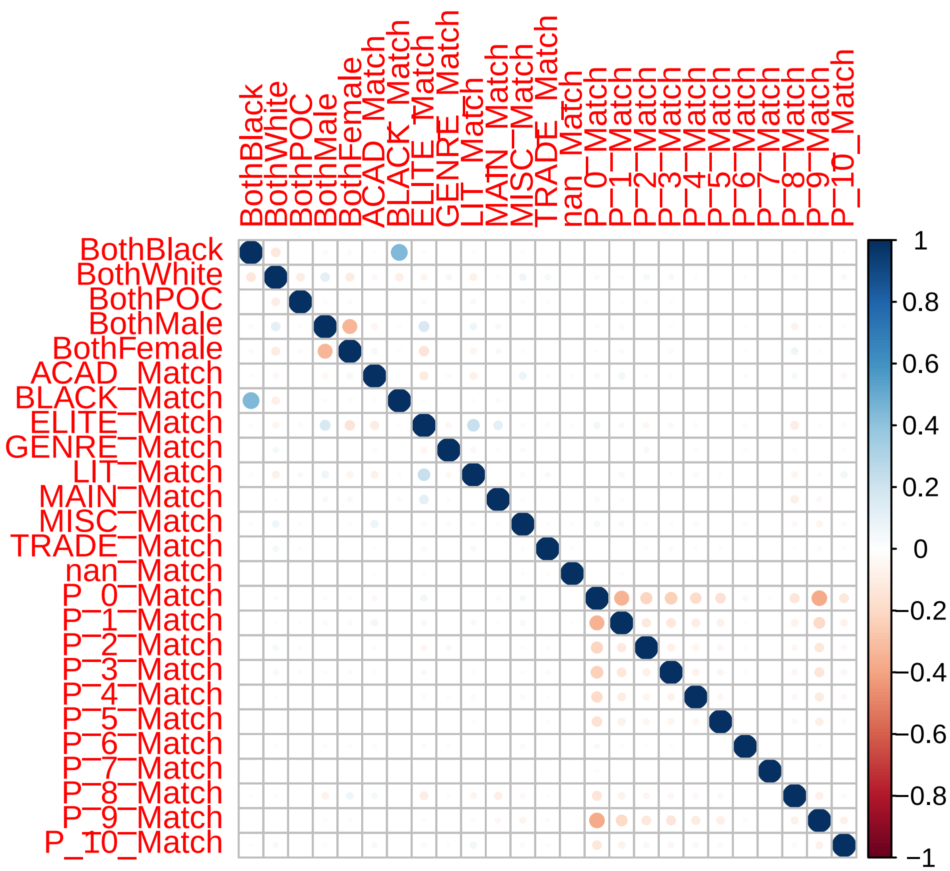
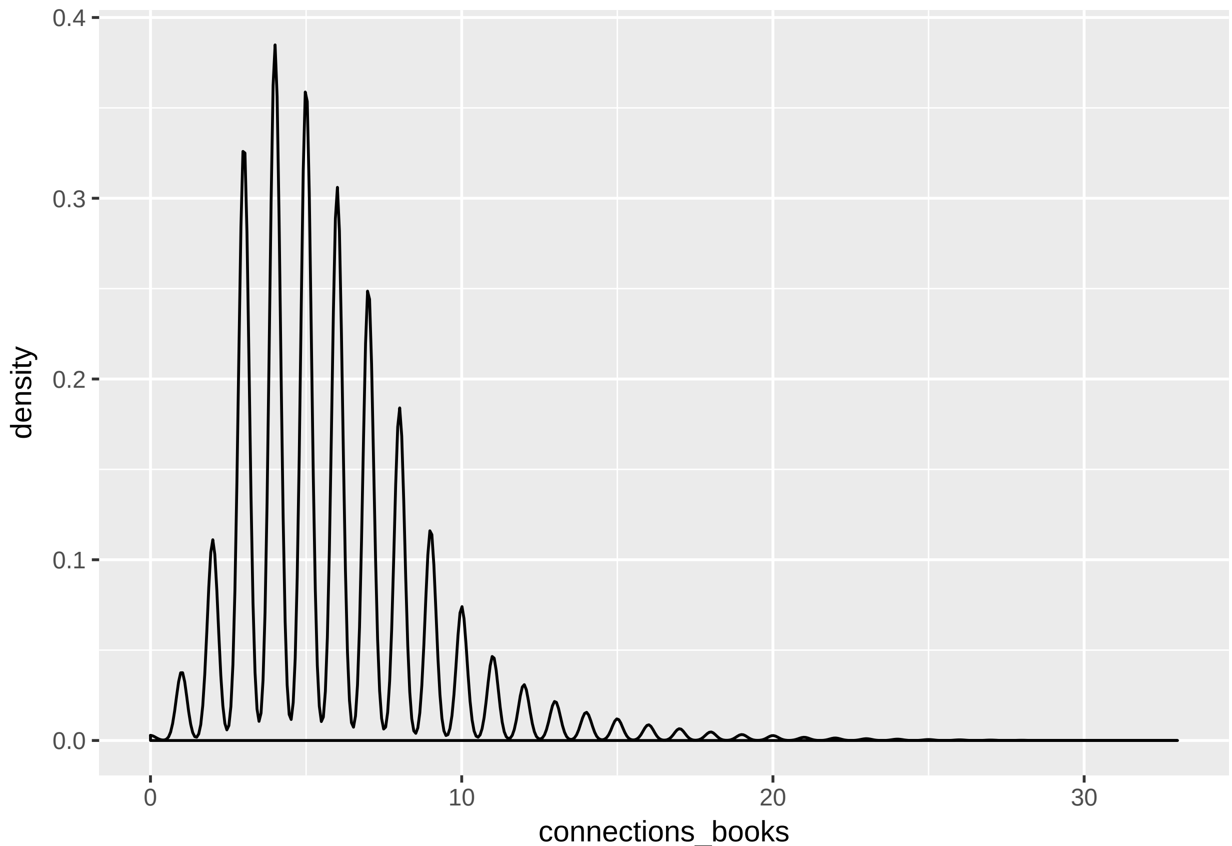
Model Comparisons- American Novels

Jessica Young

Novel Pairs Regression Analysis

```
## Warning: package 'dplyr' was built under R version 3.4.2

## Warning: package 'reshape2' was built under R version 3.4.3

## Warning: package 'foreach' was built under R version 3.4.3

## Warning: package 'doParallel' was built under R version 3.4.2

## Warning: package 'car' was built under R version 3.4.3

## Warning: package 'corrplot' was built under R version 3.4.3

## corrplot 0.84 loaded
```





Model for 1965-1970

```
a_1 = amer_novels %>%
      filter(Years=="65-70")

ggplot(a_1, aes(x=connections_books))+geom_density()

            log_model1 = train(connections_books~ BothBlack + BothWhite + BothPOC + BothMale + BothFemal
                                  BothBlack:BothMale + BothBlack:BothFemale +
                                  BothWhite:BothMale + BothWhite:BothFemale +
                                  BothPOC:BothMale + BothPOC:BothFemale +
                                  ACAD_Match + BLACK_Match + ELITE_Match + GENRE_Match + LIT_Match + MAI
                                  P_0_Match + P_1_Match + P_2_Match + P_3_Match + P_4_Match + P_5_Match
                             data = a_1,
                           method="glm",
                           family="poisson",
                           trControl=trainControl(
                             method='cv',number=10
                           ))
```

```
            summary(log_model1)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -3.4242  -0.8183  -0.0616   0.6807   3.4995
##
## Coefficients: (10 not defined because of singularities)
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)              1.763511   0.043595  40.452  < 2e-16 ***
## BothBlack               -0.003464   0.160463  -0.022 0.982777
## BothWhite                0.031887   0.025165   1.267 0.205123
## BothPOC                        NA         NA      NA       NA
## BothMale                 0.133933   0.027548   4.862 1.16e-06 ***
## BothFemale              -0.145675   0.084283  -1.728 0.083915 .
## ACAD_Match              -0.059080   0.016444  -3.593 0.000327 ***
## BLACK_Match              0.018648   0.127657   0.146 0.883857
## ELITE_Match                    NA         NA      NA       NA
## GENRE_Match             -0.056849   0.051634  -1.101 0.270895
## LIT_Match                0.232149   0.008521  27.245  < 2e-16 ***
## MAIN_Match                     NA         NA      NA       NA
## MISC_Match               0.280374   0.012394  22.622  < 2e-16 ***
## TRADE_Match              0.149858   0.022485   6.665 2.65e-11 ***
## nan_Match                      NA         NA      NA       NA
## P_0_Match                0.038556   0.028191   1.368 0.171416
## P_1_Match               -0.143606   0.029938  -4.797 1.61e-06 ***
## P_2_Match                0.112745   0.032038   3.519 0.000433 ***
## P_3_Match                0.198368   0.030041   6.603 4.02e-11 ***
## P_4_Match                0.098585   0.032597   3.024 0.002492 **
## P_5_Match                0.223730   0.032495   6.885 5.77e-12 ***
## P_6_Match               -0.117290   0.056611  -2.072 0.038279 *
## P_7_Match                      NA         NA      NA       NA
## P_8_Match                      NA         NA      NA       NA
## P_9_Match                0.075047   0.029277   2.563 0.010366 *
## P_10_Match                     NA         NA      NA       NA
## `BothBlack:BothMale`     0.072362   0.198859   0.364 0.715944
## `BothBlack:BothFemale`         NA         NA      NA       NA
## `BothWhite:BothMale`    -0.030561   0.029171  -1.048 0.294795
## `BothWhite:BothFemale`   0.020785   0.089415   0.232 0.816186
## `BothPOC:BothMale`             NA         NA      NA       NA
## `BothPOC:BothFemale`           NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 9313.2  on 6669  degrees of freedom
## Residual deviance: 7394.7  on 6648  degrees of freedom
## AIC: 35136
##
## Number of Fisher Scoring iterations: 4
```

