Logistic Regression Analysis

Anne Wagner

2024-03-02

Introduction (4 points)

[TIP FOR YOUR ANALYSIS: You should briefly discuss both the data you will analyze and the benefit of such an analysis]

```
callbackData<-read.csv("callbackData.csv")
```

[NOTE: When adding code chunks, they should take the format seen below]

```
# The line above must contain {r ...} to specify the type of code that will be included
#  and a unique identifier for the code block in place of "..."
# If repeated names are used, Rstudio will throw an error when trying to compile
```

In this analysis we look at data gathered from an experiment to assess callbacks on job applications for fictional candidates. The submitted resumes varied wildly, but specifically included names that would be perceived as either male or female, and either black or white. The original purpose of this data gathering was to assess whether there was evidence that callbacks were impacted by the perceived sex or race of the applicant.

The Variables (8 points)

[TIP FOR YOUR ANALYSIS: List all the variables, with some information to summarize their distribution of values, also specify the type of variable as either categorical or numeric. Numeric variables can be plotted in a histogram, while categorical variables can be summarized with a table, but other options are available if you wish to explore those.]

The variables available for this analysis are:

Response:

- Callback (Binary/categorical): This records whether the submitted resume received a callback for an interview. This is the variable we seek to predict with our logistic regression model.

```
table(callbackData$call)
```

```
##
##    0    1
## 4478  392
```

Predictors:

```
names(callbackData)
```

```
##  [1] "call"           "city"           "college"        "yearsexp"
##  [5] "honors"         "militaryexp"    "emailincluded"  "sex"
```

```
## [9] "race"           "computerskills" "workinschool"
```

- city (Binary/categorical): This variable looked at the city in which the job being applied to was located. There are only two cities in this study, Chicago and Boston.

```r
table(callbackData$city)
```

```
##
##  Boston Chicago
##    2166    2704
```

- college (Binary/categorical): This variable takes value 1 for resumes that list a college degree, and 0 otherwise.

```r
table(callbackData$college)
```

```
##
##    0    1
## 1366 3504
```

- yearsexp (Numeric): This variable lists the years of relevant working experience of the applicant. It has mean 7.84 years and standard deviation of 5.04 years.

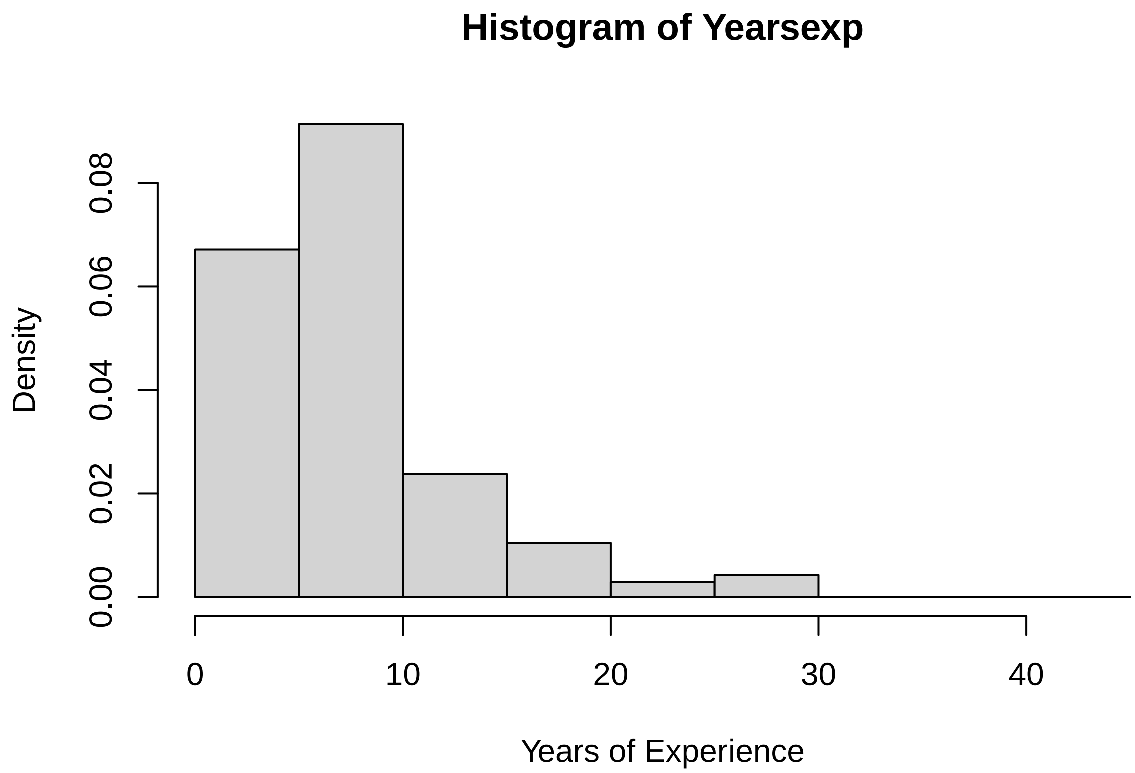
```r
mean(callbackData$yearsexp)
```

```
## [1] 7.842916
```

```r
sd(callbackData$yearsexp)
```

```
## [1] 5.044612
```

```r
hist(callbackData$yearsexp, main='Histogram of Yearsexp', xlab='Years of Experience', freq=FALSE)
```



