

Problem Set 2

Seth Harrison

02/03/2020

1. (10 points) Estimate the MSE of the model using the traditional approach. That is, fit the linear regression model using the *entire* dataset and calculate the mean squared error for the *entire* dataset. Present and discuss your results at a simple, high level.

```
# Fit Linear Model
```

```
Model1 <- lm(nes$biden~nes$female+nes$age+nes$educ+nes$dem+nes$rep)
summary(Model1)
```

```
##
## Call:
## lm(formula = nes$biden ~ nes$female + nes$age + nes$educ + nes$dem +
##     nes$rep)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -75.546 -11.295   1.018  12.776  53.977
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  58.81126    3.12444  18.823  < 2e-16 ***
## nes$female    4.10323    0.94823   4.327 1.59e-05 ***
## nes$age        0.04826    0.02825   1.708  0.0877 .
## nes$educ      -0.34533    0.19478  -1.773  0.0764 .
## nes$dem       15.42426    1.06803  14.442  < 2e-16 ***
## nes$rep      -15.84951    1.31136 -12.086  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 19.91 on 1801 degrees of freedom
## Multiple R-squared:  0.2815, Adjusted R-squared:  0.2795
## F-statistic: 141.1 on 5 and 1801 DF,  p-value: < 2.2e-16
```

```
# Calculate MSE
```

```
mse <- mean(Model1$residuals^2)
mse
```

```
## [1] 395.2702
```