Untitled

Jyun-Yu Cheng, Haokun Zhang

2023-02-15

Question 1: Saratoga house prices
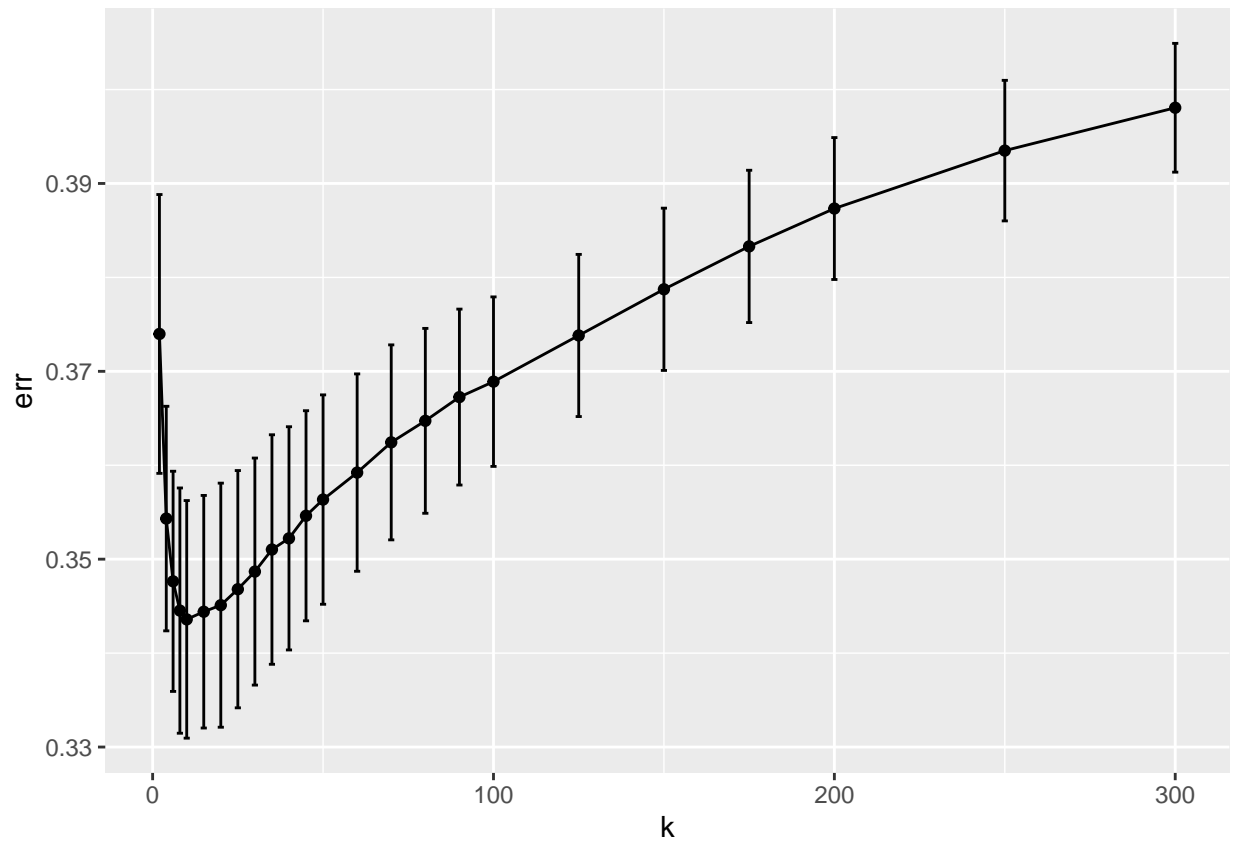
```
## [1] 0.3876865
```

```
## [1] 0.3630362
```

```
## [1] 0.3609134
```

Since lm3 has the least rmse, lm3 is the best model among these above 3 lm models

```
##            k       err     std_err
## result.1   2 0.3739806 0.01484344
## result.2   4 0.3543227 0.01195440
## result.3   6 0.3476416 0.01171929
## result.4   8 0.3445254 0.01305642
## result.5  10 0.3435925 0.01264814
## result.6  15 0.3444047 0.01238207
```

We can get the best k according to the following:

[1] 10

Then we calculate the best knn method's error

[1] 0.4013657

Then averaging the estimate of out-of-sample RMSE over many different random train/test splits for the linear models, either randomly or by cross-validation.

[1] 0.3573336 0.3268562 0.3267688

[1] 0.3468531

The error from the optimal knn model is 0.3486, larger than that of linear model 3, which is 0.29. Thus linear model seems to do better at achieving lower out-of-sample mean-squared error. Therefore as for price-modeling strategies for a local taxing authority, we should pay more attention to the linear model prediction to estimate market values for properties.

Question 2: Classification and retrospective sampling

Default Probabiltiy by Credit History
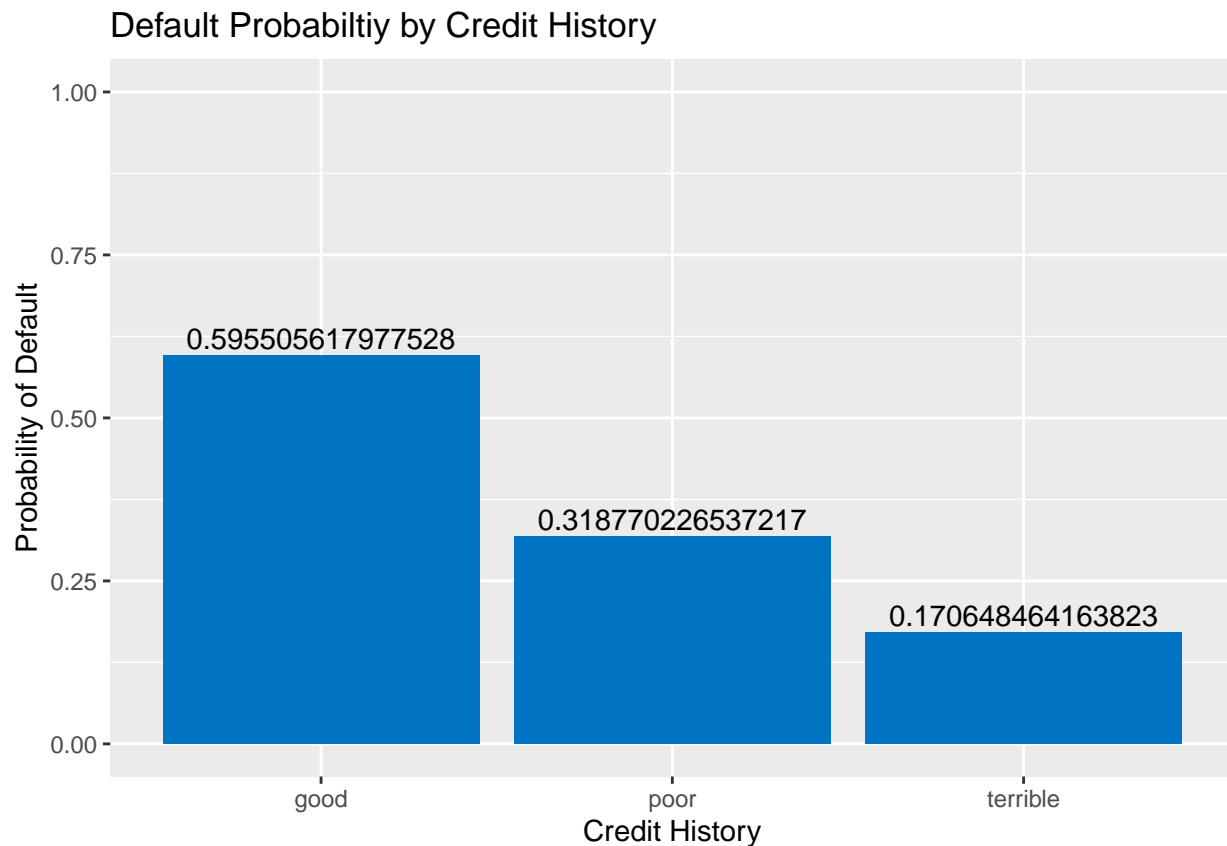
After running the regression, a out-of-sample performance is examined.

