The Effect of Beauty and Other Characteristics on Professor Evaluation Ratings

Katelyn Barbre, Samuel Shaud, Isaac Stevens, Melanie Sattler

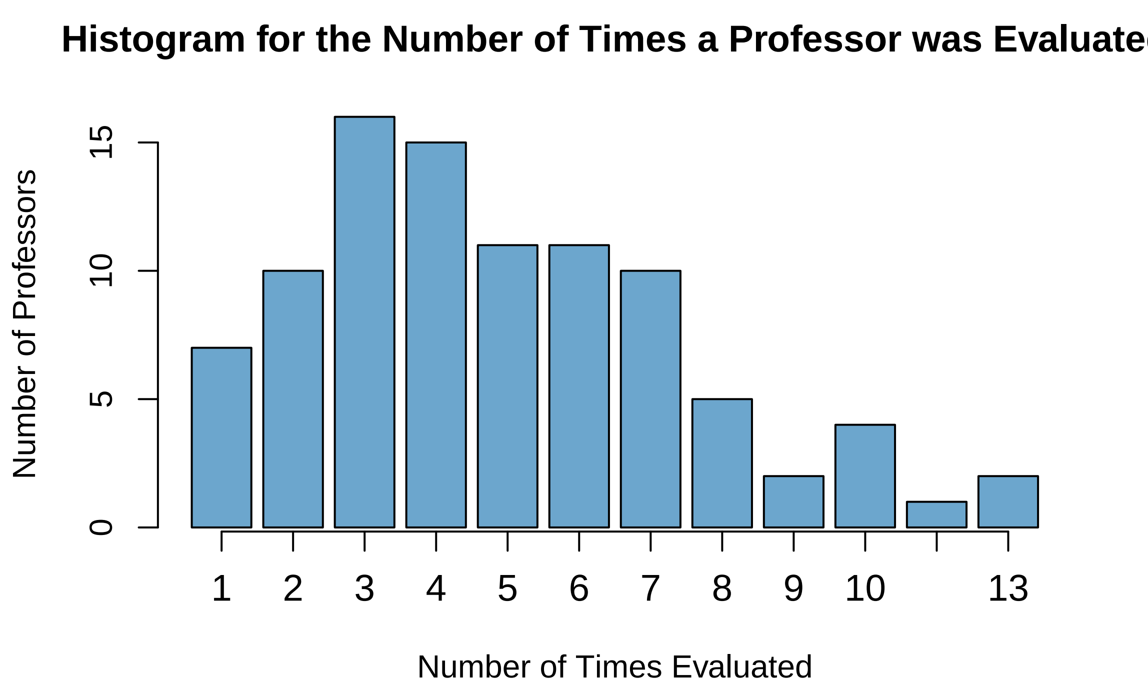
10/29/2020

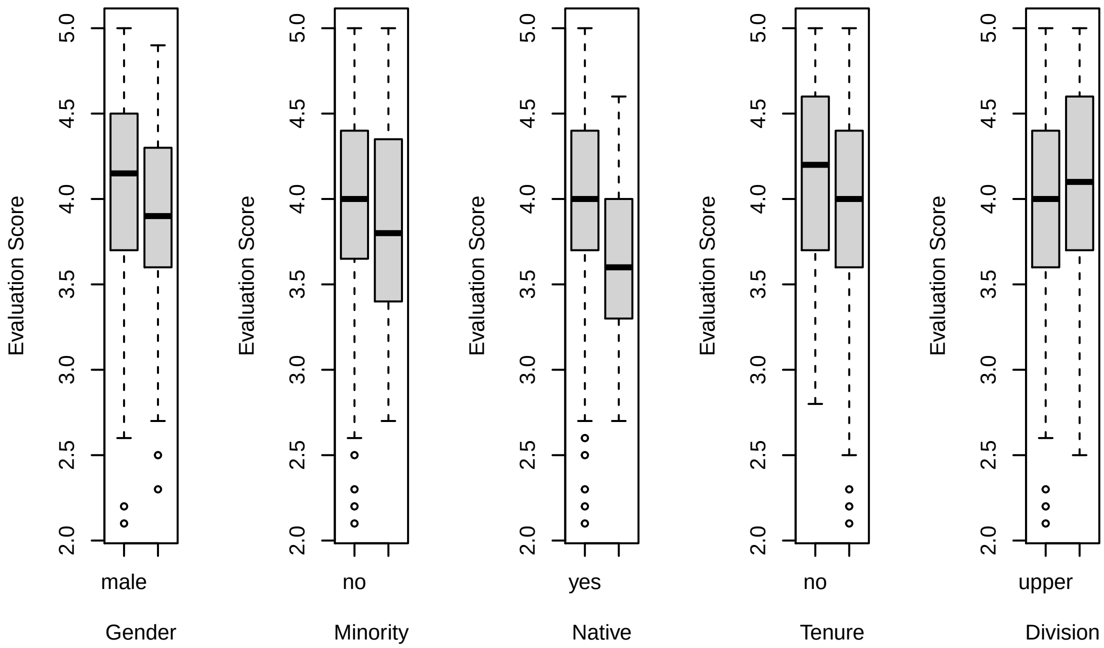
Summary

Typing words outside of chunks will be the paragraphs that show up in the PDF.