- honors (Binary/categorical): This variable records whether or not the applicant's resume listed any relevant honors or not.

```
table(callbackData$honors)
```

```
##
##    0    1
## 4613  257
```

- militaryexp (Binary/categorical): This variable records whether or not the applicant's resume listed military experience.

```
table(callbackData$militaryexp)
```

```
##
##    0    1
## 4397  473
```

- emailincluded (Binary/categorical): This variable records whether or not the resume included a contact email. This dataset is… fairly old.

```
table(callbackData$emailincluded)
```

```
##
##    0    1
## 2536 2334
```

- sex (Binary/categorical): This variable records the perceived sex of the name of the applicant. It takes values of "Male" and "Female".

```r
table(callbackData$sex)
```

```
##
## Female   Male
##   3746   1124
```

- race (Binary/categorical): This variable records the perceived race of the name of the applicant. It takes values of “White” and “Black”.

```r
table(callbackData$race)
```

```
##
## Black White
##  2435  2435
```

- computerskills (Binary/categorical): This variable records whether or not the resume included any relevant computer skills.

```r
table(callbackData$computerskills)
```

```
##
##    0    1
##  874 3996
```

- workinschool (Binary/categorical): This variable records whether or not the resume included a work history during education.

```r
table(callbackData$workinschool)
```

```
##
##    0    1
## 2145 2725
```

Model Fitting (12 points)

[NOTE: You should detail the methods used to fit the model, and any other models that were explored in the process.]