Model for 1971-1975

```r
a_2 = amer_novels %>%
      filter(Years=="71-75")

ggplot(a_2, aes(x=connections_books))+geom_density()
```

```r
## BOTH MATCH  version
        log_model2 = train(connections_books~ BothBlack + BothWhite + BothPOC + BothMale + BothFemal
                               BothBlack:BothMale + BothBlack:BothFemale +
                               BothWhite:BothMale + BothWhite:BothFemale +
                               BothPOC:BothMale + BothPOC:BothFemale +
                               ACAD_Match + BLACK_Match + ELITE_Match + GENRE_Match + LIT_Match + MAI
                               P_0_Match + P_1_Match + P_2_Match + P_3_Match + P_4_Match + P_5_Match
                          data = a_2,
                        method="glm",
                        family="poisson",
                        trControl=trainControl(
                          method='cv',number=10
                        ))
```

```r
        summary(log_model2)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.8006  -0.7779  -0.1075   0.5903   4.0896
##
## Coefficients: (8 not defined because of singularities)
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)             1.653611   0.029096  56.833  < 2e-16 ***
## BothBlack               0.083669   0.115743   0.723  0.46975
## BothWhite               0.105595   0.020123   5.247 1.54e-07 ***
## BothPOC                       NA         NA      NA       NA
## BothMale                0.197782   0.027285   7.249 4.20e-13 ***
## BothFemale             -0.066741   0.042552  -1.568  0.11677
## ACAD_Match             -0.042906   0.018536  -2.315  0.02063 *
## BLACK_Match             0.044752   0.082774   0.541  0.58874
## ELITE_Match                   NA         NA      NA       NA
## GENRE_Match             0.526062   0.196316   2.680  0.00737 **
## LIT_Match               0.251930   0.009075  27.760  < 2e-16 ***
## MAIN_Match                    NA         NA      NA       NA
## MISC_Match              0.399362   0.019737  20.234  < 2e-16 ***
## TRADE_Match                   NA         NA      NA       NA
## nan_Match                     NA         NA      NA       NA
## P_0_Match               0.183861   0.018608   9.881  < 2e-16 ***
## P_1_Match               0.020320   0.023742   0.856  0.39207
## P_2_Match               0.293336   0.032554   9.011  < 2e-16 ***
## P_3_Match               0.089481   0.021839   4.097 4.18e-05 ***
## P_4_Match               0.332234   0.024817  13.387  < 2e-16 ***
## P_5_Match               0.151391   0.022489   6.732 1.68e-11 ***
## P_6_Match               0.368565   0.034174  10.785  < 2e-16 ***
## P_7_Match              -0.158604   0.056440  -2.810  0.00495 **
## P_8_Match              -0.060775   0.081305  -0.747  0.45477
## P_9_Match               0.227060   0.020129  11.280  < 2e-16 ***
## P_10_Match                    NA         NA      NA       NA
## `BothBlack:BothMale`   -0.121645   0.202273  -0.601  0.54758
## `BothBlack:BothFemale` -0.073815   0.187772  -0.393  0.69424
```

```
## `BothWhite:BothMale`   -0.060105   0.028844  -2.084  0.03718 *
## `BothWhite:BothFemale` -0.019145   0.048110  -0.398  0.69068
## `BothPOC:BothMale`            NA         NA      NA       NA
## `BothPOC:BothFemale`          NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 8572.8  on 5355  degrees of freedom
## Residual deviance: 5599.7  on 5332  degrees of freedom
## AIC: 28273
##
## Number of Fisher Scoring iterations: 4
```

Model for 1976-1980

```
a_3 = amer_novels %>%
      filter(Years=="76-80")

ggplot(a_3, aes(x=connections_books))+geom_density()
    ## BOTH MATCH  version
            log_model3 = train(connections_books~ BothBlack + BothWhite + BothPOC + BothMale + BothFemal
                                  BothBlack:BothMale + BothBlack:BothFemale +
                                  BothWhite:BothMale + BothWhite:BothFemale +
                                  BothPOC:BothMale + BothPOC:BothFemale +
                                  ACAD_Match + BLACK_Match + ELITE_Match + GENRE_Match + LIT_Match + MA
                                  P_0_Match + P_1_Match + P_2_Match + P_3_Match + P_4_Match + P_5_Match
                             data = a_3,
                           method="glm",
                           family="poisson",
                           trControl=trainControl(
                             method='cv',number=10
                           ))
```

```
            summary(log_model3)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.1264  -0.8396  -0.1478   0.6487   3.8175
##
## Coefficients: (9 not defined because of singularities)
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)             1.112398   0.052457  21.206  < 2e-16 ***
## BothBlack               0.020742   0.146642   0.141 0.887519
## BothWhite               0.023452   0.016418   1.428 0.153182
## BothPOC                -0.136383   0.277795  -0.491 0.623464
## BothMale               -0.032639   0.029418  -1.109 0.267221
## BothFemale             -0.001648   0.027618  -0.060 0.952430
## ACAD_Match              0.015933   0.014157   1.125 0.260404
## BLACK_Match             0.073102   0.082834   0.883 0.377503
```

```
## ELITE_Match             0.509824   0.043867  11.622  < 2e-16 ***
## GENRE_Match             0.013098   0.128239   0.102 0.918648
## LIT_Match               0.573222   0.008451  67.830  < 2e-16 ***
## MAIN_Match                    NA         NA      NA       NA
## MISC_Match              0.249086   0.013950  17.856  < 2e-16 ***
## TRADE_Match                   NA         NA      NA       NA
## nan_Match                     NA         NA      NA       NA
## P_0_Match               0.256802   0.020114  12.767  < 2e-16 ***
## P_1_Match               0.162573   0.024101   6.745 1.53e-11 ***
## P_2_Match               0.050722   0.031990   1.586 0.112839
## P_3_Match               0.099954   0.027484   3.637 0.000276 ***
## P_4_Match               0.384153   0.025456  15.091  < 2e-16 ***
## P_5_Match               0.179512   0.025727   6.978 3.00e-12 ***
## P_6_Match                     NA         NA      NA       NA
## P_7_Match                     NA         NA      NA       NA
## P_8_Match               0.155745   0.034369   4.532 5.86e-06 ***
## P_9_Match               0.218360   0.020744  10.526  < 2e-16 ***
## P_10_Match                    NA         NA      NA       NA
## `BothBlack:BothMale`          NA         NA      NA       NA
## `BothBlack:BothFemale`  0.040758   0.197313   0.207 0.836350
## `BothWhite:BothMale`    0.091337   0.030986   2.948 0.003202 **
## `BothWhite:BothFemale`  0.008291   0.031056   0.267 0.789482
## `BothPOC:BothMale`            NA         NA      NA       NA
## `BothPOC:BothFemale`          NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 14470.6  on 6327  degrees of freedom
## Residual deviance:  7306.6  on 6305  degrees of freedom
## AIC: 33311
##
## Number of Fisher Scoring iterations: 4
```

Model for 1981-1985

```r
a_4 = amer_novels %>%
      filter(Years=="81-85")

ggplot(a_4, aes(x=connections_books))+geom_density()
    ## BOTH MATCH  version
            log_model4 = train(connections_books~ BothBlack + BothWhite + BothPOC + BothMale + BothFemal
                                  BothBlack:BothMale + BothBlack:BothFemale +
                                  BothWhite:BothMale + BothWhite:BothFemale +
                                  BothPOC:BothMale + BothPOC:BothFemale +
                                  ACAD_Match + BLACK_Match + ELITE_Match + GENRE_Match + LIT_Match + MAI
                                  P_0_Match + P_1_Match + P_2_Match + P_3_Match + P_4_Match + P_5_Match
                             data = a_4,
                           method="glm",
                           family="poisson",
                           trControl=trainControl(
                             method='cv',number=10
                           ))
```

```r
summary(log_model4)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.4205  -0.7128  -0.0944   0.5133   4.7335
##
## Coefficients: (7 not defined because of singularities)
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)             0.994885   0.027109  36.699  < 2e-16 ***
## BothBlack              -0.046551   0.073642  -0.632  0.52730
## BothWhite               0.047695   0.010839   4.401 1.08e-05 ***
## BothPOC                 0.289730   0.200502   1.445  0.14845
## BothMale               -0.108170   0.016407  -6.593 4.31e-11 ***
## BothFemale              0.037036   0.020190   1.834  0.06659 .
## ACAD_Match              0.021071   0.010272   2.051  0.04024 *
## BLACK_Match            -0.024965   0.050606  -0.493  0.62180
## ELITE_Match             0.391318   0.009421  41.539  < 2e-16 ***
## GENRE_Match             0.028492   0.027219   1.047  0.29520
## LIT_Match               0.318957   0.006248  51.050  < 2e-16 ***
## MAIN_Match                    NA         NA      NA       NA
## MISC_Match              0.188666   0.006133  30.765  < 2e-16 ***
## TRADE_Match             0.317433   0.020854  15.221  < 2e-16 ***
## nan_Match                     NA         NA      NA       NA
## P_0_Match               0.054890   0.010585   5.186 2.15e-07 ***
## P_1_Match               0.039330   0.012069   3.259  0.00112 **
## P_2_Match               0.021017   0.017242   1.219  0.22288
## P_3_Match               0.024096   0.016254   1.483  0.13821
## P_4_Match               0.022446   0.014313   1.568  0.11682
## P_5_Match              -0.042888   0.023791  -1.803  0.07143 .
## P_6_Match                     NA         NA      NA       NA
## P_7_Match                     NA         NA      NA       NA
## P_8_Match              -0.298245   0.076169  -3.916 9.02e-05 ***
## P_9_Match               0.088685   0.012079   7.342 2.10e-13 ***
## P_10_Match                    NA         NA      NA       NA
## `BothBlack:BothMale`    0.026129   0.183559   0.142  0.88681
## `BothBlack:BothFemale`  0.004603   0.103385   0.045  0.96449
## `BothWhite:BothMale`    0.186757   0.017393  10.737  < 2e-16 ***
## `BothWhite:BothFemale` -0.038478   0.022627  -1.701  0.08903 .
## `BothPOC:BothMale`            NA         NA      NA       NA
## `BothPOC:BothFemale`          NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 27650  on 21320  degrees of freedom
## Residual deviance: 19680  on 21296  degrees of freedom
## AIC: 99837
##
## Number of Fisher Scoring iterations: 4
```

