Readmission Analysis

Amira Burns

2023-01-24

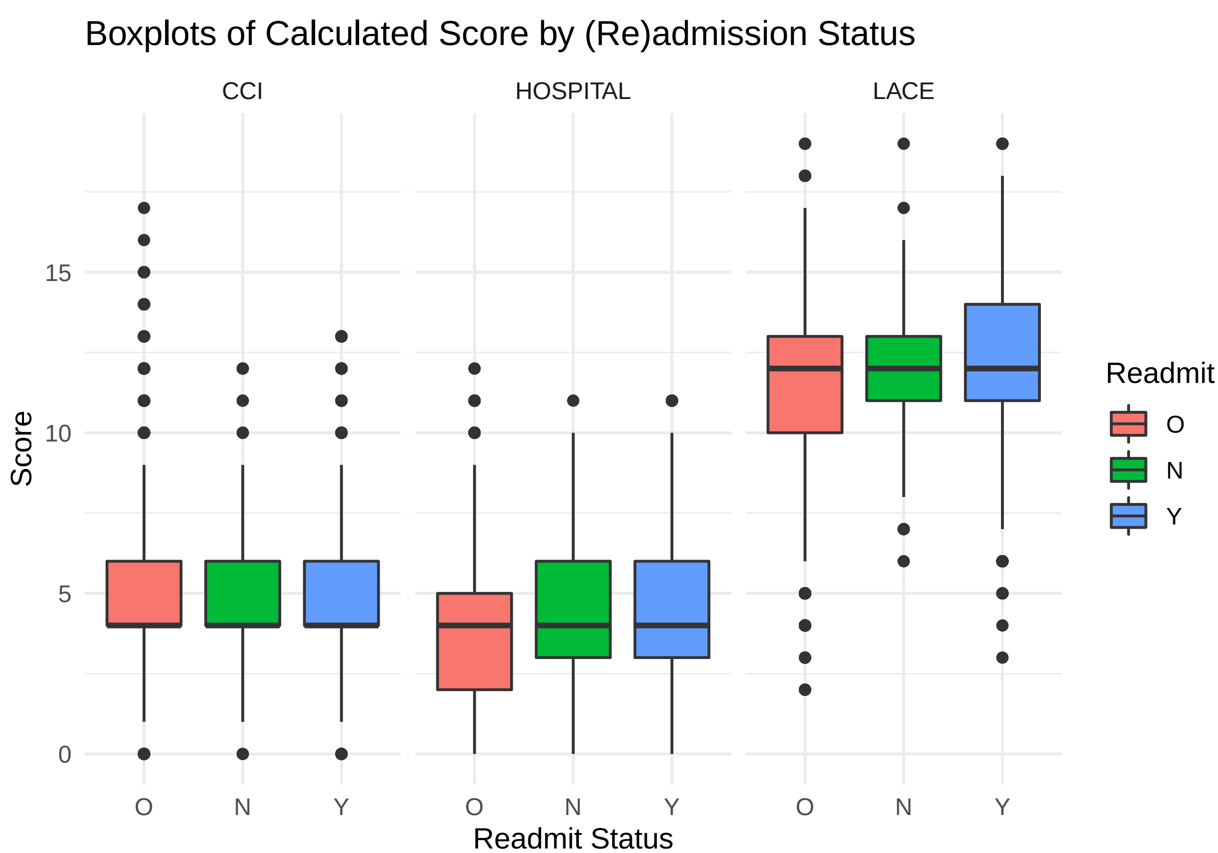
Readmission Analysis EDA

Explore any patterns related to readmission status. These records are classified as an original admission ("O"), readmission within 30 days of discharge after original hospital stay ("Y"), or readmission outside the 30-day window after original discharge ("N").

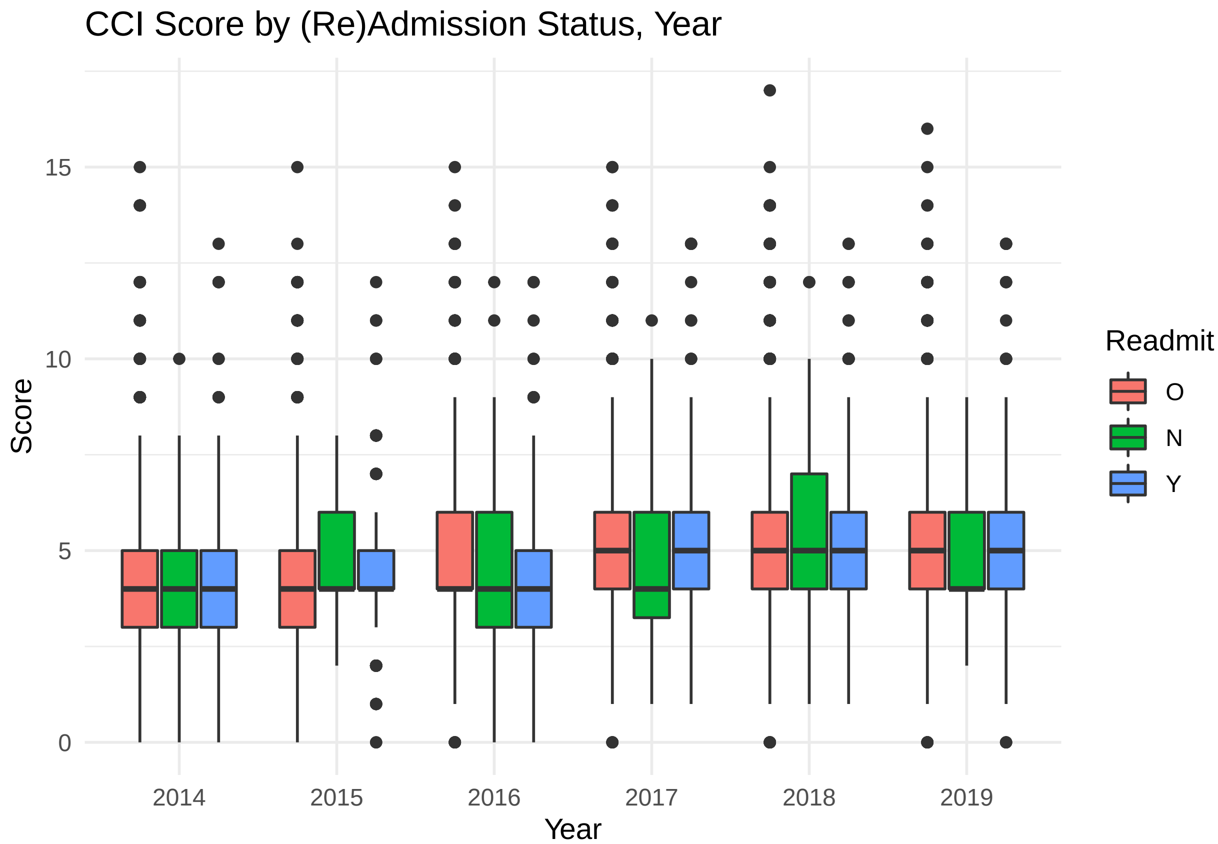
When accounting for all demographics, many sub-group samples are size N = 0. In order to preserve differences between type of admission, we can aggregate over Race/Ethnicity and note this topic as worthy of study with a future dataset.