Here we utilize the step function with AIC scoring to determine the best model. To produce and test a variety of models, we utilize both forward and backward step functions. Due to the specific purpose of this analysis, all models will include the variables ‘sex’ and ‘race’, as the impact of these variables on callbacks was of specific interest to the researchers.

```r
fullmodel=glm(call~city+college+yearsexp+honors+militaryexp+
                emailincluded+sex+race+computerskills+
                workinschool, data=callbackData, family = "binomial")
basicmodel=glm(call~race+sex, data=callbackData, family = "binomial")
```

```r
step(basicmodel,
     scope=list(lower=basicmodel, upper=fullmodel),
     direction='forward', steps=25)
```

```
## Start:  AIC=2714.95
## call ~ race + sex
##
##                 Df Deviance    AIC
## + honors         1   2689.2 2697.2
## + city           1   2691.3 2699.3
## + yearsexp       1   2692.4 2700.4
```

```
## + computerskills  1   2704.4 2712.4
## + workinschool    1   2705.1 2713.1
## + emailincluded   1   2705.8 2713.8
## <none>                2708.9 2714.9
## + militaryexp     1   2707.2 2715.2
## + college         1   2708.8 2716.8
##
## Step:  AIC=2697.15
## call ~ race + sex + honors
##
##                  Df Deviance    AIC
## + city            1   2670.9 2680.9
## + yearsexp        1   2677.3 2687.3
## + computerskills  1   2685.3 2695.3
## + workinschool    1   2685.9 2695.9
## <none>                2689.2 2697.2
## + militaryexp     1   2687.2 2697.2
## + emailincluded   1   2687.4 2697.4
## + college         1   2688.7 2698.7
##
## Step:  AIC=2680.94
## call ~ race + sex + honors + city
##
##                  Df Deviance    AIC
## + yearsexp        1   2664.5 2676.5
## + militaryexp     1   2668.2 2680.2
## + workinschool    1   2668.5 2680.5
## + emailincluded   1   2668.7 2680.7
## + computerskills  1   2668.7 2680.7
## <none>                2670.9 2680.9
## + college         1   2670.1 2682.1
##
## Step:  AIC=2676.51
## call ~ race + sex + honors + city + yearsexp
##
##                  Df Deviance    AIC
## + emailincluded   1   2662.2 2676.2
## <none>                2664.5 2676.5
## + computerskills  1   2662.7 2676.7
## + workinschool    1   2663.0 2677.0
## + militaryexp     1   2663.4 2677.4
## + college         1   2664.1 2678.1
##
## Step:  AIC=2676.24
## call ~ race + sex + honors + city + yearsexp + emailincluded
##
##                  Df Deviance    AIC
## + workinschool    1   2659.0 2675.0
## + computerskills  1   2659.0 2675.0
## + militaryexp     1   2659.6 2675.6
## <none>                2662.2 2676.2
## + college         1   2661.9 2677.9
##
## Step:  AIC=2675.01
```

```
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool
##
##                  Df Deviance    AIC
## + militaryexp     1   2656.5 2674.5
## <none>                2659.0 2675.0
## + computerskills  1   2657.1 2675.1
## + college         1   2658.8 2676.8
##
## Step:  AIC=2674.48
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp
##
##                  Df Deviance    AIC
## <none>                2656.5 2674.5
## + computerskills  1   2654.8 2674.8
## + college         1   2656.3 2676.3
```

```
##
## Call:  glm(formula = call ~ race + sex + honors + city + yearsexp +
##     emailincluded + workinschool + militaryexp, family = "binomial",
##     data = callbackData)
##
## Coefficients:
##   (Intercept)      raceWhite        sexMale         honors    cityChicago
##      -2.62363        0.44247       -0.19408        0.74917       -0.42868
##      yearsexp  emailincluded   workinschool    militaryexp
##       0.01818        0.28278       -0.19892       -0.33388
##
## Degrees of Freedom: 4869 Total (i.e. Null);  4861 Residual
## Null Deviance:       2727
## Residual Deviance: 2656  AIC: 2674
```

```r
step(fullmodel,
     scope=list(lower=basicmodel, upper=fullmodel),
     direction='backward', steps=25)
```

```
## Start:  AIC=2676.57
## call ~ city + college + yearsexp + honors + militaryexp + emailincluded +
##     sex + race + computerskills + workinschool
##
##                  Df Deviance    AIC
## - college         1   2654.8 2674.8
## - computerskills  1   2656.3 2676.3
## - workinschool    1   2656.3 2676.3
## <none>                2654.6 2676.6
## - militaryexp     1   2656.9 2676.9
## - yearsexp        1   2657.4 2677.4
## - emailincluded   1   2660.9 2680.9
## - city            1   2667.8 2687.8
## - honors          1   2669.0 2689.0
##
## Step:  AIC=2674.78
## call ~ city + yearsexp + honors + militaryexp + emailincluded +
##     sex + race + computerskills + workinschool
```

```
##
##                   Df Deviance    AIC
## - computerskills  1   2656.5 2674.5
## - workinschool    1   2656.6 2674.6
## <none>                2654.8 2674.8
## - militaryexp     1   2657.1 2675.1
## - yearsexp        1   2657.8 2675.8
## - emailincluded   1   2661.3 2679.3
## - city            1   2667.8 2685.8
## - honors          1   2669.1 2687.1
##
## Step:  AIC=2674.48
## call ~ city + yearsexp + honors + militaryexp + emailincluded +
##     sex + race + workinschool
##
##                  Df Deviance    AIC
## <none>               2656.5 2674.5
## - militaryexp    1   2659.0 2675.0
## - yearsexp       1   2659.5 2675.5
## - workinschool   1   2659.6 2675.6
## - emailincluded  1   2662.2 2678.2
## - city           1   2670.6 2686.6
## - honors         1   2671.0 2687.0
```

```
##
## Call:  glm(formula = call ~ city + yearsexp + honors + militaryexp +
##     emailincluded + sex + race + workinschool, family = "binomial",
##     data = callbackData)
##
## Coefficients:
##   (Intercept)    cityChicago       yearsexp         honors    militaryexp
##      -2.62363       -0.42868        0.01818        0.74917       -0.33388
## emailincluded        sexMale      raceWhite   workinschool
##       0.28278       -0.19408        0.44247       -0.19892
##
## Degrees of Freedom: 4869 Total (i.e. Null);  4861 Residual
## Null Deviance:       2727
## Residual Deviance: 2656  AIC: 2674
```

