The MechaCar\_mpg.csv dataset contains mpg test results for 50 prototype MechaCars.

The dataset “MechaCar\_mpg.csv” has miles per gallon (mpg) test results for 50 prototype MechaCars. These prototypes were produced using multiple design specifications to identify ideal vehicle performance. Multiple metrics, such as vehicle length, vehicle weight, spoiler angle, drivetrain, and ground clearance, were collected for each vehicle. Using my knowledge of R, I designed a linear model that predicts the mpg of MechaCar prototypes using several variables from the “MechaCar\_mpg.csv” dataset.

**Statistical Summary:**

Image

From the above output we can see that:

- The **vehicle length**, and **vehicle ground clearance** are statistically unlikely to provide random variance results to the linear model. The vehicle length and vehicle ground clearance have a significant impact on the mpg the prototypes. In contrast, the spoiler angle, **vehicle weight** and**all wheel drive (**AWD) have p-Values that point to a random amount of variance within the “MechaCar\_mpg.csv” dataset.

The p-value of 5.35e-11 is much smaller than the significance level of 0.05. There is sufficient evidence that the slope of the linear model is not 0.

With an R-squared of 0.7149 there is a moderate to strong chance that this linear model is effective at predicting mpg. There is a 71% chance that the variability of mpg is explained using this linear model.

Additional Step

When we remove the vehicle weight, spoiler angle, and AWD (less impactful independent variables), the predictability does decrease, but not drastically: the r-squared value falls from 0.7149 to 0.674.

Image

D2

##Summary Statistics on Suspension Coils

The design specifications for the MechaCar suspension coils dictate that the variance of the suspension coils must not exceed 100 pounds per square inch (PSI). Based on a review of the dataframes Lot 3 exceeds 100 PSI, therefore does not meet the design specification. The reason for this are numerous outliers (refer to plot below).

image

D3

The true mean of the sample is 1498.78. With a p-Value of 0.06028, which is higher than the common significance level of 0.05, there isn’t enough evidence to support rejecting the null hypothesis. The mean of all three of these manufacturing lots is statistically similar to the presumed population mean of 1500.

Lot 1 image

Lot 1 has a true sample mean of 1500 and a p-Value of 1. We can’t reject the null hypothesis that there is no statistical difference between the observed sample mean and the presumed population mean of 1500.

Lot 2 image

Lot 2 has a true sample mean of 1500.02 and a p-Value of 0.6072. Again we can’t reject the null hypothesis.

Lot 3 image

Lot 3 has a true sample mean of 1496.14and a p-Value of 0.04168. The p-Value is lower than the common significance level of 0.05, therefore we **should reject the null hypothesis** that this sample mean and the presumed population mean are not statistically different.

.

Some appears to have happened in the production cycle of Lot 3. The lot needs to be inspected to identify the issue.