# **Multiple Regression Analysis:**

After checking all individual variables, for a normal distribution, the following variables were selected to perform a multiple regression analysis on:

**MLR:**

MLR <- lm (mpg ~ vehicle.length + vehicle.weight + spoiler.angle + ground.clearance, data = mecha\_data)

After concluding that only two variables (vehicle.length and ground.clearance) were significant in providing variance to MPG, a second MLR(MLR1) was plotted using just the significant variables.

**MLR1:**

MLR1 <- lm (mpg ~ vehicle.length + ground.clearance , data = mecha\_data)

**---------------------------- MLR ------------------------------------**

Call:

lm (formula = mpg ~ vehicle.length + vehicle.weight + spoiler.angle +

ground.clearance, data = mecha\_data)

Residuals:

Min 1Q Median 3Q Max

-21.3395 -4.1155 -0.2094 6.8789 17.2672

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -1.076e+02 1.576e+01 -6.823 1.87e-08 \*\*\*

vehicle.length 6.240e+00 6.609e-01 9.441 3.05e-12 \*\*\*

vehicle.weight 1.277e-03 6.948e-04 1.837 0.0728 .

spoiler.angle 8.031e-02 6.656e-02 1.207 0.2339

ground.clearance 3.659e+00 5.394e-01 6.784 2.13e-08 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.853 on 45 degrees of freedom

Multiple R-squared: 0.7032, Adjusted R-squared: 0.6768

F-statistic: 26.65 on 4 and 45 DF, p-value: 2.277e-11

**------------------------------MLR 1 ----------------------------------**

Call:

lm(formula = mpg ~ vehicle.length + ground.clearance, data = mecha\_data)

Residuals:

Min 1Q Median 3Q Max

-17.493 -7.705 1.344 6.642 18.500

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -91.5573 13.9649 -6.556 3.86e-08 \*\*\*

vehicle.length 6.0811 0.6732 9.033 7.68e-12 \*\*\*

ground.clearance 3.5669 0.5401 6.604 3.26e-08 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.078 on 47 degrees of freedom

Multiple R-squared: 0.674, Adjusted R-squared: 0.6601

F-statistic: 48.59 on 2 and 47 DF, p-value: 3.637e-12

------------------------------------------------------------------

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

Intercept coefficient, vehicle.length and ground.clearance being less than the significance level of 0.05,provide a non-random amount of variance to the mpg values in the dataset.

Also, a statistically significant intercept (<0.05), means that there are other variables and factors, beyond the dataset (not included in the model), that contribute to the variation in mpg rating.

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

If there is no significant linear relationship, each dependent value would be determined by random chance and error. Therefore, our linear model would be a flat line with a slope of 0. But as we proved earlier that there are two variables with a significant linear relationship, with positive coefficients(vehicle.lenght,~6 and ground.clearance ~4)in both MLR and MLR1, it can be concluded that the slope does not seem to be zero.

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

Multiple R-squared: 0.7032, Adjusted R-squared: 0.6768

Although a significant y-intercept shows that some important variables have not been accounted for in this model. The r-squared (r2) value (coefficient of determination) represents how well the regression model approximates real-world data points, and In the above case we see a strong r2 value of ~.7 and adjusted r2 of ~0.68, which indicates that this model can predict mpg of MechaCar with a reasonable amount of effectiveness.

# **Coil Suspension Analysis:**

The design specifications for the MechaCar suspension coils dictate that the variance of the suspension coils must not exceed 100 pounds per inch. Does the current manufacturing data meet this design specification? Why or why not?

**----------------------Combined Summary------------------------**

Mean Median SD Variance

1 1499.531 1499.747 8.731242 76.23459

If you look at the combined summary of the overall data, the variance comes out to be ~76, which is within the 100 pounds per inch limit set by the design specifications.

**----------------------Summary By Lot----------------------------**

Manufacturing\_Lot Mean Median SD Variance

<chr> <dbl> <dbl> <dbl> <dbl>

1 Lot1 1500. 1500. 1.07 1.15

2 Lot2 1500. 1499. 3.18 10.1

3 Lot3 1499. 1498. 14.8 220.

But at a deeper glance when you calculate the summary by lot, it is clear that only Lot1 and Lot2 are within the design specifications and the process of Lot 3 needs a further inspection to understand why its variance jumps to a high of 220, which is well above the design limit of 100 pounds per inch.

# **Suspension Coil T-Test Analysis:**

**---------------------One Sample t-test----------------------------**

data: s\_coil$PSI

t = -0.65784, df = 149, p-value = 0.5117

alternative hypothesis: true mean is not equal to 1500

95 percent confidence interval:

1498.122 1500.940

sample estimates:

mean of x

1499.531

The high p-value of 0.5117 which is greater than the significance level of 0.05, means that we have failed to reject the null hypothesis. So, it is fair to conclude that sample mean is the same as the desired mean of 1500.

# **Study: MechaCar's Comparison with Competitors.**

To design a study that compares the performance of the MechaCar prototype vehicle to other comparable vehicles on the market in the current year, some of the questions we need to keep in mind are:

1- Price compared to other comparable models?

2- Fuel efficiency compared to other comparable models?

3- Color options and features available compared to other comparable models?

4- resale value compared to other comparable models?

5- spare part availability and after sales support and service, compared to other comparable models?

**Metrics:**

1- Cost of Purchase

2- Running cost (Fuel)

3- Options and Features available

4- Resale Value

5- Maintenance cost

**Null Hypothesis:**

1- There is no difference between the prices of MechaCar and other comparable models.

2- There is no difference between the Fuel efficiency of MechaCar and other comparable models.

3- There is no difference in Color options and features available compared to other comparable models.

4- There is no difference in Resale Value of MechaCar and other comparable models.

5- There is no difference in Maintenance cost of MechaCar and other comparable models.

**Alternate Hypothesis:**

1- There is a difference between the prices of MechaCar and other comparable models.

2- There is a difference between the Fuel efficiency of MechaCar and other comparable models.

3- There is a difference in Color options and features available compared to other comparable models.

4- There is a difference in Resale Value of MechaCar and other comparable models.

5- There is a difference in Maintenance cost of MechaCar and other comparable models.

**Statistical Tests:**

1- One-sample t-test to compare the prices of MechaCar compared to an average price of cars that are comparable.

2- One-sample t-test to compare the fuel efficiency of MechaCar compared to an average fuel efficiency of cars that are comparable.

3- One-sample t-test to compare the number of features of MechaCar compared to an average number of features, of cars that are comparable.

4- One-sample t-test to compare the resale value of MechaCar compared to an average resale value of cars that are comparable.

5- One-sample t-test to compare the maintenance cost of MechaCar compared to an average maintenance cost of cars that are comparable.

**Data to be Collected:**

Detailed data about price, resale value, fuel efficiency, number of features available and maintenance cost for all comparable models of the current year.

------------------------------------------------------------------------------------------------------------------------------------------