Logistic Regression Analysis

Anthony Wen

2024-03-13

Introduction (4 points)

```
# You can feel free to modify this line to read the data in
#  as you will only be submitting a completed PDF
loanDefaultData<-read.csv("LoanDefaultData.csv")
```

In this study we will look at the different variables or characteristics of the people that will potentially request loans to determine whether or not a bank should grant their loan. In other words, it is for predicting loan default, whether a machine will say yes or no to someone requesting a loan. The data that was submitted had information such as the person's marital status, credit score and other variables that might have an impact on the person's loan default. This study would analyze if such variables would impact a person's loan default, just as the original study has also intended to predict with a special emphasis (out of personal interest) on whether there was evidence that default were impacted by the marital status or highest education level of the applicant.

The Variables (8 points)

The variables available for this analysis are:

Response:

- Default (binary/Catagorical): This records whether or not the loan was defaulted on. It is the variable that we seek to predict in our logistic regression model。

```
table(loanDefaultData$Default)
```

```
## 
##    0    1 
## 3091  409
```

Predictors:

```
names(loanDefaultData)
```

```
##  [1] "Age"            "Income"         "LoanAmount"     "CreditScore"   
##  [5] "MonthsEmployed" "NumCreditLines" "InterestRate"   "LoanTerm"      
##  [9] "Education"      "EmploymentType" "MaritalStatus"  "HasCoSigner"   
## [13] "Default"
```

- Age (numeric): a numeric variable that tells us the applicant's age.It has mean of 43.7 years or around 44 years old and a standard deviation of 14.96 years. It shows how most of the applicants are middle aged people. The range is from 18 years old to 69 years old

```
mean(loanDefaultData$Age)
```

```
## [1] 43.74314
```

```
sd(loanDefaultData$Age)
```

```
## [1] 14.95721
```

```
range(loanDefaultData$Age)
```

```
## [1] 18 69
```

```
hist(loanDefaultData$Age, main='Histogram of Age', xlab='Age of Applicant (Years)', freq
=FALSE)
```

- Income (numeric): a numeric variable that tells us an applicant's annual income in USD. It has mean of 82351.62 or around 82 thousand USD of annual income and a standard deviation of 39177.07 USD.

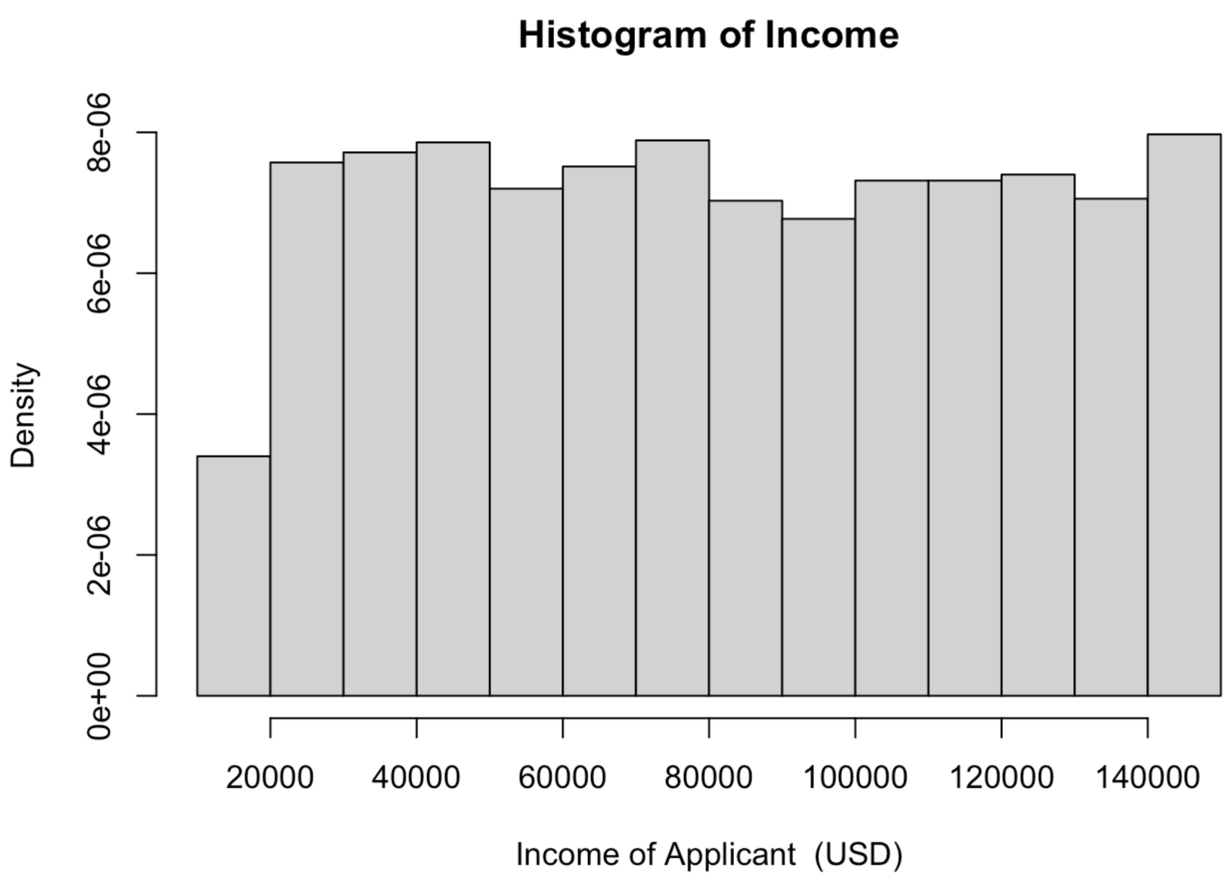
```
mean(loanDefaultData$Income)
```

```
## [1] 82351.62
```

```
sd(loanDefaultData$Income)
```

```
## [1] 39177.07
```

```
hist(loanDefaultData$Income, main='Histogram of Income', xlab='Income of Applicant  (US
D)', freq=FALSE)
```



- LoanAmount (numeric): a numeric variable that tells us the loan amount requested by the applicant in USD or the amount of money that is being borrowed. It has mean of 128733.5 or around 129 thousand USD of annual income and a standard deviation of 70830.42 USD. The mean of the loan is much larger than the mean of the applicant's annual income, which shows how people are usually borrowing more money than they are making.

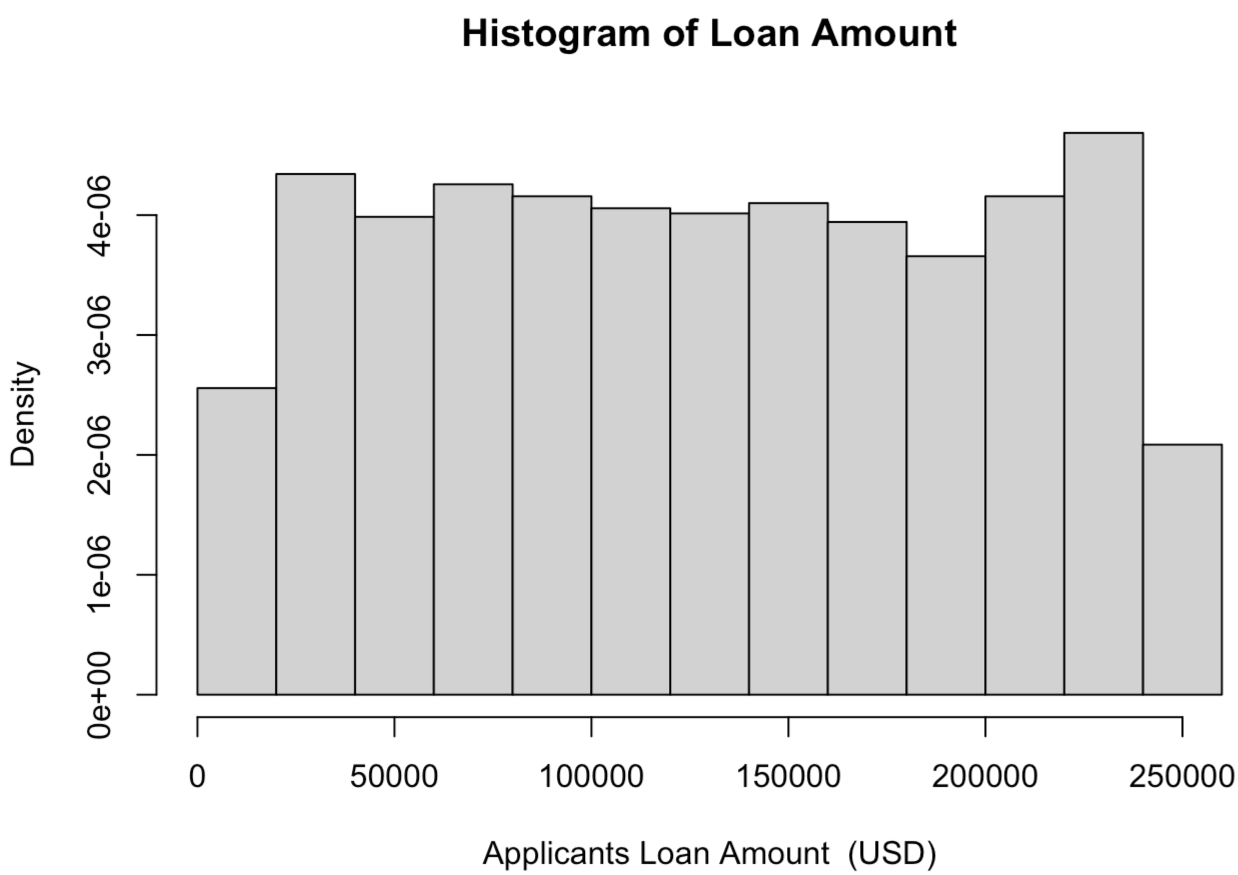
```
mean(loanDefaultData$LoanAmount)
```

```
## [1] 128733.5
```

```
sd(loanDefaultData$LoanAmount)
```

```
## [1] 70830.42
```

```
hist(loanDefaultData$LoanAmount, main='Histogram of Loan Amount', xlab='Applicants Loan
Amount  (USD)', freq=FALSE)
```



- CreditScore (numeric): a numeric variable that tells us the applicant's credit score to tell audience a applicant's creditworthiness. It has mean of 575.0686 or around 575 points/credit score and a standard deviation of 159.4175 or 159 points. The variability suggests a diverse set of credit scores among the applicants.