	coef
(Intercept)	-0.76
duration	0.03
amount	0.00
installment	0.26
age	-0.02
historypoor	-1.23
historyterrible	-2.08
purposeedu	0.73
purposegoods/repair	0.14
purposenewcar	0.94
purposeusedcar	-0.66

	coef
foreigngerman	-1.63

	0	1
0	132	12
1	40	16

The coefficients for having poor and terrible credit history are negative. They are also statsitically signicifant. Having a poor or terrible credit history multiplies odds of default, which counter-intuitively has negative effect on default. The retrospective, "case-control" method that the bank used to select samples resulted in a substantial oversampling of defaults, relative to a random sample of loans in the bank's overall portfolio. If the purpose of the model is to screen prospective borrowers to classify them into "high" versus "low" probability of default, this data set is not appropriate for building a predictive model.

Table 3: Default Sample Size

Default Type	Sample Size
0	700
1	300

To improve the sampling scheme under limited resources to get the data on all subjects in the loans, it requires a method to reduce the oversampling of defaults. Previously, the set of non-defauled loans ("controls") is matched with similar set of defaulted loans, leaving "controls" less independent of the set of defaulted loans, the "controls" are not reprensentative of the source population that produced the default cases either. Selection bias resulted. A better sampling scheme should adhere to fixing these two issues. A random sampling method will give defaulted loans and non-defaulted loans equal chance of being selected, and the "controls" is selected independently from the cases.

The second step is to increase the ratio between non-defaulted loan and defaulted loan. Since defaulted loans are rare cases and non-defaulted loans are plentiful, the statistical power of the study can be increased by enrolling more non-defaulted loans than the defaulted. However, ratio more than 4 controls over 1 case adds little impact on power. Therefore, if it is time-consuming or expensive to collect sets of controls, the ratio should not exceed 4:1. Alternatively, if it is not expensive, there is no reason to limit the number of non-defaulted loans. In this case of 300 defaulted loans, the number of non-defaulted loans can be increased from 700 to 1200 or even more. With these two steps, the sampling scheme will be improved.