As can be observed from the step functions, both models returned are identical (with some reordering of variables). Once again, as the variables 'sex' and 'race' were of interest, we will explore whether interactions with these terms are useful for the model.

[NOTE: If you wish to explore a large number of interaction terms, you can utilize standard mathematical notation of a*(b+c+d) to see interactions between the variable a and variables b, c, and d. This is done for both the race and sex variables as seen below.]

```r
current.model = glm(formula = call ~ race + sex + honors + city + yearsexp +
    emailincluded + workinschool + militaryexp, family = "binomial",
    data = callbackData)
interactions.model=glm(formula = call ~ (race + sex) * (honors + city + yearsexp +
    emailincluded + workinschool + militaryexp), family = "binomial",
    data = callbackData)
```

```r
step(current.model,
     scope=list(lower=current.model, upper=interactions.model),
```

```r
    direction='forward', steps=25)
```

```
## Start:  AIC=2674.48
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp
##
##                       Df Deviance    AIC
## + sex:city             1   2640.8 2660.8
## + sex:yearsexp         1   2651.9 2671.9
## + sex:honors           1   2652.3 2672.3
## + sex:militaryexp      1   2653.0 2673.0
## <none>                     2656.5 2674.5
## + race:emailincluded   1   2655.5 2675.5
## + race:militaryexp     1   2655.6 2675.6
## + race:workinschool    1   2656.2 2676.2
## + sex:emailincluded    1   2656.3 2676.3
## + race:honors          1   2656.3 2676.3
## + sex:workinschool     1   2656.3 2676.3
## + race:yearsexp        1   2656.4 2676.4
## + race:city            1   2656.4 2676.4
##
## Step:  AIC=2660.77
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + sex:city
##
##                       Df Deviance    AIC
## + sex:militaryexp      1   2635.5 2657.5
## + sex:honors           1   2637.7 2659.7
## <none>                     2640.8 2660.8
## + sex:yearsexp         1   2639.6 2661.6
## + race:emailincluded   1   2639.8 2661.8
## + race:militaryexp     1   2639.9 2661.9
## + race:workinschool    1   2640.5 2662.5
## + race:honors          1   2640.6 2662.6
## + race:yearsexp        1   2640.7 2662.7
## + race:city            1   2640.7 2662.7
## + sex:emailincluded    1   2640.8 2662.8
## + sex:workinschool     1   2640.8 2662.8
##
## Step:  AIC=2657.49
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + sex:city + sex:militaryexp
##
##                       Df Deviance    AIC
## + sex:honors           1   2633.4 2657.4
## <none>                     2635.5 2657.5
## + sex:emailincluded    1   2634.1 2658.1
## + race:emailincluded   1   2634.6 2658.6
## + race:militaryexp     1   2634.6 2658.6
## + race:workinschool    1   2635.2 2659.2
## + sex:yearsexp         1   2635.3 2659.3
## + race:honors          1   2635.3 2659.3
## + sex:workinschool     1   2635.4 2659.4
## + race:city            1   2635.4 2659.4
```

```
## + race:yearsexp       1   2635.4 2659.4
##
## Step:  AIC=2657.37
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + sex:city + sex:militaryexp +
##     sex:honors
##
##                      Df Deviance    AIC
## <none>                    2633.4 2657.4
## + sex:emailincluded   1   2631.8 2657.8
## + race:militaryexp    1   2632.5 2658.5
## + race:emailincluded  1   2632.6 2658.6
## + sex:yearsexp        1   2632.8 2658.8
## + race:workinschool   1   2633.1 2659.1
## + race:honors         1   2633.3 2659.3
## + sex:workinschool    1   2633.3 2659.3
## + race:yearsexp       1   2633.3 2659.3
## + race:city           1   2633.3 2659.3
```

```
