Rena Herman

**Part 1:**

Tableau:

Sheet 1 – Engine performance through time:

The top, left graph shows different model years and the horsepower and miles per gallon throughout the years. This graph can be filtered by three regions to see the data for each one – America, Asia, and Europe. The graph also shows that horsepower went down from year 79 till 82 and that MPG went up, meaning that newer cars have more miles per gallon of gas.

Sheet 2 – Car names and MPG

The top, right graph shows each car’s name and how much MPG each car holds. I only chose to present the highest ones because there were too many car names in the data.

Toyota Corolla is at the top with the highest mpg at 98.60.

Sheet 3 - Horsepower-to-Weight vs MPG:

The bottom, left graph shows the engines power to weight and how it effects the MPG – the fuel efficiency. Each origin has a different shape, making it clear as to which origin that data is talking about. USA is a circle, Asia is a plus sign, and Europe is a square. It is also colored by cylinder, to show how many are in each car. The cylinders show engine size, which shows if car is heavy or light. You can see from this graph that cars with higher horsepower to weight have lower MPG. You can also see that the fewer cylinders the car has, the more fuel efficient it is. Lastly, you can see that Asia has more fuel-efficient cars as the higher the MPG, the more plus signs there are.

\*\*For the information below, I needed to convert the Horsepower column into numeric values so it can be used to do regression. If it would not be converted into a numeric value, it got an NA and all NA values were then removed from dataset.

**Simple Linear Regression:**

Muliple r squared: 0.7714

Adjusted r squared: 0.7706

Complete linear Regression Equation: mpg=40.562−0.00629⋅weight

**Multiple Linear Regresssion:**

Muliple r squared: 0.8158

Adjusted r squared: 0.812

Complete linear Regression Equation: mpg=7.12−0.402(cylinder)+0.0026(displacement)−0.0078(horsepower)−0.0055(weight)−0.0417(acceleration)+0.459(model.year)

The residual plot shows how my predictions compare to the car’s actual reported MPG.

I calculated my predictions and then took the actual MPG and subtracted that from my predictions to get the residuals.

In the Residual Plot, the points are scattered randomly, mostly clustered around the 0 horizontal line. This shows that the model fits the data well.

The histogram helps visualize how the residuals are distributed. One can see from here if the residuals follow a pattern.

The Histogram is an approximate bell shape, which means the data is normally distributed. Most of the values are towards the middle, with fewer values as you go out from the middle. This means that this model fits the data well.

**PART 2:**

I first analyzed “How does response time impact customer satisfaction (CSAT)?" First, I made a box plot to help me analyze this question. There are three SLA options, Below SLA, Within SLA, and Above SLA. Below SLA took faster than agreed upon time, Within SLA was within the agreed upon time and Above SLA took longer than the agreed upon time. As shown on boxplot, the faster the response time, the happier the customer is, Below SLA = high CSAT. The opposite was also true, the slower the response time the more upset customers get, Above SLA = low CSAT. There are some outliers and some people who got a longer response time were still happy customers. This shows that response time is not the only factor in seeing how happy a customer is.

Next,I analyzed the question “Which reasons for calls are the most frequent?” I made a bar chart showing the different reasons and how many calls each reason had. It shows the 3 reasons that people would call – billing, payments, and service outage. Billing questions has the most calls (23,462), followed by payments (4749) and service outages (4730), which have basically the same number of calls. This is important for the company to know because then they can focus on why so many people call regarding billing and train their employees accordingly.

Thelast question I analyzed was “Which state received the most calls?” I made a bar graph to help see clearly which state got most calls. You can see from graph that California and Florida have the highest number of calls. This might be because they have high population numbers. Maine, Wyoming, and New Hampshire have lowest number of calls.