Question 3: Children and hotel reservations

```
##              hotel lead_time stays_in_weekend_nights stays_in_week_nights adults
## 1     City_Hotel       217                         1                    3      2
## 2     City_Hotel         2                         0                    1      2
## 3   Resort_Hotel        95                         2                    5      2
## 4   Resort_Hotel       143                         2                    6      2
## 5   Resort_Hotel       136                         1                    4      2
## 6     City_Hotel        67                         2                    2      2
##    children meal market_segment distribution_channel is_repeated_guest
## 1         0   BB    Offline_TA/TO                TA/TO                 0
## 2         0   BB           Direct               Direct                 0
## 3         0   BB        Online_TA                TA/TO                 0
## 4         0   HB        Online_TA                TA/TO                 0
## 5         0   HB           Direct               Direct                 0
## 6         0   SC        Online_TA                TA/TO                 0
##    previous_cancellations previous_bookings_not_canceled reserved_room_type
## 1                       0                              0                  A
## 2                       0                              0                  D
## 3                       0                              0                  A
## 4                       0                              0                  A
## 5                       0                              0                  F
## 6                       0                              0                  A
##    assigned_room_type booking_changes deposit_type days_in_waiting_list
## 1                   A               0   No_Deposit                    0
## 2                   K               0   No_Deposit                    0
## 3                   A               2   No_Deposit                    0
## 4                   A               0   No_Deposit                    0
## 5                   F               0   No_Deposit                    0
```

```
## 6                      A          0   No_Deposit                     0
##     customer_type average_daily_rate required_car_parking_spaces
## 1 Transient-Party              80.75                        none
## 2       Transient             170.00                        none
## 3       Transient               8.00                        none
## 4       Transient              81.00                        none
## 5       Transient             157.60                        none
## 6       Transient              49.09                        none
##   total_of_special_requests arrival_date
## 1                         1   2016-09-01
## 2                         3   2017-08-25
## 3                         2   2016-11-19
## 4                         1   2016-04-26
## 5                         4   2016-12-28
## 6                         1   2016-03-13


##           hotel lead_time stays_in_weekend_nights stays_in_week_nights adults
## 1 Resort_Hotel        47                       0                    2      2
## 2 Resort_Hotel        46                       0                    2      2
## 3   City_Hotel        22                       1                    2      2
## 4 Resort_Hotel       209                       2                    5      2
## 5   City_Hotel         1                       1                    0      2
## 6 Resort_Hotel       171                       2                    5      2
##   children meal market_segment distribution_channel is_repeated_guest
## 1        1   BB         Direct               Direct                 0
## 2        0   BB  Offline_TA/TO                TA/TO                 0
## 3        0   BB      Online_TA                TA/TO                 0
## 4        0   BB  Offline_TA/TO                TA/TO                 0
## 5        0   SC      Online_TA                TA/TO                 0
## 6        0   BB      Online_TA                TA/TO                 0
##   previous_cancellations previous_bookings_not_canceled reserved_room_type
## 1                      0                              0                  C
## 2                      0                              0                  D
## 3                      0                              0                  D
## 4                      0                              0                  A
## 5                      0                              0                  A
## 6                      0                              0                  D
##   assigned_room_type booking_changes deposit_type days_in_waiting_list
## 1                  C               0   No_Deposit                    0
## 2                  D               0   No_Deposit                    0
## 3                  D               0   No_Deposit                    0
## 4                  A               0   No_Deposit                    0
## 5                  B               0   No_Deposit                    0
## 6                  D               2   No_Deposit                    0
##     customer_type average_daily_rate required_car_parking_spaces
## 1       Transient             289.00                        none
## 2       Transient             162.00                        none
## 3       Transient             121.33                        none
## 4       Transient              76.22                        none
## 5       Transient              98.00                        none
## 6 Transient-Party             182.86                        none
##   total_of_special_requests arrival_date
## 1                         1   2017-08-23
## 2                         0   2016-12-30
```
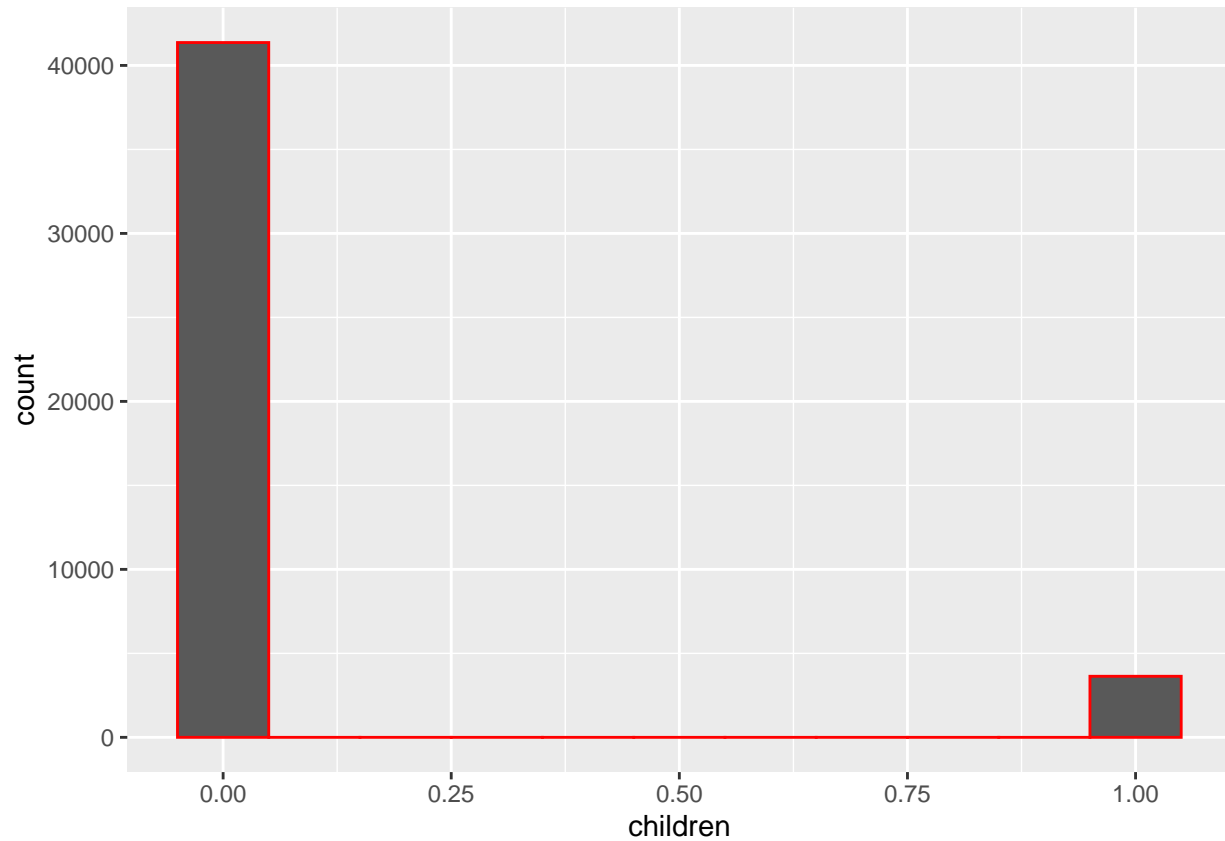
```
## 3                               0    2017-03-13
## 4                               0    2016-07-12
## 5                               0    2016-08-07
## 6                               1    2017-07-25
```



```
## fold 1,2,3,4,5,6,7,8,9,10,done.


## fold 1,2,3,4,5,6,7,8,9,10,done.


## 49 x 1 sparse Matrix of class "dgCMatrix"
##                                          seg100
## intercept                          -4.7956085151
## hotelCity_Hotel                     0.6897506739
## hotelResort_Hotel                  -0.0011376561
## lead_time                           0.0002685365
## stays_in_weekend_nights             0.0397920703
## stays_in_week_nights               -0.0106634164
## adults                             -0.4891192486
## mealFB                              0.6203768343
## mealHB                              0.0096185470
## mealSC                             -1.0882648838
## mealUndefined                           .
## market_segmentComplementary             .
## market_segmentCorporate            -1.0848601982
## market_segmentDirect                    .
```