##
## Call:  glm(formula = call ~ race + sex + honors + city + yearsexp +
##     emailincluded + workinschool + militaryexp + sex:city + sex:militaryexp +
##     sex:honors, family = "binomial", data = callbackData)
##
## Coefficients:
##         (Intercept)            raceWhite              sexMale
##            -2.52982              0.44599             -0.76549
##              honors          cityChicago             yearsexp
##             0.56003             -0.62918              0.02224
##       emailincluded         workinschool          militaryexp
##             0.22916             -0.13321             -0.62861
## sexMale:cityChicago  sexMale:militaryexp       sexMale:honors
##             1.13371              0.87511              0.64238
##
## Degrees of Freedom: 4869 Total (i.e. Null);  4858 Residual
## Null Deviance:       2727
## Residual Deviance: 2633  AIC: 2657
```

```r
step(interactions.model,
     scope=list(lower=current.model, upper=interactions.model),
     direction='backward', steps=25)
```

```
## Start:  AIC=2671.61
## call ~ (race + sex) * (honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp)
##
##                      Df Deviance    AIC
## - sex:workinschool    1   2629.6 2669.6
## - race:yearsexp       1   2629.7 2669.7
## - race:workinschool   1   2629.7 2669.7
## - race:honors         1   2629.7 2669.7
## - race:city           1   2629.7 2669.7
## - race:militaryexp    1   2629.9 2669.9
## - race:emailincluded  1   2630.0 2670.0
## - sex:yearsexp        1   2630.1 2670.1
```

```
## - sex:emailincluded   1   2631.2 2671.2
## <none>                    2629.6 2671.6
## - sex:honors          1   2632.2 2672.2
## - sex:militaryexp     1   2634.3 2674.3
## - sex:city            1   2643.1 2683.1
##
## Step:  AIC=2669.64
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + race:honors + race:city + race:yearsexp +
##     race:emailincluded + race:workinschool + race:militaryexp +
##     sex:honors + sex:city + sex:yearsexp + sex:emailincluded +
##     sex:militaryexp
##
##                      Df Deviance    AIC
## - race:workinschool   1   2629.7 2667.7
## - race:yearsexp       1   2629.7 2667.7
## - race:honors         1   2629.7 2667.7
## - race:city           1   2629.7 2667.7
## - race:militaryexp    1   2630.0 2668.0
## - race:emailincluded  1   2630.0 2668.0
## - sex:yearsexp        1   2630.2 2668.2
## - sex:emailincluded   1   2631.4 2669.4
## <none>                    2629.6 2669.6
## - sex:honors          1   2632.2 2670.2
## - sex:militaryexp     1   2634.3 2672.3
## - sex:city            1   2643.7 2681.7
##
## Step:  AIC=2667.69
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + race:honors + race:city + race:yearsexp +
##     race:emailincluded + race:militaryexp + sex:honors + sex:city +
##     sex:yearsexp + sex:emailincluded + sex:militaryexp
##
##                      Df Deviance    AIC
## - race:yearsexp       1   2629.8 2665.8
## - race:honors         1   2629.8 2665.8
## - race:city           1   2629.8 2665.8
## - race:militaryexp    1   2630.0 2666.0
## - race:emailincluded  1   2630.2 2666.2
## - sex:yearsexp        1   2630.2 2666.2
## - sex:emailincluded   1   2631.4 2667.4
## <none>                    2629.7 2667.7
## - sex:honors          1   2632.2 2668.2
## - sex:militaryexp     1   2634.3 2670.3
## - sex:city            1   2643.8 2679.8
##
## Step:  AIC=2665.76
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + race:honors + race:city + race:emailincluded +
##     race:militaryexp + sex:honors + sex:city + sex:yearsexp +
##     sex:emailincluded + sex:militaryexp
##
##                      Df Deviance    AIC
## - race:city           1   2629.8 2663.8
```

```
## - race:honors         1   2629.8 2663.8
## - race:militaryexp    1   2630.2 2664.2
## - race:emailincluded  1   2630.2 2664.2
## - sex:yearsexp        1   2630.3 2664.3
## - sex:emailincluded   1   2631.5 2665.5
## <none>                    2629.8 2665.8
## - sex:honors          1   2632.3 2666.3
## - sex:militaryexp     1   2634.4 2668.4
## - sex:city            1   2643.8 2677.8
##
## Step:  AIC=2663.81
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + race:honors + race:emailincluded +
