Here in this dataset I will investigate the impact of a number of automobile engine factors on the vehicle’s mpg.

In every model I will make the Linear Regression model and normal curve.

1. Mpg and cylinder: Here r2= 60.12%, correlation is negative and strong relationship between mpg and cylinder.

cor(autompg$cylinder,autompg$mpg)

[1] -0.7753963

Residual standard error: 4.942 on 396 degrees of freedom

Multiple R-squared: 0.6012, Adjusted R-squared: 0.6002

F-statistic: 597.1 on 1 and 396 DF, p-value: < 2.2e-16

1. Mpg and displacement: Here r2= 64.6%, correlation is negative and strong relationship between mpg and displacement.

cor(autompg$displacement,autompg$mpg)

[1] -0.8042028

Residual standard error: 4.651 on 396 degrees of freedom

R-squared: 0.6467, Adjusted R-squared: 0.6459

F-statistic: 725 on 1 and 396 DF, p-value: < 2.2e-16

1. Mpg and weight: Here r2= 69.1%, correlation is negative and very strong relationship between mpg and weight.

cor(autompg$weight,autompg$mpg)

[1] -0.8317409

Residual standard error: 4.345 on 396 degrees of freedom

Multiple R-squared: 0.6918, Adjusted R-squared: 0.691

F-statistic: 888.9 on 1 and 396 DF, p-value: < 2.2e-16

1. Mpg and acceleration: Here r2= 17.66%, correlation is positive and moderate strong relationship between mpg and acceleration.

cor(autompg$acceleration,autompg$mpg)

[1] 0.4202889

Residual standard error: 7.101 on 396 degrees of freedom

Multiple R-squared: 0.1766, Adjusted R-squared: 0.1746

F-statistic: 84.96 on 1 and 396 DF, p-value: < 2.2e-16

1. Mpg and model. Year: Here r2=33.56%, correlation is positive and moderate strong relationship between mpg and model.year

cor(autompg$model.year,autompg$mpg)

[1] 0.5792671

Residual standard error: 6.379 on 396 degrees of freedom

Multiple R-squared: 0.3356, Adjusted R-squared: 0.3339

F-statistic: 200 on 1 and 396 DF, p-value: < 2.2e-16

1. 6.Mpg and origin: Here r2= 31.75%, correlation is positive and moderate strong relationship between mpg and origin.

cor(autompg$origin,autompg$mpg)

[1] 0.5634504

Residual standard error: 6.465 on 396 degrees of freedom

Multiple R-squared: 0.3175, Adjusted R-squared: 0.3158

F-statistic: 184.2 on 1 and 396 DF, p-value: < 2.2e-16

1. 7.Mpg and horsepower: Here r2= 20.27%, correlation is positive and moderate strong relationship between mpg and horsepower.

cor(autompg\_1$horsepower,autompg\_1$mpg)

[1] 0.4502621

Residual standard error: 6.978 on 390 degrees of freedom

Multiple R-squared: 0.2027, Adjusted R-squared: 0.2007

F-statistic: 99.17 on 1 and 390 DF, p-value: < 2.2e-16

1. 8.Mpg and car.name: Here r2= 7.505%, correlation is positive and very weak relationship between mpg and car name.

cor(autompg$car.name,autompg$mpg)

[1] 0.2739516

Residual standard error: 7.526 on 396 degrees of freedom

Multiple R-squared: 0.07505, Adjusted R-squared: 0.07271

F-statistic: 32.13 on 1 and 396 DF, p-value: 2.78e-08

So after analysis best model is:

* Mpg and weight
* Mpg and displacement
* Mpg and cylinder