Model for 1986-1990

```r
a_5 = amer_novels %>%
      filter(Years=="86-90")

ggplot(a_5, aes(x=connections_books))+geom_density()
    ## BOTH MATCH  version
            log_model5 = train(connections_books~ BothBlack + BothWhite + BothPOC + BothMale + BothFema
                                  BothBlack:BothMale + BothBlack:BothFemale +
                                  BothWhite:BothMale + BothWhite:BothFemale +
                                  BothPOC:BothMale + BothPOC:BothFemale +
                                  ACAD_Match + BLACK_Match + ELITE_Match + GENRE_Match + LIT_Match + MA
                                  P_0_Match + P_1_Match + P_2_Match + P_3_Match + P_4_Match + P_5_Match
                             data = a_5,
                           method="glm",
                           family="poisson",
                           trControl=trainControl(
                             method='cv',number=10
                           ))
```

```r
            summary(log_model5)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.9030  -0.7163  -0.1698   0.5416   5.4067
##
## Coefficients: (2 not defined because of singularities)
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -0.1077924  0.0454828  -2.370 0.017790 *
## BothBlack             -0.0137430  0.0867125  -0.158 0.874072
## BothWhite             -0.0374810  0.0078777  -4.758 1.96e-06 ***
## BothPOC               -0.0613331  0.1304179  -0.470 0.638154
## BothMale              -0.0850174  0.0168853  -5.035 4.78e-07 ***
## BothFemale            -0.0042409  0.0121158  -0.350 0.726318
## ACAD_Match             0.0872029  0.0047304  18.435  < 2e-16 ***
## BLACK_Match           -0.0268322  0.0925137  -0.290 0.771789
## ELITE_Match            0.3313752  0.0051998  63.729  < 2e-16 ***
## GENRE_Match           -0.0236769  0.0145269  -1.630 0.103130
## LIT_Match              0.2136146  0.0050567  42.244  < 2e-16 ***
## MAIN_Match             1.2307353  0.0406065  30.309  < 2e-16 ***
## MISC_Match             0.1011594  0.0046931  21.555  < 2e-16 ***
## TRADE_Match            0.1894146  0.0092406  20.498  < 2e-16 ***
## nan_Match              0.1701802  0.1460182   1.165 0.243828
## P_0_Match              0.1180476  0.0176890   6.673 2.50e-11 ***
## P_1_Match              0.0938307  0.0179523   5.227 1.73e-07 ***
## P_2_Match              0.0698203  0.0194379   3.592 0.000328 ***
## P_3_Match              0.0471938  0.0189177   2.495 0.012607 *
## P_4_Match              0.0332429  0.0192033   1.731 0.083433 .
## P_5_Match              0.1542084  0.0204109   7.555 4.18e-14 ***
## P_6_Match              0.2778896  0.0282022   9.853  < 2e-16 ***
## P_7_Match                     NA         NA      NA       NA
```

```
## P_8_Match               0.0110849  0.0251588   0.441 0.659506
## P_9_Match               0.0885085  0.0181543   4.875 1.09e-06 ***
## P_10_Match                     NA         NA      NA       NA
## `BothBlack:BothMale`    0.1506867  0.2909568   0.518 0.604528
## `BothBlack:BothFemale` -0.0539082  0.1089932  -0.495 0.620881
## `BothWhite:BothMale`    0.1412930  0.0174198   8.111 5.02e-16 ***
## `BothWhite:BothFemale`  0.0008485  0.0134045   0.063 0.949531
## `BothPOC:BothMale`     -0.2652961  0.5169942  -0.513 0.607846
## `BothPOC:BothFemale`    0.4585971  0.1679728   2.730 0.006330 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 53259  on 49769  degrees of freedom
## Residual deviance: 43344  on 49740  degrees of freedom
## AIC: 220191
##
## Number of Fisher Scoring iterations: 4
```

Model for 1991-1995

```r
a_6 = amer_novels %>%
      filter(Years=="91-95")

ggplot(a_6, aes(x=connections_books))+geom_density()
    ## BOTH MATCH  version
            log_model6 = train(connections_books~ BothBlack + BothWhite + BothPOC + BothMale + BothFemal
                                  BothBlack:BothMale + BothBlack:BothFemale +
                                  BothWhite:BothMale + BothWhite:BothFemale +
                                  BothPOC:BothMale + BothPOC:BothFemale +
                                  ACAD_Match + BLACK_Match + ELITE_Match + GENRE_Match + LIT_Match + MA
                                  P_0_Match + P_1_Match + P_2_Match + P_3_Match + P_4_Match + P_5_Match
                             data = a_6,
                           method="glm",
                           family="poisson",
                           trControl=trainControl(
                             method='cv',number=10
                           ))
```

```r
            summary(log_model6)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.7431  -0.6604  -0.1439   0.4571   5.5494
##
## Coefficients: (3 not defined because of singularities)
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)             0.328925   0.024110  13.643  < 2e-16 ***
## BothBlack              -0.035209   0.028870  -1.220 0.222626
## BothWhite               0.016712   0.004467   3.741 0.000183 ***
```

```
## BothPOC                  0.173435   0.030638   5.661 1.51e-08 ***
## BothMale                 0.064509   0.006822   9.456  < 2e-16 ***
## BothFemale              -0.026893   0.006575  -4.090 4.31e-05 ***
## ACAD_Match               0.103902   0.003804  27.314  < 2e-16 ***
## BLACK_Match              0.117301   0.029830   3.932 8.41e-05 ***
## ELITE_Match              0.333527   0.003250 102.620  < 2e-16 ***
## GENRE_Match              0.131747   0.007045  18.701  < 2e-16 ***
## LIT_Match                0.201102   0.003344  60.145  < 2e-16 ***
## MAIN_Match               0.451008   0.020002  22.548  < 2e-16 ***
## MISC_Match               0.127056   0.002839  44.758  < 2e-16 ***
## TRADE_Match              0.305689   0.007077  43.193  < 2e-16 ***
## nan_Match               -0.060140   0.011720  -5.131 2.88e-07 ***
## P_0_Match                0.227989   0.011369  20.053  < 2e-16 ***
## P_1_Match                0.179949   0.011698  15.383  < 2e-16 ***
## P_2_Match                0.152850   0.012280  12.447  < 2e-16 ***
## P_3_Match                0.218654   0.012204  17.917  < 2e-16 ***
## P_4_Match                0.201500   0.012786  15.759  < 2e-16 ***
## P_5_Match                0.226475   0.012955  17.482  < 2e-16 ***
## P_6_Match                      NA         NA      NA       NA
## P_7_Match                      NA         NA      NA       NA
## P_8_Match                0.246598   0.013484  18.289  < 2e-16 ***
## P_9_Match                0.182680   0.011540  15.830  < 2e-16 ***
## P_10_Match                     NA         NA      NA       NA
## `BothBlack:BothMale`    -0.018547   0.052340  -0.354 0.723069
## `BothBlack:BothFemale`  -0.071736   0.039510  -1.816 0.069423 .
## `BothWhite:BothMale`    -0.035354   0.007682  -4.602 4.18e-06 ***
## `BothWhite:BothFemale`   0.082836   0.007690  10.772  < 2e-16 ***
## `BothPOC:BothMale`       0.038270   0.058747   0.651 0.514773
## `BothPOC:BothFemale`    -0.074830   0.050922  -1.470 0.141695
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 109458  on 107415  degrees of freedom
## Residual deviance:  84589  on 107387  degrees of freedom
## AIC: 454816
##
## Number of Fisher Scoring iterations: 4
```

Model for 1996-2000

```
a_7 = amer_novels %>%
      filter(Years=="96-00")

ggplot(a_7, aes(x=connections_books))+geom_density()
    ## BOTH MATCH  version
            log_model7 = train(connections_books~ BothBlack + BothWhite + BothPOC + BothMale + BothFemal
                                  BothBlack:BothMale + BothBlack:BothFemale +
                                  BothWhite:BothMale + BothWhite:BothFemale +
                                  BothPOC:BothMale + BothPOC:BothFemale +
                                  ACAD_Match + BLACK_Match + ELITE_Match + GENRE_Match + LIT_Match + MAI
                                  P_0_Match + P_1_Match + P_2_Match + P_3_Match + P_4_Match + P_5_Match
                             data = a_7,
                           method="glm",
```