```
## market_segmentGroups              -1.0176854274
## market_segmentOffline_TA/TO            .
## market_segmentOnline_TA              0.0779494280
## distribution_channelDirect          0.1529181154
## distribution_channelGDS            -1.3769640426
## distribution_channelTA/TO               .
## is_repeated_guest                 -0.9098917297
## previous_cancellations                  .
## previous_bookings_not_canceled    -0.1046358854
## reserved_room_typeB                1.7947794204
## reserved_room_typeC                2.6326786811
## reserved_room_typeD               -1.0328999229
## reserved_room_typeE               -0.0686080615
## reserved_room_typeF                1.7164159281
## reserved_room_typeG                2.4189060003
## reserved_room_typeH                3.3134757000
## assigned_room_typeB                0.1446359244
## assigned_room_typeC                1.5937943111
## assigned_room_typeD                0.9718280946
## assigned_room_typeE                0.5515719951
## assigned_room_typeF                0.6979563600
## assigned_room_typeG                0.9139488697
## assigned_room_typeH                1.2846092279
## assigned_room_typeI                1.0890703841
## assigned_room_typeK                0.2103499212
## booking_changes                   0.2262499243
## deposit_typeNon_Refund                  .
## deposit_typeRefundable                  .
## days_in_waiting_list                    .
## customer_typeGroup                      .
## customer_typeTransient             0.5998673915
## customer_typeTransient-Party      -0.0729396579
## average_daily_rate                 0.0108359585
## required_car_parking_spacesparking 0.0467396601
## total_of_special_requests          0.4742417935


##                              strong_interaction_name strong_interaction_beta
## 1           reserved_room_typeH:assigned_room_typeF         4.09873398871384
## 2       market_segmentOnline_TA:reserved_room_typeB        -3.93136324342353
## 3          market_segmentGroups:reserved_room_typeG         3.84710711628772
## 4   market_segmentComplementary:reserved_room_typeC         3.32479215779641
## 5              hotelResort_Hotel:reserved_room_typeB         2.93631897530884
## 6          market_segmentGroups:assigned_room_typeH         2.86214867810281
## 7       market_segmentOnline_TA:assigned_room_typeG          2.7476190075697
## 8   market_segmentOffline_TA/TO:reserved_room_typeH         2.73420967934745
## 9                   mealHB:reserved_room_typeF         2.68324234254356
## 10          reserved_room_typeD:assigned_room_typeB         2.62856949536017
## 11            mealUndefined:assigned_room_typeD        -2.61492096120591
## 12                  mealFB:assigned_room_typeI        -2.57181425344971
## 13          reserved_room_typeB:assigned_room_typeG         2.51791185120615
## 14          reserved_room_typeF:assigned_room_typeK         2.42192995406018
## 15         market_segmentDirect:reserved_room_typeF         2.33719353758446
## 16          reserved_room_typeB:assigned_room_typeB        -2.30846492323049
## 17 market_segmentComplementary:reserved_room_typeF         2.27818040322205
```

```
## 18                        mealSC:is_repeated_guest              2.2094392216579
## 19            reserved_room_typeB:assigned_room_typeE           2.18474462691401
## 20            adults:previous_bookings_not_canceled            -2.17696298519968
## 21                   mealUndefined:reserved_room_typeB         -2.0071697242497
## 22              hotelResort_Hotel:reserved_room_typeE          -1.92793866713012
## 23                        mealSC:reserved_room_typeF           -1.91022735496279
## 24            reserved_room_typeD:assigned_room_typeC           1.86921687922732
## 25     market_segmentCorporate:assigned_room_typeK            -1.85292423595847
## 26                              reserved_room_typeG             1.77972634154438
## 27                              reserved_room_typeF             1.75917856316537
## 28            reserved_room_typeF:assigned_room_typeB           1.74846565287662
## 29                              reserved_room_typeE             1.74760646427176
## 30              mealUndefined:reserved_room_typeE             -1.71399934638127
##     abs_beta
## 1   4.098734
## 2   3.931363
## 3   3.847107
## 4   3.324792
## 5   2.936319
## 6   2.862149
## 7   2.747619
## 8   2.734210
## 9   2.683242
## 10  2.628569
## 11  2.614921
## 12  2.571814
## 13  2.517912
## 14  2.421930
## 15  2.337194
## 16  2.308465
## 17  2.278180
## 18  2.209439
## 19  2.184745
## 20  2.176963
## 21  2.007170
## 22  1.927939
## 23  1.910227
## 24  1.869217
## 25  1.852924
## 26  1.779726
## 27  1.759179
## 28  1.748466
## 29  1.747606
## 30  1.713999


##                                   (Intercept)
##                                  -6.228235e+10
##                               hotelResort_Hotel
##                                  -2.786083e+04
##                                      lead_time
##                                  -5.621157e+01
##                           stays_in_weekend_nights
##                                   2.751483e+03
##                            stays_in_week_nights
```

```
##                                        -1.391210e+03
##                                                adults
##                                        -2.973252e+04
##                                                mealFB
##                                         4.942971e+03
##                                                mealHB
##                                         6.166438e+03
##                                                mealSC
##                                         1.219592e+04
##                                         mealUndefined
##                                        -1.487736e+04
##                         market_segmentComplementary
##                                         6.228246e+10
##                            market_segmentCorporate
##                                         6.228242e+10
##                              market_segmentDirect
##                                         6.228244e+10
##                              market_segmentGroups
##                                         6.228244e+10
##                        market_segmentOffline_TA/TO
##                                         6.228244e+10
##                           market_segmentOnline_TA
##                                         6.228243e+10
##                         distribution_channelDirect
##                                        -2.115981e+04
##                            distribution_channelGDS
##                                         2.380113e+07
##                          distribution_channelTA/TO
##                                        -5.270205e+03
##                                    is_repeated_guest
##                                        -2.170360e+04
##                              previous_cancellations
##                                        -4.764640e+03
##                    previous_bookings_not_canceled
##                                         1.642889e+04
##                                  reserved_room_typeB
##                                         1.724107e+04
##                                  reserved_room_typeC
##                                        -6.103588e+04
##                                  reserved_room_typeD
##                                         2.002369e+08
##                                  reserved_room_typeE
##                                        -2.004031e+08
##                                  reserved_room_typeF
##                                         3.571274e+04
##                                  reserved_room_typeG
##                                         1.451776e+05
##                                  reserved_room_typeH
##                                         1.566875e+20
##                                  assigned_room_typeB
##                                        -3.747781e+02
##                                  assigned_room_typeC
##                                        -8.823320e+03
##                                  assigned_room_typeD
```

```
##                                          -2.196610e+03
##                          assigned_room_typeE
##                                          -1.781955e+02
##                          assigned_room_typeF
##                                          -1.682686e+04
##                          assigned_room_typeG
##                                           1.575799e+05
##                          assigned_room_typeH
##                                          -3.397461e+05
##                          assigned_room_typeI
##                                          -2.203846e+04
##                          assigned_room_typeK
##                                          -9.025049e+03
##                             booking_changes
##                                          -7.175800e+02
##                        days_in_waiting_list
##                                          -2.446837e+00
##                           customer_typeGroup
##                                           4.291725e+02
##                       customer_typeTransient
##                                           2.151316e+08
##                 customer_typeTransient-Party
##                                           1.115705e+03
##                           average_daily_rate
##                                           7.598556e+00
##          required_car_parking_spacesparking
##                                           3.365266e+03
##                      total_of_special_requests
##                                          -8.288166e+01
##                  mealFB:reserved_room_typeB
##                                                     NA
##                  mealHB:reserved_room_typeB
##                                           1.412856e+02
##                  mealSC:reserved_room_typeB
##                                          -2.680307e+03
##           mealUndefined:reserved_room_typeB
##                                                     NA
##                  mealFB:reserved_room_typeC
##                                          -4.623081e+03
##                  mealHB:reserved_room_typeC
##                                          -1.959330e+03
##                  mealSC:reserved_room_typeC
##                                          -3.316696e+04
##           mealUndefined:reserved_room_typeC
##                                          -6.544432e+03
##                  mealFB:reserved_room_typeD
##                                          -1.038748e+03
##                  mealHB:reserved_room_typeD
##                                           1.247994e+03
##                  mealSC:reserved_room_typeD
##                                          -2.935301e+03
##           mealUndefined:reserved_room_typeD
##                                          -5.036284e+03
##                  mealFB:reserved_room_typeE
```

```
##                                        -2.192141e+07
##                  mealHB:reserved_room_typeE
##                                         1.816411e+02
##                  mealSC:reserved_room_typeE
##                                         4.529659e+05
##           mealUndefined:reserved_room_typeE
##                                        -1.111833e+07
##                  mealFB:reserved_room_typeF
##                                        -1.453917e+07
##                  mealHB:reserved_room_typeF
##                                         2.011346e+03
##                  mealSC:reserved_room_typeF
##                                        -5.629505e+14
##           mealUndefined:reserved_room_typeF
##                                        -1.563671e+06
##                  mealFB:reserved_room_typeG
##                                         1.368626e+06
##                  mealHB:reserved_room_typeG
##                                         9.179114e+03
##                  mealSC:reserved_room_typeG
##                                                   NA
##           mealUndefined:reserved_room_typeG
##                                        -5.836883e+08
##                  mealFB:reserved_room_typeH
##                                                   NA
##                  mealHB:reserved_room_typeH
##                                        -6.922090e+04
##                  mealSC:reserved_room_typeH
##                                                   NA
##           mealUndefined:reserved_room_typeH
##                                                   NA
##       reserved_room_typeB:assigned_room_typeB
##                                        -1.519797e+03
##       reserved_room_typeC:assigned_room_typeB
##                                        -4.513753e+06
##       reserved_room_typeD:assigned_room_typeB
##                                        -3.388686e+06
##       reserved_room_typeE:assigned_room_typeB
##                                        -5.693129e+05
##       reserved_room_typeF:assigned_room_typeB
##                                        -3.201530e+04
##       reserved_room_typeG:assigned_room_typeB
##                                        -2.596510e+07
##       reserved_room_typeH:assigned_room_typeB
##                                                   NA
##       reserved_room_typeB:assigned_room_typeC
##                                                   NA
##       reserved_room_typeC:assigned_room_typeC
##                                         2.831779e+04
##       reserved_room_typeD:assigned_room_typeC
##                                        -3.343853e+03
##       reserved_room_typeE:assigned_room_typeC
##                                         5.943476e+06
##       reserved_room_typeF:assigned_room_typeC
```

