```r
#loading the data
auto_data <- read.csv(file.choose())

#View the data in a table format
View(auto_data)

#summary of the data
summary(auto_data)

#structure of the data
str(auto_data)

#removes any missing values
auto_data <- na.omit(auto_data[, c('mpg', 'horsepower', 'weight', 'acceleration',
'displacement')])

#checks to see if any missing values
sum(is.na(auto_data))

#converting data type
auto_data$horsepower <- as.numeric(auto_data$horsepower)

#checks if any na values were created
sum(is.na(auto_data$horsepower))

#splitting the data
train_data <- auto_data[1:300, ]
test_data <- auto_data[301:398, ]

#run simple linear regression
model <- lm(mpg ~ weight, data = train_data)

#summary of model
summary(model)

#creates a scatter plot
plot(train_data$weight, train_data$mpg,
     xlab = "Weight",
     ylab = "MPG",
     main = "Scatter Plot of Weight vs Mpg")

#fitting a regression line to the scatter plot

abline(lm(train_data$mpg ~ train_data$weight), col ="red")

#multiple linear regression model
model2 <- lm(mpg ~ horsepower + weight + acceleration + displacement, data = train_data)

#summary of the model
summary(model2)

#linear regression using test data
model3 <- lm(mpg ~ horsepower + weight + acceleration + displacement, data = test_data)

#model summary
summary(model3)

#predict mpg for test data
predictions <- predict(model3, test_data)

#compares the actual vs predicted
comparison <- data.frame(Actual = test_data$mpg, Predicted = predictions)
residuals <- comparison$Actual - comparison$Predicted
```

```
#prints the residuals in a data frame
print(comparison)



#plot residual chart
plot(predictions, residuals,
     main = "Residual Plot",
     xlab = "Predictied mpg",
     ylab = " Residuals",
    pch = 19, col = 'red'
    )
abline( h= 0, col = 'blue', lwd = 2)



#histogram
hist(residuals,
     main = "Histogram of Residuals",
     xlab = "Residuals",
     col = "lightblue",
     breaks = 20)

#comparison
comparison <- data.frame(Actual_MPG = actual_mpg, Predicted_MPG = predictions, Residuals =
residuals)
head(comparison)
```