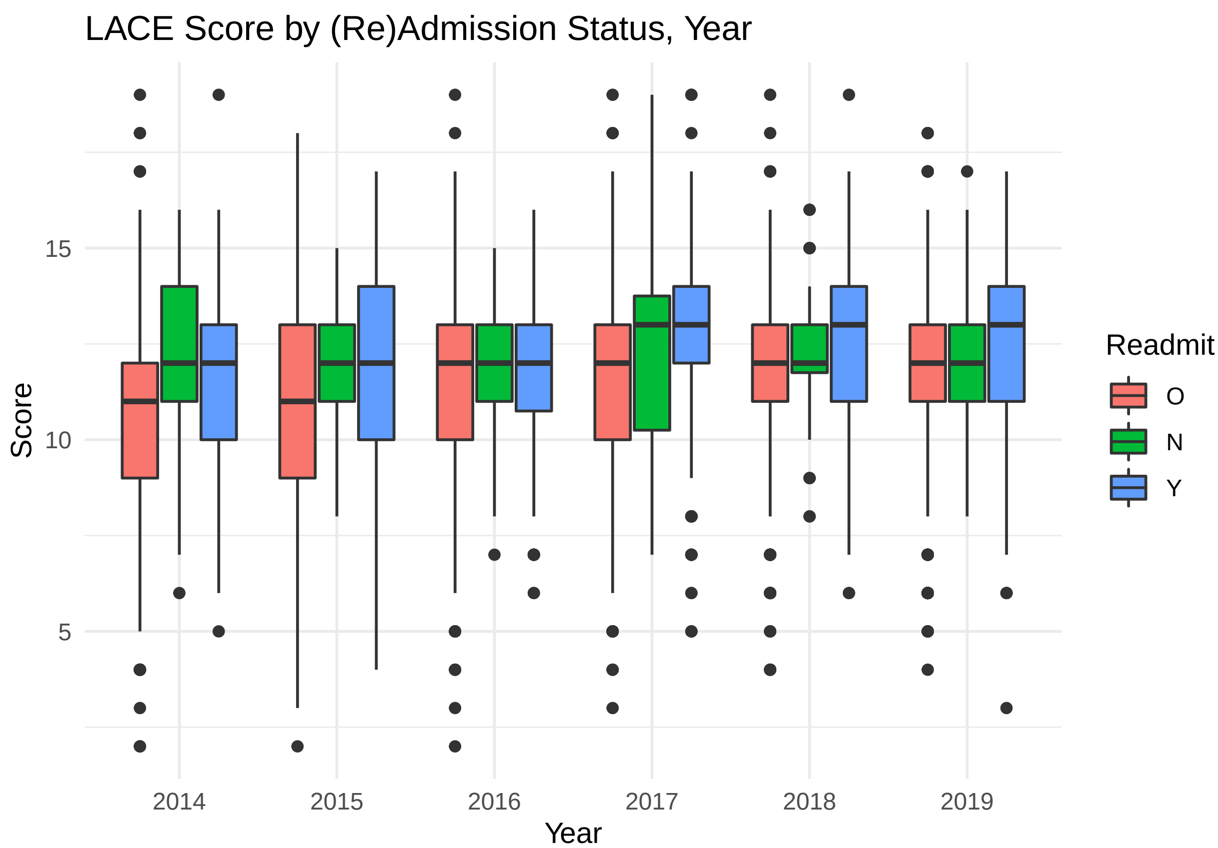
The only group with N=0 sample size is Men readmitted outside 30 days without Medicare/Medicaid in 2017.

Let's look at visualizations to explore any patterns between readmission status and calculated scores or demographic factors.