```
##                                                                  NA
##                  reserved_room_typeG:assigned_room_typeC
##                                                                  NA
##                  reserved_room_typeH:assigned_room_typeC
##                                                                  NA
##                  reserved_room_typeB:assigned_room_typeD
##                                                        2.346405e+03
##                  reserved_room_typeC:assigned_room_typeD
##                                                       -2.841793e+07
##                  reserved_room_typeD:assigned_room_typeD
##                                                       -3.496949e+03
##                  reserved_room_typeE:assigned_room_typeD
##                                                       -2.833702e+04
##                  reserved_room_typeF:assigned_room_typeD
##                                                                  NA
##                  reserved_room_typeG:assigned_room_typeD
##                                                                  NA
##                  reserved_room_typeH:assigned_room_typeD
##                                                       -1.566875e+20
##                  reserved_room_typeB:assigned_room_typeE
##                                                        9.600398e+03
##                  reserved_room_typeC:assigned_room_typeE
##                                                        3.056477e+07
##                  reserved_room_typeD:assigned_room_typeE
##                                                       -1.605508e+04
##                  reserved_room_typeE:assigned_room_typeE
##                                                       -2.739045e+04
##                  reserved_room_typeF:assigned_room_typeE
##                                                       -1.957363e+04
##                  reserved_room_typeG:assigned_room_typeE
##                                                       -3.563052e+06
##                  reserved_room_typeH:assigned_room_typeE
##                                                                  NA
##                  reserved_room_typeB:assigned_room_typeF
##                                                                  NA
##                  reserved_room_typeC:assigned_room_typeF
##                                                                  NA
##                  reserved_room_typeD:assigned_room_typeF
##                                                        6.650135e+03
##                  reserved_room_typeE:assigned_room_typeF
##                                                       -1.667911e+04
##                  reserved_room_typeF:assigned_room_typeF
##                                                       -1.192349e+04
##                  reserved_room_typeG:assigned_room_typeF
##                                                       -8.023830e+04
##                  reserved_room_typeH:assigned_room_typeF
##                                                                  NA
##                  reserved_room_typeB:assigned_room_typeG
##                                                       -3.685367e+06
##                  reserved_room_typeC:assigned_room_typeG
##                                                       -1.910872e+06
##                  reserved_room_typeD:assigned_room_typeG
##                                                       -1.805556e+05
##                  reserved_room_typeE:assigned_room_typeG
```

```
##                                                   -2.045499e+05
##                 reserved_room_typeF:assigned_room_typeG
##                                                   -2.306201e+05
##                 reserved_room_typeG:assigned_room_typeG
##                                                   -2.936659e+05
##                 reserved_room_typeH:assigned_room_typeG
##                                                   -1.566875e+20
##                 reserved_room_typeB:assigned_room_typeH
##                                                             NA
##                 reserved_room_typeC:assigned_room_typeH
##                                                   -3.391627e+03
##                 reserved_room_typeD:assigned_room_typeH
##                                                    2.828127e+03
##                 reserved_room_typeE:assigned_room_typeH
##                                                   -1.010622e+04
##                 reserved_room_typeF:assigned_room_typeH
##                                                   -1.093550e+04
##                 reserved_room_typeG:assigned_room_typeH
##                                                    6.394773e+03
##                 reserved_room_typeH:assigned_room_typeH
##                                                   -1.566875e+20
##                 reserved_room_typeB:assigned_room_typeI
##                                                             NA
##                 reserved_room_typeC:assigned_room_typeI
##                                                    5.258488e+03
##                 reserved_room_typeD:assigned_room_typeI
##                                                    1.426127e+03
##                 reserved_room_typeE:assigned_room_typeI
##                                                   -2.833793e+03
##                 reserved_room_typeF:assigned_room_typeI
##                                                    2.564215e+03
##                 reserved_room_typeG:assigned_room_typeI
##                                                    9.857770e+03
##                 reserved_room_typeH:assigned_room_typeI
##                                                   -1.566875e+20
##                 reserved_room_typeB:assigned_room_typeK
##                                                    1.951905e+04
##                 reserved_room_typeC:assigned_room_typeK
##                                                             NA
##                 reserved_room_typeD:assigned_room_typeK
##                                                    1.870230e+04
##                 reserved_room_typeE:assigned_room_typeK
##                                                    1.733403e+04
##                 reserved_room_typeF:assigned_room_typeK
##                                                    1.074325e+04
##                 reserved_room_typeG:assigned_room_typeK
##                                                    2.816802e+04
##                 reserved_room_typeH:assigned_room_typeK
##                                                             NA
##                    hotelResort_Hotel:reserved_room_typeB
##                                                             NA
##                    hotelResort_Hotel:reserved_room_typeC
##                                                   -5.370823e+03
##                    hotelResort_Hotel:reserved_room_typeD
```