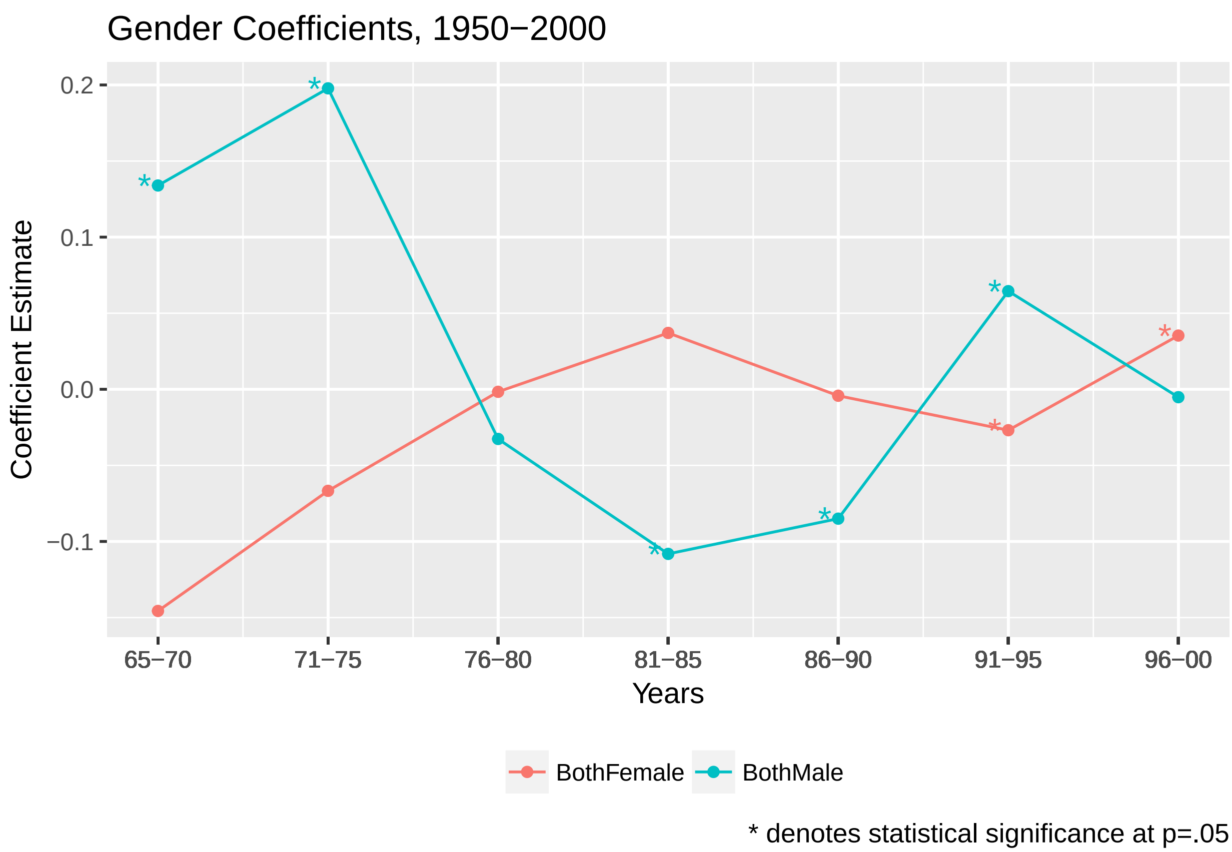
```r
                         family="poisson",
                         trControl=trainControl(
                           method='cv',number=10
                         ))
```

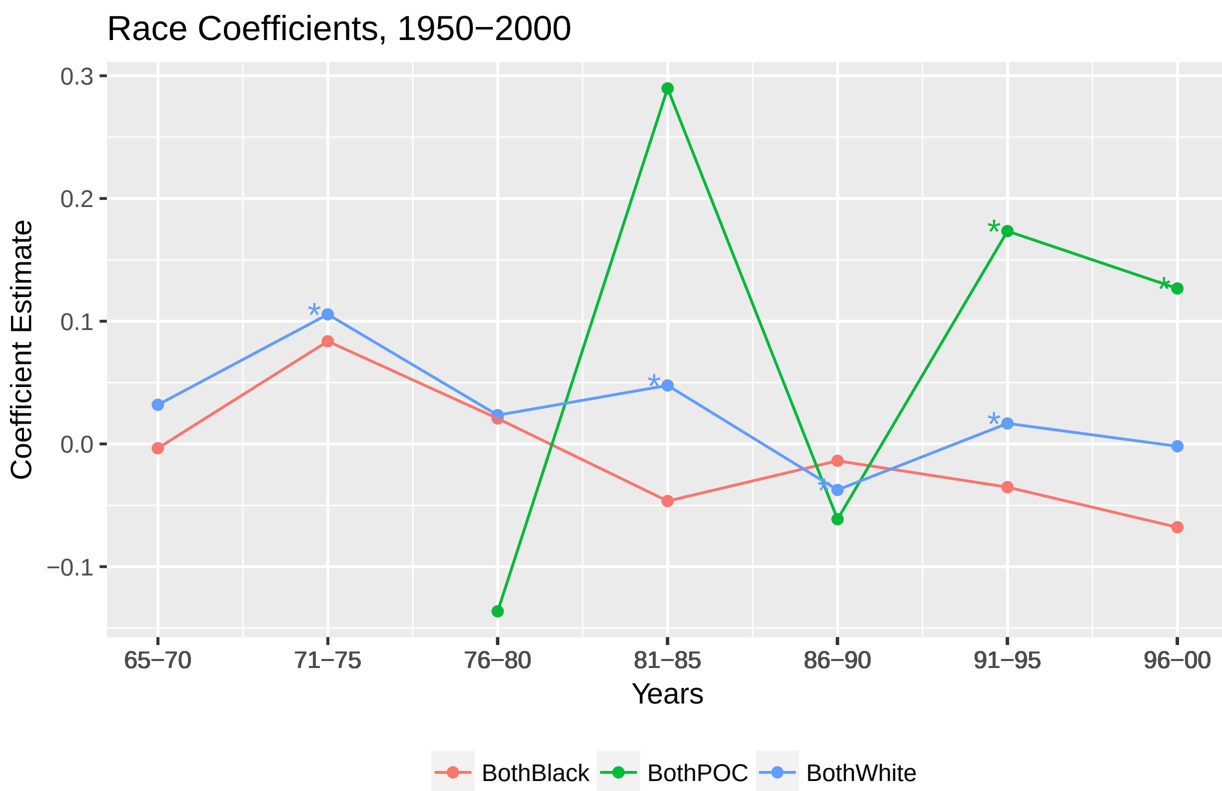
```r
summary(log_model7)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.2682  -0.5940  -0.2597   0.4906   5.4582
##
## Coefficients: (3 not defined because of singularities)
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)             0.488399   0.026639  18.334  < 2e-16 ***
## BothBlack              -0.067894   0.050101  -1.355  0.17537
## BothWhite              -0.001899   0.006514  -0.292  0.77066
## BothPOC                 0.126671   0.046237   2.740  0.00615 **
## BothMale               -0.005190   0.011359  -0.457  0.64777
## BothFemale              0.035297   0.009054   3.898 9.69e-05 ***
## ACAD_Match              0.296043   0.008489  34.873  < 2e-16 ***
## BLACK_Match                   NA         NA      NA       NA
## ELITE_Match             0.344087   0.004123  83.458  < 2e-16 ***
## GENRE_Match             0.111874   0.037027   3.021  0.00252 **
## LIT_Match               0.153056   0.006076  25.189  < 2e-16 ***
## MAIN_Match              0.413726   0.020015  20.671  < 2e-16 ***
## MISC_Match              0.091916   0.004203  21.867  < 2e-16 ***
## TRADE_Match             0.263965   0.009508  27.763  < 2e-16 ***
## nan_Match               0.168044   0.020711   8.114 4.91e-16 ***
## P_0_Match               0.164998   0.014597  11.304  < 2e-16 ***
## P_1_Match               0.110228   0.015236   7.235 4.67e-13 ***
## P_2_Match               0.133778   0.015821   8.456  < 2e-16 ***
## P_3_Match               0.132555   0.015370   8.624  < 2e-16 ***
## P_4_Match               0.157813   0.017179   9.186  < 2e-16 ***
## P_5_Match               0.089720   0.018601   4.823 1.41e-06 ***
## P_6_Match                     NA         NA      NA       NA
## P_7_Match              -0.404534   0.186247  -2.172  0.02985 *
## P_8_Match               0.161849   0.017555   9.219  < 2e-16 ***
## P_9_Match               0.131481   0.015043   8.740  < 2e-16 ***
## P_10_Match                    NA         NA      NA       NA
## `BothBlack:BothMale`   -0.085108   0.104938  -0.811  0.41735
## `BothBlack:BothFemale`  0.127733   0.075072   1.701  0.08886 .
## `BothWhite:BothMale`   -0.005423   0.012427  -0.436  0.66253
## `BothWhite:BothFemale`  0.046934   0.010590   4.432 9.34e-06 ***
## `BothPOC:BothMale`      0.020046   0.139544   0.144  0.88577
## `BothPOC:BothFemale`   -0.107626   0.060076  -1.792  0.07321 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
```