Exploratory Data Analysis

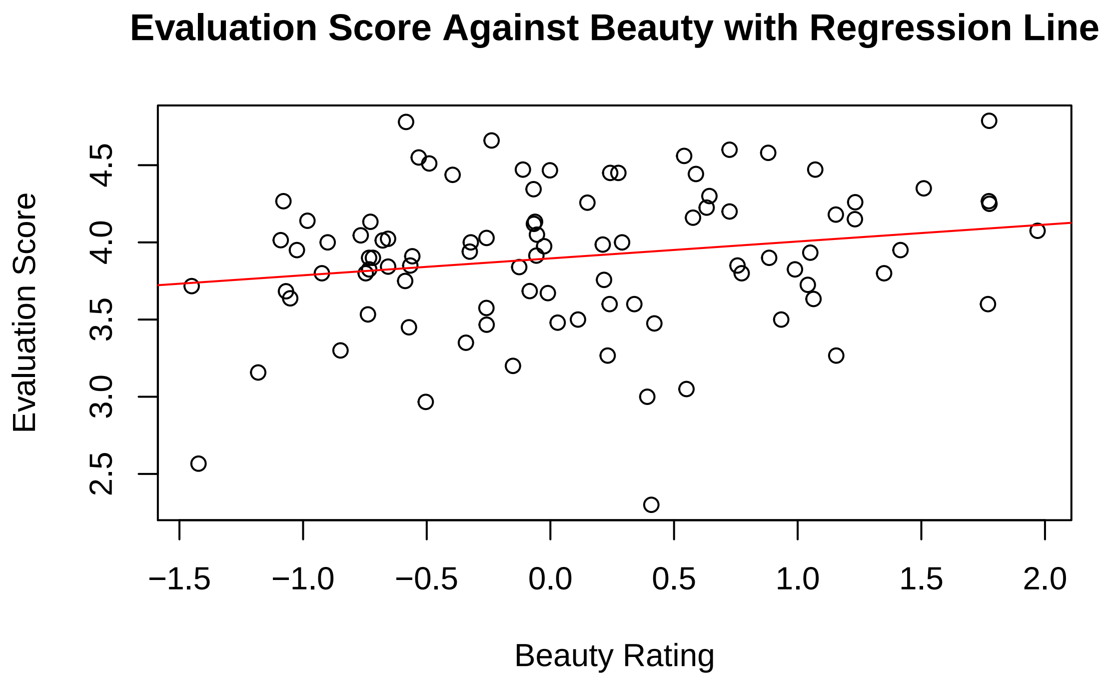
Typing words outside of chunks will be the paragraphs that show up in the PDF.





Simple Linear Regression

Our first goal for the data was to determine whether or not there was a linear relationship between a professors perceived beauty and their class evaluation scores. When looking at the scatter plot and corresponding regression line, there appears to be a weak relationship.



The summary statistics for the regression line show that the coefficient is only significant at a 90% confidence level. That would be an acceptable confidence level for some industries, but the more disturbing statistic is the R^2 value. The closer the R^2 value is to zero, the more likely it is that the regression line does not fit the data well. Since the R^2 value is 3.85%, we can conclude that the regression model does not predict the variation in the data very well at all.

```
##
## Call:
## lm(formula = evals ~ beaut, data = group.data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.64092 -0.27265  0.06915  0.24221  0.94757
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.89627    0.04716   82.62   <2e-16 ***
## beaut        0.10940    0.05697    1.92   0.0579 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4546 on 92 degrees of freedom
```

```
## Multiple R-squared:  0.03854,    Adjusted R-squared:  0.02809
## F-statistic: 3.688 on 1 and 92 DF,  p-value: 0.0579
```

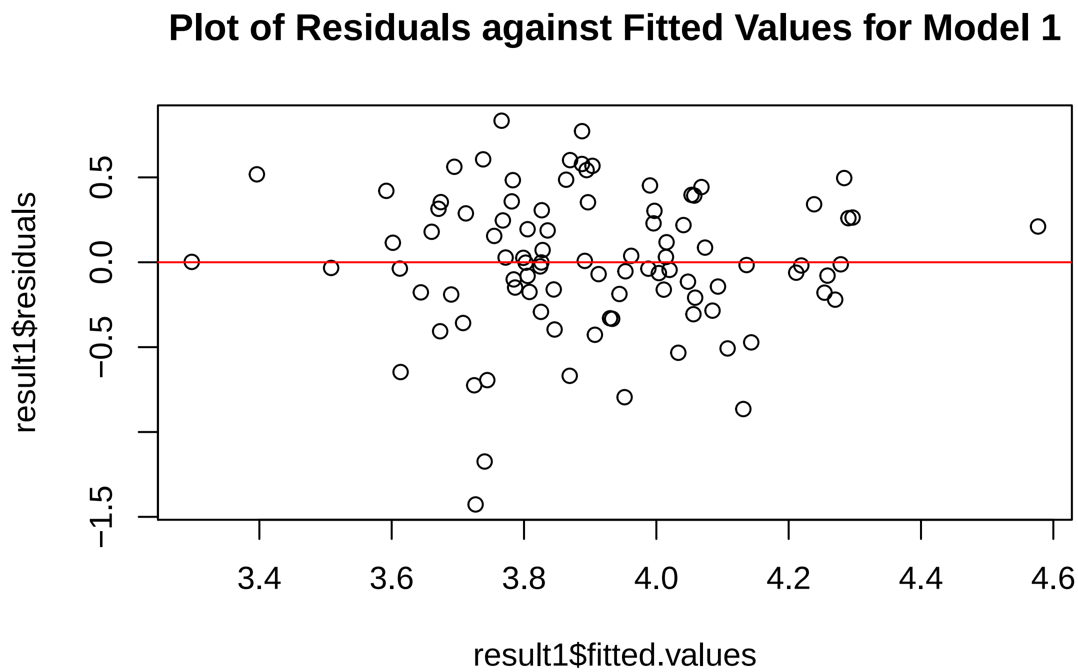
Since the original question did not produce a satisfactory model, we were curious if there were any simple linear regression models that predicted the evaluation score better than beauty rating. After running all subsets of simple linear regression, only two models had significant predictors: evaluation score against the percentage of upper division classes and number of students that took the evaluation. However, similarly to the beauty model, the R^2 value shows that the regression equation accounts for almost none of the variation in the data.