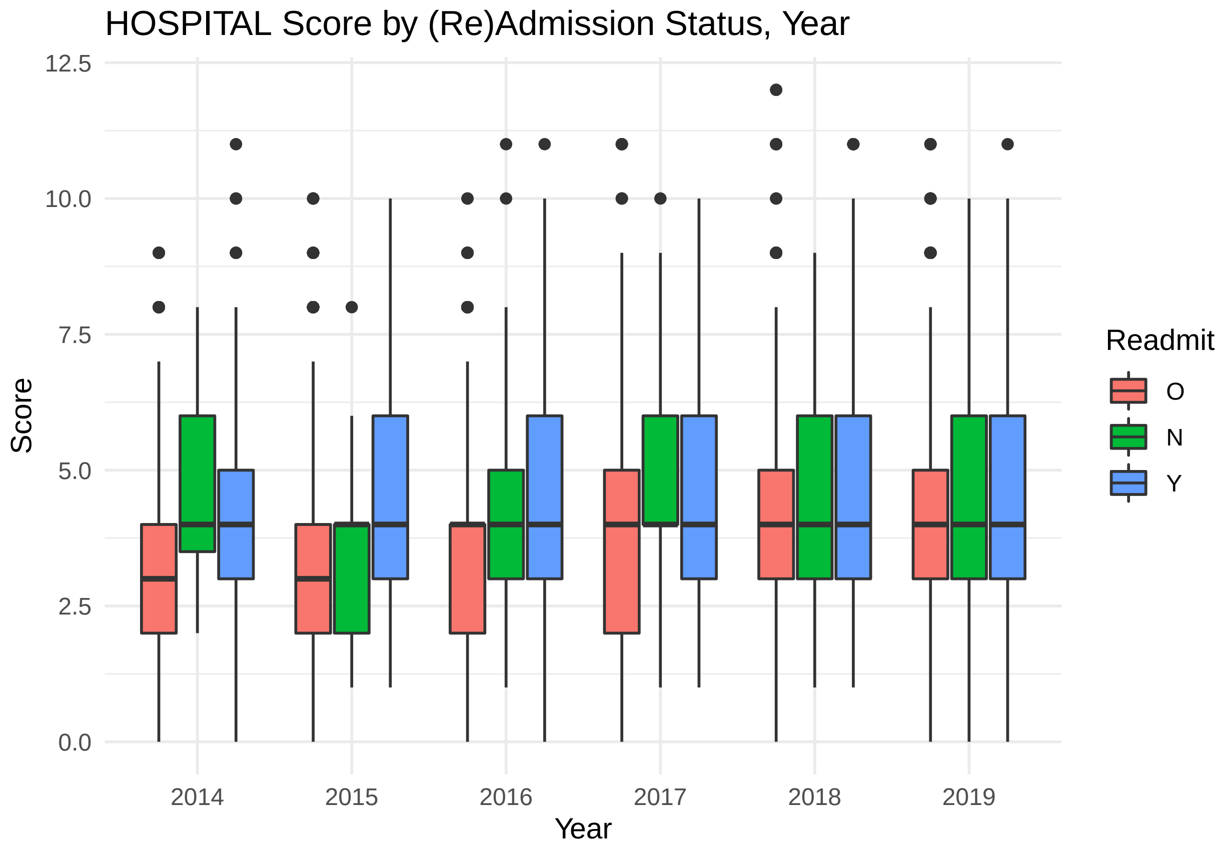


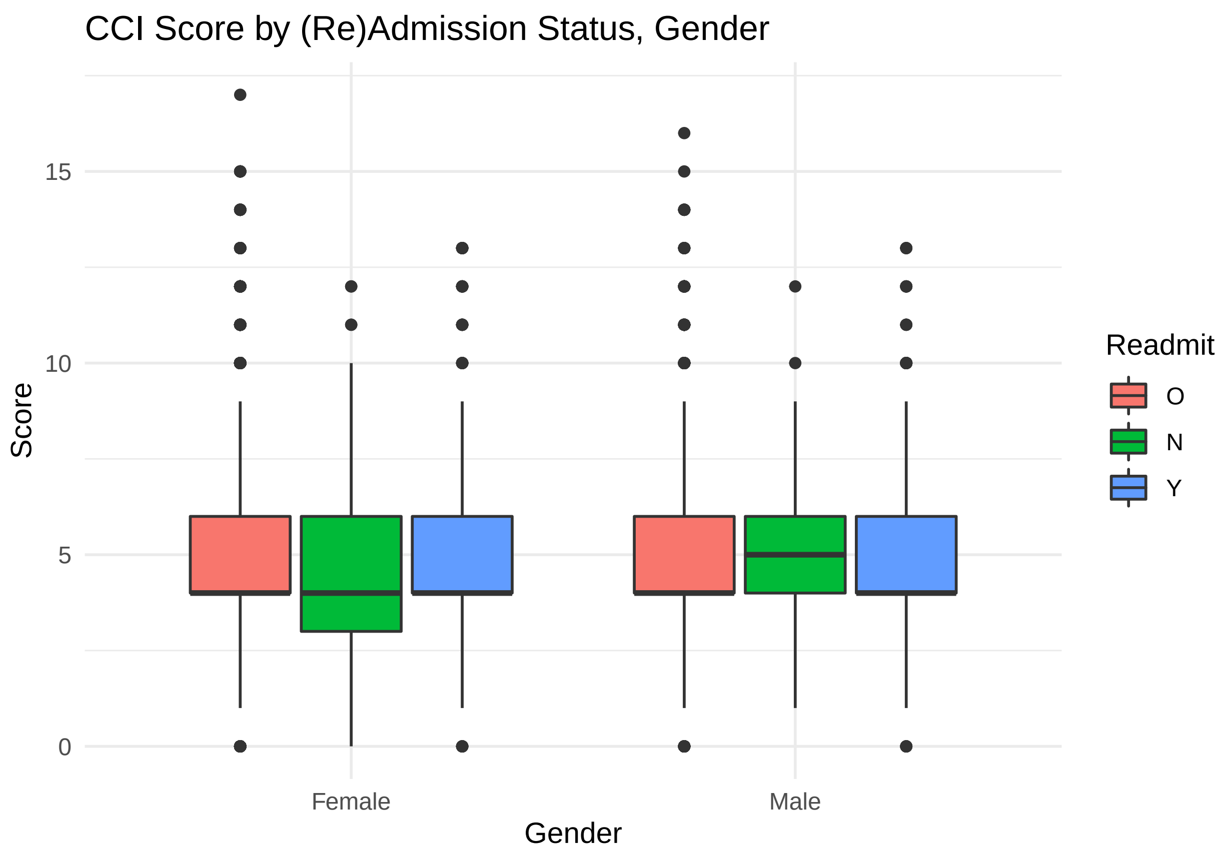
When aggregating across all demographics, there are no visually striking differences in readmission status across score.

