**Project Goal :**

AutosRUs’ is the newest prototype, the MechaCar, s suffering from production troubles. The goal of this analysis is to see what design changes should a company do, in order to make changes in its production plan and solve the production troubles they are facing.

**Data Source:**

MechaCar\_mpg.csv

Suspension\_coil.csv

We will be importing data from these CSV files containing information about the various prototype design specification Mechacar is considering.

Multiple metrics such as vehicle length, vehicle weight, spoiler angle, drivetrain, and ground clearance was collected for each vehicle.

Suspension coil provide information about multiple suspension coils, the data is collected in order to see the manufacturing process is consistent across the lots.

**MPG Regression :**

Multiple the linear regression model is used to design a linear model that will predict the mpg of mechanical prototypes.

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

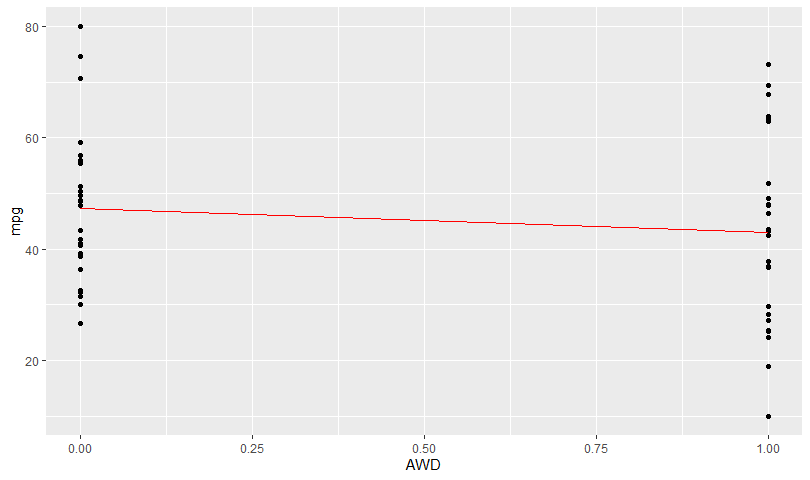
AWD and Spoiler\_analgles coefficients provided a non-random amount of variance to the mpg values in the dataset, the fact can be determined by looking at the visualization where the slope is 0

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

After running lm function on  the various variable  of the dataset we found out  the p-value is much less than 0.05, we reject the null hypothesis that β = 0. Hence there is a significant relationship between the variables in the linear regression model.

A picture containing sky, photo, outdoor

Description automatically generatedA picture containing photo, showing, sky, different

Description automatically generatedA picture containing sky, outdoor, photo

Description automatically generatedA picture containing photo, sky, showing

Description automatically generated

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

The linear the model does not predict the mpg of Mechacar prototype effectively. , we established that our regression model provides unbiased predictions of the observed values.

Even though Vehicle\_length, weight and  ground clearance p-value  shows that these are the significant variables and adjusted R -squared is nearly 0.7 which is closer to 1, it doesn’t address the precision of those predictions. Precision measures how close the predictions are to the observed values. We want the predictions to be both unbiased and close to the actual values. Predictions are precise when the observed values cluster close to the predicted values.

Regression predictions are for the mean of the dependent variable and there is variation around the mean. The same applies to the predicted mean of the dependent variable. In the fitted line plot, the regression line is nicely in the center of the data points. However, there is a spread of data points around the line. We need to quantify that spread to know how close the predictions are to the observed values. If the spread is too large, the predictions won’t provide useful information..

**Suspension Coil Summary.**

**Summary Statistics table for the suspension coil**

The design specifications for the MechaCar suspension coils dictate that the variance of the suspension coils must not exceed 100 pounds per inch.  On analyzing the  suspension **coil\_summary\_table**  we found out that the current manufacturing data meets this design specification with a variance of 62.29356.

A screenshot of a cell phone

Description automatically generated

**Welch Two Sample t-test**

A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related to certain features.  Similarly, T set using a single sample and  two sample (welch two sample test)

A picture containing wall

Description automatically generatedA group of people in a room

Description automatically generated

data:  log10(sample\_table$PSI) and log10(sample\_table2$PSI)

t = -1.3472, df = 97.896, p-value = 0.181

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

95 percent confidence interval:

 -0.0015607458  0.0002985442

sample estimates:

mean of x mean of y

 3.175500  3.176131

Based on the result, we can say: *at 95% confidence level, there is no significant difference (p-value = 0.18) of the two means*. Here we  should accept the null hypothesis that the two means are equal because the p-value is larger than 0.05 and hence We determined that the suspension coil’s pound-per-inch results are **not** statistically different from the mean population results of 1,500 pounds per inch.

Design Your own study.

From  Manufacturing Point Of View  production house should take into Consideration few KPI like yield time, cycle time, customer fill time, etc.

Some examples

**Manufacturing cycle time.** Cycle time is the total time from the beginning to the end of a process. In manufacturing, it measures the time taken for a product to pass through all machines, processes, and cycles to become a finished product. The total time an item spends in the manufacturing system between the order release and completion is the “total manufacturing cycle time”. Reducing this time can deliver reduced costs, better response to customers and increased flexibility.

**Customer fill rate, on-time delivery, perfect order percentage**: This metric is key to the order-management process and will ultimately determine your customer relationships. It shows you the percentage of your orders that are shipped in full and on time as a percentage of all your orders. In other words, it tells you the likelihood that you will effectively serve your customers. The higher your fill rate, the more likely your customers are to trust you and choose you over your competitors. It also helps show how efficient your production line is when it comes to getting the product out of the door, and how successfully you are keeping to production schedules.

      The  hypothesis to be checked :

Analysis of the product, process, and system-level factor that impact productivity in the automotive sector.

Analysis  and  determination  of the factors that  most  effectively  influence  the commonly  used productivity  measures  and  their  variation  in  terms  of  car ownership demographics.

To analyze the various factor we can use Multiple liner regression to predict dependent variable using a linear model and values from the independent variable, along with this t-test to compare the cost of various production methods which can help in reducing the time and hence helping in getting the advantage.