2.9

Sai Ananthula - emw832

```
Auto <- read.csv("Auto.csv", header = T, na.string = "?", stringsAsFactors =  T)
head(Auto)
```

The predictors that are quantitative are Cylinder, displacement, horsepower, weight, and acceleration. The only qualitative predictor is year since it is a category.

Ranges for Predictors

```
range(Auto$mpg)
```

```
[1]  9.0 46.6
```

```
range(Auto$cylinders)
```

```
[1] 3 8
```

```
range(Auto$displacement)
```

```
[1]  68 455
```

```
range(Auto$horsepower)
```

```
[1]  46 230
```

```
range(Auto$weight)
```

```
[1] 1613 5140
```

```
range(Auto$acceleration)
```

```
[1]  8.0 24.8
```

Standard Deviation of Predictors

```
sd(Auto$mpg)
```

```
[1] 7.805007
```

```
  sd(Auto$cylinders)
```

```
[1] 1.705783
```

```
  sd(Auto$displacement)
```

```
[1] 104.644
```

```
  sd(Auto$horsepower)
```

```
[1] 38.49116
```

```
  sd(Auto$weight)
```

```
[1] 849.4026
```

```
  sd(Auto$acceleration)
```

```
[1] 2.758864
```

Mean of Predictors

```
  mean(Auto$mpg)
```

```
[1] 23.44592
```

```
  mean(Auto$cylinders)
```

```
[1] 5.471939
```

```
  mean(Auto$displacement)
```

```
[1] 194.412
```

```
  mean(Auto$horsepower)
```

```
[1] 104.4694
```

```
  mean(Auto$weight)
```

```
[1] 2977.584
```

```
  mean(Auto$acceleration)
```

```
[1] 15.54133
```

Removing rows 10 - 85

```
  Auto <- Auto[-(10:85),]
```

Range of Predictors post Row Removal

```
  range(Auto$mpg)
```

```
[1] 11.0 46.6
```

```
  range(Auto$cylinders)
```

```
[1] 3 8
```

```
  range(Auto$displacement)
```

```
[1]  68 455
```

```
  range(Auto$horsepower)
```

```
[1]  46 230
```

```
  range(Auto$weight)
```

```
[1] 1649 4997
```

```
  range(Auto$acceleration)
```

```
[1]  8.5 24.8
```

Standard Deviation of Predictors post Row Removal

```
  sd(Auto$mpg)
```

```
[1] 7.867283
```

```
  sd(Auto$cylinders)
```

```
[1] 1.654179
```

```
  sd(Auto$displacement)
```

```
[1] 99.67837
```

```
  sd(Auto$horsepower)
```

```
[1] 35.70885
```

```
  sd(Auto$weight)
```

```
[1] 811.3002
```

```
  sd(Auto$acceleration)
```

```
[1] 2.693721
```

Mean of Predictors post Row Removal

```
  mean(Auto$mpg)
```

```
[1] 24.40443
```

```
  mean(Auto$cylinders)
```

```
[1] 5.373418
```

```
  mean(Auto$displacement)
```

```
[1] 187.2405
```

```
  mean(Auto$horsepower)
```

```
[1] 100.7215
```

```
  mean(Auto$weight)
```

```
[1] 2935.972
```

```
  mean(Auto$acceleration)
```

```
[1] 15.7269
```

Scatter plots looking at the relationship between MPG and certain key predictors

```
  plot(Auto$weight,Auto$mpg)
```

```
  plot(Auto$displacement,Auto$mpg)
```

```
  plot(Auto$horsepower,Auto$mpg)
```

It seems weight has the strongest relationship with MPG since as weight increase the mpg decrease. Increased displacement and horsepower are also related to lower MPG. All three of these predictors are more common on automobiles that perform less efficiently so its not much of a surprise that they correlate negatively with MPG.

Influence of year of make on MPG

```
plot(Auto$year,Auto$mpg)
```

It seems the newer cars have better has mileage which is no surprise but there is potential for deeper analysis.

Other variables that would help is if each observation had the manufacturer and type of car. It would allow for deeper analysis since a sedan 9/10 times has better MPG than a pick up or SUV. Also, if you wanted to drill down deeper listing the model and trim of each observation to make some very precise insights since two trims of one model can have different gas mileage.