

biomarkers

Parinitha Kompala

9/1/2021

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
library(healthcareai)
```

```
## healthcareai version 2.5.0
```

```
## Please visit https://docs.healthcare.ai for full documentation and vignettes. Join the community at 1
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.1.4      v dplyr  1.0.7
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   2.0.1      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(readxl)
```

getting the data

```
setwd("/Users/parinithakompala/Desktop")
```

```
data<-read_excel("/Users/parinithakompala/Library/Containers/com.microsoft.Excel/Data/Downloads/Adult_B")
getwd()
```

```
## [1] "/Users/parinithakompala/Desktop"
```

```
head(data)
```

```
## # A tibble: 6 x 260
```

```
##   readmit30d_yn_state  dead obleed  ocva oleginf  otia   orf opneu oafib2 return
##           <dbl> <dbl>   <dbl> <dbl>   <dbl> <dbl> <dbl> <dbl>   <dbl>   <dbl>
## 1             1     0     0     0     0     0     0     0     0     1     0
## 2             1     0     0     0     0     0     0     0     0     1     0
## 3             0     0     0     0     0     0     0     0     0     0     0
## 4             0     0     0     0     0     0     0     0     0     0     0
## 5             0     0     0     0     0     0     0     0     0     1     0
## 6             0     0     0     0     0     0     0     0     0     0     0
```

```
## # ... with 250 more variables: akin <dbl>, lm2cat <dbl>, rfcr_2 <dbl>,
```

```
## # iabppre_2 <dbl>, atfibyn_2 <dbl>, hyperyn_2 <dbl>, ef50_neg <dbl>,
## # ef50_neg_d <dbl>, priormi_21 <dbl>, priormi_22 <dbl>, priormi_23 <dbl>,
## # priormi_24 <dbl>, age_d <dbl>, ua_nmi7 <dbl>, creatpre_2 <dbl>, rf_1 <dbl>,
## # rf_2 <dbl>, rf_3 <dbl>, priority_21 <dbl>, priority_22 <dbl>, dm3cat <dbl>,
## # dm3cat_21 <dbl>, dm3cat_22 <dbl>, dm3cat_21_2 <dbl>, dm3cat_22_2 <dbl>,
## # pci_ta <dbl>, efvalue_r <dbl>, ef50_neg_r <dbl>, ef50_neg_d_r <dbl>, ...
```

```
view(data)
```

analysing the data

```
data
```

```
## # A tibble: 1,651 x 260
```

```
##   readmit30d_yn_state  dead obleed  ocva oleginf  otia  orf opneu oafib2 return
##           <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1             1      0      0      0      0      0      0      0      1      0
## 2             1      0      0      0      0      0      0      0      1      0
## 3             0      0      0      0      0      0      0      0      0      0
## 4             0      0      0      0      0      0      0      0      0      0
## 5             0      0      0      0      0      0      0      0      1      0
## 6             0      0      0      0      0      0      0      0      0      0
## 7             1      0      0      0      0      0      0      0      1      0
## 8             0      0      0      0      0      0      0      0      0      0
## 9             0      0      0      1      0      0      0      0      1      0
## 10            0      0      0      0      0      0      0      0      1      0
```

```
## # ... with 1,641 more rows, and 250 more variables: akin <dbl>, lm2cat <dbl>,
## # rfcr_2 <dbl>, iabppre_2 <dbl>, atfibyn_2 <dbl>, hyperyn_2 <dbl>,
## # ef50_neg <dbl>, ef50_neg_d <dbl>, priormi_21 <dbl>, priormi_22 <dbl>,
## # priormi_23 <dbl>, priormi_24 <dbl>, age_d <dbl>, ua_nmi7 <dbl>,
## # creatpre_2 <dbl>, rf_1 <dbl>, rf_2 <dbl>, rf_3 <dbl>, priority_21 <dbl>,
## # priority_22 <dbl>, dm3cat <dbl>, dm3cat_21 <dbl>, dm3cat_22 <dbl>,
## # dm3cat_21_2 <dbl>, dm3cat_22_2 <dbl>, pci_ta <dbl>, efvalue_r <dbl>, ...
```

