition by cost, city or highway fuel efficiency and horse power while alternative hypothesis the Competition will out perform MechaCar by cost, city or highway fuel efficiency and horse power. T-test will be used determine which hypothesis is most likely to be true. The data needed are based of vehicles with similar body style, trim levels, model year, engine size, cost/maintenance cost to name a few.