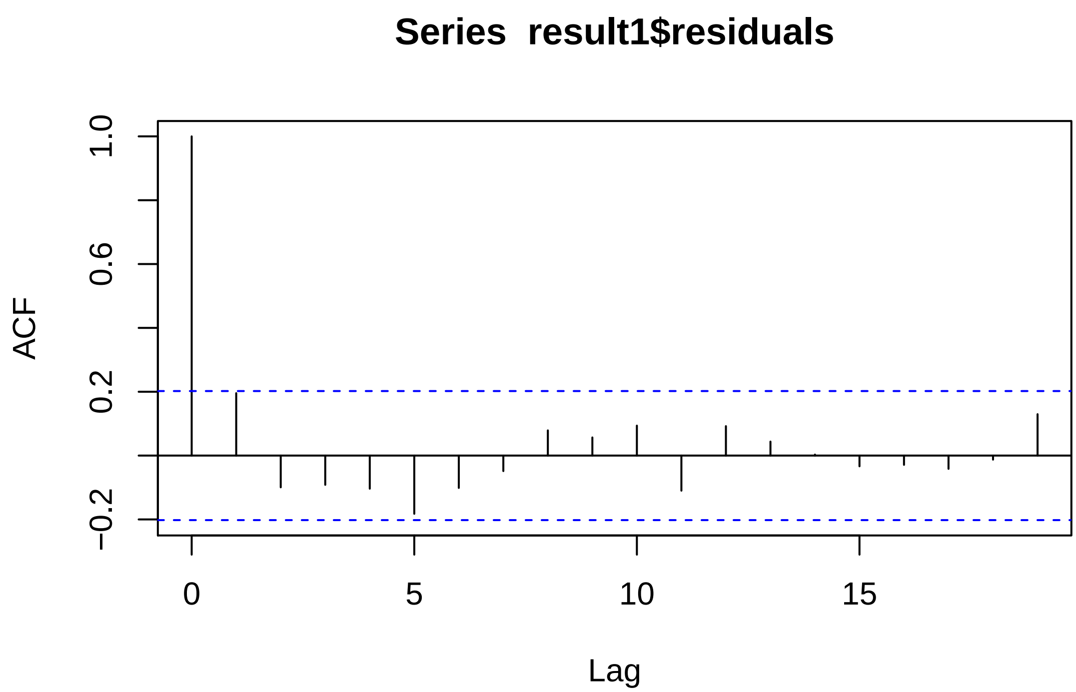
Multiple Linear Regression

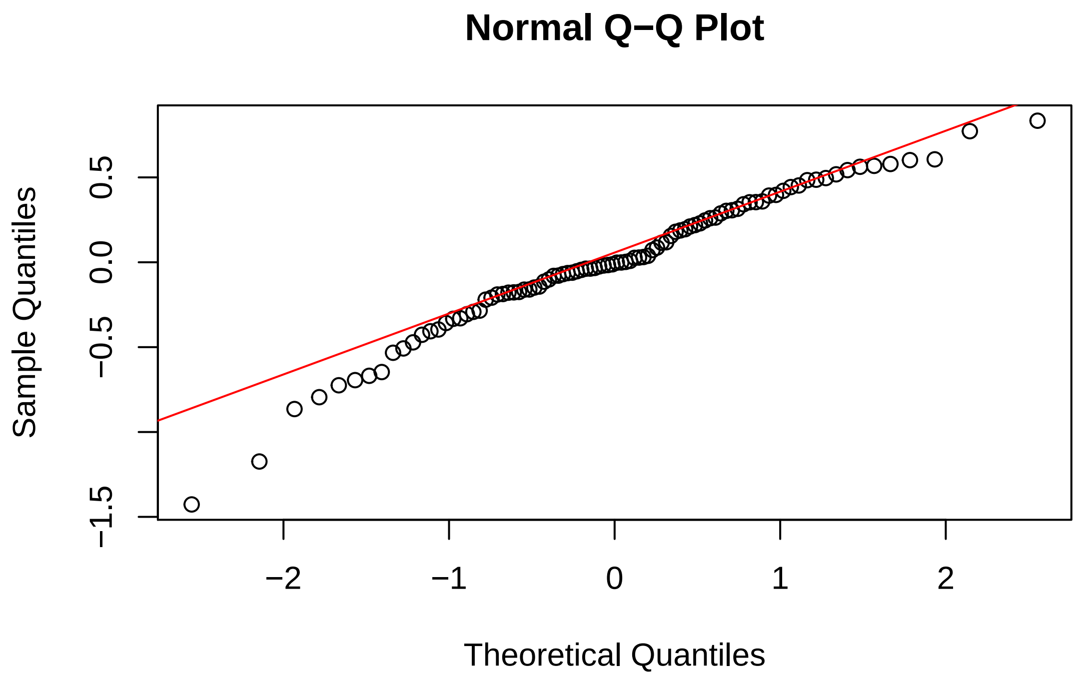
Typing words outside of chunks will be the paragraphs that show up in the PDF.

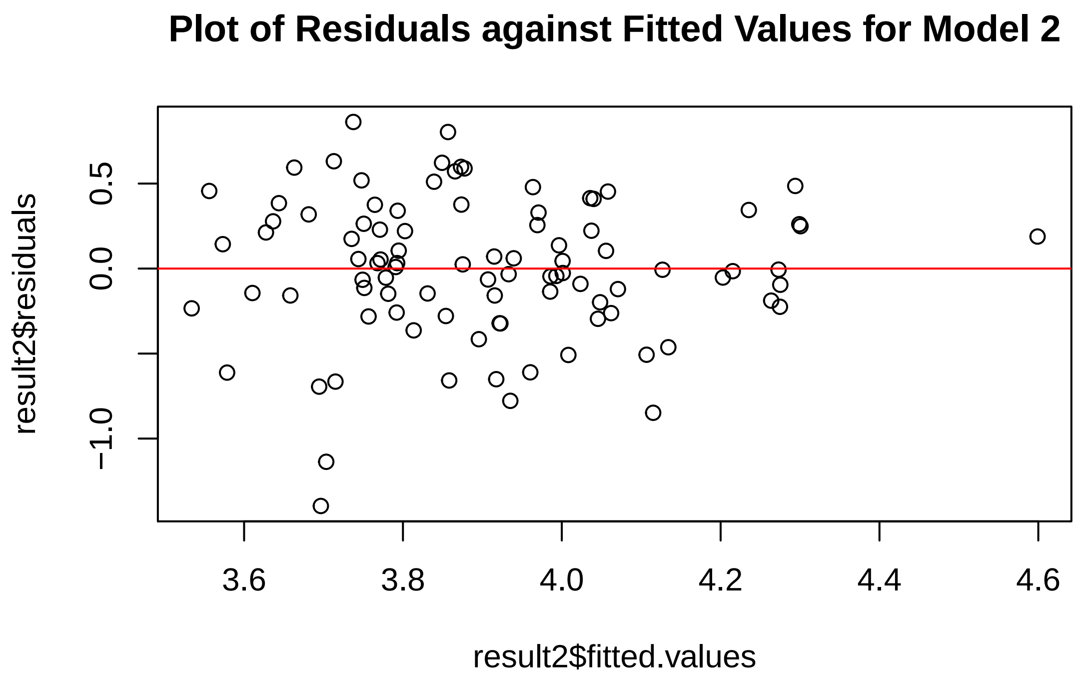
```
## Analysis of Variance Table
##
## Model 1: evals ~ division + beaut + gender + tenure + native
## Model 2: evals ~ division + beaut + gender + tenure + native + age + students +
##     allstudents + minority
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     88 15.499
## 2     84 14.470  4    1.0286 1.4928 0.2117
```

