DATA 603 Porject - Group 6

Group 6

2025-03-28

```
boston_data= read.csv("./BostonHousing.csv")
head(boston_data)
```

```
##      crim zn indus chas   nox    rm  age    dis rad tax ptratio      b lstat
## 1 0.00632 18  2.31    0 0.538 6.575 65.2 4.0900   1 296    15.3 396.90  4.98
## 2 0.02731  0  7.07    0 0.469 6.421 78.9 4.9671   2 242    17.8 396.90  9.14
## 3 0.02729  0  7.07    0 0.469 7.185 61.1 4.9671   2 242    17.8 392.83  4.03
## 4 0.03237  0  2.18    0 0.458 6.998 45.8 6.0622   3 222    18.7 394.63  2.94
## 5 0.06905  0  2.18    0 0.458 7.147 54.2 6.0622   3 222    18.7 396.90  5.33
## 6 0.02985  0  2.18    0 0.458 6.430 58.7 6.0622   3 222    18.7 394.12  5.21
##   medv
## 1 24.0
## 2 21.6
## 3 34.7
## 4 33.4
## 5 36.2
## 6 28.7
```

```
boston_additive = lm(medv~crim+zn+indus+chas+nox+rm+age+dis+rad+tax+ptratio+b+lstat, data= boston_data)
summary(boston_additive)
```

```
##
## Call:
## lm(formula = medv ~ crim + zn + indus + chas + nox + rm + age +
##     dis + rad + tax + ptratio + b + lstat, data = boston_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -15.595  -2.730  -0.518   1.777  26.199
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.646e+01  5.103e+00   7.144 3.28e-12 ***
## crim        -1.080e-01  3.286e-02  -3.287 0.001087 **
## zn           4.642e-02  1.373e-02   3.382 0.000778 ***
## indus        2.056e-02  6.150e-02   0.334 0.738288
## chas         2.687e+00  8.616e-01   3.118 0.001925 **
## nox         -1.777e+01  3.820e+00  -4.651 4.25e-06 ***
## rm           3.810e+00  4.179e-01   9.116  < 2e-16 ***
## age          6.922e-04  1.321e-02   0.052 0.958229
## dis         -1.476e+00  1.995e-01  -7.398 6.01e-13 ***
```

```
## rad          3.060e-01  6.635e-02   4.613 5.07e-06 ***
## tax         -1.233e-02  3.760e-03  -3.280 0.001112 **
## ptratio     -9.527e-01  1.308e-01  -7.283 1.31e-12 ***
## b            9.312e-03  2.686e-03   3.467 0.000573 ***
## lstat       -5.248e-01  5.072e-02 -10.347  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.745 on 492 degrees of freedom
## Multiple R-squared:  0.7406, Adjusted R-squared:  0.7338
## F-statistic: 108.1 on 13 and 492 DF,  p-value: < 2.2e-16
```

Individual T-test:

Hypothesis:

Based off the values indus and age do not seem relevent.

We will verify utilizing a globabal anova f-test:

```
boston_additive_dropped = lm(medv~crim+zn+chas+nox+rm+dis+rad+tax+ptratio+b+lstat, data= boston_data)
anova(boston_additive, boston_additive_dropped)
```

```
## Analysis of Variance Table
##
## Model 1: medv ~ crim + zn + indus + chas + nox + rm + age + dis + rad +
##     tax + ptratio + b + lstat
## Model 2: medv ~ crim + zn + chas + nox + rm + dis + rad + tax + ptratio +
##     b + lstat
##   Res.Df   RSS Df Sum of Sq      F Pr(>F)
## 1    492 11079
## 2    494 11081 -2   -2.5794 0.0573 0.9443
```

Fail to reject null, they are not relevent.

Comparison Of R^2 and RSE

Chas is a factor.

Final Additive Model:

```
confint(boston_additive_dropped)
```

```
##                    2.5 %       97.5 %
## (Intercept)  26.384649126  46.29764088
## crim         -0.172817670  -0.04400902
## zn            0.019275889   0.07241397
## chas          1.040324913   4.39710769
## nox         -24.321990312 -10.43005655
## rm            3.003258393   4.59989929
## dis          -1.857631161  -1.12779176
## rad           0.175037411   0.42417950
## tax          -0.018403857  -0.00515209
## ptratio      -1.200109823  -0.69293932
## b             0.004037216   0.01454447
## lstat        -0.615731781  -0.42937513
```