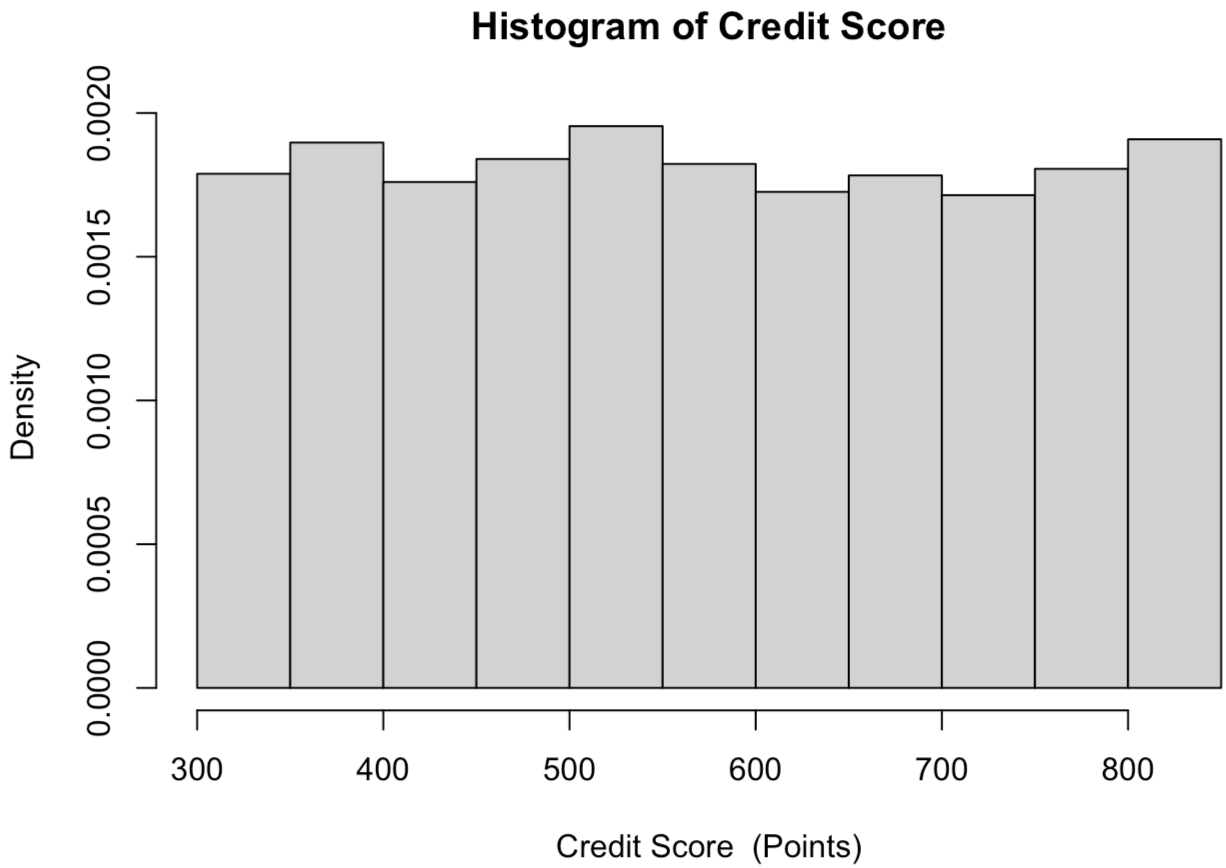
```
mean(loanDefaultData$CreditScore)
```

```
## [1] 575.0686
```

```
sd(loanDefaultData$CreditScore)
```

```
## [1] 159.4175
```

```
hist(loanDefaultData$CreditScore, main='Histogram of Credit Score', xlab='Credit Score
(Points)', freq=FALSE)
```



- MonthsEmployed (numeric): a numeric variable that tells us the number of months the applicant has been employed. It has mean of 60.15971 months or around 5 years of being employed and a standard deviation of 34.10338 months or almost 3 years.

```
mean(loanDefaultData$MonthsEmployed)
```

```
## [1] 60.15971
```

```
sd(loanDefaultData$MonthsEmployed)
```

```
## [1] 34.10338
```

```
hist(loanDefaultData$MonthsEmployed, main='Histogram of MonthsEmployed', xlab='Months Em
ployed of Applicant', freq=FALSE)
```

- NumCreditLines (numeric): a numeric variable that tells us the number of credit lines the applicant has opened. It has mean of 2.47 lines (around 2 lines opened) and a standard deviation of 1.115464 lines (basically 1 line).

```
mean(loanDefaultData$NumCreditLines)
```

```
## [1] 2.469143
```

```
sd(loanDefaultData$NumCreditLines)
```

```
## [1] 1.115464
```

```
table(loanDefaultData$NumCreditLines)
```

```
##
##   1   2   3   4
## 901 896 863 840
```

- InterestRate (numeric): a numeric variable that tells us the interest rate of the loan.It has mean of 13.61331% interest rate and a standard deviation of 6.661623% interest.

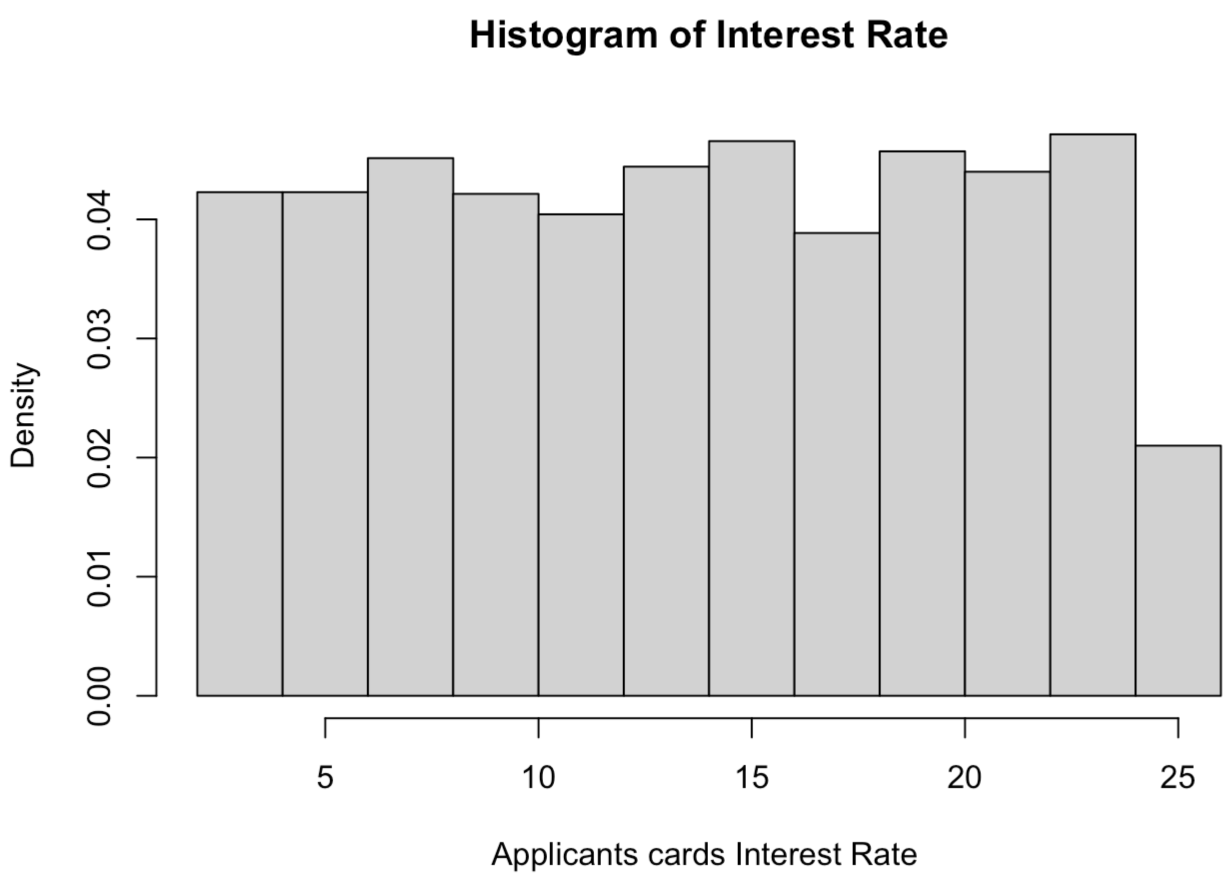
```
mean(loanDefaultData$InterestRate)
```

```
## [1] 13.61331
```

```
sd(loanDefaultData$InterestRate)
```

```
## [1] 6.661623
```

```
hist(loanDefaultData$InterestRate, main='Histogram of Interest Rate', xlab='Applicants c
ards Interest Rate ', freq=FALSE)
```



- LoanTerm (numeric): a numeric variable that tells us the length of term of the loan in months. It has mean of 36.04457 months or around 3 years of length for the loan terms and a standard deviation of 16.94259 months or almost 1.5 years.

