**AUTO RU’S / MECHA CAR**

**Introduction:**

We will assist Jeremy and the data analytics team with the following tasks during this project. Determine which characteristics in the dataset predict the mpg of Mecha Car prototypes using multiple linear regression analysis. Compile a summary of the suspension coils' pounds per square inch (PSI) data from the production lots. T-tests should be run to see if the manufacturing lots deviate significantly from the average population. Create a statistical analysis to compare the Mecha Car. Cars performance to that of other manufacturers' vehicles. We will produce a concise summary interpretation of the results for each statistical analysis.

**Objectives:**

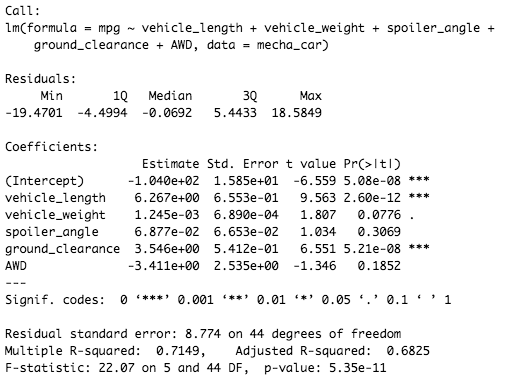
We have some important objectives to cover in this Project and those are as follow.

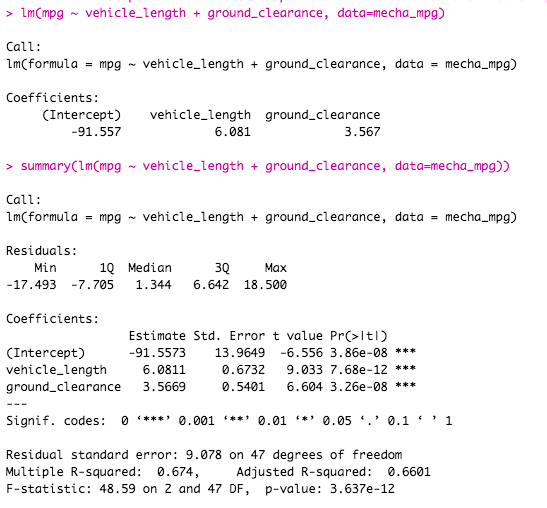
* Linear Regression to Predict MPG.
* Summary Statistics on Suspension Coils.
* T-Test on Suspension Coils.
* Design a Study Comparing the Mecha Car to the Competition.

