In class Week 5

Will Doyle

2/5/2019

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------- tidyverse 1.2.1 --
## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  2.0.1     v dplyr   0.7.8
## v tidyr   0.8.2     v stringr 1.3.1
## v readr   1.3.1     v forcats 0.3.0
## Warning: package 'tibble' was built under R version 3.5.2
## -- Conflicts ------------------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(modelr)
```

```
load("els_train.Rdata")
load("els_test.Rdata")
```

Run model against training dataset

Create model, test against testing dataset.

```
mod1<-lm(bynels2r~byses1,data=els_train)
summary(mod1)
```

```
##
## Call:
## lm(formula = bynels2r ~ byses1, data = els_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.4572  -6.1499   0.2996   6.4041  24.3580
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 29.39203    0.09808  299.68   <2e-16 ***
## byses1       5.51700    0.13176   41.87   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.562 on 7640 degrees of freedom
##   (502 observations deleted due to missingness)
## Multiple R-squared:  0.1867, Adjusted R-squared:  0.1866
## F-statistic:  1753 on 1 and 7640 DF,  p-value: < 2.2e-16
```

```
## Calculate root mean squared error
rmse_1<-modelr::rmse(mod1,els_test);rmse_1
```

```
## [1] 8.538205
```

```
mod2<-lm(bynels2r~
           byses1+
           bynels2m+
           byincome+
           female
         ,data=els_train)
summary(mod2)
```

```
##
## Call:
## lm(formula = bynels2r ~ byses1 + bynels2m + byincome + female,
##     data = els_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -31.6052  -3.8478   0.1042   4.1257  21.3352
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.320876   0.437439  12.164  < 2e-16 ***
## byses1      1.309229   0.136658   9.580  < 2e-16 ***
## bynels2m    0.491268   0.005627  87.312  < 2e-16 ***
## byincome    0.105392   0.040088   2.629  0.00858 **
## female      2.026642   0.138251  14.659  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.019 on 7599 degrees of freedom
##   (540 observations deleted due to missingness)
## Multiple R-squared:  0.5988, Adjusted R-squared:  0.5986
## F-statistic:  2835 on 4 and 7599 DF,  p-value: < 2.2e-16
```

```
## Calculate root mean squared error
rmse_2<-modelr::rmse(mod2,els_test);rmse_2
```

```
## [1] 6.160667
```