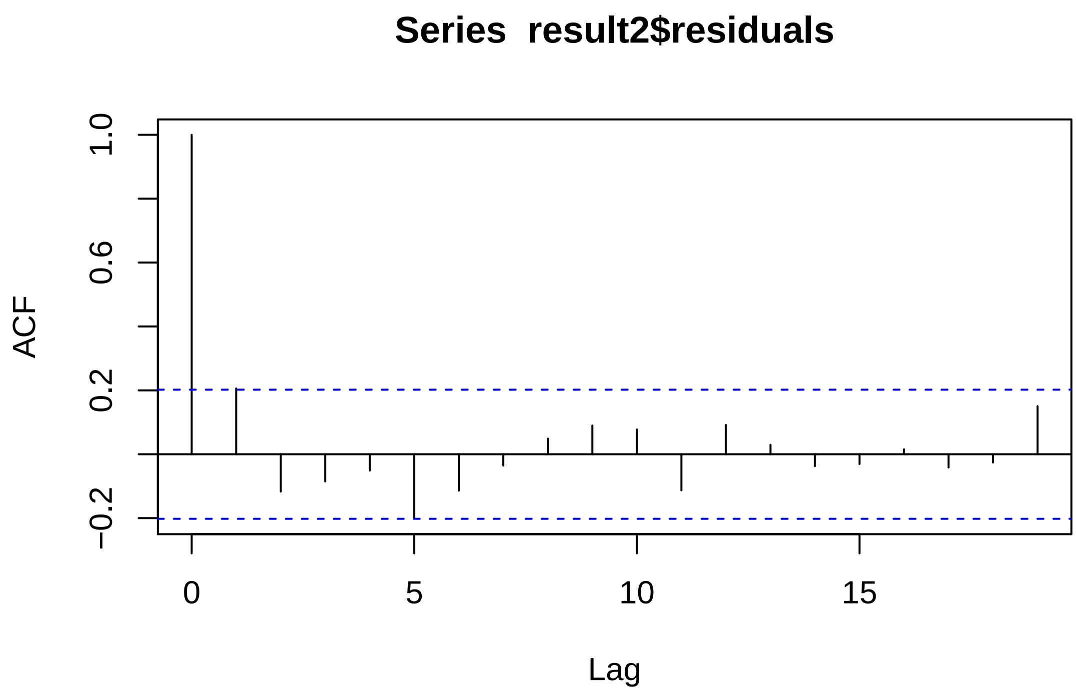
```
## Analysis of Variance Table
##
## Model 1: evals ~ division + beaut + gender + tenure
## Model 2: evals ~ division + beaut + gender + tenure + native + age + students +
##     allstudents + minority
##   Res.Df    RSS Df Sum of Sq     F Pr(>F)
## 1     89 15.964
## 2     84 14.470  5    1.4944 1.735 0.1355
```

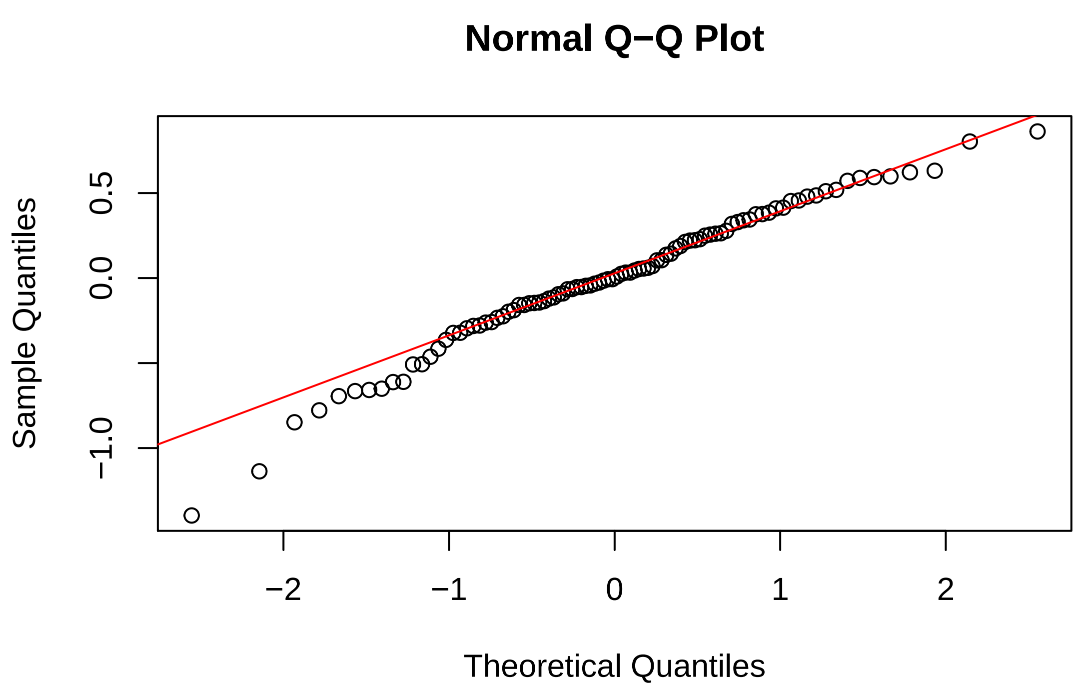

Plot of Residuals against Fitted Values for Model 1
result1$residuals
0.5
0.0
−0.5
−1.5
3.4
3.6
3.8
4.0
4.2
4.4
4.6
result1$fitted.values