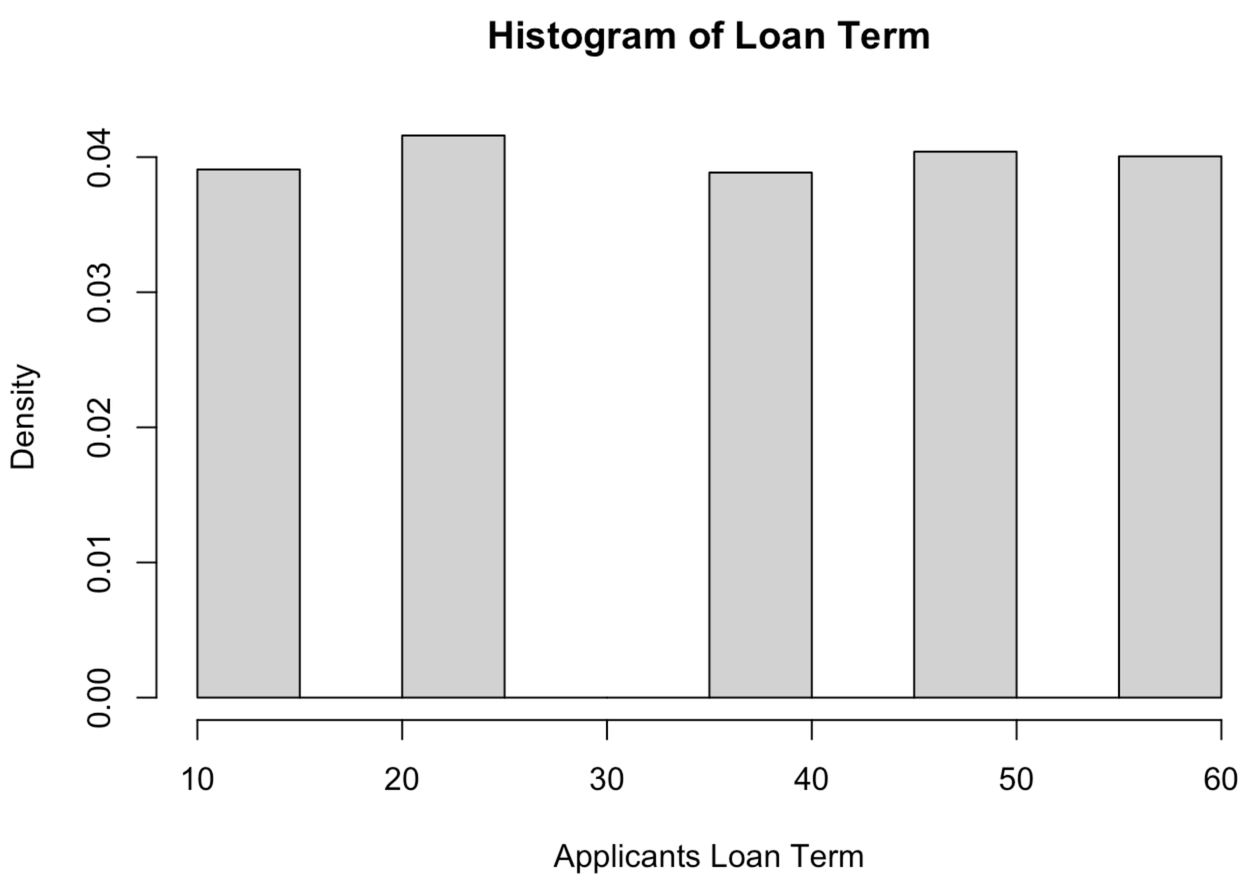
```
mean(loanDefaultData$LoanTerm)
```

```
## [1] 36.04457
```

```
sd(loanDefaultData$LoanTerm)
```

```
## [1] 16.94259
```

```
hist(loanDefaultData$LoanTerm, main='Histogram of Loan Term', xlab='Applicants Loan Ter
m', freq=FALSE)
```



- Education (categorical): a categorical variable that tells us the highest level of education attained by the applicant. The number of applicants in each catagory are pretty similar, aka the standard deviation is not very high.

```
table(loanDefaultData$Education)
```

```
##
##  Bachelor's High School    Master's         PhD
##         902         849         893         856
```

- EmploymentType (Categorical): a catagorical variable that tells us the type of employment of the applicant. The number of applicants in each catagory are pretty similar, aka the standard deviation is not very high.

```
table(loanDefaultData$EmploymentType)
```

```
##
##     Full-time     Part-time Self-employed    Unemployed
##           872           873           908           847
```

- MaritalStatus (Categorical): a catagorical variable that tells us the marital status of the applicant. It's split into single, married, or divorced. The number of applicants in each catagory are pretty similar, aka the standard deviation is not very high.

```
table(loanDefaultData$MaritalStatus)
```

```
##
## Divorced  Married   Single
##     1176     1126     1198
```

- HasCoSigner (binary/Categorical): a catagorical variable that tells us whether the loan application has a cosigner. It is a yes or no value.

```
table(loanDefaultData$HasCoSigner)
```

```
##
##   No  Yes
## 1744 1756
```

Model Fitting (12 points)

Instructions

- Show code used to explore and fit a model to the data
- Explain between code chunks why each step is being taken
- Detail the choices made to determine which variables belong in your model
- Note that there is no 'correct' answer here, you will be assessed on your ability to justify choices made in fitting your model