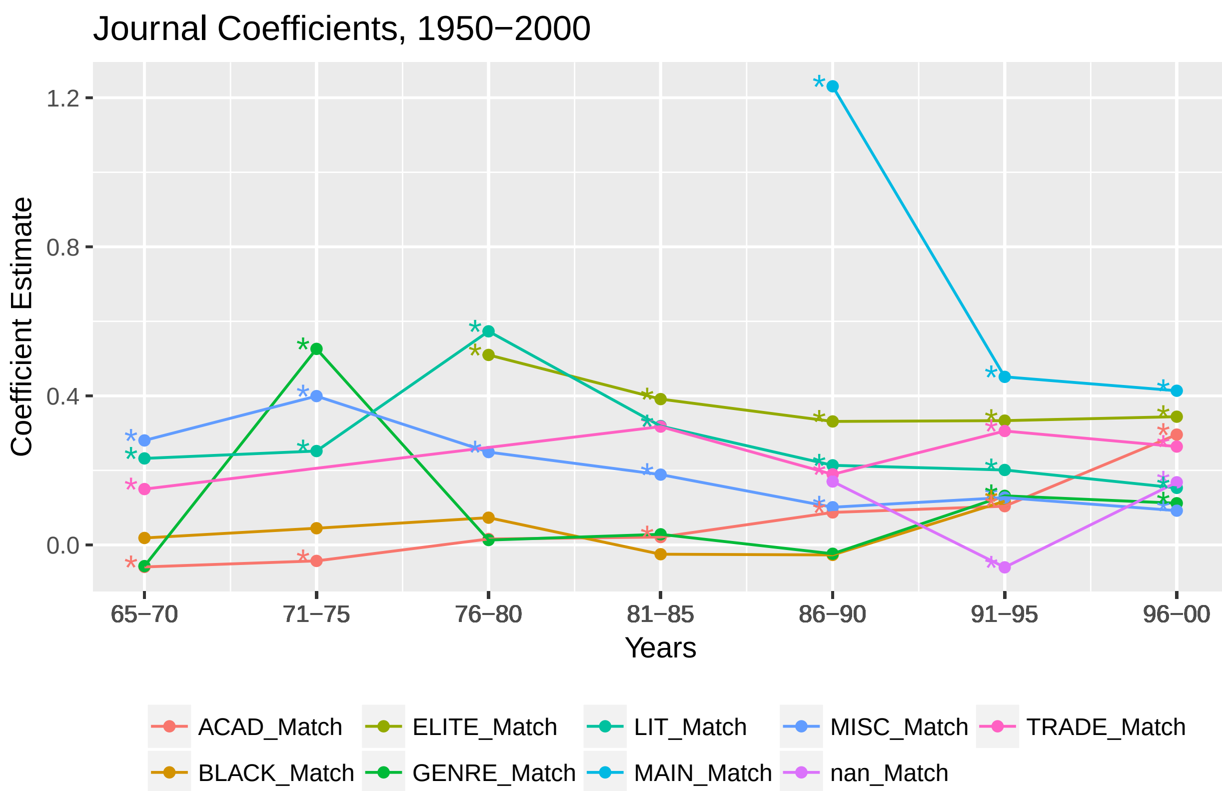
```
##     Null deviance: 55432  on 58995  degrees of freedom
## Residual deviance: 43633  on 58967  degrees of freedom
## AIC: 239763
##
## Number of Fisher Scoring iterations: 4
```

```
# add all model info into one df
model_over_time =
  bind_rows(cbind(cbind(cbind(Coefficient=rownames(summary(log_model1)$coefficients),data.frame(summary
            cbind(cbind(cbind(Coefficient=rownames(summary(log_model2)$coefficients),data.frame(summary
            cbind(cbind(cbind(Coefficient=rownames(summary(log_model3)$coefficients),data.frame(summary
            cbind(cbind(cbind(Coefficient=rownames(summary(log_model4)$coefficients),data.frame(summary
            cbind(cbind(cbind(Coefficient=rownames(summary(log_model5)$coefficients),data.frame(summary
            cbind(cbind(cbind(Coefficient=rownames(summary(log_model6)$coefficients),data.frame(summary
            cbind(cbind(cbind(Coefficient=rownames(summary(log_model7)$coefficients),data.frame(summary

model_over_time$Significance_5 = ifelse(model_over_time$Pr...z..<.05,'*','')
model_over_time$Significance_5_Meaning = ifelse(model_over_time$Pr...z..<.05,'Significant','Not Signifi
model_over_time$Significance_10 = ifelse(model_over_time$Pr...z..<.1,'*','')
model_over_time$Significance_10_Meaning = ifelse(model_over_time$Pr...z..<.1,'Significant','Not Signifi

# create a type column
model_over_time$Type = ifelse(model_over_time$Coefficient=='BothMale'|model_over_time$Coefficient=='Both
                            ifelse(model_over_time$Coefficient=='BothBlack'|model_over_time$Coefficient=
                                   ifelse(model_over_time$Coefficient=='(Intercept)','Intercept',
                                                 ifelse(model_over_time$Coefficient=='ACAD_Match'|model_
                                                          model_over_time$Coefficient=='ELITE_Match'|mod
                                                          model_over_time$Coefficient=='LIT_Match'|model
                                                          model_over_time$Coefficient=='MISC_Match'|mode
                                                          model_over_time$Coefficient=='nan_Match','Jou
                                                        ifelse(model_over_time$Coefficient=='P_0_Match'
                                                                 model_over_time$Coefficient=='P_2_Matc
                                                                 model_over_time$Coefficient=='P_4_Matc
                                                                 model_over_time$Coefficient=='P_6_Matc
                                                                 model_over_time$Coefficient=='P_8_Matc
                                                                 model_over_time$Coefficient=='P_10_Mat
```

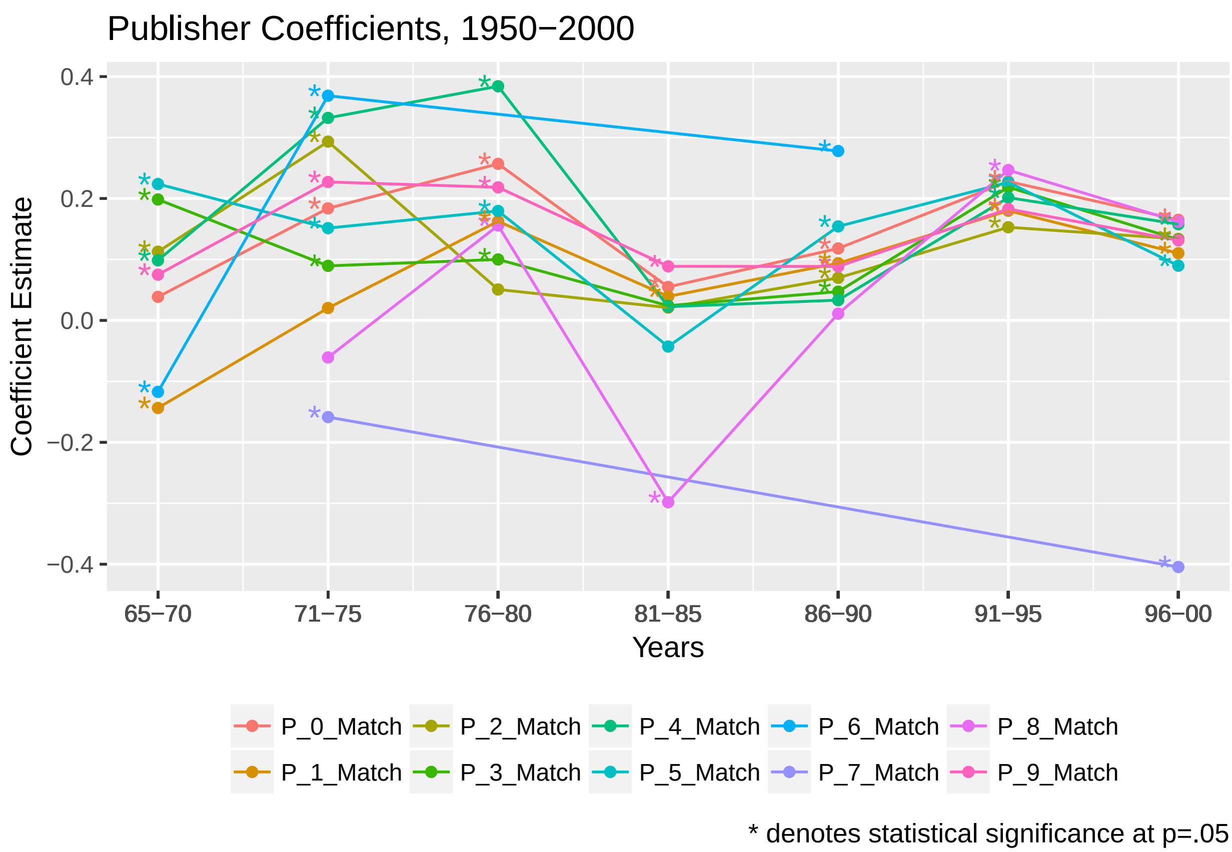

Gender Coefficients, 1950−2000
Coefficient Estimate
0.2
0.1
0.0
−0.1
65−70
71−75
76−80
81−85
86−90
91−95
96−00
Years
BothFemale
BothMale
* denotes statistical significance at p=.05



* denotes statistical significance at p=.05



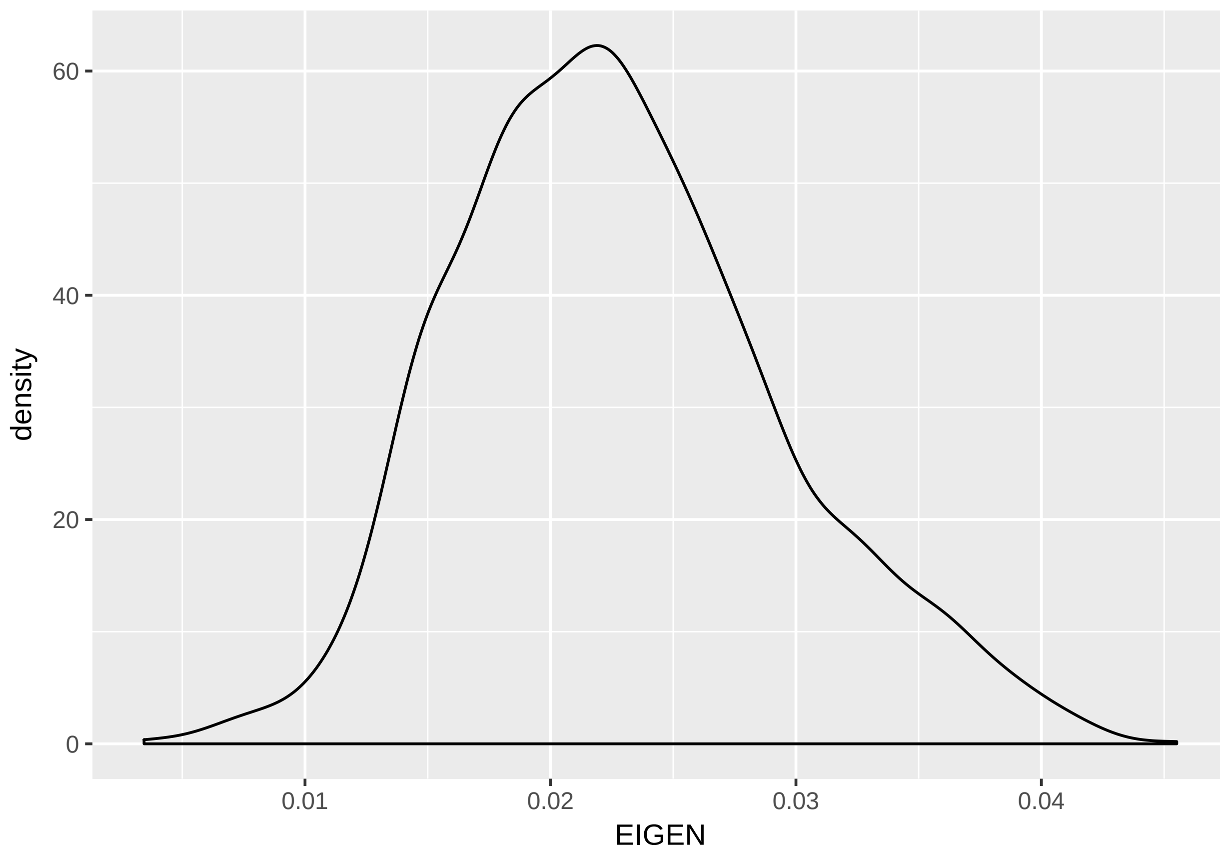
* denotes statistical significance at p=.05