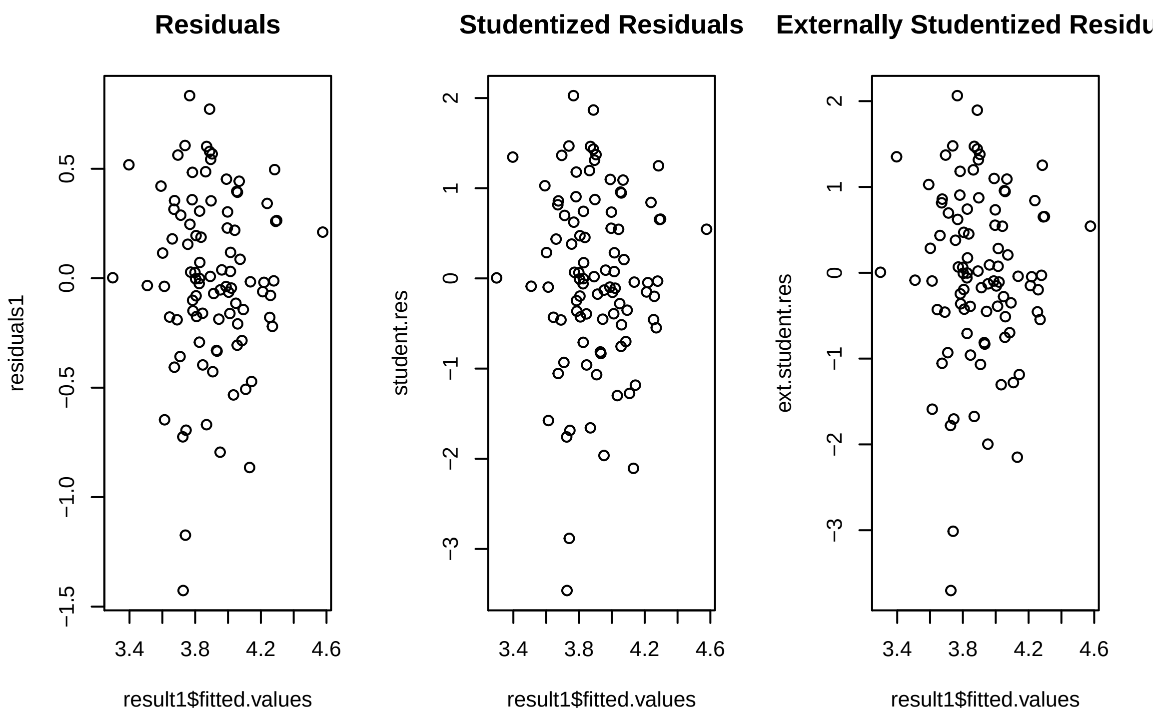

Series result1$residuals
ACF
1.0
0.6
0.2
−0.2
0
5
10
15
Lag

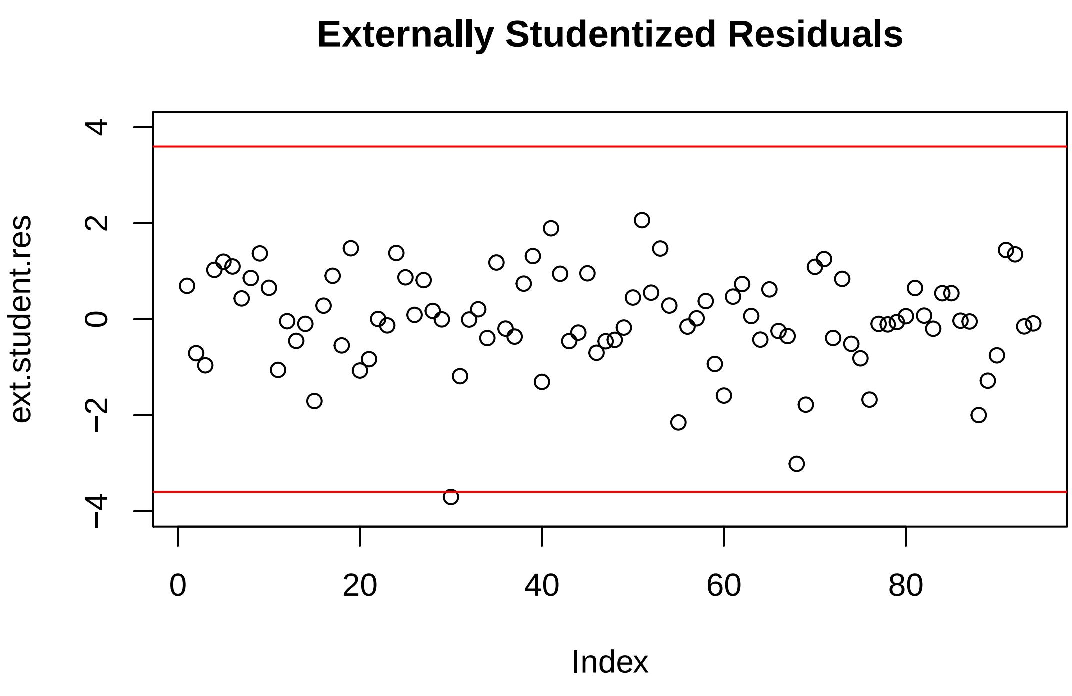

Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
−1.5
−0.5
0.0
0.5
−2
−1
0
1
2