Here I utilize the step function with AIC scoring to determine the best model. To produce and test a variety of models, I utilize both forward and backward step functions. Since I predict based on the data initially that some variables will be more impactful than others, all models will include the variables 'Education' and 'MaritalStatus'. In addition, I focus on the AIC score/value a lot when looking into whether removing or adding a variable has an impact on the odds of defaulting.

```
# Initialize the full model with all variables
full_model = glm(Default ~ Age + Income + LoanAmount + CreditScore + MonthsEmployed +
                 NumCreditLines + InterestRate + LoanTerm + Education + EmploymentType +
                 MaritalStatus + HasCoSigner, data=loanDefaultData, family = "binomial")

# Initialize the basic model with key variables identified as essential
basic_model = glm(Default ~ MaritalStatus + Education, data=loanDefaultData, family = "b
inomial")
```

```
step(basic_model,
     scope=list(lower=basic_model, upper=full_model),
     direction='forward', steps=25)
```

```
## Start:  AIC=2528.39
## Default ~ MaritalStatus + Education
##
##                  Df Deviance    AIC
## + Age             1   2425.9 2439.9
## + InterestRate    1   2460.0 2474.0
## + MonthsEmployed  1   2484.4 2498.4
## + Income          1   2498.2 2512.2
## + LoanAmount      1   2503.1 2517.1
## + CreditScore     1   2506.8 2520.8
## + HasCoSigner     1   2507.9 2521.9
## + NumCreditLines  1   2508.1 2522.1
## + EmploymentType  3   2505.6 2523.6
## <none>                2516.4 2528.4
## + LoanTerm        1   2515.7 2529.7
##
## Step:  AIC=2439.89
## Default ~ MaritalStatus + Education + Age
##
##                  Df Deviance    AIC
## + InterestRate    1   2367.0 2383.0
## + MonthsEmployed  1   2392.7 2408.7
## + Income          1   2408.0 2424.0
## + LoanAmount      1   2411.4 2427.4
## + CreditScore     1   2415.6 2431.6
## + NumCreditLines  1   2417.0 2433.0
## + HasCoSigner     1   2417.8 2433.8
## + EmploymentType  3   2415.8 2435.8
## <none>                2425.9 2439.9
## + LoanTerm        1   2425.1 2441.1
##
## Step:  AIC=2382.99
## Default ~ MaritalStatus + Education + Age + InterestRate
##
##                  Df Deviance    AIC
## + MonthsEmployed  1   2334.1 2352.1
## + Income          1   2349.2 2367.2
## + LoanAmount      1   2352.9 2370.9
## + NumCreditLines  1   2357.4 2375.4
## + CreditScore     1   2357.4 2375.4
## + EmploymentType  3   2355.0 2377.0
## + HasCoSigner     1   2360.6 2378.6
## <none>                2367.0 2383.0
## + LoanTerm        1   2366.4 2384.4
##
## Step:  AIC=2352.09
## Default ~ MaritalStatus + Education + Age + InterestRate + MonthsEmployed
##
##                  Df Deviance    AIC
## + Income          1   2315.0 2335.0
## + LoanAmount      1   2320.1 2340.1
## + NumCreditLines  1   2325.0 2345.0
```

```
## + CreditScore     1   2325.2 2345.2
## + EmploymentType  3   2321.9 2345.9
## + HasCoSigner     1   2327.9 2347.9
## <none>                2334.1 2352.1
## + LoanTerm        1   2333.4 2353.4
##
## Step:  AIC=2334.98
## Default ~ MaritalStatus + Education + Age + InterestRate + MonthsEmployed +
##     Income
##
##                  Df Deviance    AIC
## + LoanAmount      1   2300.4 2322.4
## + NumCreditLines  1   2305.6 2327.6
## + CreditScore     1   2306.1 2328.1
## + EmploymentType  3   2303.0 2329.0
## + HasCoSigner     1   2307.7 2329.7
## <none>                2315.0 2335.0
## + LoanTerm        1   2314.2 2336.2
##
## Step:  AIC=2322.36
## Default ~ MaritalStatus + Education + Age + InterestRate + MonthsEmployed +
##     Income + LoanAmount
##
##                  Df Deviance    AIC
## + CreditScore     1   2290.9 2314.9
## + NumCreditLines  1   2291.5 2315.5
## + EmploymentType  3   2287.5 2315.5
## + HasCoSigner     1   2293.2 2317.2
## <none>                2300.4 2322.4
## + LoanTerm        1   2299.7 2323.7
##
## Step:  AIC=2314.88
## Default ~ MaritalStatus + Education + Age + InterestRate + MonthsEmployed +
##     Income + LoanAmount + CreditScore
##
##                  Df Deviance    AIC
## + EmploymentType  3   2277.1 2307.1
## + NumCreditLines  1   2281.7 2307.7
## + HasCoSigner     1   2283.2 2309.2
## <none>                2290.9 2314.9
## + LoanTerm        1   2290.2 2316.2
##
## Step:  AIC=2307.14
## Default ~ MaritalStatus + Education + Age + InterestRate + MonthsEmployed +
##     Income + LoanAmount + CreditScore + EmploymentType
##
##                  Df Deviance    AIC
## + NumCreditLines  1   2268.6 2300.6
## + HasCoSigner     1   2269.1 2301.1
## <none>                2277.1 2307.1
## + LoanTerm        1   2276.3 2308.3
##
```

```
## Step:  AIC=2300.6
## Default ~ MaritalStatus + Education + Age + InterestRate + MonthsEmployed +
##     Income + LoanAmount + CreditScore + EmploymentType + NumCreditLines
##
##               Df Deviance    AIC
## + HasCoSigner  1   2260.5 2294.5
## <none>             2268.6 2300.6
## + LoanTerm     1   2267.8 2301.8
##
## Step:  AIC=2294.53
## Default ~ MaritalStatus + Education + Age + InterestRate + MonthsEmployed +
##     Income + LoanAmount + CreditScore + EmploymentType + NumCreditLines +
##     HasCoSigner
##
##            Df Deviance    AIC
## <none>          2260.5 2294.5
## + LoanTerm  1   2259.7 2295.7
```

```
##
## Call:  glm(formula = Default ~ MaritalStatus + Education + Age + InterestRate +
##     MonthsEmployed + Income + LoanAmount + CreditScore + EmploymentType +
##     NumCreditLines + HasCoSigner, family = "binomial", data = loanDefaultData)
##
## Coefficients:
##                (Intercept)        MaritalStatusMarried
##                 -7.303e-01                  -3.035e-01
##        MaritalStatusSingle        EducationHigh School
##                 -4.785e-02                   2.588e-01
##          EducationMaster's                EducationPhD
##                  1.559e-01                   4.146e-02
##                        Age                InterestRate
##                 -3.655e-02                   6.379e-02
##             MonthsEmployed                      Income
##                 -9.292e-03                  -6.567e-06
##                 LoanAmount                 CreditScore
##                  3.088e-06                  -1.168e-03
##     EmploymentTypePart-time  EmploymentTypeSelf-employed
##                  2.748e-01                   4.415e-01
##    EmploymentTypeUnemployed              NumCreditLines
##                  5.492e-01                   1.450e-01
##             HasCoSignerYes
##                 -3.152e-01
##
## Degrees of Freedom: 3499 Total (i.e. Null);  3483 Residual
## Null Deviance:       2524
## Residual Deviance: 2261  AIC: 2295
```

```
step(full_model,
     scope=list(lower=basic_model, upper=full_model),
     direction='backward', steps=25)
```

```
## Start:  AIC=2295.69
## Default ~ Age + Income + LoanAmount + CreditScore + MonthsEmployed +
##     NumCreditLines + InterestRate + LoanTerm + Education + EmploymentType +
##     MaritalStatus + HasCoSigner
##
##                  Df Deviance    AIC
## - LoanTerm        1   2260.5 2294.5
## <none>                2259.7 2295.7
## - HasCoSigner     1   2267.8 2301.8
## - NumCreditLines  1   2268.3 2302.3
## - EmploymentType  3   2273.2 2303.2
## - CreditScore     1   2270.9 2304.9
## - LoanAmount      1   2275.1 2309.1
## - Income          1   2281.1 2315.1
## - MonthsEmployed  1   2292.4 2326.4
## - InterestRate    1   2317.2 2351.2
## - Age             1   2354.1 2388.1
##
## Step:  AIC=2294.53
## Default ~ Age + Income + LoanAmount + CreditScore + MonthsEmployed +
##     NumCreditLines + InterestRate + Education + EmploymentType +
##     MaritalStatus + HasCoSigner
##
##                  Df Deviance    AIC
## <none>                2260.5 2294.5
## - HasCoSigner     1   2268.6 2300.6
## - NumCreditLines  1   2269.1 2301.1
## - EmploymentType  3   2273.9 2301.9
## - CreditScore     1   2271.7 2303.7
## - LoanAmount      1   2275.9 2307.9
## - Income          1   2281.9 2313.9
## - MonthsEmployed  1   2293.2 2325.2
## - InterestRate    1   2318.0 2350.0
## - Age             1   2355.0 2387.0
```

```
##
## Call:  glm(formula = Default ~ Age + Income + LoanAmount + CreditScore +
##     MonthsEmployed + NumCreditLines + InterestRate + Education +
##     EmploymentType + MaritalStatus + HasCoSigner, family = "binomial",
##     data = loanDefaultData)
##
## Coefficients:
##                 (Intercept)                          Age
##                  -7.303e-01                   -3.655e-02
##                      Income                   LoanAmount
##                  -6.567e-06                    3.088e-06
##                 CreditScore               MonthsEmployed
##                  -1.168e-03                   -9.292e-03
##              NumCreditLines                 InterestRate
##                   1.450e-01                    6.379e-02
##         EducationHigh School           EducationMaster's
##                   2.588e-01                    1.559e-01
##                EducationPhD      EmploymentTypePart-time
##                   4.146e-02                    2.748e-01
## EmploymentTypeSelf-employed     EmploymentTypeUnemployed
##                   4.415e-01                    5.492e-01
##        MaritalStatusMarried          MaritalStatusSingle
##                  -3.035e-01                   -4.785e-02
##              HasCoSignerYes
##                  -3.152e-01
##
## Degrees of Freedom: 3499 Total (i.e. Null);  3483 Residual
## Null Deviance:       2524
## Residual Deviance: 2261  AIC: 2295
```

As can be observed from the step functions, both models returned the same models, which shows that the two main variables of interest – maritial status and education – do play an important role in the prediction. Now, we will explore whether interactions with these terms are useful for the model. I started initially with only marital status and education as they were the variable of interest for me.

```
current.model = glm(formula = Default ~ MaritalStatus + Age + Income + LoanAmount + Cred
itScore + MonthsEmployed + NumCreditLines + InterestRate + LoanTerm + Education + Employ
mentType + HasCoSigner, family = "binomial", data = loanDefaultData)

interactions.model=glm(formula = Default ~ (MaritalStatus + Education) * (Age + Income +
LoanAmount + CreditScore + MonthsEmployed + NumCreditLines + InterestRate + LoanTerm + E
mploymentType + HasCoSigner), family = "binomial", data = loanDefaultData)
```

In this model, I began with a base model that started with marital status and education (like I mentioned before) and added all the predictors. The stepwise regression shows that interaction between marital status and number of credit lines really increased the model's fit as the AIC value got lower.

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

```
## Start:  AIC=2295.69
## Default ~ MaritalStatus + Age + Income + LoanAmount + CreditScore +
##     MonthsEmployed + NumCreditLines + InterestRate + LoanTerm +
##     Education + EmploymentType + HasCoSigner
##
##                                Df Deviance    AIC
## + MaritalStatus:NumCreditLines  2   2248.8 2288.8
## <none>                              2259.7 2295.7
## + Education:Age                 3   2256.2 2298.2
## + MaritalStatus:LoanAmount      2   2258.4 2298.4
## + MaritalStatus:CreditScore     2   2258.5 2298.5
## + MaritalStatus:LoanTerm        2   2258.7 2298.7
## + MaritalStatus:HasCoSigner     2   2258.7 2298.7
## + Education:NumCreditLines      3   2256.8 2298.8
## + Education:CreditScore         3   2257.2 2299.2
## + MaritalStatus:Age             2   2259.2 2299.2
## + MaritalStatus:MonthsEmployed  2   2259.4 2299.4
## + MaritalStatus:Income          2   2259.5 2299.5
## + MaritalStatus:InterestRate    2   2259.6 2299.6
## + Education:Income              3   2257.8 2299.8
## + Education:LoanAmount          3   2258.3 2300.3
## + Education:InterestRate        3   2258.6 2300.6
## + Education:LoanTerm            3   2259.0 2301.0
## + Education:MonthsEmployed      3   2259.0 2301.0
## + Education:HasCoSigner         3   2259.6 2301.6
## + MaritalStatus:EmploymentType  6   2255.2 2303.2
## + Education:EmploymentType      9   2250.7 2304.7
##
## Step:  AIC=2288.81
## Default ~ MaritalStatus + Age + Income + LoanAmount + CreditScore +
##     MonthsEmployed + NumCreditLines + InterestRate + LoanTerm +
##     Education + EmploymentType + HasCoSigner + MaritalStatus:NumCreditLines
##
##                                Df Deviance    AIC
## <none>                              2248.8 2288.8
## + Education:Age                 3   2244.8 2290.8
## + MaritalStatus:LoanAmount      2   2247.4 2291.4
## + MaritalStatus:LoanTerm        2   2247.6 2291.6
## + MaritalStatus:CreditScore     2   2247.7 2291.7
## + MaritalStatus:HasCoSigner     2   2247.9 2291.9
## + Education:NumCreditLines      3   2246.2 2292.2
## + Education:CreditScore         3   2246.2 2292.2
## + MaritalStatus:Age             2   2248.4 2292.4
## + MaritalStatus:MonthsEmployed  2   2248.5 2292.5
## + MaritalStatus:Income          2   2248.6 2292.6
## + MaritalStatus:InterestRate    2   2248.8 2292.8
## + Education:Income              3   2247.0 2293.0
## + Education:LoanAmount          3   2247.6 2293.6
## + Education:InterestRate        3   2247.8 2293.8
## + Education:MonthsEmployed      3   2248.0 2294.0
## + Education:LoanTerm            3   2248.2 2294.2
## + Education:HasCoSigner         3   2248.7 2294.7
```

```
## + MaritalStatus:EmploymentType  6   2244.2 2296.2
## + Education:EmploymentType      9   2239.8 2297.8
```

```
##
## Call:  glm(formula = Default ~ MaritalStatus + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + Education + EmploymentType + HasCoSigner + MaritalStatus:NumCreditLine
s,
##     family = "binomial", data = loanDefaultData)
##
## Coefficients:
##                         (Intercept)                 MaritalStatusMarried
##                          -3.415e-01                           -3.093e-01
##                  MaritalStatusSingle                                  Age
##                          -9.849e-01                           -3.647e-02
##                               Income                           LoanAmount
##                          -6.552e-06                            3.163e-06
##                          CreditScore                       MonthsEmployed
##                          -1.137e-03                           -9.297e-03
##                       NumCreditLines                         InterestRate
##                           3.009e-02                            6.376e-02
##                             LoanTerm                  EducationHigh School
##                          -3.349e-03                            2.521e-01
##                   EducationMaster's                          EducationPhD
##                           1.338e-01                            3.652e-02
##              EmploymentTypePart-time          EmploymentTypeSelf-employed
##                           2.664e-01                            4.365e-01
##             EmploymentTypeUnemployed                       HasCoSignerYes
##                           5.542e-01                           -3.092e-01
## MaritalStatusMarried:NumCreditLines   MaritalStatusSingle:NumCreditLines
##                          -1.395e-03                            3.510e-01
##
## Degrees of Freedom: 3499 Total (i.e. Null);  3480 Residual
## Null Deviance:       2524
## Residual Deviance: 2249  AIC: 2289
```

