**MPG Regression**

A screenshot of a cell phone

Description automatically generated

A screenshot of text

Description automatically generated

The code above predicts the mpg of Mechacar prototypes using a number of variables within the dataset.

The result provides the coefficients of each variables and the intercept of linear regression. In addition the summary function provides the estimated p-value for each variable. As we can see, Pr(>|t|) means probability that contributes a random amount of variance to the linear model. These three factors (**Intercept**, **vehicle.length** and **ground. clearance**) are the only three(\*\* significant) that are statistically unlikely to provide random amounts of variance to the linear model. When intercept is statistically significant, it means there are other variables and factors that contribute to the variation in mpg that have not been included in the model.

A picture containing outdoor, photo, side, white

Description automatically generatedA picture containing photo, white, dog

Description automatically generated

A picture containing photo, different, white, group

Description automatically generatedA picture containing outdoor, photo, white, side

Description automatically generated

A close up of a mans face

Description automatically generated

From all the plots above, we can see both the vehicle length and ground-clearance factor has best correlation with mpg. The rest has almost zero regression value since the line was flat with almost zero slope. the relationship between some variables is statistically significant, this linear model is pretty ideal base on the multiple R-squared value of 0.7149. If we square root this value we will have ~ 0.846 of r-value. Based on the table when r is greater or equal to 0.7, the strength of correlation is strong.

**Suspension Coil Summary**

summary\_table <-sus\_coil %>% summarize (PSI\_Mean=mean(PSI),PSI\_Median=median(PSI),PSI\_SD=sd(PSI))



The summary\_table code provide was used to calculate the PSI mean, median, standard deviation and variance. The output data above was computed result. As we can see the mean is approximate 1498.78 with 1500 mean. This indicate that the dataset is in normal distribuction, since these two values were close, there is no skewed.

The design specification for the MechaCar suspension coils dictate that the variance of the suspension coils must not exceed 100 pounds per inch. The current manufacturing data meet this specification with only 62.3 PSI variance.

#### Suspension Coil T-Test

> t.test(sus\_coil\_samp$PSI,sus\_coil\_samp2$PSI)

Welch Two Sample t-test

data: sus\_coil\_samp$PSI and sus\_coil\_samp2$PSI

t = 0.29794, df = 73.573, p-value = 0.7666

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-2.389081 3.229081

sample estimates:

mean of x mean of y

1499.68 1499.26

As we runs through the one-sample t-test, the result shows the p-value was so high, that we cannot reject the null hypothesis. Therefore, we randomly select two samples (n=50) from the whole dataset.. And runs through two-sample t-test. The p-value gets even higher, it above our significance level. We do not have enough evidence to reject the null hypothesis, we would state that the two means are statistically similar.

#### Design Your Own Study

#### 

As we are trying to see which metrics of the car will be the most interest to a consumer in US market. To answer this, we have to have the dataset of the sales amount of the car.

* Interest to a consumer: the color of the cars, whether is electrical or gas or hybrid car, the price of the car, how many seats does the car have, and how fast it can run.
* The null hypothesis (example): the price of the car has no effect on amount of sales.
* Alternative hypothesis: the price of the car does impact on the amount of sales.
* If we can set up the price into an interval, and use Chi-squared to run the data type as a categorical. We can figure out is there a difference in categorical frequencies between groups.