##     race:militaryexp + sex:honors + sex:city + sex:yearsexp +
##     sex:emailincluded + sex:militaryexp
##
##                      Df Deviance    AIC
## - race:honors         1   2629.9 2661.9
## - race:emailincluded  1   2630.2 2662.2
## - race:militaryexp    1   2630.3 2662.3
## - sex:yearsexp        1   2630.4 2662.4
## - sex:emailincluded   1   2631.5 2663.5
## <none>                    2629.8 2663.8
## - sex:honors          1   2632.3 2664.3
## - sex:militaryexp     1   2634.4 2666.4
## - sex:city            1   2643.8 2675.8
##
## Step:  AIC=2661.87
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + race:emailincluded + race:militaryexp +
##     sex:honors + sex:city + sex:yearsexp + sex:emailincluded +
##     sex:militaryexp
##
##                      Df Deviance    AIC
## - race:militaryexp    1   2630.3 2660.3
## - race:emailincluded  1   2630.3 2660.3
## - sex:yearsexp        1   2630.4 2660.4
## - sex:emailincluded   1   2631.6 2661.6
## <none>                    2629.9 2661.9
## - sex:honors          1   2632.4 2662.4
## - sex:militaryexp     1   2634.5 2664.5
## - sex:city            1   2643.9 2673.9
##
## Step:  AIC=2660.34
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + race:emailincluded + sex:honors +
##     sex:city + sex:yearsexp + sex:emailincluded + sex:militaryexp
##
##                      Df Deviance    AIC
## - sex:yearsexp        1   2630.9 2658.9
## - race:emailincluded  1   2631.2 2659.2
## - sex:emailincluded   1   2632.0 2660.0
## <none>                    2630.3 2660.3
## - sex:honors          1   2632.9 2660.9
```

```
## - sex:militaryexp     1   2635.0 2663.0
## - sex:city            1   2644.3 2672.3
##
## Step:  AIC=2658.92
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + race:emailincluded + sex:honors +
##     sex:city + sex:emailincluded + sex:militaryexp
##
##                      Df Deviance    AIC
## - race:emailincluded  1   2631.8 2657.8
## - sex:emailincluded   1   2632.6 2658.6
## <none>                    2630.9 2658.9
## - sex:honors          1   2633.2 2659.2
## - sex:militaryexp     1   2636.8 2662.8
## - sex:city            1   2648.6 2674.6
##
## Step:  AIC=2657.76
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + sex:honors + sex:city + sex:emailincluded +
##     sex:militaryexp
##
##                     Df Deviance    AIC
## - sex:emailincluded  1   2633.4 2657.4
## <none>                   2631.8 2657.8
## - sex:honors         1   2634.1 2658.1
## - sex:militaryexp    1   2637.6 2661.6
## - sex:city           1   2649.5 2673.5
##
## Step:  AIC=2657.37
## call ~ race + sex + honors + city + yearsexp + emailincluded +
##     workinschool + militaryexp + sex:honors + sex:city + sex:militaryexp
##
##                   Df Deviance    AIC
## <none>                 2633.4 2657.4
## - sex:honors       1   2635.5 2657.5
## - sex:militaryexp  1   2637.7 2659.7
## - sex:city         1   2649.8 2671.8

##
## Call:  glm(formula = call ~ race + sex + honors + city + yearsexp +
##     emailincluded + workinschool + militaryexp + sex:honors +
##     sex:city + sex:militaryexp, family = "binomial", data = callbackData)
##
## Coefficients:
##        (Intercept)            raceWhite              sexMale
##           -2.52982              0.44599             -0.76549
##             honors          cityChicago             yearsexp
##            0.56003             -0.62918              0.02224
##      emailincluded         workinschool          militaryexp
##            0.22916             -0.13321             -0.62861
##     sexMale:honors  sexMale:cityChicago  sexMale:militaryexp
##            0.64238              1.13371              0.87511
##
## Degrees of Freedom: 4869 Total (i.e. Null);  4858 Residual
```

```
## Null Deviance:       2727
## Residual Deviance: 2633  AIC: 2657
```