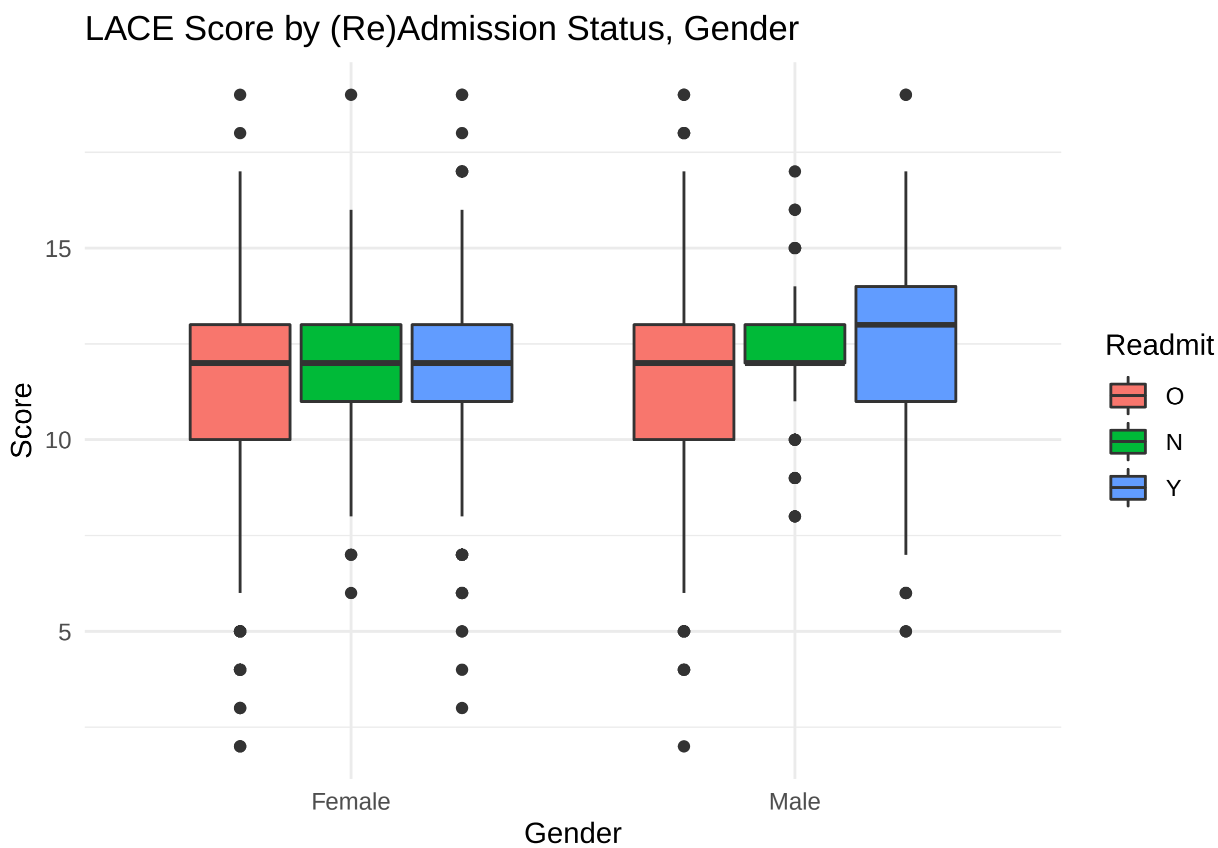




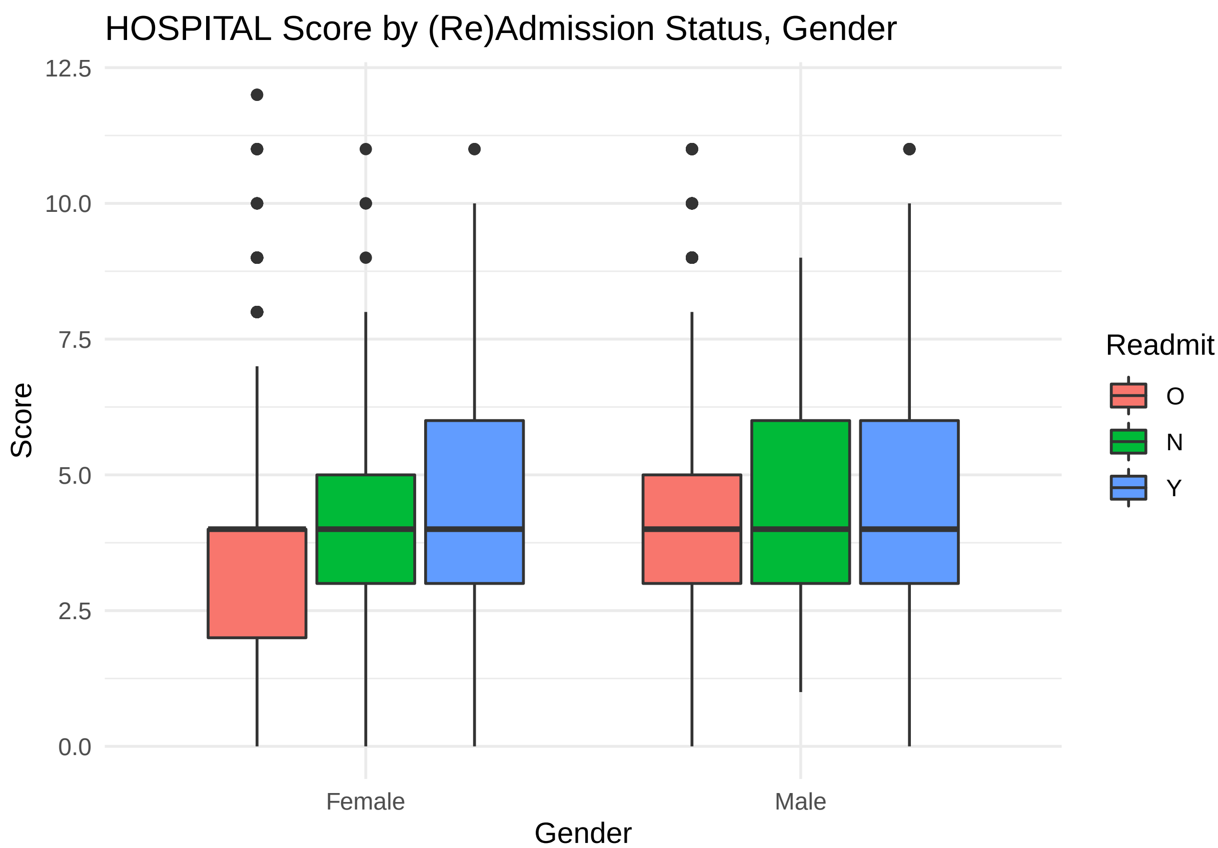
LACE Score by (Re)Admission Status, Year
Score
15
10
5
2014
2015
2016
2017
2018
2019
Year
Readmit
O
N
Y


HOSPITAL Score by (Re)Admission Status, Year
Score
Year
Readmit
O
N
Y
12.5
10.0
7.5
5.0
2.5
0.0
2014
2015
2016
2017
2018
2019