```
##                                                    -6.205441e+02
##              hotelResort_Hotel:reserved_room_typeE
##                                                    -4.894219e+02
##              hotelResort_Hotel:reserved_room_typeF
##                                                    -5.398084e+02
##              hotelResort_Hotel:reserved_room_typeG
##                                                    -6.666635e+02
##              hotelResort_Hotel:reserved_room_typeH
##                                                               NA
##      market_segmentComplementary:reserved_room_typeB
##                                                    -6.769705e+02
##         market_segmentCorporate:reserved_room_typeB
##                                                    -9.400512e+02
##            market_segmentDirect:reserved_room_typeB
##                                                    -1.296267e+02
##            market_segmentGroups:reserved_room_typeB
##                                                    -4.210158e+02
##      market_segmentOffline_TA/TO:reserved_room_typeB
##                                                     2.284690e+03
##         market_segmentOnline_TA:reserved_room_typeB
##                                                               NA
##      market_segmentComplementary:reserved_room_typeC
##                                                     1.168530e+03
##         market_segmentCorporate:reserved_room_typeC
##                                                     7.873938e+01
##            market_segmentDirect:reserved_room_typeC
##                                                    -7.335379e+02
##            market_segmentGroups:reserved_room_typeC
##                                                    -2.078793e+03
##      market_segmentOffline_TA/TO:reserved_room_typeC
##                                                     1.731833e+03
##         market_segmentOnline_TA:reserved_room_typeC
##                                                               NA
##      market_segmentComplementary:reserved_room_typeD
##                                                    -2.002220e+08
##         market_segmentCorporate:reserved_room_typeD
##                                                    -2.002224e+08
##            market_segmentDirect:reserved_room_typeD
##                                                    -2.002221e+08
##            market_segmentGroups:reserved_room_typeD
##                                                    -2.332484e+08
##      market_segmentOffline_TA/TO:reserved_room_typeD
##                                                    -2.002206e+08
##         market_segmentOnline_TA:reserved_room_typeD
##                                                    -2.002219e+08
##      market_segmentComplementary:reserved_room_typeE
##                                                     2.004219e+08
##         market_segmentCorporate:reserved_room_typeE
##                                                     2.004209e+08
##            market_segmentDirect:reserved_room_typeE
##                                                     2.004218e+08
##            market_segmentGroups:reserved_room_typeE
##                                                     2.008413e+08
##      market_segmentOffline_TA/TO:reserved_room_typeE
```

```
##                                                                2.004234e+08
##                     market_segmentOnline_TA:reserved_room_typeE
##                                                                2.004220e+08
##                 market_segmentComplementary:reserved_room_typeF
##                                                                3.814159e+02
##                     market_segmentCorporate:reserved_room_typeF
##                                                                3.832278e+03
##                        market_segmentDirect:reserved_room_typeF
##                                                               -6.636473e+02
##                        market_segmentGroups:reserved_room_typeF
##                                                               -1.489083e+03
##                  market_segmentOffline_TA/TO:reserved_room_typeF
##                                                                1.080828e+03
##                     market_segmentOnline_TA:reserved_room_typeF
##                                                                          NA
##                 market_segmentComplementary:reserved_room_typeG
##                                                               -1.314664e+02
##                     market_segmentCorporate:reserved_room_typeG
##                                                                5.853523e+02
##                        market_segmentDirect:reserved_room_typeG
##                                                               -4.356045e+02
##                        market_segmentGroups:reserved_room_typeG
##                                                               -1.131754e+05
##                  market_segmentOffline_TA/TO:reserved_room_typeG
##                                                                1.418517e+03
##                     market_segmentOnline_TA:reserved_room_typeG
##                                                                          NA
##                 market_segmentComplementary:reserved_room_typeH
##                                                                          NA
##                     market_segmentCorporate:reserved_room_typeH
##                                                                          NA
##                        market_segmentDirect:reserved_room_typeH
##                                                               -9.042098e+03
##                        market_segmentGroups:reserved_room_typeH
##                                                                          NA
##                  market_segmentOffline_TA/TO:reserved_room_typeH
##                                                                          