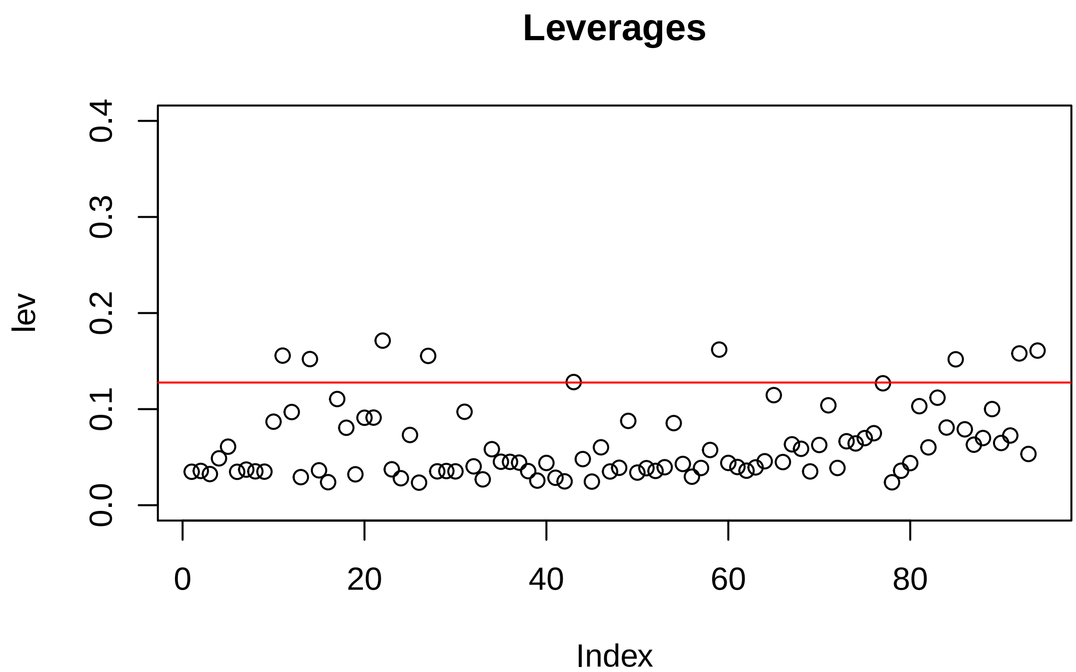

Plot of Residuals against Fitted Values for Model 2
result2$residuals
result2$fitted.values
−1.0
0.0
0.5
3.6
3.8
4.0
4.2
4.4
4.6


Series result2$residuals
ACF
1.0
0.6
0.2
−0.2
0
5
10
15
Lag


Normal Q−Q Plot
Sample Quantiles
0.5
0.0
−1.0
−2
−1
0
1
2
Theoretical Quantiles


Residuals
Studentized Residuals
Externally Studentized Residu
residuals1
student.res
ext.student.res
result1$fitted.values
result1$fitted.values
result1$fitted.values


Externally Studentized Residuals
ext.student.res
Index



Logistic Regression

Another one of our questions was whether or not age, beauty rating, and the percentage of upper level classes taught could predict the gender of a professor. Since gender is a binomial response variable, we used a logistic regression model. However when fitting the model, the regression coefficients do not seem to be significant. To test this, we used a ______ test with $H_0 : \beta_{age} = \beta_{beauty} = \beta_{division} = 0$ and H_1 : at least β_{age}, β_{beauty}, or $\beta_{division}$ does not equal zero. The p-value for the test came out to be 0.1078. Thus we fail to reject the null hypothesis with 95% confidence and conclude that the three coefficients are zero.

```
##
## Call:
## glm(formula = gender ~ age + beaut + division, family = "binomial",
##     data = group.data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.497  -1.025  -0.743   1.222   1.598
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.65116    1.14037   1.448   0.1476
```

```
## age          -0.03965    0.02302  -1.723   0.0849 .
## beaut         0.28459    0.27694   1.028   0.3041
## division     -0.36475    0.56350  -0.647   0.5174
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 128.22  on 93  degrees of freedom
## Residual deviance: 122.14  on 90  degrees of freedom
## AIC: 130.14
##
## Number of Fisher Scoring iterations: 4
```

With this result, we were not satisfied. Therefore we came up with a new query regarding the ability to predict a professor's tenure status using the other predictors in the model. The summary output for this model is slightly more promising with some predictors appearing to be significant. We ran a _____ test to be sure at least one of the coefficients was non-zero and a p-value of 0.0078 confirms that.

```
##
## Call:
## glm(formula = tenure ~ evals + age + gender + minority + native +
##     division, family = "binomial", data = group.data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.4308   0.1718   0.4325   0.6216   1.5194
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)    11.00143    4.37762   2.513   0.0120 *
## evals          -1.54289    0.85496  -1.805   0.0711 .
## age            -0.04713    0.03432  -1.374   0.1696
## genderfemale   -1.32408    0.70682  -1.873   0.0610 .
## minorityyes     1.05987    1.24879   0.849   0.3960
## nativeno       15.86342 2166.94907   0.007   0.9942
## division       -1.11048    0.78815  -1.409   0.1588
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 82.525  on 93  degrees of freedom
```

```
## Residual deviance: 70.657  on 87  degrees of freedom
## AIC: 84.657
##
## Number of Fisher Scoring iterations: 17
```

To help narrow down the necessary coefficients, we used forward, backward, and step-wise model selection techniques. All three techniques came up with the same model: evaluation being predicted by age and evaluation score. Although this is an interesting result, the summary shows that the age predictor may not be significant. Therefore we ran a ______ test to see whether or not the age coefficient was zero. The corresponding p-value for the test was 0.4333, meaning we can drop age from the regression model since it is zero.

```
##
## Call:
## glm(formula = tenure ~ evals + age, family = "binomial", data = group.data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.1605   0.3265   0.5179   0.6271   0.9297
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  8.63082    3.44632   2.504   0.0123 *
## evals       -1.47301    0.74877  -1.967   0.0492 *
## age         -0.02233    0.02863  -0.780   0.4353
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 82.525  on 93  degrees of freedom
## Residual deviance: 77.558  on 91  degrees of freedom
## AIC: 83.558
##
## Number of Fisher Scoring iterations: 5
```