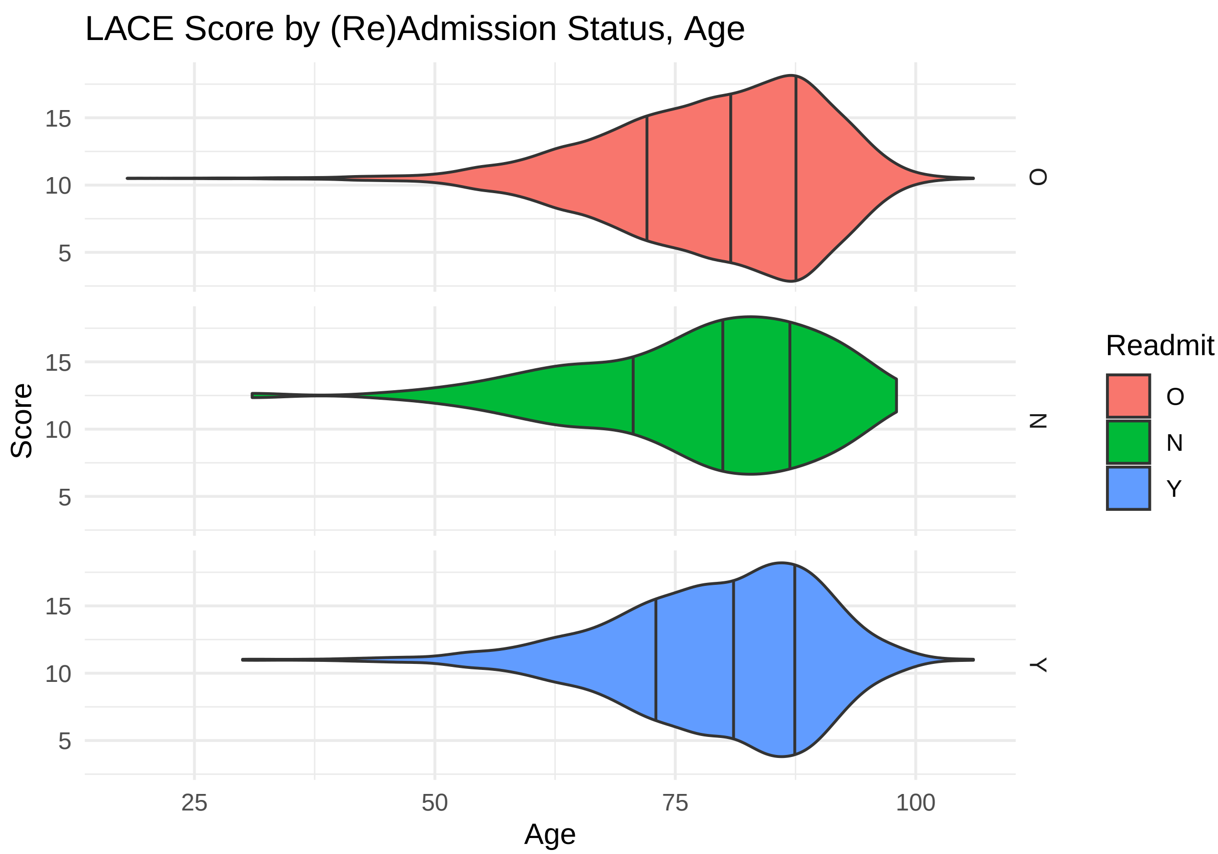

CCI Score by (Re)Admission Status, Gender
Score
15
10
5
0
Female
Male
Gender
Readmit
O
N
Y