Once again both methods return the same model, and have included interaction terms between sex and city, sex and military experience, and sex and honors.

```r
current.model = glm(formula = call ~ race + sex + honors + city + yearsexp +
    emailincluded + workinschool + militaryexp + sex:city + sex:militaryexp +
    sex:honors, family = "binomial", data = callbackData)
```

Lastly, with agreement between both forward and backward step procedures, we assess the significance of terms in the model.

```r
summary(current.model)
```

```
##
## Call:
## glm(formula = call ~ race + sex + honors + city + yearsexp +
##     emailincluded + workinschool + militaryexp + sex:city + sex:militaryexp +
##     sex:honors, family = "binomial", data = callbackData)
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -2.52982    0.16218 -15.599  < 2e-16 ***
## raceWhite            0.44599    0.10839   4.115 3.88e-05 ***
## sexMale             -0.76549    0.19062  -4.016 5.92e-05 ***
## honors               0.56003    0.21353   2.623  0.00872 **
## cityChicago         -0.62918    0.12375  -5.084 3.69e-07 ***
## yearsexp             0.02224    0.01033   2.153  0.03132 *
## emailincluded        0.22916    0.11880   1.929  0.05375 .
## workinschool        -0.13321    0.11447  -1.164  0.24455
## militaryexp         -0.62861    0.29265  -2.148  0.03171 *
## sexMale:cityChicago  1.13371    0.27328   4.149 3.35e-05 ***
## sexMale:militaryexp  0.87511    0.41820   2.093  0.03639 *
## sexMale:honors       0.64238    0.43087   1.491  0.13599
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2726.9  on 4869  degrees of freedom
## Residual deviance: 2633.4  on 4858  degrees of freedom
## AIC: 2657.4
##
## Number of Fisher Scoring iterations: 5
```

The variable recording whether the applicant worked in school and the interaction between sex and honors both appear non-significant. As neither of these variables are of particular interest to our model, we remove them both to produce our final model.

```r
final.model = glm(formula = call ~ race + sex + honors + city + yearsexp +
    emailincluded + militaryexp + sex:city + sex:militaryexp,
    family = "binomial", data = callbackData)
summary(final.model)
```

```
##
## Call:
```

```
## glm(formula = call ~ race + sex + honors + city + yearsexp +
##     emailincluded + militaryexp + sex:city + sex:militaryexp,
##     family = "binomial", data = callbackData)
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -2.59625    0.15454 -16.800  < 2e-16 ***
## raceWhite            0.45009    0.10832   4.155 3.25e-05 ***
## sexMale             -0.74521    0.18788  -3.966 7.30e-05 ***
## honors               0.70585    0.18688   3.777 0.000159 ***
## cityChicago         -0.64256    0.12340  -5.207 1.92e-07 ***
## yearsexp             0.02342    0.01021   2.293 0.021827 *
## emailincluded        0.18127    0.11340   1.598 0.109945
## militaryexp         -0.63199    0.29296  -2.157 0.030984 *
## sexMale:cityChicago  1.21013    0.26833   4.510 6.49e-06 ***
## sexMale:militaryexp  0.95232    0.41199   2.311 0.020805 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2726.9  on 4869  degrees of freedom
## Residual deviance: 2636.9  on 4860  degrees of freedom
## AIC: 2656.9
##
## Number of Fisher Scoring iterations: 5
```