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

```
## Start:  AIC=2371.01
## Default ~ (MaritalStatus + Education) * (Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner)
##
##                                Df Deviance    AIC
## - Education:EmploymentType      9   2223.5 2361.5
## - MaritalStatus:EmploymentType  6   2219.2 2363.2
## - Education:HasCoSigner         3   2215.3 2365.3
## - Education:LoanTerm            3   2215.5 2365.5
## - Education:LoanAmount          3   2216.0 2366.0
## - Education:MonthsEmployed      3   2216.1 2366.1
## - Education:InterestRate        3   2216.2 2366.2
## - Education:Income              3   2216.9 2366.9
## - MaritalStatus:InterestRate    2   2215.0 2367.0
## - MaritalStatus:MonthsEmployed  2   2215.2 2367.2
## - MaritalStatus:Income          2   2215.2 2367.2
## - MaritalStatus:Age             2   2215.6 2367.6
## - Education:CreditScore         3   2217.7 2367.7
## - MaritalStatus:CreditScore     2   2215.8 2367.8
## - Education:NumCreditLines      3   2217.9 2367.9
## - MaritalStatus:HasCoSigner     2   2216.1 2368.1
## - MaritalStatus:LoanAmount      2   2216.1 2368.1
## - MaritalStatus:LoanTerm        2   2216.5 2368.5
## - Education:Age                 3   2219.0 2369.0
## <none>                              2215.0 2371.0
## - MaritalStatus:NumCreditLines  2   2226.3 2378.3
##
## Step:  AIC=2361.48
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:InterestRate + MaritalStatus:LoanTerm + MaritalStatus:EmploymentTyp
e +
##     MaritalStatus:HasCoSigner + Education:Age + Education:Income +
##     Education:LoanAmount + Education:CreditScore + Education:MonthsEmployed +
##     Education:NumCreditLines + Education:InterestRate + Education:LoanTerm +
##     Education:HasCoSigner
##
##                                Df Deviance    AIC
## - MaritalStatus:EmploymentType  6   2228.1 2354.1
## - Education:HasCoSigner         3   2223.7 2355.7
## - Education:LoanTerm            3   2224.1 2356.1
## - Education:InterestRate        3   2224.5 2356.5
## - Education:MonthsEmployed      3   2224.6 2356.6
## - Education:LoanAmount          3   2224.7 2356.7
## - MaritalStatus:InterestRate    2   2223.5 2357.5
## - Education:Income              3   2225.6 2357.6
## - MaritalStatus:MonthsEmployed  2   2223.7 2357.7
## - MaritalStatus:Income          2   2223.8 2357.8
```

```
## - Education:CreditScore        3   2225.9 2357.9
## - MaritalStatus:Age            2   2224.1 2358.1
## - Education:NumCreditLines     3   2226.4 2358.4
## - MaritalStatus:HasCoSigner    2   2224.4 2358.4
## - MaritalStatus:CreditScore    2   2224.4 2358.4
## - MaritalStatus:LoanAmount     2   2224.6 2358.6
## - MaritalStatus:LoanTerm       2   2225.1 2359.1
## - Education:Age                3   2227.3 2359.3
## <none>                             2223.5 2361.5
## - MaritalStatus:NumCreditLines  2   2234.7 2368.7
##
## Step:  AIC=2354.05
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:InterestRate + MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner +
##     Education:Age + Education:Income + Education:LoanAmount +
##     Education:CreditScore + Education:MonthsEmployed + Education:NumCreditLines +
##     Education:InterestRate + Education:LoanTerm + Education:HasCoSigner
##
##                                Df Deviance    AIC
## - Education:HasCoSigner        3   2228.3 2348.3
## - Education:LoanTerm           3   2228.7 2348.7
## - Education:InterestRate       3   2229.1 2349.1
## - Education:MonthsEmployed     3   2229.1 2349.1
## - Education:LoanAmount         3   2229.3 2349.3
## - MaritalStatus:InterestRate   2   2228.1 2350.1
## - Education:Income             3   2230.1 2350.1
## - MaritalStatus:MonthsEmployed  2   2228.2 2350.2
## - MaritalStatus:Income         2   2228.3 2350.3
## - Education:CreditScore        3   2230.6 2350.6
## - MaritalStatus:Age            2   2228.7 2350.7
## - Education:NumCreditLines     3   2230.8 2350.8
## - MaritalStatus:HasCoSigner    2   2229.0 2351.0
## - MaritalStatus:CreditScore    2   2229.1 2351.1
## - MaritalStatus:LoanAmount     2   2229.2 2351.2
## - MaritalStatus:LoanTerm       2   2229.6 2351.6
## - Education:Age                3   2231.9 2351.9
## <none>                             2228.1 2354.1
## - MaritalStatus:NumCreditLines  2   2239.2 2361.2
##
## Step:  AIC=2348.26
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:InterestRate + MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner +
##     Education:Age + Education:Income + Education:LoanAmount +
##     Education:CreditScore + Education:MonthsEmployed + Education:NumCreditLines +
```

```
##     Education:InterestRate + Education:LoanTerm
##
##                                Df Deviance    AIC
## - Education:LoanTerm            3   2228.9 2342.9
## - Education:MonthsEmployed      3   2229.3 2343.3
## - Education:InterestRate        3   2229.3 2343.3
## - Education:LoanAmount          3   2229.4 2343.4
## - MaritalStatus:InterestRate    2   2228.3 2344.3
## - Education:Income              3   2230.3 2344.3
## - MaritalStatus:MonthsEmployed  2   2228.4 2344.4
## - MaritalStatus:Income          2   2228.5 2344.5
## - Education:CreditScore         3   2230.7 2344.7
## - MaritalStatus:Age             2   2228.9 2344.9
## - Education:NumCreditLines      3   2230.9 2344.9
## - MaritalStatus:HasCoSigner     2   2229.3 2345.3
## - MaritalStatus:CreditScore     2   2229.4 2345.4
## - MaritalStatus:LoanAmount      2   2229.4 2345.4
## - MaritalStatus:LoanTerm        2   2229.8 2345.8
## - Education:Age                 3   2232.1 2346.1
## <none>                              2228.3 2348.3
## - MaritalStatus:NumCreditLines  2   2239.4 2355.4
##
## Step:  AIC=2342.89
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:InterestRate + MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner +
##     Education:Age + Education:Income + Education:LoanAmount +
##     Education:CreditScore + Education:MonthsEmployed + Education:NumCreditLines +
##     Education:InterestRate
##
##                                Df Deviance    AIC
## - Education:MonthsEmployed      3   2229.9 2337.9
## - Education:InterestRate        3   2229.9 2337.9
## - Education:LoanAmount          3   2230.1 2338.1
## - MaritalStatus:InterestRate    2   2228.9 2338.9
## - Education:Income              3   2231.0 2339.0
## - MaritalStatus:MonthsEmployed  2   2229.1 2339.1
## - MaritalStatus:Income          2   2229.1 2339.1
## - Education:CreditScore         3   2231.4 2339.4
## - MaritalStatus:Age             2   2229.5 2339.5
## - Education:NumCreditLines      3   2231.6 2339.6
## - MaritalStatus:HasCoSigner     2   2229.9 2339.9
## - MaritalStatus:CreditScore     2   2230.0 2340.0
## - MaritalStatus:LoanAmount      2   2230.0 2340.0
## - MaritalStatus:LoanTerm        2   2230.4 2340.4
## - Education:Age                 3   2232.7 2340.7
## <none>                              2228.9 2342.9
## - MaritalStatus:NumCreditLines  2   2240.2 2350.2
##
```

```
## Step:  AIC=2337.93
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:InterestRate + MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner +
##     Education:Age + Education:Income + Education:LoanAmount +
##     Education:CreditScore + Education:NumCreditLines + Education:InterestRate
## 
##                                Df Deviance    AIC
## - Education:InterestRate        3   2230.9 2332.9
## - Education:LoanAmount          3   2231.1 2333.1
## - Education:Income              3   2231.9 2333.9
## - MaritalStatus:InterestRate    2   2229.9 2333.9
## - MaritalStatus:MonthsEmployed  2   2230.2 2334.2
## - MaritalStatus:Income          2   2230.2 2334.2
## - Education:CreditScore         3   2232.4 2334.4
## - MaritalStatus:Age             2   2230.5 2334.5
## - Education:NumCreditLines      3   2232.6 2334.6
## - MaritalStatus:HasCoSigner     2   2230.9 2334.9
## - MaritalStatus:CreditScore     2   2231.0 2335.0
## - MaritalStatus:LoanAmount      2   2231.1 2335.1
## - MaritalStatus:LoanTerm        2   2231.4 2335.4
## - Education:Age                 3   2233.7 2335.7
## <none>                              2229.9 2337.9
## - MaritalStatus:NumCreditLines  2   2241.1 2345.1
## 
## Step:  AIC=2332.94
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:InterestRate + MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner +
##     Education:Age + Education:Income + Education:LoanAmount +
##     Education:CreditScore + Education:NumCreditLines
## 
##                                Df Deviance    AIC
## - Education:LoanAmount          3   2232.1 2328.1
## - Education:Income              3   2232.9 2328.9
## - MaritalStatus:InterestRate    2   2231.0 2329.0
## - MaritalStatus:Income          2   2231.2 2329.2
## - MaritalStatus:MonthsEmployed  2   2231.2 2329.2
## - Education:CreditScore         3   2233.3 2329.3
## - MaritalStatus:Age             2   2231.5 2329.5
## - Education:NumCreditLines      3   2233.7 2329.7
## - MaritalStatus:HasCoSigner     2   2232.0 2330.0
## - MaritalStatus:CreditScore     2   2232.1 2330.1
## - MaritalStatus:LoanAmount      2   2232.1 2330.1
## - MaritalStatus:LoanTerm        2   2232.5 2330.5
## - Education:Age                 3   2234.7 2330.7
```

```
## <none>                              2230.9 2332.9
## - MaritalStatus:NumCreditLines  2   2242.2 2340.2
##
## Step:  AIC=2328.13
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:InterestRate + MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner +
##     Education:Age + Education:Income + Education:CreditScore +
##     Education:NumCreditLines
##
##                                 Df Deviance    AIC
## - Education:Income               3   2234.1 2324.1
## - MaritalStatus:InterestRate     2   2232.2 2324.2
## - MaritalStatus:Income           2   2232.3 2324.3
## - MaritalStatus:MonthsEmployed   2   2232.4 2324.4
## - Education:CreditScore          3   2234.5 2324.5
## - MaritalStatus:Age              2   2232.7 2324.7
## - Education:NumCreditLines       3   2234.9 2324.9
## - MaritalStatus:HasCoSigner      2   2233.2 2325.2
## - MaritalStatus:CreditScore      2   2233.3 2325.3
## - MaritalStatus:LoanAmount       2   2233.4 2325.4
## - MaritalStatus:LoanTerm         2   2233.7 2325.7
## - Education:Age                  3   2235.8 2325.8
## <none>                              2232.1 2328.1
## - MaritalStatus:NumCreditLines  2   2243.4 2335.4
##
## Step:  AIC=2324.06
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:InterestRate + MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner +
##     Education:Age + Education:CreditScore + Education:NumCreditLines
##
##                                 Df Deviance    AIC
## - MaritalStatus:InterestRate     2   2234.1 2320.1
## - MaritalStatus:MonthsEmployed   2   2234.3 2320.3
## - MaritalStatus:Income           2   2234.3 2320.3
## - Education:CreditScore          3   2236.6 2320.6
## - MaritalStatus:Age              2   2234.7 2320.7
## - Education:NumCreditLines       3   2236.8 2320.8
## - MaritalStatus:HasCoSigner      2   2235.1 2321.1
## - MaritalStatus:CreditScore      2   2235.2 2321.2
## - MaritalStatus:LoanAmount       2   2235.3 2321.3
## - MaritalStatus:LoanTerm         2   2235.5 2321.5
## - Education:Age                  3   2237.6 2321.6
## <none>                              2234.1 2324.1
## - MaritalStatus:NumCreditLines  2   2245.5 2331.5
```

```
##
## Step:  AIC=2320.1
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:MonthsEmployed + MaritalStatus:NumCreditLines +
##     MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner + Education:Age +
##     Education:CreditScore + Education:NumCreditLines
##
##                                Df Deviance    AIC
## - MaritalStatus:MonthsEmployed  2   2234.3 2316.3
## - MaritalStatus:Income          2   2234.3 2316.3
## - Education:CreditScore         3   2236.6 2316.6
## - MaritalStatus:Age             2   2234.7 2316.7
## - Education:NumCreditLines      3   2236.8 2316.8
## - MaritalStatus:HasCoSigner     2   2235.1 2317.1
## - MaritalStatus:CreditScore     2   2235.2 2317.2
## - MaritalStatus:LoanAmount      2   2235.4 2317.4
## - MaritalStatus:LoanTerm        2   2235.6 2317.6
## - Education:Age                 3   2237.6 2317.6
## <none>                              2234.1 2320.1
## - MaritalStatus:NumCreditLines  2   2245.6 2327.6
##
## Step:  AIC=2316.3
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:Income + MaritalStatus:LoanAmount + MaritalStatus:CreditScore +
##     MaritalStatus:NumCreditLines + MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner
+
##     Education:Age + Education:CreditScore + Education:NumCreditLines
##
##                                Df Deviance    AIC
## - MaritalStatus:Income          2   2234.5 2312.5
## - Education:CreditScore         3   2236.8 2312.8
## - MaritalStatus:Age             2   2234.9 2312.9
## - Education:NumCreditLines      3   2237.0 2313.0
## - MaritalStatus:HasCoSigner     2   2235.3 2313.3
## - MaritalStatus:CreditScore     2   2235.4 2313.4
## - MaritalStatus:LoanAmount      2   2235.6 2313.6
## - MaritalStatus:LoanTerm        2   2235.8 2313.8
## - Education:Age                 3   2237.8 2313.8
## <none>                              2234.3 2316.3
## - MaritalStatus:NumCreditLines  2   2245.8 2323.8
##
## Step:  AIC=2312.5
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:LoanAmount + MaritalStatus:CreditScore + MaritalStatus:NumCreditLin
es +
```

```
##     MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner + Education:Age +
##     Education:CreditScore + Education:NumCreditLines
##
##                                 Df Deviance    AIC
## - Education:CreditScore         3   2237.0 2309.0
## - MaritalStatus:Age             2   2235.1 2309.1
## - Education:NumCreditLines      3   2237.2 2309.2
## - MaritalStatus:HasCoSigner     2   2235.5 2309.5
## - MaritalStatus:CreditScore     2   2235.6 2309.6
## - MaritalStatus:LoanAmount      2   2235.8 2309.8
## - MaritalStatus:LoanTerm        2   2236.0 2310.0
## - Education:Age                 3   2238.1 2310.1
## <none>                              2234.5 2312.5
## - MaritalStatus:NumCreditLines  2   2245.9 2319.9
##
## Step:  AIC=2309
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:Age +
##     MaritalStatus:LoanAmount + MaritalStatus:CreditScore + MaritalStatus:NumCreditLin
es +
##     MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner + Education:Age +
##     Education:NumCreditLines
##
##                                 Df Deviance    AIC
## - MaritalStatus:Age             2   2237.6 2305.6
## - Education:NumCreditLines      3   2239.6 2305.6
## - MaritalStatus:HasCoSigner     2   2238.1 2306.1
## - MaritalStatus:CreditScore     2   2238.2 2306.2
## - MaritalStatus:LoanAmount      2   2238.3 2306.3
## - MaritalStatus:LoanTerm        2   2238.5 2306.5
## - Education:Age                 3   2240.7 2306.7
## <none>                              2237.0 2309.0
## - MaritalStatus:NumCreditLines  2   2248.4 2316.4
##
## Step:  AIC=2305.57
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:LoanAmount +
##     MaritalStatus:CreditScore + MaritalStatus:NumCreditLines +
##     MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner + Education:Age +
##     Education:NumCreditLines
##
##                                 Df Deviance    AIC
## - Education:NumCreditLines      3   2240.1 2302.1
## - MaritalStatus:HasCoSigner     2   2238.7 2302.7
## - MaritalStatus:CreditScore     2   2238.8 2302.8
## - MaritalStatus:LoanAmount      2   2238.9 2302.9
## - MaritalStatus:LoanTerm        2   2239.0 2303.0
## - Education:Age                 3   2241.2 2303.2
## <none>                              2237.6 2305.6
## - MaritalStatus:NumCreditLines  2   2249.1 2313.1
```

```
##
## Step:  AIC=2302.11
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:LoanAmount +
##     MaritalStatus:CreditScore + MaritalStatus:NumCreditLines +
##     MaritalStatus:LoanTerm + MaritalStatus:HasCoSigner + Education:Age
##
##                                 Df Deviance    AIC
## - MaritalStatus:HasCoSigner      2   2241.1 2299.1
## - MaritalStatus:CreditScore      2   2241.2 2299.2
## - MaritalStatus:LoanAmount       2   2241.4 2299.4
## - MaritalStatus:LoanTerm         2   2241.4 2299.4
## - Education:Age                  3   2244.1 2300.1
## <none>                               2240.1 2302.1
## - MaritalStatus:NumCreditLines  2   2251.9 2309.9
##
## Step:  AIC=2299.13
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:LoanAmount +
##     MaritalStatus:CreditScore + MaritalStatus:NumCreditLines +
##     MaritalStatus:LoanTerm + Education:Age
##
##                                 Df Deviance    AIC
## - MaritalStatus:CreditScore      2   2242.2 2296.2
## - MaritalStatus:LoanAmount       2   2242.4 2296.4
## - MaritalStatus:LoanTerm         2   2242.5 2296.5
## - Education:Age                  3   2245.1 2297.1
## <none>                               2241.1 2299.1
## - MaritalStatus:NumCreditLines  2   2252.9 2306.9
##
## Step:  AIC=2296.18
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:LoanAmount +
##     MaritalStatus:NumCreditLines + MaritalStatus:LoanTerm + Education:Age
##
##                                 Df Deviance    AIC
## - MaritalStatus:LoanAmount       2   2243.4 2293.4
## - MaritalStatus:LoanTerm         2   2243.6 2293.6
## - Education:Age                  3   2246.2 2294.2
## <none>                               2242.2 2296.2
## - MaritalStatus:NumCreditLines  2   2253.9 2303.9
##
## Step:  AIC=2293.41
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:NumCreditLines +
##     MaritalStatus:LoanTerm + Education:Age
##
##                                 Df Deviance    AIC
```

```
## - MaritalStatus:LoanTerm        2   2244.8 2290.8
## - Education:Age                 3   2247.6 2291.6
## <none>                              2243.4 2293.4
## - MaritalStatus:NumCreditLines  2   2255.1 2301.1
##
## Step:  AIC=2290.83
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:NumCreditLines +
##     Education:Age
##
##                                Df Deviance    AIC
## - Education:Age                 3   2248.8 2288.8
## <none>                              2244.8 2290.8
## - MaritalStatus:NumCreditLines  2   2256.2 2298.2
##
## Step:  AIC=2288.81
## Default ~ MaritalStatus + Education + Age + Income + LoanAmount +
##     CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
##     LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:NumCreditLines
##
##                                Df Deviance    AIC
## <none>                              2248.8 2288.8
## - MaritalStatus:NumCreditLines  2   2259.7 2295.7
```

```
##
## Call:  glm(formula = Default ~ MaritalStatus + Education + Age + Income +
##     LoanAmount + CreditScore + MonthsEmployed + NumCreditLines +
##     InterestRate + LoanTerm + EmploymentType + HasCoSigner +
##     MaritalStatus:NumCreditLines, family = "binomial", data = loanDefaultData)
##
## Coefficients:
##                       (Intercept)                MaritalStatusMarried
##                        -3.415e-01                          -3.093e-01
##               MaritalStatusSingle                 EducationHigh School
##                        -9.849e-01                           2.521e-01
##                EducationMaster's                          EducationPhD
##                         1.338e-01                           3.652e-02
##                               Age                                Income
##                        -3.647e-02                          -6.552e-06
##                        LoanAmount                           CreditScore
##                         3.163e-06                          -1.137e-03
##                    MonthsEmployed                        NumCreditLines
##                        -9.297e-03                           3.009e-02
##                      InterestRate                              LoanTerm
##                         6.376e-02                          -3.349e-03
##            EmploymentTypePart-time          EmploymentTypeSelf-employed
##                         2.664e-01                           4.365e-01
##           EmploymentTypeUnemployed                        HasCoSignerYes
##                         5.542e-01                          -3.092e-01
## MaritalStatusMarried:NumCreditLines   MaritalStatusSingle:NumCreditLines
##                        -1.395e-03                           3.510e-01
##
## Degrees of Freedom: 3499 Total (i.e. Null);  3480 Residual
## Null Deviance:       2524
## Residual Deviance: 2249  AIC: 2289
```