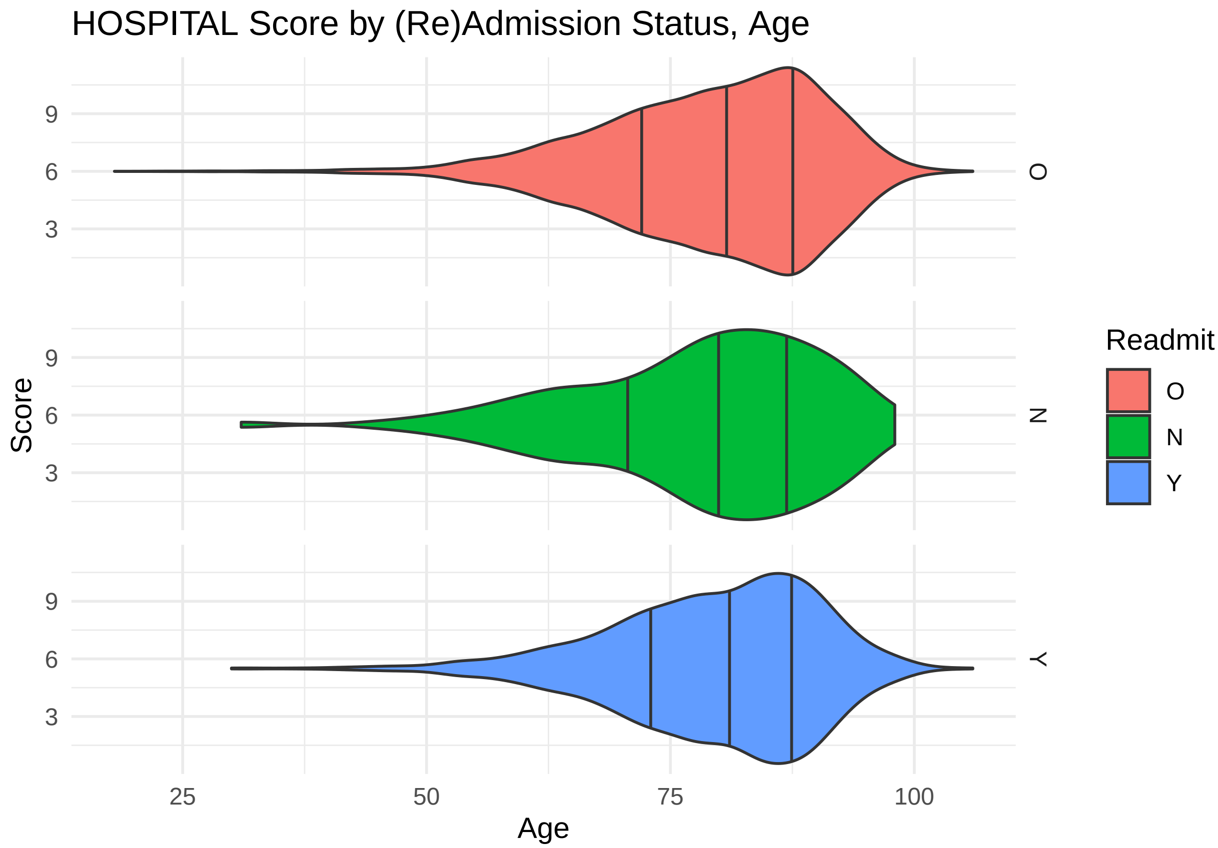


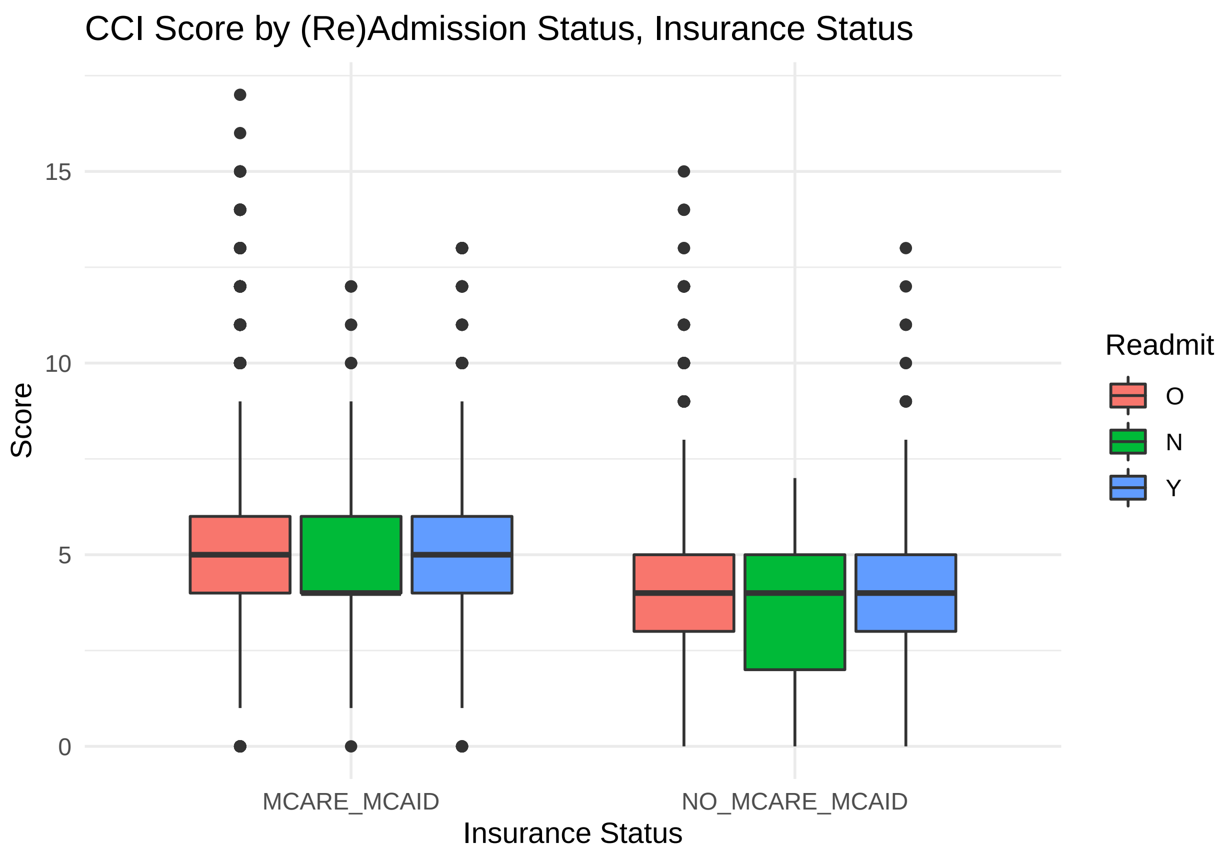
LACE Score by (Re)Admission Status, Gender
Score
15
10
5
Female
Male
Gender
Readmit
O
N
Y

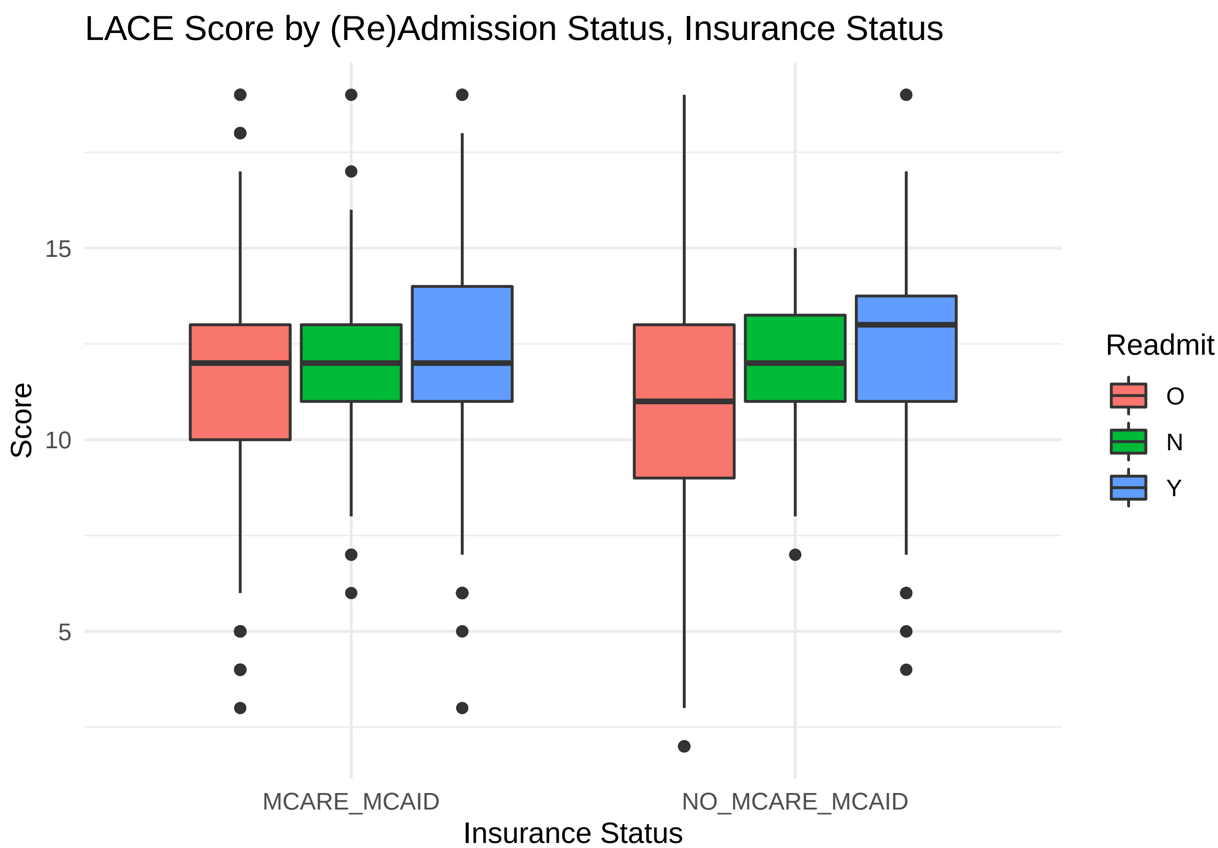

HOSPITAL Score by (Re)Admission Status, Gender
12.5
10.0
7.5
5.0
2.5
0.0
Score
Female
Male
Gender
Readmit
O
N
Y