Eigen Regression Analysis

```
# read in node data
n = read.csv('C:/Users/jyoung22.ADND/Documents/Network Analysis/REVIEWS_1965_2000_NODES_METADATA_EIGENS

# convert levels
n$GENDER = factor(ifelse(n$GENDER==0, 'Male',
                  ifelse(n$GENDER==1, 'Female',NA)),
                  levels=c('Male','Female'))

n$RACE = factor(ifelse(n$RACE==1, 'Black',
                  ifelse(n$RACE==2, 'POC',
                         ifelse(n$RACE==0, 'White',NA))),
                levels=c('White','Black','POC'))

n$P3 = factor(n$P3)

# create Year categories
n$HalfDecades = factor(ifelse(n$YEAR>64 & n$YEAR<71, '65-70',
                       ifelse(n$YEAR>70 & n$YEAR<76, '71-75',
                            ifelse(n$YEAR>75 & n$YEAR<81,'76-80',
                                    ifelse(n$YEAR>80 & n$YEAR<86, '81-85',
                                        ifelse(n$YEAR>85 & n$YEAR<91,'86-90',
                                               ifelse(n$YEAR>90 & n$YEAR<96, '91-95',
```

```r
                                                            ifelse(n$YEAR>95 | n$YEAR==00,'96-00',NA))))))))

n$Decades = factor(ifelse(n$YEAR<70, '60-69',
                       ifelse(n$YEAR>69 & n$YEAR<80, '70-79',
                            ifelse(n$YEAR>79 & n$YEAR<90,'80-89',
                                    ifelse(n$YEAR>89 & n$YEAR<=99, '90-99',
                                        ifelse(n$YEAR==100 | n$YEAR==00,'00-09',NA))))))

# RandomHouse titles
rh = c("are you there god? it's me, margaret", 'a tidewater morning', 'red square (kandinsky). audio ve

# Major publishers
mp = c('Random House', 'Random Hous', 'Random House audio',
       'Alfred A. Knopf', 'Knopf', 'Knopf Doubleday' ,'Knopf Doubleday Publishing','Knopf,',
       'Ballantine', 'Ballantine Books',
       'Modern Library',
       'Library of America',
       'Pantheon','Pantheon Books')

n$RandomHouse = factor(ifelse(n$TITLE%in%rh,'RH','Other'), levels=c('Other','RH'))
n$MajorPub = factor(ifelse(tolower(trimws(n$PUBLISHER))%in%tolower(mp),'Major','Other'), levels=c('Othe



ggplot(n, aes(x=EIGEN))+geom_density()
```



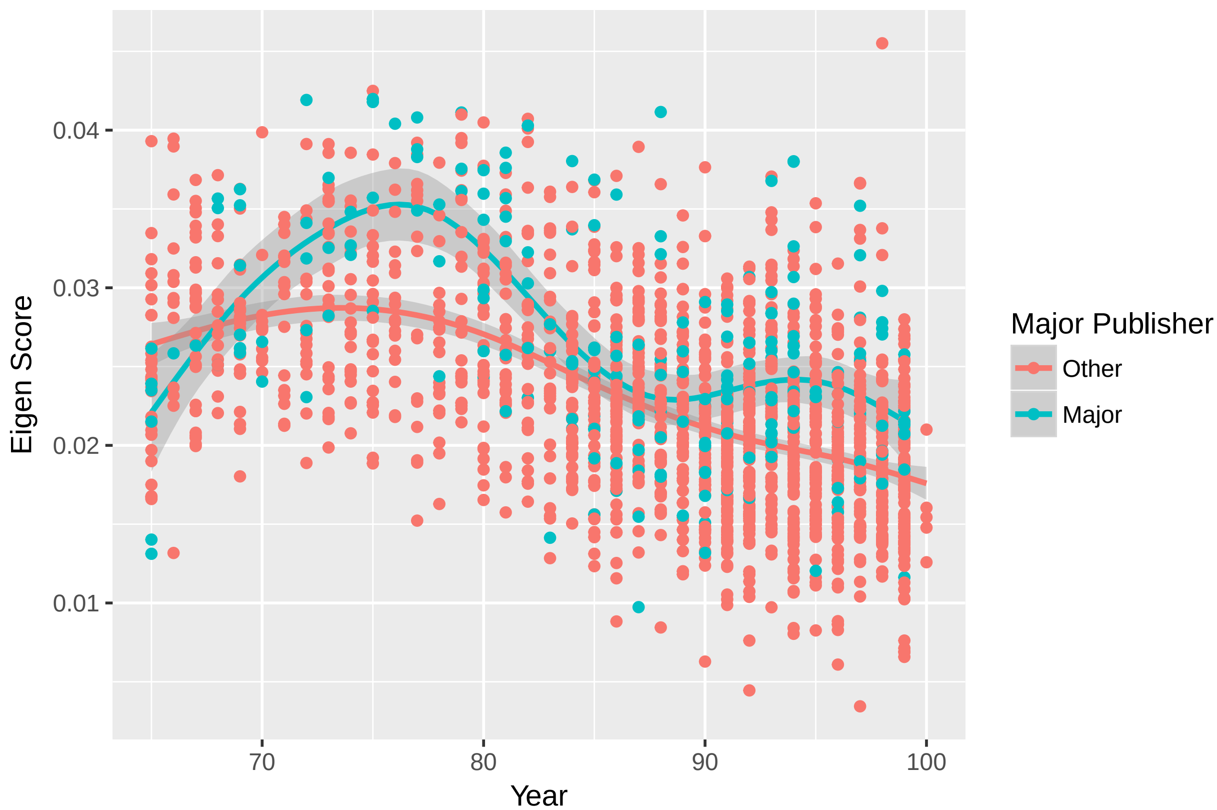
```r
# normal assumption met
```

```r
ggplot(na.omit(n), aes(x=YEAR, y=EIGEN, colour=GENDER))+
  facet_wrap(~RACE)+
  geom_smooth()+
  geom_point()+
  labs(main='Eigen Scores by Race and Gender',
       caption='LOESS smoothing used',
       x='Year',
       y='Eigen Score',
       colour='Gender')
```

```
## `geom_smooth()` using method = 'loess'
```


```r
ggplot(na.omit(n), aes(x=YEAR, y=EIGEN, colour=MajorPub))+
#  facet_wrap(~GENDER)+
  geom_smooth()+
  geom_point()+
  labs(main='Eigen Scores by Publisher',
       caption='GAM smoothing used',
       x='Year',
       y='Eigen Score',
       colour='Major Publisher')
```

```
## `geom_smooth()` using method = 'gam'
```



GAM smoothing used

```
ggplot(na.omit(n), aes(x=YEAR, y=EIGEN, colour=MajorPub))+
  facet_wrap(~GENDER)+
  geom_smooth()+
  geom_point()+
  labs(main='Eigen Scores by Publisher',
       caption='LOESS smoothing used',
       x='Year',
       y='Eigen Score',
       colour='Major Publisher')
```

```
## `geom_smooth()` using method = 'loess'
```