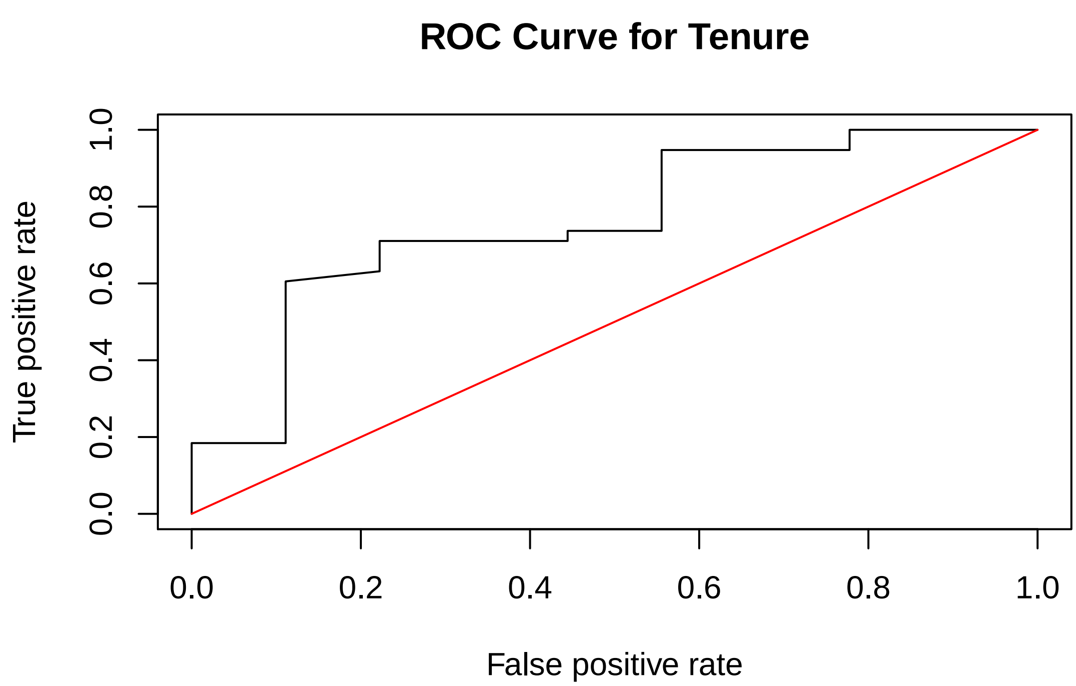
With the zero coefficient removed, our final log regression model for predicting the tenure of a professor was down to the evaluation score. The summary output shows the coefficient is significant at a 90% confidence level

```
##
## Call:
## glm(formula = tenure ~ evals, family = "binomial", data = group.data)
##
```

```
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2368   0.3583   0.5198   0.6439   0.8965
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   7.4655     3.0580   2.441   0.0146 *
## evals        -1.4510     0.7458  -1.946   0.0517 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 82.525  on 93  degrees of freedom
## Residual deviance: 78.169  on 92  degrees of freedom
## AIC: 82.169
##
## Number of Fisher Scoring iterations: 5
```