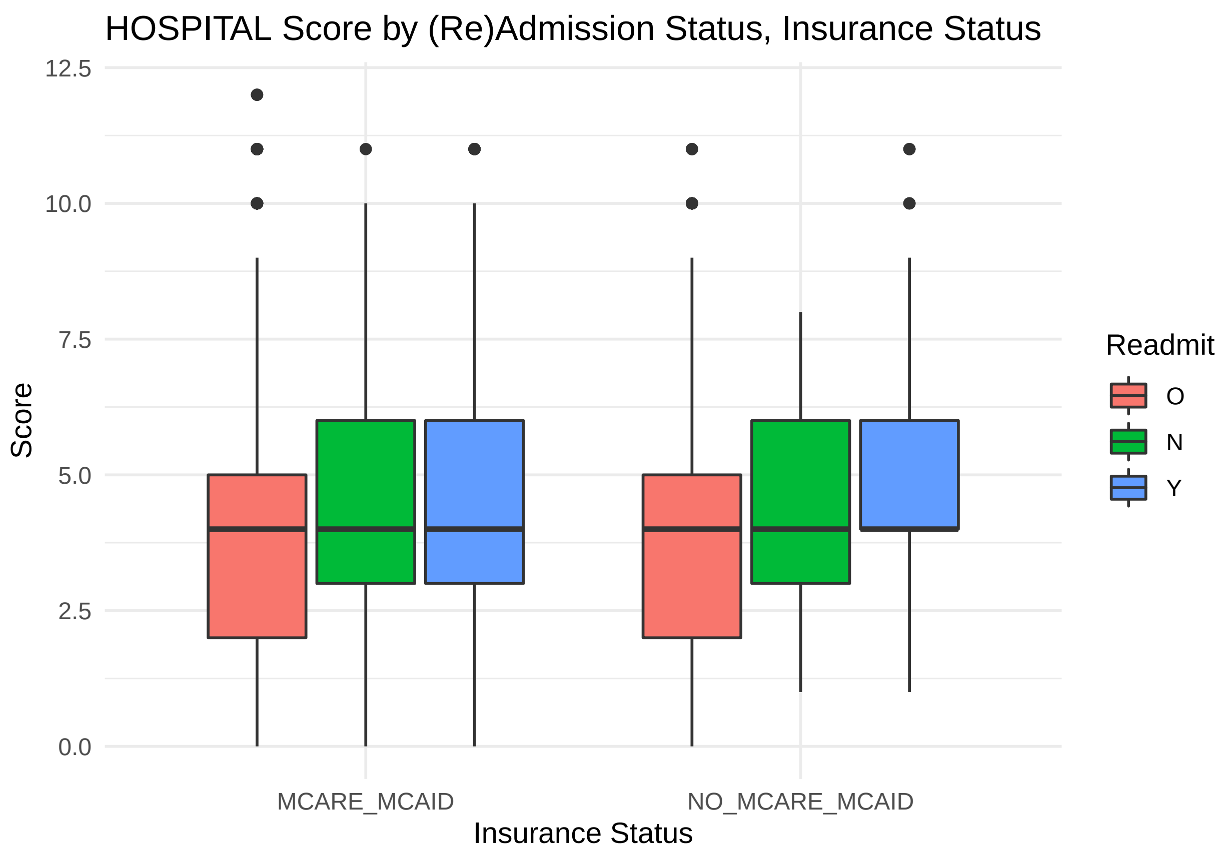

CCI Score by (Re)Admission Status, Age
Score
Age
25
50
75
100
0
5
10
15
O
N
Y
Readmit
O
N
Y


LACE Score by (Re)Admission Status, Age
Score
15
10
5
Age
25
50
75
100
O
N
Y
Readmit
O
N
Y


HOSPITAL Score by (Re)Admission Status, Age
Score
Age
9
6
3
25
50
75
100
O
N
Y
Readmit
O
N
Y


CCI Score by (Re)Admission Status, Insurance Status
Score
15
10
5
0
MCARE_MCAID
NO_MCARE_MCAID
Insurance Status
Readmit
O
N
Y


LACE Score by (Re)Admission Status, Insurance Status
15
10
5
Score
MCARE_MCAID
NO_MCARE_MCAID
Insurance Status
Readmit
O
N
Y



Model Fitting

Correlation check among scores does not cause any concern regarding including all scores in a given model. Since these calculated scores are the main predictors of interest, they will remain in all models assessed.

```
##                 CCI      LACE   HOSPITAL
## CCI      1.00000000 0.3782846 0.03486513
## LACE     0.37828456 1.0000000 0.44171172
## HOSPITAL 0.03486513 0.4417117 1.00000000
```

Two possible model classes were explored. If all records that are not readmissions within 30 days are categorized together, binomial logistic regression models were fit. If different categories for readmissions outisde 30 days and original admissions were preserved, multinomial logistic regression models were fit.

Binomial Logistic Regression

We may consider there to be no meaningful clinical difference between an original admission and a readmission outside 30 days. Consequently, combine the "O" and "N" factor levels of Readmit and treat the response as a binary variable. The appropriate model class is logistic regression.

```
## # A tibble: 24 x 14
## # Groups:   Year, Gender, Insurance [24]
##     Year Gender Insurance      n_N   n_Y PropR~1 meanA~2 meanA~3 meanC~4 meanC~5
##    <dbl> <fct>  <fct>        <int> <int>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
```

```
## 1 2014 Female MCARE_MCAID    822   134 0.140     80.6    80.9    4.52    4.41
## 2 2014 Female NO_MCARE_MC~   253    23 0.0833    66.5    72.5    3.25    4.35
## 3 2014 Male   MCARE_MCAID    434   106 0.196     80.4    80.6    4.56    4.42
## 4 2014 Male   NO_MCARE_MC~   167    32 0.161     68.0    68.9    3.59    4.06
## 5 2015 Female MCARE_MCAID   1010   163 0.139     80.8    81.1    4.60    4.62
## 6 2015 Female NO_MCARE_MC~   235    26 0.0996    68.6    70.3    3.76    3.42
## 7 2015 Male   MCARE_MCAID    482   105 0.179     80.8    82.5    4.73    4.75
## 8 2015 Male   NO_MCARE_MC~   148    18 0.108     67.3    72.4    3.61    4.11
## 9 2016 Female MCARE_MCAID   1105   159 0.126     81.4    81.8    4.80    4.82
## 10 2016 Female NO_MCARE_MC~  216    39 0.153     71.9    67.2    4.04    3.33
## # ... with 14 more rows, 4 more variables: meanLACE_N <dbl>, meanLACE_Y <dbl>,
## #   meanHOS_N <dbl>, meanHOS_Y <dbl>, and abbreviated variable names
## #   1: PropReadmit, 2: meanAge_N, 3: meanAge_Y, 4: meanCCI_N, 5: meanCCI_Y
```