```
dim(data)
```

```
## [1] 1651 260
```

```
summary(data)
```

```
##   readmit30d_yn_state      dead      obleed      ocva
##   Min.   :0.00000      Min.   :0.0000      Min.   :0.00000      Min.   :0.00000
##   1st Qu.:0.00000      1st Qu.:0.0000      1st Qu.:0.00000      1st Qu.:0.00000
##   Median :0.00000      Median :0.0000      Median :0.00000      Median :0.00000
##   Mean    :0.09873      Mean    :0.1381      Mean    :0.01514      Mean    :0.01272
##   3rd Qu.:0.00000      3rd Qu.:0.0000      3rd Qu.:0.00000      3rd Qu.:0.00000
##   Max.    :1.00000      Max.    :1.0000      Max.    :1.00000      Max.    :1.00000
##
##   oleginf      otia      orf      opneu
##   Min.   :0.000000      Min.   :0.000000      Min.   :0.000000      Min.   :0.00000
##   1st Qu.:0.000000      1st Qu.:0.000000      1st Qu.:0.000000      1st Qu.:0.00000
##   Median :0.000000      Median :0.000000      Median :0.000000      Median :0.00000
##   Mean    :0.009085      Mean    :0.003028      Mean    :0.004845      Mean    :0.01757
##   3rd Qu.:0.000000      3rd Qu.:0.000000      3rd Qu.:0.000000      3rd Qu.:0.00000
##   Max.    :1.000000      Max.    :1.000000      Max.    :1.000000      Max.    :1.00000
##
##   oafib2      return      akin      lm2cat
```

```

## Min. :0.0000 Min. :0.00000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :0.0000 Median :0.00000 Median :0.0000 Median :0.0000
## Mean :0.2114 Mean :0.02847 Mean :0.3895 Mean :0.3368
## 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:1.0000
## Max. :1.0000 Max. :1.00000 Max. :3.0000 Max. :1.0000
##
## rfcr_2 iabppre_2 atfibyn_2 hyperyn_2
## Min. :0.00000 Min. :0.00000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:1.0000
## Median :0.00000 Median :0.00000 Median :0.0000 Median :1.0000
## Mean :0.02938 Mean :0.04422 Mean :0.0697 Mean :0.8065
## 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:1.0000
## Max. :1.00000 Max. :1.00000 Max. :1.0000 Max. :1.0000
## NA's :17 NA's :1 NA's :2
## ef50_neg ef50_neg_d priormi_21 priormi_22
## Min. : 0.000 Min. :1.00 Min. :0.0000 Min. :0.0000
## 1st Qu.: 0.000 1st Qu.:1.00 1st Qu.:0.0000 1st Qu.:0.0000
## Median : 0.000 Median :1.00 Median :0.0000 Median :0.0000
## Mean : 3.256 Mean :1.42 Mean :0.4403 Mean :0.4228
## 3rd Qu.: 3.000 3rd Qu.:2.00 3rd Qu.:1.0000 3rd Qu.:1.0000
## Max. :37.000 Max. :5.00 Max. :1.0000 Max. :1.0000
## NA's :75 NA's :75
## priormi_23 priormi_24 age_d ua_nmi7
## Min. :0.0000 Min. :0.0000 Min. : 3.000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.: 6.000 1st Qu.:0.0000
## Median :0.0000 Median :0.0000 Median : 7.000 Median :0.0000
## Mean :0.2374 Mean :0.1393 Mean : 7.003 Mean :0.4239
## 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.: 8.000 3rd Qu.:1.0000
## Max. :1.0000 Max. :1.0000 Max. :10.000 Max. :1.0000
## NA's :16
## creatpre_2 rf_1 rf_2 rf_3
## Min. : 0.400 Min. :0.000000 Min. :0.0000 Min. :0.00000
## 1st Qu.: 0.900 1st Qu.:0.000000 1st Qu.:0.0000 1st Qu.:0.00000
## Median : 1.000 Median :0.000000 Median :0.0000 Median :0.00000
## Mean : 1.138 Mean :0.009085 Mean :0.1581 Mean :0.03867
## 3rd Qu.: 1.200 3rd Qu.:0.000000 3rd Qu.:0.2000 3rd Qu.:0.00000
## Max. :11.000 Max. :1.000000 Max. :4.0000 Max. :3.50000
## NA's :17 NA's :17 NA's :17
## priority_21 priority_22 dm3cat dm3cat_21
## Min. :0.0000 Min. :0.00000 Min. :1.000 Min. :0.000
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:1.000 1st Qu.:0.000
## Median :1.0000 Median :0.00000 Median :1.000 Median :0.000
## Mean :0.6923 Mean :0.01878 Mean :1.482 Mean :0.379
## 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:2.000 3rd Qu.:1.000
## Max. :1.0000 Max. :1.00000 Max. :3.000 Max. :1.000
## NA's :2 NA's :2
## dm3cat_22 dm3cat_21_2 dm3cat_22_2 pci_ta
## Min. :0.0000 Min. :0.000 Min. :0.0000 Min. :0.00000
## 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.00000
## Median :0.0000 Median :0.000 Median :0.0000 Median :0.00000
## Mean :0.1025 Mean :0.379 Mean :0.1025 Mean :0.03634
## 3rd Qu.:0.0000 3rd Qu.:1.000 3rd Qu.:0.0000 3rd Qu.:0.00000
## Max. :1.0000 Max. :1.000 Max. :1.0000 Max. :1.00000

```