Log Regression Model Validation

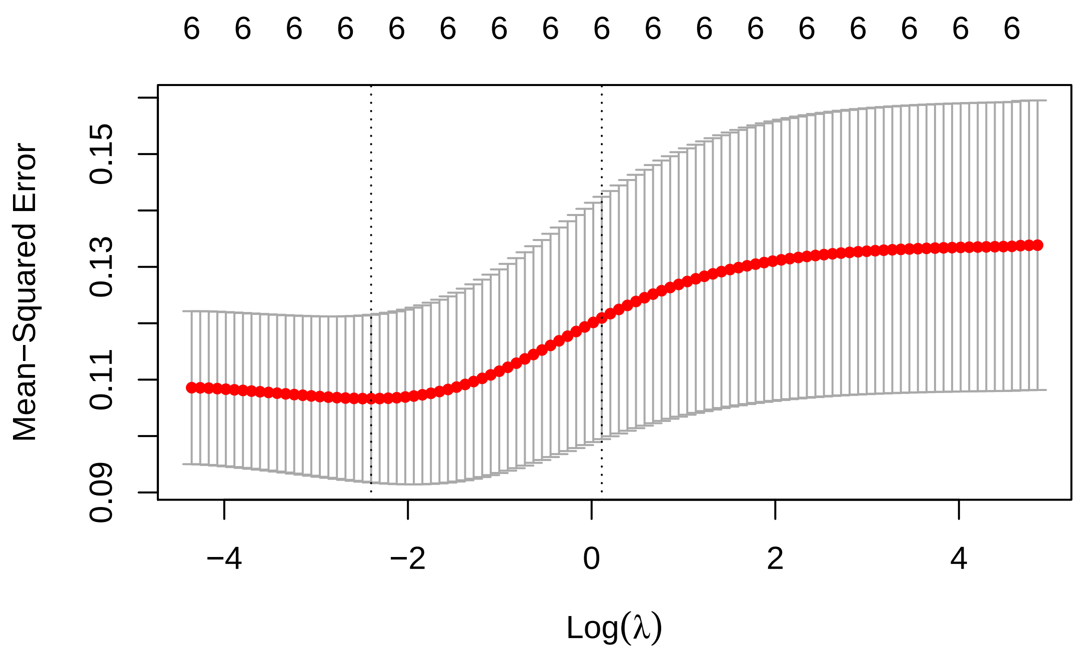
Since we have a regression model

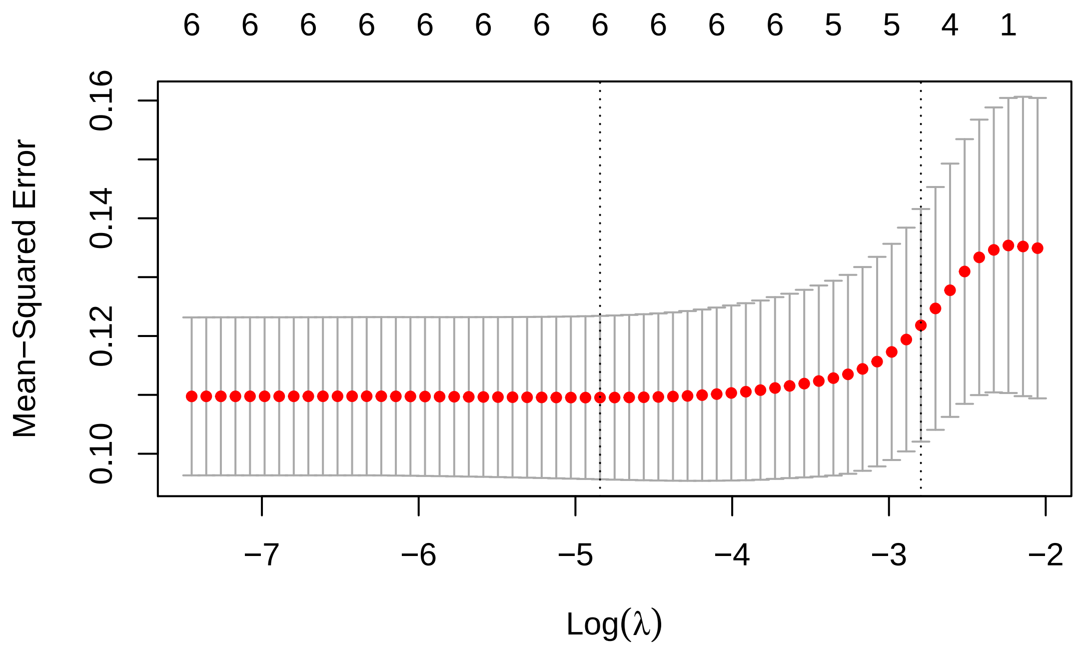


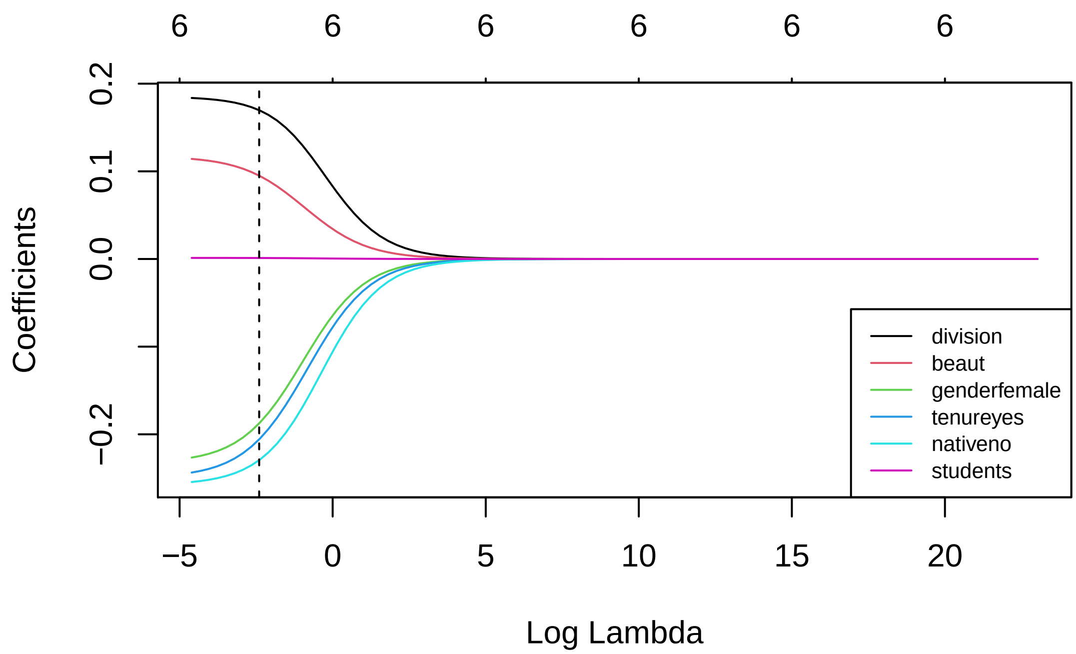
```
##
##        TRUE
##   no      9
##   yes    38
```

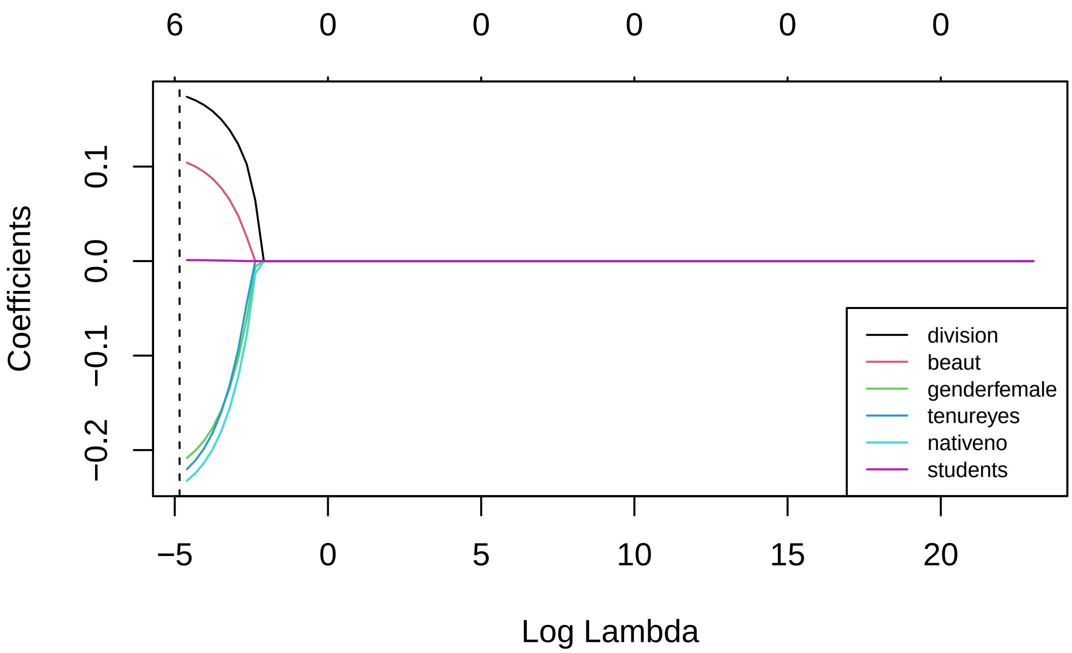
```
##
##        TRUE
##   no      9
##   yes    38
```

Ridge/Lasso Regression




6 6 6 6 6 6 6 6 6 6 6 5 5 4 1
Mean−Squared Error
0.10
0.12
0.14
0.16
−7
−6
−5
−4
−3
−2
Log(λ)


6 6 6 6 6 6
Coefficients
−0.2
0.0
0.1
0.2
−5
0
5
10
15
20
Log Lambda
division
beaut
genderfemale
tenureyes
nativeno
students


6
0
0
0
0
0
Coefficients
0.1
0.0
−0.1
−0.2
division
beaut
genderfemale
tenureyes
nativeno
students
−5
0
5
10
15
20
Log Lambda