LOESS smoothing used

```
# Major Pub only
n_model = train(EIGEN~ GENDER + RACE  + HalfDecades + MajorPub + # main effects
                  RACE:GENDER + RACE:HalfDecades + RACE:MajorPub + GENDER:HalfDecades + GENDER:MajorPub
                  GENDER:RACE:HalfDecades + GENDER:RACE:MajorPub,# + GENDER:HalfDecades:MajorPub + RACE
                  #GENDER:RACE:HalfDecades:MajorPub, # 3rd level interactions # not significant
                   data = na.omit(n),
                   method="glm",
                   family="gaussian",
                   trControl=trainControl(
                     method='cv',number=10
                   ))
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading
```

```
            summary(n_model)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
##        Min          1Q      Median          3Q         Max
## -0.0167914  -0.0035985  -0.0004183   0.0031519   0.0275451
##
## Coefficients: (2 not defined because of singularities)
##                                             Estimate Std. Error t value
## (Intercept)                                2.751e-02  5.655e-04  48.649
## GENDERFemale                              -1.778e-03  1.166e-03  -1.525
## RACEBlack                                  1.409e-03  2.467e-03   0.571
## RACEPOC                                   -5.465e-03  5.399e-03  -1.012
## `HalfDecades71-75`                         2.255e-03  8.529e-04   2.644
## `HalfDecades76-80`                         1.757e-03  8.600e-04   2.043
## `HalfDecades81-85`                        -1.568e-03  7.454e-04  -2.104
## `HalfDecades86-90`                        -4.438e-03  7.014e-04  -6.327
## `HalfDecades91-95`                        -6.570e-03  6.791e-04  -9.675
## `HalfDecades96-00`                        -9.033e-03  7.137e-04 -12.656
## MajorPubMajor                             -1.002e-03  1.342e-03  -0.747
## `GENDERFemale:RACEBlack`                   7.275e-04  5.996e-03   0.121
## `GENDERFemale:RACEPOC`                    -1.728e-04  2.705e-03  -0.064
## `RACEBlack:HalfDecades71-75`              -1.331e-03  4.013e-03  -0.332
## `RACEPOC:HalfDecades71-75`                 5.203e-04  7.641e-03   0.068
## `RACEBlack:HalfDecades76-80`               1.547e-03  5.944e-03   0.260
## `RACEPOC:HalfDecades76-80`                 5.107e-03  7.641e-03   0.668
## `RACEBlack:HalfDecades81-85`              -4.648e-03  3.824e-03  -1.216
## `RACEPOC:HalfDecades81-85`                 4.317e-03  6.049e-03   0.714
## `RACEBlack:HalfDecades86-90`              -6.172e-03  3.467e-03  -1.780
## `RACEPOC:HalfDecades86-90`                 1.997e-03  6.613e-03   0.302
## `RACEBlack:HalfDecades91-95`              -1.628e-03  2.878e-03  -0.566
## `RACEPOC:HalfDecades91-95`                 8.020e-03  5.621e-03   1.427
## `RACEBlack:HalfDecades96-00`              -1.488e-04  3.112e-03  -0.048
## `RACEPOC:HalfDecades96-00`                 8.555e-03  5.881e-03   1.455
## `RACEBlack:MajorPubMajor`                 -8.516e-03  4.264e-03  -1.997
## `RACEPOC:MajorPubMajor`                   -4.058e-03  4.112e-03  -0.987
## `GENDERFemale:HalfDecades71-75`           -1.254e-03  1.693e-03  -0.741
```

```
## `GENDERFemale:HalfDecades76-80`           -1.671e-03  1.544e-03  -1.082
## `GENDERFemale:HalfDecades81-85`           -1.177e-03  1.420e-03  -0.829
## `GENDERFemale:HalfDecades86-90`            8.271e-05  1.319e-03   0.063
## `GENDERFemale:HalfDecades91-95`           -4.758e-04  1.277e-03  -0.373
## `GENDERFemale:HalfDecades96-00`            1.268e-03  1.314e-03   0.964
## `GENDERFemale:MajorPubMajor`               1.479e-03  9.124e-04   1.621
## `HalfDecades71-75:MajorPubMajor`           5.016e-03  2.009e-03   2.497
## `HalfDecades76-80:MajorPubMajor`           6.707e-03  1.933e-03   3.469
## `HalfDecades81-85:MajorPubMajor`           3.424e-03  1.697e-03   2.017
## `HalfDecades86-90:MajorPubMajor`           5.444e-04  1.639e-03   0.332
## `HalfDecades91-95:MajorPubMajor`           4.933e-03  1.566e-03   3.150
## `HalfDecades96-00:MajorPubMajor`           4.601e-03  1.665e-03   2.763
## `GENDERFemale:RACEBlack:HalfDecades71-75` -5.349e-03  7.367e-03  -0.726
## `GENDERFemale:RACEPOC:HalfDecades71-75`           NA         NA       NA
## `GENDERFemale:RACEBlack:HalfDecades76-80` -2.783e-03  8.581e-03  -0.324
## `GENDERFemale:RACEPOC:HalfDecades76-80`    1.115e-03  7.279e-03   0.153
## `GENDERFemale:RACEBlack:HalfDecades81-85`  2.968e-03  6.976e-03   0.425
## `GENDERFemale:RACEPOC:HalfDecades81-85`   -1.835e-03  6.634e-03  -0.277
## `GENDERFemale:RACEBlack:HalfDecades86-90`  2.258e-03  6.713e-03   0.336
## `GENDERFemale:RACEPOC:HalfDecades86-90`    9.950e-03  5.280e-03   1.884
## `GENDERFemale:RACEBlack:HalfDecades91-95`  8.567e-04  6.285e-03   0.136
## `GENDERFemale:RACEPOC:HalfDecades91-95`    7.074e-04  3.285e-03   0.215
## `GENDERFemale:RACEBlack:HalfDecades96-00` -1.770e-03  6.491e-03  -0.273
## `GENDERFemale:RACEPOC:HalfDecades96-00`           NA         NA       NA
## `GENDERFemale:RACEBlack:MajorPubMajor`     9.543e-03  5.211e-03   1.831
## `GENDERFemale:RACEPOC:MajorPubMajor`       2.431e-03  4.818e-03   0.505
##                                           Pr(>|t|)
## (Intercept)                                < 2e-16 ***
## GENDERFemale                              0.127551
## RACEBlack                                 0.568079
## RACEPOC                                   0.311539
## `HalfDecades71-75`                        0.008273 **
## `HalfDecades76-80`                        0.041248 *
## `HalfDecades81-85`                        0.035558 *
## `HalfDecades86-90`                        3.17e-10 ***
## `HalfDecades91-95`                         < 2e-16 ***
## `HalfDecades96-00`                         < 2e-16 ***
## MajorPubMajor                             0.455234
## `GENDERFemale:RACEBlack`                  0.903445
## `GENDERFemale:RACEPOC`                    0.949060
## `RACEBlack:HalfDecades71-75`              0.740236
## `RACEPOC:HalfDecades71-75`                0.945713
## `RACEBlack:HalfDecades76-80`              0.794633
## `RACEPOC:HalfDecades76-80`                0.504027
## `RACEBlack:HalfDecades81-85`              0.224270
## `RACEPOC:HalfDecades81-85`                0.475471
## `RACEBlack:HalfDecades86-90`              0.075248 .
## `RACEPOC:HalfDecades86-90`                0.762747
## `RACEBlack:HalfDecades91-95`              0.571696
## `RACEPOC:HalfDecades91-95`                0.153825
## `RACEBlack:HalfDecades96-00`              0.961867
## `RACEPOC:HalfDecades96-00`                0.145926
## `RACEBlack:MajorPubMajor`                 0.045961 *
## `RACEPOC:MajorPubMajor`                   0.323758
```

```
## `GENDERFemale:HalfDecades71-75`           0.458960
## `GENDERFemale:HalfDecades76-80`           0.279394
## `GENDERFemale:HalfDecades81-85`           0.407189
## `GENDERFemale:HalfDecades86-90`           0.950001
## `GENDERFemale:HalfDecades91-95`           0.709392
## `GENDERFemale:HalfDecades96-00`           0.335010
## `GENDERFemale:MajorPubMajor`              0.105100
## `HalfDecades71-75:MajorPubMajor`          0.012606 *
## `HalfDecades76-80:MajorPubMajor`          0.000535 ***
## `HalfDecades81-85:MajorPubMajor`          0.043821 *
## `HalfDecades86-90:MajorPubMajor`          0.739825
## `HalfDecades91-95:MajorPubMajor`          0.001662 **
## `HalfDecades96-00:MajorPubMajor`          0.005787 **
## `GENDERFemale:RACEBlack:HalfDecades71-75` 0.467897
## `GENDERFemale:RACEPOC:HalfDecades71-75`         NA
## `GENDERFemale:RACEBlack:HalfDecades76-80` 0.745772
## `GENDERFemale:RACEPOC:HalfDecades76-80`   0.878303
## `GENDERFemale:RACEBlack:HalfDecades81-85` 0.670569
## `GENDERFemale:RACEPOC:HalfDecades81-85`   0.782101
## `GENDERFemale:RACEBlack:HalfDecades86-90` 0.736692
## `GENDERFemale:RACEPOC:HalfDecades86-90`   0.059679 .
## `GENDERFemale:RACEBlack:HalfDecades91-95` 0.891596
## `GENDERFemale:RACEPOC:HalfDecades91-95`   0.829558
## `GENDERFemale:RACEBlack:HalfDecades96-00` 0.785099
## `GENDERFemale:RACEPOC:HalfDecades96-00`         NA
## `GENDERFemale:RACEBlack:MajorPubMajor`    0.067223 .
## `GENDERFemale:RACEPOC:MajorPubMajor`      0.613928
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 2.882639e-05)
##
##     Null deviance: 0.076543  on 1771  degrees of freedom
## Residual deviance: 0.049581  on 1720  degrees of freedom
## AIC: -13443
##
## Number of Fisher Scoring iterations: 2
```

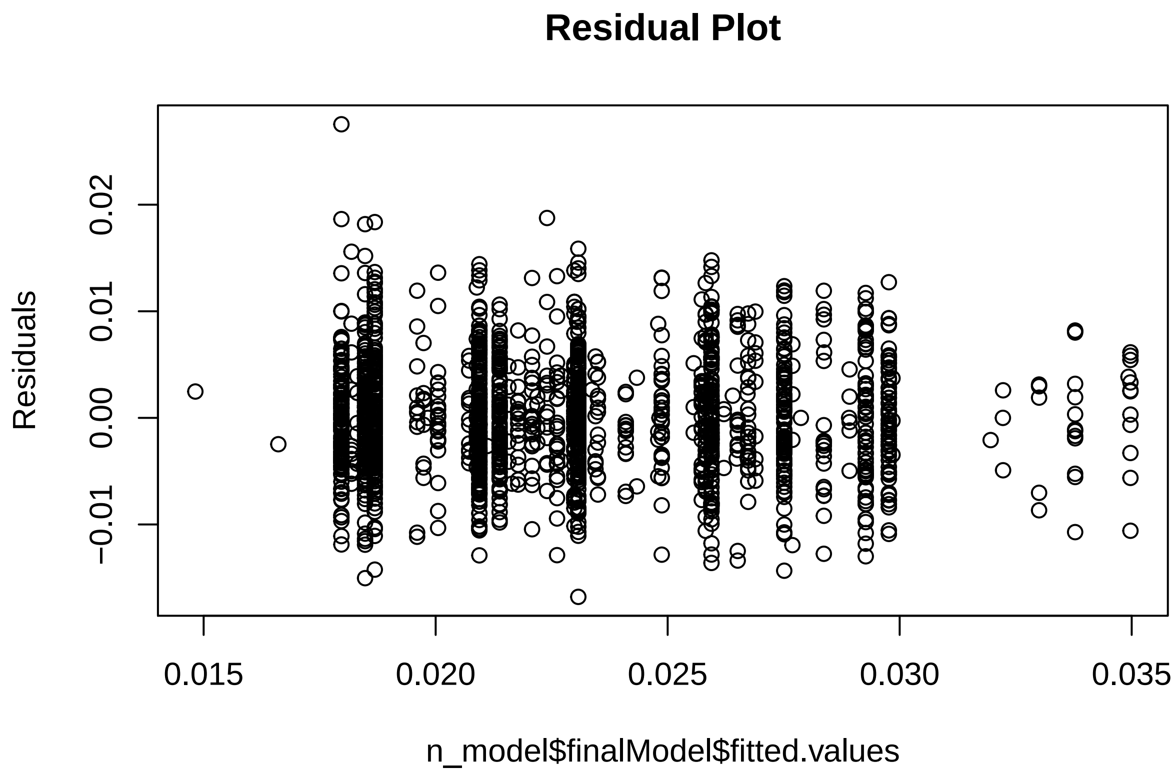
```
            n_model
```

```
## Generalized Linear Model
##
## 1772 samples
##    4 predictor
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 1594, 1595, 1594, 1596, 1596, 1596, ...
## Resampling results:
##
##   RMSE         Rsquared   MAE
##   0.005424452  0.3227488  0.004250186
```