```

## NA's :2      NA's :2      NA's :2
## efvalue_r      ef50_neg_r      ef50_neg_d_r      atfibyn_2_r
## Min. :13.00    Min. : 0.000    Min. :1.000    Min. :0.00000
## 1st Qu.:48.00    1st Qu.: 0.000    1st Qu.:1.000    1st Qu.:0.00000
## Median :57.00    Median : 0.000    Median :1.000    Median :0.00000
## Mean :53.96     Mean : 3.177     Mean :1.411     Mean :0.06965
## 3rd Qu.:60.00    3rd Qu.: 2.000    3rd Qu.:2.000    3rd Qu.:0.00000
## Max. :83.00     Max. :37.000     Max. :5.000     Max. :1.00000
##
## priormi_r      priormi_21_r      priormi_22_r      priormi_23_r
## Min. :0.00     Min. :0.0000     Min. :0.0000     Min. :0.0000
## 1st Qu.:0.00    1st Qu.:0.0000    1st Qu.:0.0000    1st Qu.:0.0000
## Median :0.00    Median :0.0000    Median :0.0000    Median :0.0000
## Mean :1.24     Mean :0.4403     Mean :0.4228     Mean :0.2374
## 3rd Qu.:2.00    3rd Qu.:1.0000    3rd Qu.:1.0000    3rd Qu.:0.0000
## Max. :4.00     Max. :1.0000     Max. :1.0000     Max. :1.0000
##
## priormi_24_r      age_r      age_d_r      ua_r
## Min. :0.0000     Min. :25.88     Min. : 3.000     Min. :0.0000
## 1st Qu.:0.0000    1st Qu.:58.14    1st Qu.: 6.000    1st Qu.:0.0000
## Median :0.0000    Median :65.14    Median : 7.000    Median :1.0000
## Mean :0.1393     Mean :65.19     Mean : 7.003     Mean :0.5324
## 3rd Qu.:0.0000    3rd Qu.:73.15    3rd Qu.: 8.000    3rd Qu.:1.0000
## Max. :1.0000     Max. :92.92     Max. :10.000     Max. :1.0000
##
## ua_nmi7_r      chf_r      chf2cat_r      rf_r
## Min. :0.0000     Min. :0.0000     Min. :0.0000     Min. :0.000000
## 1st Qu.:0.0000    1st Qu.:0.0000    1st Qu.:0.0000    1st Qu.:0.000000
## Median :0.0000    Median :0.0000    Median :0.0000    Median :0.000000
## Mean :0.4197     Mean :0.2126     Mean :0.1078     Mean :0.009085
## 3rd Qu.:1.0000    3rd Qu.:0.0000    3rd Qu.:0.0000    3rd Qu.:0.000000
## Max. :1.0000     Max. :4.0000     Max. :1.0000     Max. :1.000000
##
## creatpre_2_r      rf_1_r      rf_2_r      rf_3_r
## Min. : 0.400     Min. :0.000000     Min. :0.0000     Min. :0.00000
## 1st Qu.: 0.900    1st Qu.:0.000000    1st Qu.:0.0000    1st Qu.:0.00000
## Median : 1.000    Median :0.000000    Median :0.0000    Median :0.00000
## Mean : 1.138     Mean :0.009085     Mean :0.1564     Mean :0.03827
## 3rd Qu.: 1.200    3rd Qu.:0.000000    3rd Qu.:0.2000    3rd Qu.:0.00000
## Max. :11.000     Max. :1.000000     Max. :4.0000     Max. :3.50000
## NA's :17
## priority_r      priority_21_r      priority_22_r      sex_r
## Min. :1.000     Min. :0.0000     Min. :0.00000     Min. :0.0000
## 1st Qu.:2.000    1st Qu.:0.0000    1st Qu.:0.00000    1st Qu.:0.0000
## Median :2.000    Median :1.0000    Median :0.00000    Median :0.0000
## Mean :2.289     Mean :0.6923     Mean :0.01878     Mean :0.2265
## 3rd Qu.:3.000    3rd Qu.:1.0000    3rd Qu.:0.00000    3rd Qu.:0.0000
## Max. :3.000     Max. :1.0000     Max. :1.00000     Max. :1.0000
##
## prcabg_r      copd_r      anydm_r      dmtx_r
## Min. :0.00000     Min. :0.0000     Min. :0.0000     Min. :0.000
## 1st Qu.:0.00000    1st Qu.:0.0000    1st Qu.:0.0000    1st Qu.:1.000
## Median :0.00000    Median :0.0000    Median :0.0000    Median :2.000
## Mean :0.02181     Mean :0.1302     Mean :0.3798     Mean :1.624

```