Once again both methods return the same model, and have included only the interaction terms between MaritalStatus and NumCreditLines (marital status and number of credit line an applicant has)

```
current.model = glm(formula = Default ~ MaritalStatus + Education + Age + Income +
    LoanAmount + CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
    LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:NumCreditLines,
    family = "binomial", data = loanDefaultData)
```

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

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

```
##
## Call:
## glm(formula = Default ~ MaritalStatus + Education + Age + Income +
##     LoanAmount + CreditScore + MonthsEmployed + NumCreditLines +
##     InterestRate + LoanTerm + EmploymentType + HasCoSigner +
##     MaritalStatus:NumCreditLines, family = "binomial", data = loanDefaultData)
##
## Coefficients:
##                                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)                         -3.415e-01  4.424e-01  -0.772 0.440125
## MaritalStatusMarried                -3.093e-01  3.299e-01  -0.938 0.348488
## MaritalStatusSingle                 -9.849e-01  3.476e-01  -2.833 0.004607 **
## EducationHigh School                 2.521e-01  1.550e-01   1.626 0.103947
## EducationMaster's                    1.338e-01  1.568e-01   0.853 0.393397
## EducationPhD                         3.652e-02  1.603e-01   0.228 0.819788
## Age                                 -3.647e-02  3.911e-03  -9.326  < 2e-16 ***
## Income                              -6.552e-06  1.433e-06  -4.570 4.87e-06 ***
## LoanAmount                           3.163e-06  7.946e-07   3.980 6.88e-05 ***
## CreditScore                         -1.137e-03  3.520e-04  -3.231 0.001235 **
## MonthsEmployed                      -9.297e-03  1.650e-03  -5.635 1.75e-08 ***
## NumCreditLines                       3.009e-02  8.136e-02   0.370 0.711446
## InterestRate                         6.376e-02  8.628e-03   7.390 1.47e-13 ***
## LoanTerm                            -3.349e-03  3.270e-03  -1.024 0.305766
## EmploymentTypePart-time              2.664e-01  1.661e-01   1.603 0.108828
## EmploymentTypeSelf-employed          4.365e-01  1.612e-01   2.708 0.006772 **
## EmploymentTypeUnemployed             5.542e-01  1.620e-01   3.421 0.000624 ***
## HasCoSignerYes                      -3.092e-01  1.116e-01  -2.772 0.005578 **
## MaritalStatusMarried:NumCreditLines -1.395e-03  1.219e-01  -0.011 0.990873
## MaritalStatusSingle:NumCreditLines   3.510e-01  1.205e-01   2.913 0.003583 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2524.3  on 3499  degrees of freedom
## Residual deviance: 2248.8  on 3480  degrees of freedom
## AIC: 2288.8
##
## Number of Fisher Scoring iterations: 5
```

The variable that tells us that any characteristic that has to do with an applicant's marital status being married is insignificant on its own except for Marital Status Single and its interaction with number of credit lines. For example, the MaritalStatusMarried, MaritalStatusMarried:NumCreditLines. In addition, the Education levels, LoanTerm, EmploymentTypePart-time appears to be insignificant as well. This proves that my initial guess was wrong, that these terms do not have a effect on default as I thought they would be. Thus, I removed them one by one to produce our final model.

```
final.model = glm(formula = Default ~ Age + Income + LoanAmount + CreditScore +
    MonthsEmployed + NumCreditLines + InterestRate + LoanTerm +
    EmploymentType + Education, family = "binomial", data = loanDefaultData)
summary(final.model)
```

```
##
## Call:
## glm(formula = Default ~ Age + Income + LoanAmount + CreditScore +
##     MonthsEmployed + NumCreditLines + InterestRate + LoanTerm +
##     EmploymentType + Education, family = "binomial", data = loanDefaultData)
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 -9.322e-01  4.019e-01  -2.319 0.020379 *
## Age                         -3.626e-02  3.880e-03  -9.345  < 2e-16 ***
## Income                      -6.401e-06  1.427e-06  -4.487 7.22e-06 ***
## LoanAmount                   3.062e-06  7.882e-07   3.884 0.000103 ***
## CreditScore                 -1.127e-03  3.497e-04  -3.224 0.001265 **
## MonthsEmployed              -9.467e-03  1.642e-03  -5.766 8.10e-09 ***
## NumCreditLines               1.466e-01  4.935e-02   2.971 0.002972 **
## InterestRate                 6.457e-02  8.593e-03   7.515 5.71e-14 ***
## LoanTerm                    -2.612e-03  3.244e-03  -0.805 0.420670
## EmploymentTypePart-time      2.846e-01  1.652e-01   1.723 0.084963 .
## EmploymentTypeSelf-employed  4.327e-01  1.603e-01   2.699 0.006944 **
## EmploymentTypeUnemployed     5.357e-01  1.613e-01   3.321 0.000897 ***
## EducationHigh School         2.546e-01  1.542e-01   1.651 0.098812 .
## EducationMaster's            1.483e-01  1.557e-01   0.953 0.340757
## EducationPhD                 1.943e-02  1.593e-01   0.122 0.902938
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2524.3  on 3499  degrees of freedom
## Residual deviance: 2273.7  on 3485  degrees of freedom
## AIC: 2303.7
##
## Number of Fisher Scoring iterations: 5
```

After removing them, I realized that with the removal, the AIC value got even worse, which was what I focused on. I also noticed there were insignificant terms like the highest education levels that the applicant has obtained and also the loan term. Thus, I removed those terms as well.

```
final.model = glm(formula = Default ~ Age + Income + LoanAmount + CreditScore +
    MonthsEmployed + NumCreditLines + InterestRate + LoanTerm, family = "binomial", data
= loanDefaultData)
summary(final.model)
```

```
##
## Call:
## glm(formula = Default ~ Age + Income + LoanAmount + CreditScore +
##     MonthsEmployed + NumCreditLines + InterestRate + LoanTerm,
##     family = "binomial", data = loanDefaultData)
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -5.086e-01  3.699e-01  -1.375 0.169130
## Age            -3.606e-02  3.861e-03  -9.341  < 2e-16 ***
## Income         -6.490e-06  1.422e-06  -4.563 5.05e-06 ***
## LoanAmount      2.961e-06  7.852e-07   3.771 0.000163 ***
## CreditScore    -1.065e-03  3.475e-04  -3.064 0.002182 **
## MonthsEmployed -9.460e-03  1.632e-03  -5.796 6.80e-09 ***
## NumCreditLines  1.503e-01  4.914e-02   3.058 0.002225 **
## InterestRate    6.313e-02  8.552e-03   7.382 1.56e-13 ***
## LoanTerm       -2.664e-03  3.222e-03  -0.827 0.408368
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2524.3  on 3499  degrees of freedom
## Residual deviance: 2289.7  on 3491  degrees of freedom
## AIC: 2307.7
##
## Number of Fisher Scoring iterations: 5
```