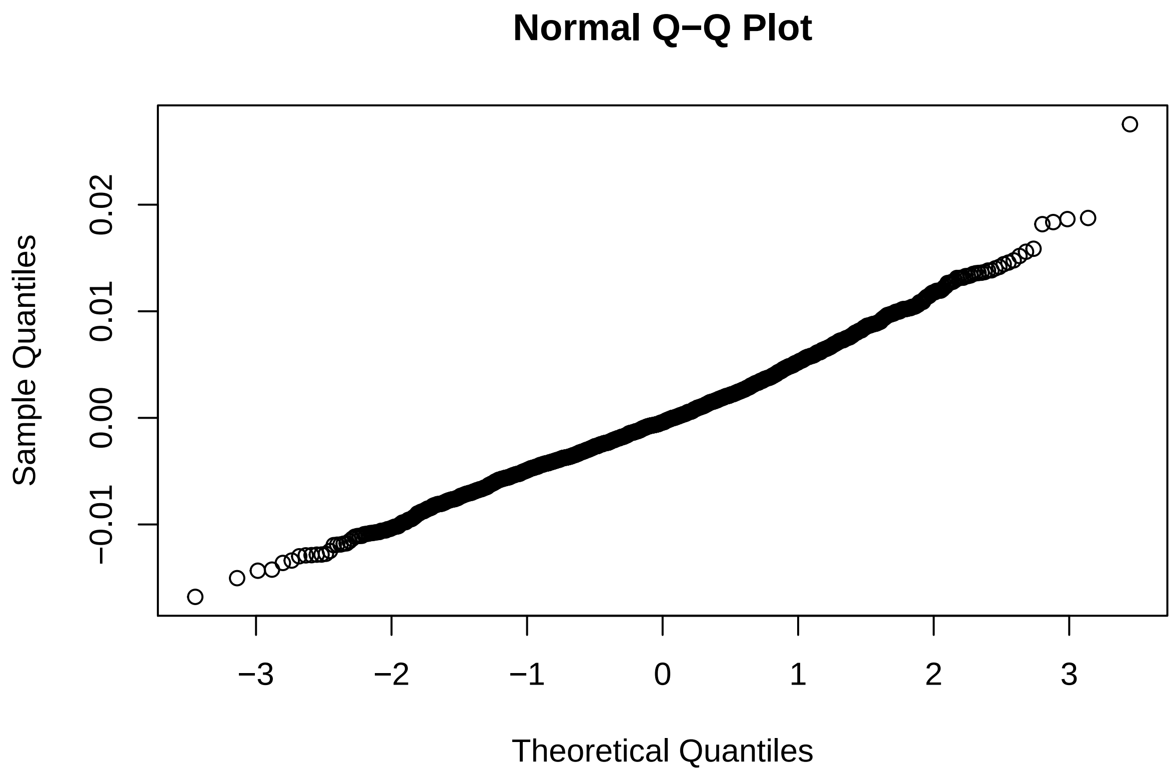
```r
varImp(n_model)
```

```
## glm variable importance
##
##   only 20 most important variables shown (out of 51)
##
##                                          Overall
## `HalfDecades96-00`                       100.000
## `HalfDecades91-95`                        76.359
## `HalfDecades86-90`                        49.805
## `HalfDecades76-80:MajorPubMajor`          27.136
## `HalfDecades91-95:MajorPubMajor`          24.603
## `HalfDecades96-00:MajorPubMajor`          21.536
## `HalfDecades71-75`                        20.590
## `HalfDecades71-75:MajorPubMajor`          19.428
## `HalfDecades81-85`                        16.305
## `HalfDecades76-80`                        15.821
## `HalfDecades81-85:MajorPubMajor`          15.621
## `RACEBlack:MajorPubMajor`                 15.461
## `GENDERFemale:RACEPOC:HalfDecades86-90`   14.567
## `GENDERFemale:RACEBlack:MajorPubMajor`    14.146
## `RACEBlack:HalfDecades86-90`              13.739
## `GENDERFemale:MajorPubMajor`              12.481
## GENDERFemale                              11.713
## `RACEPOC:HalfDecades96-00`                11.159
## `RACEPOC:HalfDecades91-95`                10.937
## `RACEBlack:HalfDecades81-85`               9.263
```

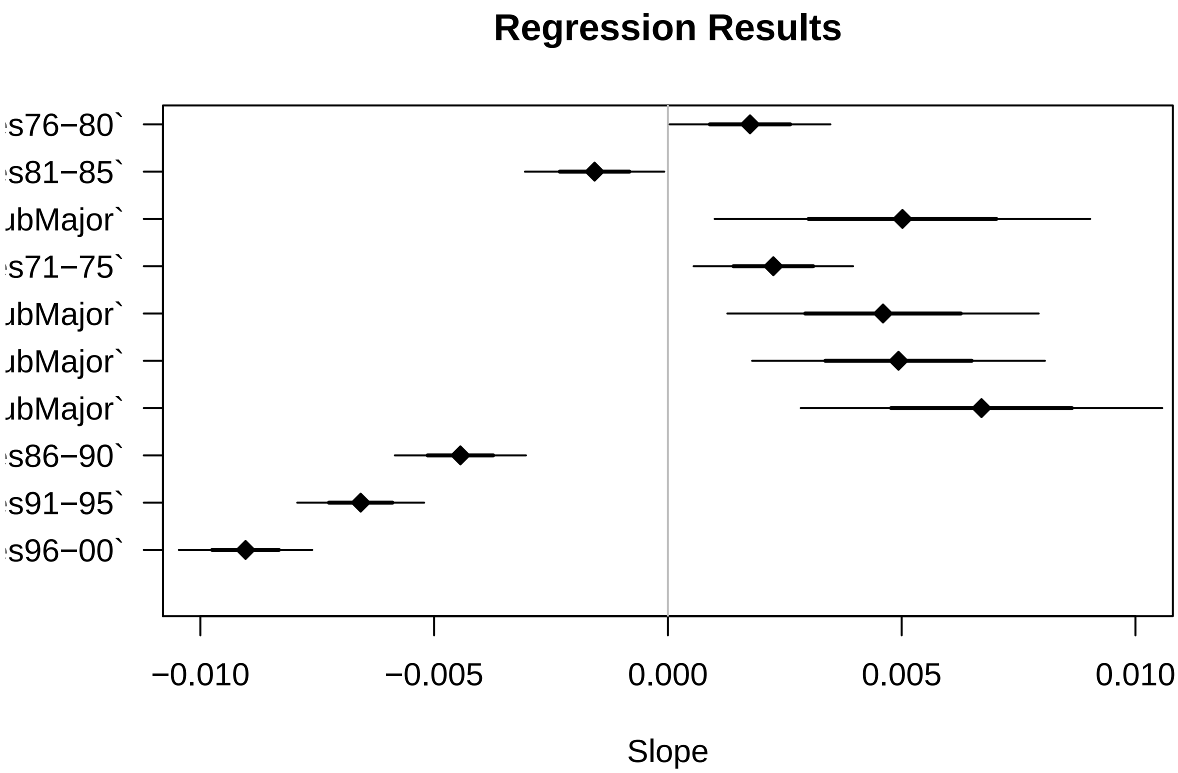
```r
plot(y=n_model$finalModel$residuals, x=n_model$finalModel$fitted.values, ylab='Residuals', 
```




```
qqnorm(n_model$finalModel$residuals)# qq
```




```
regplot(n_model)
```




```
# Random House only
n1_model = train(EIGEN~ GENDER + RACE  + HalfDecades + RandomHouse + # main effects
                  RACE:GENDER + RACE:HalfDecades + RACE:RandomHouse + GENDER:HalfDecades + GENDER:Random
                  GENDER:RACE:HalfDecades + GENDER:RACE:RandomHouse,# + GENDER:HalfDecades:RandomHouse +
                  #GENDER:RACE:HalfDecades:RandomHouse, # 3rd level interactions # not significant
                   data = na.omit(n),
                   method="glm",
                   family="gaussian",
                   trControl=trainControl(
                     method='cv',number=10
                   ))

            summary(n1_model)
            n1_model

            varImp(n1_model)

            plot(y=n1_model$finalModel$residuals, x=n1_model$finalModel$fitted.values, ylab='Residuals'

            qqnorm(n1_model$finalModel$residuals)# qq


# P3 only
n3_model = train(EIGEN~ GENDER + RACE  + HalfDecades + P3 + # main effects
                  RACE:GENDER + RACE:HalfDecades + RACE:P3 + GENDER:HalfDecades + GENDER:P3 + HalfDecade
                  #GENDER:RACE:HalfDecades + GENDER:RACE:P3,# + GENDER:HalfDecades:P3,# + RACE:HalfDeca
```

```
               #GENDER:RACE:HalfDecades:P3, # 3rd level interactions # not significant
                data = na.omit(n),
                method="glm",
                family="gaussian",
                trControl=trainControl(
                  method='cv',number=10
                ))

summary(n3_model)
n3_model

varImp(n3_model)

plot(y=n3_model$finalModel$residuals, x=n3_model$finalModel$fitted.values, ylab='Residuals'
qqnorm(n3_model$finalModel$residuals)# qq
```

The above varImp function automatically scales the variable importance out of 100 to make it easier to interpret.