```

## 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:2.000
## Max. :1.00000 Max. :1.0000 Max. :1.0000 Max. :3.000
## NA's :864
## dm3cat_r dm3cat_21_r dm3cat_22_r iabppre_r
## Min. :1.000 Min. :0.0000 Min. :0.0000 Min. :0.00000
## 1st Qu.:1.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000
## Median :1.000 Median :0.0000 Median :0.0000 Median :0.00000
## Mean :1.482 Mean :0.3798 Mean :0.1024 Mean :0.04422
## 3rd Qu.:2.000 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:0.00000
## Max. :3.000 Max. :1.0000 Max. :1.0000 Max. :1.00000
##
## anyvad_r vad_r hyper_r hyperyn_2_r
## Min. :0.0000 Min. :0.0000 Min. :0.000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:1.000 1st Qu.:1.0000
## Median :0.0000 Median :0.0000 Median :1.000 Median :1.0000
## Mean :0.2701 Mean :0.5379 Mean :1.327 Mean :0.8068
## 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:2.000 3rd Qu.:1.0000
## Max. :1.0000 Max. :5.0000 Max. :3.000 Max. :1.0000
##
## prptca6_r pci_ta_r lm3cat_r lm2cat_r
## Min. :0.0000 Min. :0.00000 Min. :1.000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:1.000 1st Qu.:0.0000
## Median :0.0000 Median :0.00000 Median :1.000 Median :0.0000
## Mean :0.3828 Mean :0.03634 Mean :1.379 Mean :0.3368
## 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:2.000 3rd Qu.:1.0000
## Max. :4.0000 Max. :1.00000 Max. :3.000 Max. :1.0000
##
## aortic_insuff aortic_insuff_r aortic_sten aortic_sten_r
## Min. :0.0000 Min. :0.000000 Min. :0.0000 Min. :0.00000
## 1st Qu.:0.0000 1st Qu.:0.000000 1st Qu.:0.0000 1st Qu.:0.00000
## Median :0.0000 Median :0.000000 Median :0.0000 Median :0.00000
## Mean :0.0178 Mean :0.006663 Mean :0.0418 Mean :0.01575
## 3rd Qu.:0.0000 3rd Qu.:0.000000 3rd Qu.:0.0000 3rd Qu.:0.00000
## Max. :1.0000 Max. :1.000000 Max. :1.0000 Max. :1.00000
## NA's :1033 NA's :1029
## chf_nyha_iv chf_nyha_ltiv chf_nyha_iv_r chf_nyha_ltiv_r
## Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000
## 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000
## Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000
## Mean :0.01735 Mean :0.03533 Mean :0.01696 Mean :0.03392
## 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000
## Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000
## NA's :37 NA's :66
## smoker_r cvd cvd_r htc_m_d
## Min. :0.000 Min. :0.0000 Min. :0.0000 Min. :12.00
## 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:16.00
## Median :0.000 Median :0.0000 Median :0.0000 Median :17.00
## Mean :0.232 Mean :0.1775 Mean :0.1775 Mean :16.76
## 3rd Qu.:0.000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:17.00
## Max. :1.000 Max. :1.0000 Max. :1.0000 Max. :20.00
## NA's :7
## htc_m_r htc_m_d_r mitral_insuff mitral_insuff_r
## Min. :122.1 Min. :12.00 Min. :0.0000 Min. :0.00000
## 1st Qu.:165.1 1st Qu.:16.00 1st Qu.:0.0000 1st Qu.:0.00000

```