With the removal of these two terms, it appears the term 'emailincluded' is also not significant. As once again this does not represent an important variable for this analysis, we choose to remove it from the model.

```r
final.model = glm(formula = call ~ race + sex + honors + city + yearsexp +
    militaryexp + sex:city + sex:militaryexp,
    family = "binomial", data = callbackData)
summary(final.model)
```

```
##
## Call:
## glm(formula = call ~ race + sex + honors + city + yearsexp +
##     militaryexp + sex:city + sex:militaryexp, family = "binomial",
##     data = callbackData)
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -2.53589    0.14848 -17.078  < 2e-16 ***
## raceWhite            0.45064    0.10828   4.162 3.16e-05 ***
## sexMale             -0.77368    0.18665  -4.145 3.40e-05 ***
## honors               0.73747    0.18560   3.973 7.09e-05 ***
## cityChicago         -0.63039    0.12298  -5.126 2.96e-07 ***
## yearsexp             0.02524    0.01007   2.506   0.0122 *
## militaryexp         -0.53003    0.28618  -1.852   0.0640 .
## sexMale:cityChicago  1.23952    0.26748   4.634 3.58e-06 ***
## sexMale:militaryexp  0.97334    0.41170   2.364   0.0181 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2726.9  on 4869  degrees of freedom
## Residual deviance: 2639.5  on 4861  degrees of freedom
## AIC: 2657.5
##
## Number of Fisher Scoring iterations: 5
```

Model summary (8 points)

[NOTE: This section should detail the final model and each of the terms in the model, as well as their effects on the odds.]

The final resulting model found was:

log-odds(call) ~ -2.536 + 0.451 * (race=White) + -0.774 * (sex=Male) + 0.737 * honors + -0.630 * (city=Chicago) + 0.0252 * yearsexp + -0.530 * militaryexp + 1.240 * (sex=Male:city=Chicago) + 0.973 * (sex=Male:militaryexp)

Each term in the model is statistically significant except for military experience, but the inclusion of the interaction term which is significant indicates that we keep the military experience term as well.

The effect on the odds of each of the terms are listed below.

```r
oddsimpact = exp(final.model$coefficients)
oddsimpact
```

```
##         (Intercept)           raceWhite             sexMale              honors
##          0.07919125          1.56932103          0.46131323          2.09063026
##         cityChicago            yearsexp         militaryexp sexMale:cityChicago
##          0.53238399          1.02555762          0.58858603          3.45394196
## sexMale:militaryexp
##          2.64677596
```

Note: The baseline individual in this study is a black female without military experience applying to a job in Boston.

- race=White: White respondents saw a 1.57 times increase in the odds of receiving a callback.
- honors: Resumes with listed honors saw a 2.09 times increase in the odds of receiving a callback.
- city=Chicago: Compared with applications in Boston, female applicants in Chicago saw a 0.53 times reduction in odds of receiving a callback.
- yearexp: All other variables held constant, a one year increase in experience was equivalent to a 1.03 times increase in the odds of receiving a callback.
- militaryexp: Resumes listing military experience for female applicants saw a 0.589 times reduction in the odds of receiving a callback.
- sex=Male: Male applicants without military experience applying in Chicago saw a 0.46 times decrease in the odds of receiving a callback.
- sex=Male:city=Chicago: Male applicants applying in Chicago saw a 3.45 times increase in their odds of receiving a callback compared to male applicants in Boston.
- sex=Male:militaryexp: Male applicants with military experience saw a 2.65 times increase in their odds of receiving a callback compared to female applicants with military experience.

Conclusion (8 points)

[NOTE: This section should note the conclusions of the analysis, including identifying which variables positively and negatively impacted the odds for this data]

This analysis demonstrated that the variables of perceived race, perceived sex, honors, years experience, and military experience of the applicant impacted their odds of a callback. The city of the job being applied to also had a strong impact on callbacks. It was also shown that military service and the city being applied to impacted the odds of callbacks differently for men and women applying to jobs.

The variables that positively impacted callback odds were race being white, honors, years of experience, and for specifically male applicants the job being in Chicago or military experience.

The variables that negatively impacted callback odds were sex being male for applicants in Boston, applications in Chicago, and military experience for women.

Overall this analysis indicates a preference for perceived white candidates over perceived black candidates. There were also interesting relationships between the perceived sex of the candidates and both the city being applied to or military service.