NA
##                     market_segmentOnline_TA:reserved_room_typeH
##                                                                          NA
##                                           mealFB:is_repeated_guest
##                                                               -2.912301e+03
##                                           mealHB:is_repeated_guest
##                                                                4.008593e+02
##                                           mealSC:is_repeated_guest
##                                                                6.776472e+02
##                                    mealUndefined:is_repeated_guest
##                                                               -3.448842e+03
##                         adults:previous_bookings_not_canceled
##                                                                2.793033e+02
##                         mealFB:previous_bookings_not_canceled
##                                                                          NA
##                         mealHB:previous_bookings_not_canceled
##                                                                3.370834e+01
##                         mealSC:previous_bookings_not_canceled
```

```
##                                                   1.082813e+09
##            mealUndefined:previous_bookings_not_canceled
##                                                  -3.310646e+05
##        market_segmentComplementary:customer_typeGroup
##                                                  -9.850521e+06
##            market_segmentCorporate:customer_typeGroup
##                                                   7.522260e+03
##               market_segmentDirect:customer_typeGroup
##                                                   7.745992e+02
##               market_segmentGroups:customer_typeGroup
##                                                  -1.196862e+03
##        market_segmentOffline_TA/TO:customer_typeGroup
##                                                  -1.051335e+01
##            market_segmentOnline_TA:customer_typeGroup
##                                                              NA
##     market_segmentComplementary:customer_typeTransient
##                                                  -2.151307e+08
##         market_segmentCorporate:customer_typeTransient
##                                                  -2.151289e+08
##            market_segmentDirect:customer_typeTransient
##                                                  -2.151323e+08
##            market_segmentGroups:customer_typeTransient
##                                                  -2.151332e+08
##     market_segmentOffline_TA/TO:customer_typeTransient
##                                                  -2.151341e+08
##         market_segmentOnline_TA:customer_typeTransient
##                                                  -2.151327e+08
## market_segmentComplementary:customer_typeTransient-Party
##                                                              NA
##     market_segmentCorporate:customer_typeTransient-Party
##                                                   3.586564e+03
##       market_segmentDirect:customer_typeTransient-Party
##                                                  -1.333592e+03
##       market_segmentGroups:customer_typeTransient-Party
##                                                  -7.768697e+02
## market_segmentOffline_TA/TO:customer_typeTransient-Party
##                                                  -2.482825e+03
##     market_segmentOnline_TA:customer_typeTransient-Party
##                                                              NA
##                   is_repeated_guest:assigned_room_typeB
##                                                   1.332976e+03
##                   is_repeated_guest:assigned_room_typeC
##                                                   1.937405e+02
##                   is_repeated_guest:assigned_room_typeD
##                                                   1.772995e+02
##                   is_repeated_guest:assigned_room_typeE
##                                                  -3.696172e+03
##                   is_repeated_guest:assigned_room_typeF
##                                                  -1.130500e+03
##                   is_repeated_guest:assigned_room_typeG
##                                                  -2.340076e+03
##                   is_repeated_guest:assigned_room_typeH
##                                                   3.225794e+04
##                   is_repeated_guest:assigned_room_typeI
```

```
##                                                  -2.201986e+03
##                     is_repeated_guest:assigned_room_typeK
##                                                   1.460226e+04
##     assigned_room_typeB:required_car_parking_spacesparking
##                                                   6.578777e+02
##     assigned_room_typeC:required_car_parking_spacesparking
##                                                  -8.675481e+01
##     assigned_room_typeD:required_car_parking_spacesparking
##                                                   1.442037e+02
##     assigned_room_typeE:required_car_parking_spacesparking
##                                                   1.579681e+02
##     assigned_room_typeF:required_car_parking_spacesparking
##                                                   2.889318e+01
##     assigned_room_typeG:required_car_parking_spacesparking
##                                                   5.117089e+02
##     assigned_room_typeH:required_car_parking_spacesparking
##                                                   3.238187e+03
##     assigned_room_typeI:required_car_parking_spacesparking
##                                                   1.114010e+03
##     assigned_room_typeK:required_car_parking_spacesparking
##                                                   1.090075e+07


##                       [,1]       [,2]      [,3]      [,4]
## measurement           "Deviance" "TPR"     "FPR"     "FDR"
## eval_baseline1        "4979.002" "0"       "0"       "NaN"
## eval_baseline2        "3506.812" "0.342"   "0.013"   "0.312"
## eval_lasso_selected   "3453.253" "0.367"   "0.013"   "0.285"
```
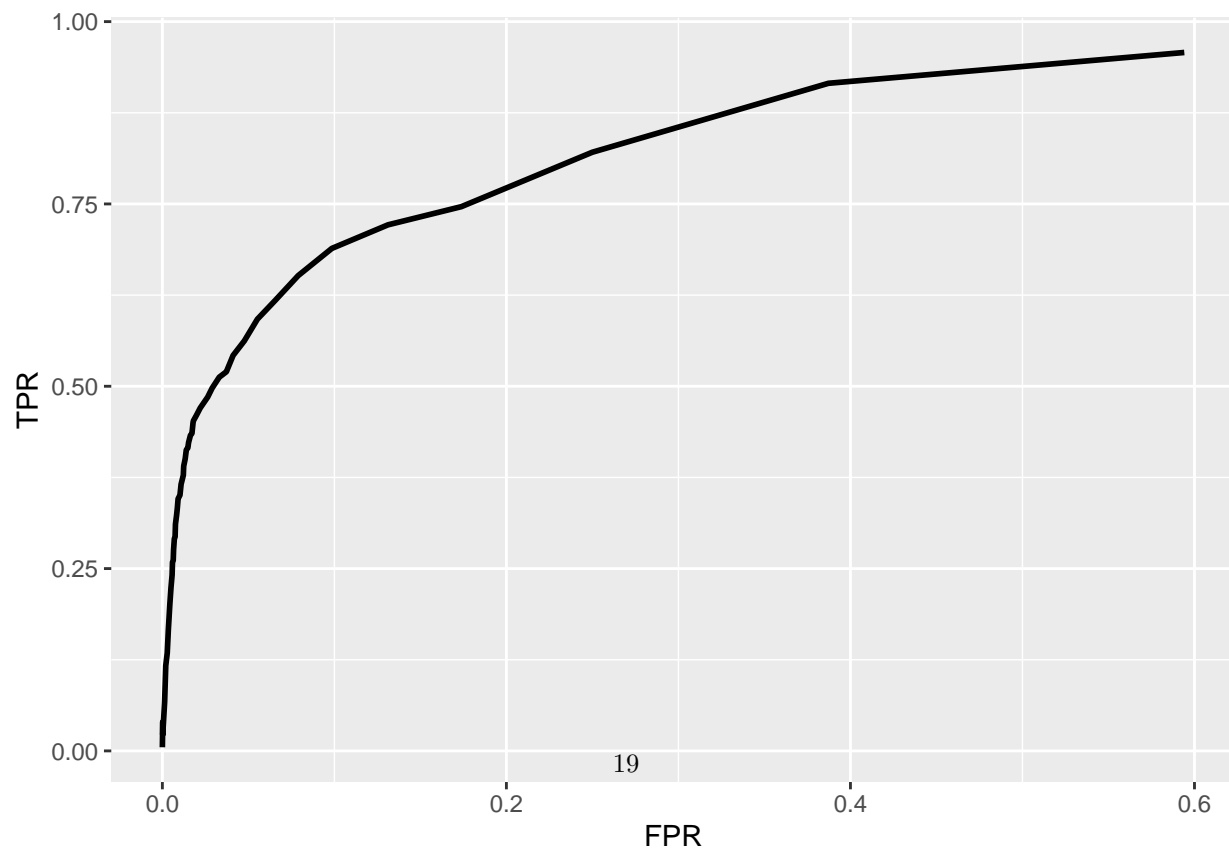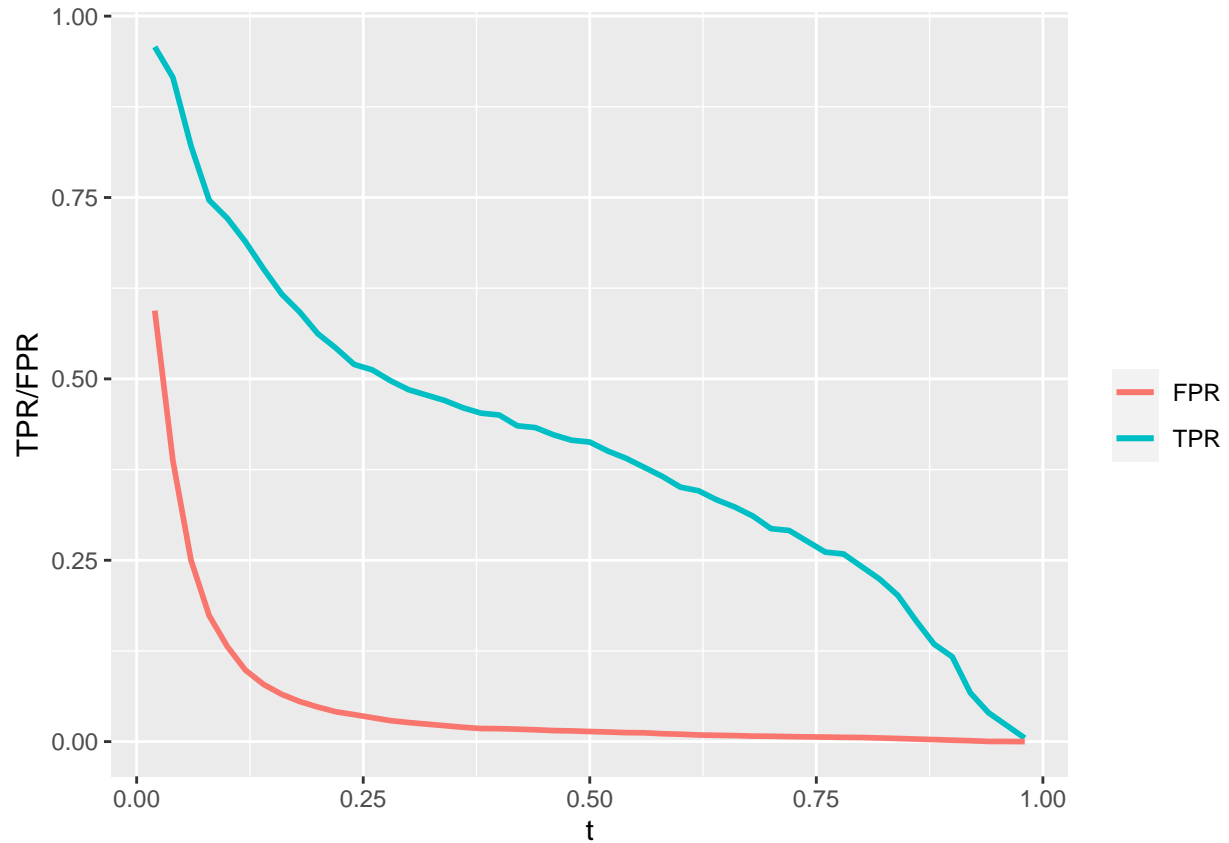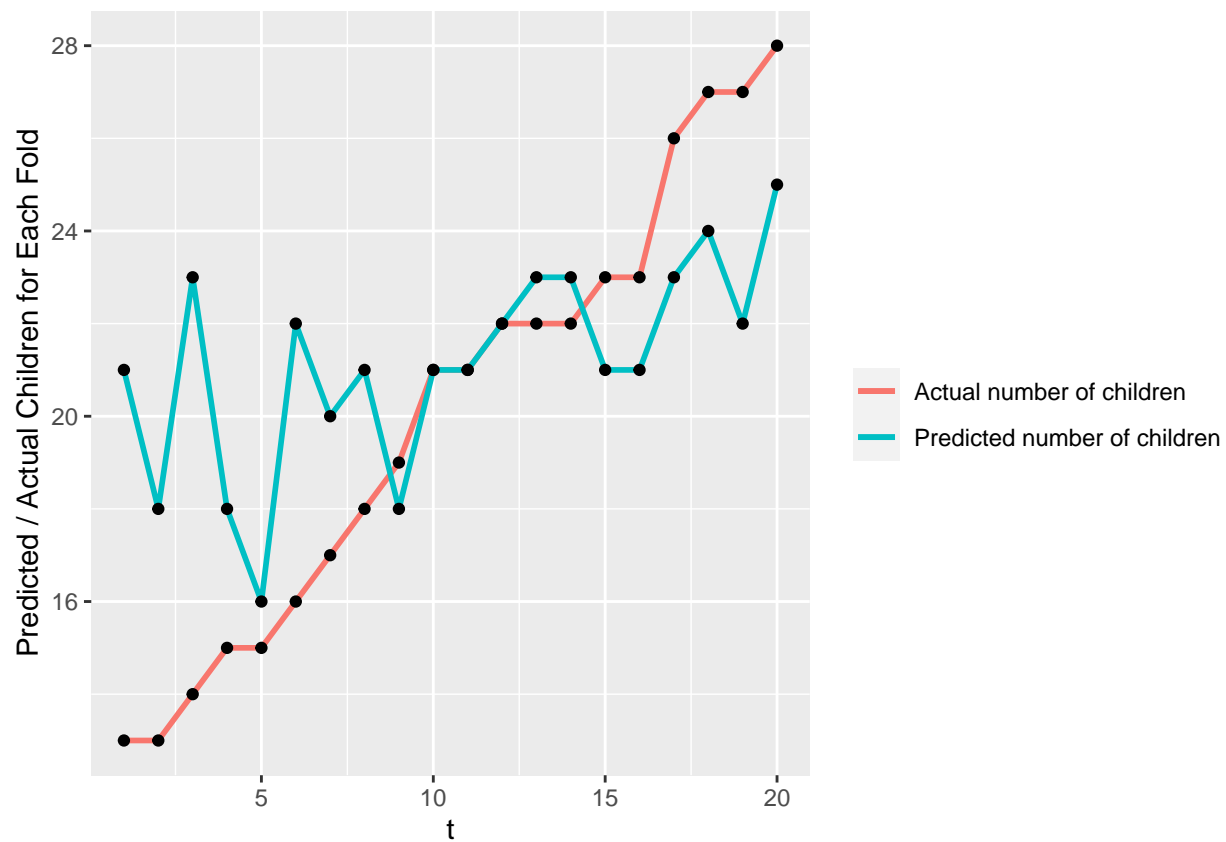
We can see that the lasso model has the smallest deviance, the highest TPR(higher is better)

and the lowest FPR & FDR (Both 2 are "lower is better")

19

The ROC curve of our best model

We can see that our model doesn't perform well ## The prediction isn't accurate. The predict numbers wiggle more than actual number.