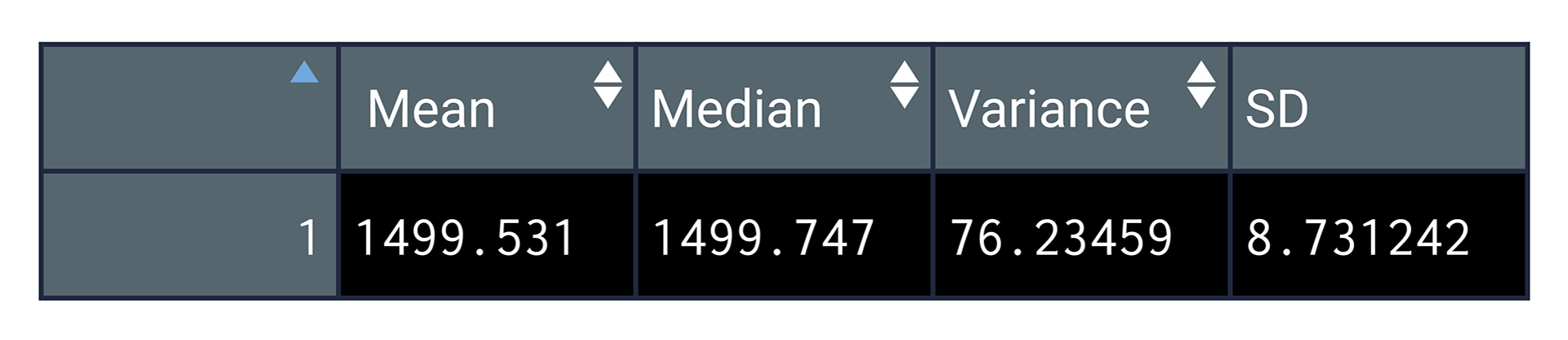
**Code:**

|  |
| --- |
| ### CHALLENGE 15: DELIVERABLE 1 |
|  |  |
|  | # 3. Use the library() function to load the dplyr package |
|  | library(dplyr) |
|  |  |
|  | #4. Import and read in the MechaCar\_mpg.csv file as a dataframe. |
|  | library(tidyverse) |
|  | mecha\_mpg <- read.csv(file='./Resources/MechaCar\_mpg.csv',check.names=F,stringsAsFactors = F) |
|  |  |
|  | #5. Perform linear regression using the lm() function |
|  | lm(mpg ~ vehicle\_length + vehicle\_weight + spoiler\_angle + ground\_clearance + AWD, data=mecha\_mpg) |
|  |  |
|  | #6. Using the summary() function, determine the p-value and the r-squared value for the linear regression model. |
|  | summary(lm(mpg ~ vehicle\_length + vehicle\_weight + spoiler\_angle + ground\_clearance + AWD, data=mecha\_mpg)) |
|  |  |
|  | #Additional Step: eliminate the independent variables that have little impact on predicting mpg to see impact: |
|  | lm(mpg ~ vehicle\_length + ground\_clearance, data=mecha\_mpg) |
|  | summary(lm(mpg ~ vehicle\_length + ground\_clearance, data=mecha\_mpg)) |
|  |  |
|  |  |
|  |  |
|  | ### CHALLENGE 15: DELIVERABLE 2 |
|  |  |
|  | #2. Import and read in the Suspension\_Coil.csv file as a table |
|  | mecha\_coil <- read.csv(file='./Resources/Suspension\_Coil.csv',check.names=F,stringsAsFactors = F) |
|  |  |
|  | #3. Create a total\_summary dataframe using the summarize() function to get the mean, median, variance, and standard deviation of the suspension coil’s PSI column. |
|  | total\_summary <- mecha\_coil %>% summarize(Mean\_PSI=mean(PSI), |
|  | Median\_PSI=median(PSI), |
|  | Var\_PSI=var(PSI), |
|  | Std\_Dev\_PSI=sd(PSI), |
|  | Num\_Coil=n(), .groups = 'keep') |
|  | #4. Create a lot\_summary dataframe using the group\_by() and the summarize() functions to group each manufacturing lot. |
|  | lot\_summary <- mecha\_coil %>% group\_by(Manufacturing\_Lot) %>% summarize(Mean\_PSI=mean(PSI), |
|  | Median\_PSI=median(PSI), |
|  | Var\_PSI=var(PSI), |
|  | Std\_Dev\_PSI=sd(PSI), |
|  | Num\_Coil=n(), .groups = 'keep') |
|  |  |
|  | #box plot: PSI Whole lot |
|  | plt1 <- ggplot(mecha\_coil,aes(y=PSI)) #import dataset into ggplot2 |
|  | plt1 + geom\_boxplot() #add boxplot |
|  |  |
|  | #box plot: PSI each indicdiual Lot |
|  | plt2 <- ggplot(mecha\_coil,aes(x=Manufacturing\_Lot,y=PSI)) #import dataset into ggplot2 |
|  | plt2 + geom\_boxplot() |
|  |  |
|  | ### CHALLENGE 15: DELIVERABLE 3 |
|  |  |
|  | #1. use t.test() to determine if the PSI across ALL lots is statistically different from the pop. mean of 1,500 PSI. |
|  | t.test(mecha\_coil$PSI,mu=1500) |
|  |  |
|  |  |
|  | #2. Use t.test() function 3 more times with subset() to determine if PSI for each manufacturing lot is statistically different from the pop. mean of 1,500 PSI |
|  | lot1 <- subset(mecha\_coil, Manufacturing\_Lot=="Lot1") |
|  | lot2 <- subset(mecha\_coil, Manufacturing\_Lot=="Lot2") |
|  | lot3 <- subset(mecha\_coil, Manufacturing\_Lot=="Lot3") |
|  |  |
|  | t.test(lot1$PSI,mu=1500) |
|  | t.test(lot2$PSI,mu=1500) |
|  | t.test(lot3$PSI,mu=1500) |
|  |  |
|  | # log10 was not used as the data was not skewed and did not need any smoothing |

**Results:**

****

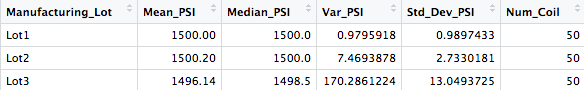




Table

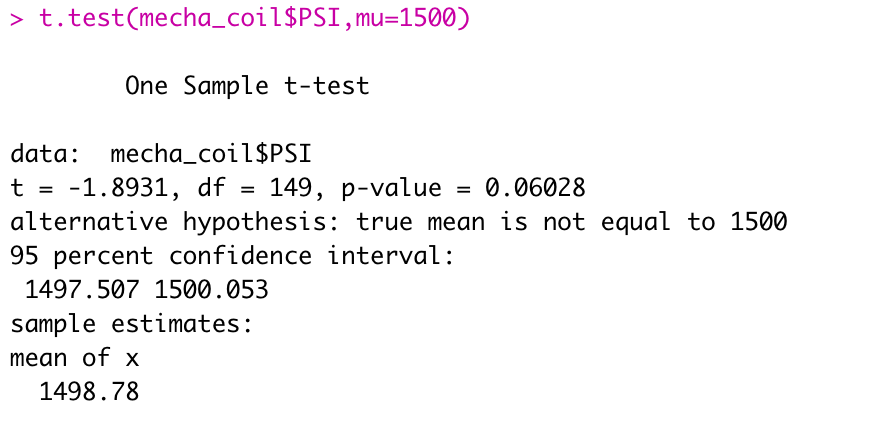
Description automatically generated





Chart, box and whisker chart

Description automatically generated



Text

Description automatically generated

**Conclusions:**

Subsequent to figuring out which variables are key for the Mecha Car's class.

**Invalid Theory (Ho):** Mecha Car is valued accurately founded on its presentation of key elements for its type.

**Elective Speculation (Ha):** Mecha Car isn't valued accurately founded on the presentation of key variables for its class.

As we have performed all the three deliverables and we have concluded through codes and Outputs. We are now Well aware of this Auto Rus Project.