Even after the removal, the model still got worse as its AIC model value got higher. Since I was focusing mainly on this score, this made me realized that I have probably removed a variable that was important. It might have been because the variable that I removed was not so important on its own but really important if you took into consideration its interaction with other terms. I tested again by adding back some variables one by one or multiple at one time, but the model just keeps getting worse and worse. This made me question if it is because all the variables that I have thought before were actually all significant factors. Maybe the initial model with everything was actually the best model already (looking mostly at AIC value). Thus, I tried some more combinations by observing the initial dataset again. In the end, I came to the conclusion that not removing any of the variables but taking into considering the marriage and number of credit line’s interaction is by far the best model in terms of AIC.

```
final.model = glm(formula = Default ~ MaritalStatus + Education + Age + Income +
    LoanAmount + CreditScore + MonthsEmployed + NumCreditLines + InterestRate +
    LoanTerm + EmploymentType + HasCoSigner + MaritalStatus:NumCreditLines,
    family = "binomial", data = loanDefaultData)
summary(final.model)
```

```
##
## Call:
## glm(formula = Default ~ MaritalStatus + Education + Age + Income +
##     LoanAmount + CreditScore + MonthsEmployed + NumCreditLines +
##     InterestRate + LoanTerm + EmploymentType + HasCoSigner +
##     MaritalStatus:NumCreditLines, family = "binomial", data = loanDefaultData)
##
## Coefficients:
##                                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)                         -3.415e-01  4.424e-01  -0.772 0.440125
## MaritalStatusMarried                -3.093e-01  3.299e-01  -0.938 0.348488
## MaritalStatusSingle                 -9.849e-01  3.476e-01  -2.833 0.004607 **
## EducationHigh School                 2.521e-01  1.550e-01   1.626 0.103947
## EducationMaster's                    1.338e-01  1.568e-01   0.853 0.393397
## EducationPhD                         3.652e-02  1.603e-01   0.228 0.819788
## Age                                 -3.647e-02  3.911e-03  -9.326  < 2e-16 ***
## Income                              -6.552e-06  1.433e-06  -4.570 4.87e-06 ***
## LoanAmount                           3.163e-06  7.946e-07   3.980 6.88e-05 ***
## CreditScore                         -1.137e-03  3.520e-04  -3.231 0.001235 **
## MonthsEmployed                      -9.297e-03  1.650e-03  -5.635 1.75e-08 ***
## NumCreditLines                       3.009e-02  8.136e-02   0.370 0.711446
## InterestRate                         6.376e-02  8.628e-03   7.390 1.47e-13 ***
## LoanTerm                            -3.349e-03  3.270e-03  -1.024 0.305766
## EmploymentTypePart-time              2.664e-01  1.661e-01   1.603 0.108828
## EmploymentTypeSelf-employed          4.365e-01  1.612e-01   2.708 0.006772 **
## EmploymentTypeUnemployed             5.542e-01  1.620e-01   3.421 0.000624 ***
## HasCoSignerYes                      -3.092e-01  1.116e-01  -2.772 0.005578 **
## MaritalStatusMarried:NumCreditLines -1.395e-03  1.219e-01  -0.011 0.990873
## MaritalStatusSingle:NumCreditLines   3.510e-01  1.205e-01   2.913 0.003583 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2524.3  on 3499  degrees of freedom
## Residual deviance: 2248.8  on 3480  degrees of freedom
## AIC: 2288.8
##
## Number of Fisher Scoring iterations: 5
```

Model summary (8 points)

The final resulting model found was:

log-odds(Default) = -0.342 + -0.985 * MaritalStatusSingle - 0.309 * MaritalStatusMarried + 0.252 * EducationHighSchool + 0.134 * EducationMaster’s + 0.0365 * EducationPhD - 0.0365 * Age - 0.00000655 * Income + 0.00000316 * LoanAmount - 0.00114 * CreditScore - 0.00930 * MonthsEmployed + 0.0301 * NumCreditLines + 0.0648 * InterestRate - 0.00335 * LoanTerm + 0.266 * EmploymentTypePart-time + 0.437 * EmploymentTypeSelf-employed + 0.554 * EmploymentTypeUnemployed - 0.309 * HasCoSignerYes - 0.001395 * MaritalStatusMarried:NumCreditLines + 0.3510 * MaritalStatusSingle:NumCreditLines

Each term in the model is statistically significant except for marital status, education level and loanTerm, but the inclusion of the interaction term MaritalStatusSingle and NumCreditLines which is significant indicates that we keep the marital status term as well.

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

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

```
##                        (Intercept)              MaritalStatusMarried
##                          0.7106688                         0.7339390
##                MaritalStatusSingle             EducationHigh School
##                          0.3734649                         1.2867114
##                  EducationMaster's                      EducationPhD
##                          1.1431698                         1.0371929
##                                Age                            Income
##                          0.9641866                         0.9999934
##                         LoanAmount                       CreditScore
##                          1.0000032                         0.9988635
##                     MonthsEmployed                    NumCreditLines
##                          0.9907461                         1.0305523
##                       InterestRate                          LoanTerm
##                          1.0658363                         0.9966565
##           EmploymentTypePart-time       EmploymentTypeSelf-employed
##                          1.3052814                         1.5473174
##          EmploymentTypeUnemployed                      HasCoSignerYes
##                          1.7406012                         0.7340144
## MaritalStatusMarried:NumCreditLines  MaritalStatusSingle:NumCreditLines
##                          0.9986061                         1.4204487
```

Note: The baseline individual in this study is a divorced individual with a Bachelor's level of education and average values for continuous predictors like Age and Income.

- Marital Status: is not a significant factor that effects our model a lot individually as its p-value is greater than 0.05, but I kept it for reasons explained above (AIC value of the model, interaction with others, as well as since its the topic of interest).
- MaritalStatus = Single: Being single - as compared to being another baseline catagory - divorced, saw a decrease in the probability of defaulting by about 0.373 times or 62.7% decrease.
- MaritalStatus = Married: Being married, as compared to being divorced, has caused the probability to default by about 0.734 times as well.
- Education = HighSchool, Phd or Masters: Having a education level that is not Bachelors (the other baseline catagory) all saw about the same increase as the baseline in the probability of defaulting. With high school having a increase of 1.29 times, with PhD of having a increase of 1.04 times, and Masters having a increase of 1.14 times. Thus, it is not very significant impact to the model, but for the same reason I kept marital status, I kept education as well.
- Age: each year older the applicant is, the probability of default actually does not decrease much (only about 3%) however, if taken as a whole, it decreases by about 0.964 times. In other words, this was included as it impacted the AIC value of the model in subtle ways.
- CreditScore: decreases the odds of default by a factor of approximately 0.999 times.This shows that each point increase does not effect much, but if you take into consideration multiple points, the effects might stack up. Thus, I kept this value as well.

- LoanTerm: Each additional term month decreases the odds by about 0.997 times.
- HasCoSignerYes: Having a co-signer as compared to the baseline of having no co-signer decreases the odds of default by a factor of about 0.73.
- 1 USD dollar higher income or loan amount have a very small effect on the odds, close to 1, which indicates a negligible change. However, if taken all together in large sums (like 1 million USD increase each time), it is statistically significant. Thus, it is included in the model.
- Being employed as either part time or self-employed, as compared to the baseline of full time, increases the odds of default by about 1.55 times or 1.31 times.
- Each additional month employed decreases the probability of default by about 0.991 times. Thus, not really significant but I kept it for its effect on the AIC value.
- Additional credit lines very slightly increase the probability of default by a factor of approximately 1.03 times.
- interest rates: Each percentage point increase increases the probability about 1.07 times
- MaritalStatusMarried:NumCreditLines: Being single and having more credit lines decrease the probability of defaulting more than if someone is divorced, about 0.999 times more.
- MaritalStatusSingle:NumCreditLines: Being single and having more credit lines increases the probability of defaulting more than if someone is divorced after accounting for the interaction with the number of credit lines, about 1.42 times more. Compared to being married though, it does increase the probability of default a lot more as well.

Conclusion (8 points)

This analysis demonstrated that almost all variables in the original dataset had a significant impact on defualt if solely defining a good model with its AIC value and its interactions and if marital status and highest level of education was focused on more in the study. After some trial and error and steps mentioned in the above report, an applicant’s personal demographics, financial status, and loan terms are all significant predictors of loan default, with marital status and employment type showing particularly strong associations (also including their interaction terms). The presence of a co-signer also impacted the odds of default.

The variable that positively impact the odds are the number of credit lines, being employed as part time or self-employed, having higher loan amount, and having more months being employed or not a bachelor as the highest degree.

The variable that negatively impact the odds are having a co-signer, higher income, marital status, age, credit score, loan term and being unemployed.

Overall this analysis indicates that marital status and education levels do not impact the odds of default as other variables when considered individually, but considering the relationships it had with the number of credit lines and through trial and error of taking out variables in the AIC models and checking its effect on the AIC value, it was kept in the final model. It shows that there is a preference for a divorced, long term employee that does not have a bachelors degree and is requesting a higher loan amount holding more credit lines as well. It seems that someone who might be counted as “more in need” would be more likely to be defaulted.