Model Selection: perform drop-in-deviance tests to compare nested models

```
## # A tibble: 12 x 6
##    Model              logLik    AIC    BIC deviance df.residual
##    <chr>               <dbl>  <dbl>  <dbl>    <dbl>       <int>
##  1 Null Model         -5980. 11963. 11971.   11961.       13809
##  2 Scores             -5762. 11531. 11562.   11523.       13806
##  3 Year               -5978. 11960. 11975.   11956.       13808
##  4 Scores + Year      -5745. 11499. 11537.   11489.       13805
##  5 Age                -5979. 11962. 11978.   11958.       13808
##  6 Scores + Age       -5758. 11527. 11565.   11517.       13805
##  7 Gender             -5949. 11903. 11918.   11899.       13808
##  8 Scores + Gender    -5745. 11499. 11537.   11489.       13805
##  9 Insurance          -5978. 11960. 11975.   11956.       13808
## 10 Scores + Insurance -5758. 11526. 11564.   11516.       13805
## 11 Additive Model     -5738. 11492. 11553.   11476.       13802
## 12 Saturated Model    -5649. 11554. 12518.   11298.       13682
```

```
## Analysis of Deviance Table
##
## Model 1: Readmit ~ CCI + LACE + HOSPITAL + Year + PatientAge + Gender +
##     Insurance
## Model 2: Readmit ~ CCI * LACE * HOSPITAL * Year * PatientAge * Gender *
##     Insurance
##   Resid. Df Resid. Dev  Df Deviance  Pr(>Chi)
## 1     13802      11476
## 2     13682      11298 120   178.64 0.0004156 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

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

Despite the drop-in-deviance test determining that the saturated model better fits the data, the additive model has a lower AIC. The additional benefit of parsimony (8 parameters in the additive model instead of 128 parameters in the full model) allows us to proceed to model diagnostics for the additive model.

```
##
## Call:
## glm(formula = Readmit ~ CCI + LACE + HOSPITAL + Year + PatientAge +
##     Gender + Insurance, family = binomial(link = "logit"), data = snf_data_b)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.3390  -0.6227  -0.5190  -0.4060   2.5204
##
## Coefficients:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)             14.027524  29.273082   0.479 0.631800
## CCI                     -0.056634   0.014700  -3.853 0.000117 ***
## LACE                     0.094405   0.012976   7.275 3.46e-13 ***
## HOSPITAL                 0.205246   0.014577  14.081  < 2e-16 ***
## Year                    -0.008882   0.014524  -0.611 0.540871
## PatientAge               0.005238   0.002612   2.005 0.044944 *
## GenderMale               0.293986   0.048475   6.065 1.32e-09 ***
## InsuranceNO_MCARE_MCAID -0.140956   0.068054  -2.071 0.038336 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 11961  on 13809  degrees of freedom
## Residual deviance: 11476  on 13802  degrees of freedom
## AIC: 11492
##
## Number of Fisher Scoring iterations: 4
```

```
## Analysis of Deviance Table
##
## Model 1: Readmit ~ Gender + CCI + LACE + HOSPITAL
## Model 2: Readmit ~ CCI + LACE + HOSPITAL + Year + PatientAge + Gender +
##     Insurance
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1     13805      11489
## 2     13802      11476  3   12.938 0.004772 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on adjusted p-values, compare the additive model to the model including just scores and gender. The simpler model has higher AIC, and the drop-in-deviance test favors the larger model, so we'll stick with the additive model. The model can be written as:

$$\begin{aligned} log(\frac{p_i}{1-p_i}) &= \beta_0 + \beta_1 * CCI_i + \beta_2 * LACE_i + \beta_3 * HOSPITAL_i + \beta_4 * Year_i + \beta_5 * Age_i + \beta_6 * Gender_{Male} + \beta_7 * Insura \\ &= 14.028 - 0.057 * CCI_i + 0.094 * LACE_i + 0.202 * HOSPITAL_i - 0.009 * Year_i + 0.005 * Age_i + 0.294 * Gend \end{aligned}$$

$$p_i = \frac{exp(\beta_0 + \beta_1 * CCI_i + \beta_2 * LACE_i + \beta_3 * HOSPITAL_i + \beta_4 * Year_i + \beta_5 * Age_i + \beta_6 * Gender_{Male} + \beta_7 *}{1 + exp(\beta_0 + \beta_1 * CCI_i + \beta_2 * LACE_i + \beta_3 * HOSPITAL_i + \beta_4 * Year_i + \beta_5 * Age_i + \beta_6 * Gender_{Male} + \beta_7}$$

where p_i is the probability that a given record in the dataset is a readmission within 30 days.

AB: need to perform model diagnostics to check assumptions, create visual of model with empirical logits

Multinomial Logistic Regression

```
## # weights:  6 (2 variable)
## initial  value 15171.835707
## iter  10 value 7287.505124
## final  value 7184.477659
## converged
```

```
## # weights:  15 (8 variable)
## initial  value 15171.835707
## iter  10 value 8564.536062
## iter  20 value 6970.830740
## iter  30 value 6941.291910
## iter  30 value 6941.291907
## iter  30 value 6941.291907
## final  value 6941.291907
## converged
```

```
## # weights:  9 (4 variable)
## initial  value 15171.835707
## iter  10 value 7184.528898
## iter  20 value 7182.158962
## iter  30 value 7181.665933
## iter  40 value 7181.402783
## final  value 7181.285110
## converged
```

```
## # weights:  18 (10 variable)
## initial  value 15171.835707
## iter  10 value 6973.092127
## iter  20 value 6941.239907
## iter  30 value 6941.187688
## iter  40 value 6941.136828
## final  value 6941.136710
## converged
```

```
## # weights:  9 (4 variable)
## initial  value 15171.835707
## iter  10 value 7182.148717
## iter  10 value 7182.148716
## final  value 7182.148716
## converged
```

```
## # weights:  18 (10 variable)
## initial  value 15171.835707
## iter  10 value 7168.608150
## final  value 6937.171883
## converged
```

```
## # weights:  9 (4 variable)
## initial  value 15171.835707
## iter  10 value 7511.346643
## final  value 7151.416012
## converged
```

```
## # weights:  18 (10 variable)
## initial  value 15171.835707
## iter  10 value 8783.684715
## iter  20 value 7043.483235
## iter  30 value 6927.899616
## final  value 6923.394610
## converged
```

```
## # weights:  9 (4 variable)
## initial  value 15171.835707
## iter  10 value 7215.731043
## final  value 7181.923967
## converged
```

```
## # weights:  18 (10 variable)
## initial  value 15171.835707
## iter  10 value 8898.295813
## iter  20 value 6982.006718
## iter  30 value 6939.734375
## final  value 6937.540011
## converged
```

AB: need to perform model selection and model diagnostics