```

## Median :173.0 Median :17.00 Median :0.0000 Median :0.00000
## Mean :171.8 Mean :16.76 Mean :0.0787 Mean :0.02968
## 3rd Qu.:178.0 3rd Qu.:17.00 3rd Qu.:0.0000 3rd Qu.:0.00000
## Max. :203.0 Max. :20.00 Max. :1.0000 Max. :1.00000
## NA's :1028
## carotid_sten carotid_sten_r pvd pvd_r
## Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :0.0000 Median :0.0000 Median :0.0000 Median :0.0000
## Mean :0.1239 Mean :0.0854 Mean :0.1411 Mean :0.1411
## 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000
## Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000
## NA's :513
## tricuspid_insuff tricuspid_insuff_r bmi_squared bmi_r
## Min. :0.0000 Min. :0.000000 Min. : 184.2 Min. :13.57
## 1st Qu.:0.0000 1st Qu.:0.000000 1st Qu.: 674.4 1st Qu.:25.97
## Median :0.0000 Median :0.000000 Median : 843.6 Median :29.04
## Mean :0.0066 Mean :0.002423 Mean : 914.1 Mean :29.72
## 3rd Qu.:0.0000 3rd Qu.:0.000000 3rd Qu.:1065.1 3rd Qu.:32.62
## Max. :1.0000 Max. :1.000000 Max. :3339.2 Max. :57.79
## NA's :1041 NA's :7
## bmi_squared_r novsl_r readmit_1y_yn_state anyakin
## Min. : 184.2 Min. :1.000 Min. :0.0000 Min. :0.0000
## 1st Qu.: 674.4 1st Qu.:2.000 1st Qu.:0.0000 1st Qu.:0.0000
## Median : 843.5 Median :3.000 Median :0.0000 Median :0.0000
## Mean : 913.8 Mean :2.434 Mean :0.2429 Mean :0.3337
## 3rd Qu.:1064.4 3rd Qu.:3.000 3rd Qu.:0.0000 3rd Qu.:1.0000
## Max. :3339.2 Max. :3.000 Max. :1.0000 Max. :1.0000
## creatcat lm50 anymssd lowoutput
## Min. :0.0000 Min. :0.0000 Min. :0.00000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.0000
## Median :0.0000 Median :0.0000 Median :0.00000 Median :0.0000
## Mean :0.2235 Mean :0.3368 Mean :0.00848 Mean :0.0321
## 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:0.0000
## Max. :3.0000 Max. :1.0000 Max. :1.00000 Max. :1.0000
## emerg urg elec bmicat
## Min. :0.00000 Min. :0.0000 Min. :0.0000 Min. :1.000
## 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:3.000
## Median :0.00000 Median :1.0000 Median :0.0000 Median :3.000
## Mean :0.01878 Mean :0.6735 Mean :0.3077 Mean :3.431
## 3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:4.000
## Max. :1.00000 Max. :1.0000 Max. :1.0000 Max. :6.000
## bmi1 bmi2 bmi3 bmi4
## Min. :0.000000 Min. :0.0000 Min. :0.0000 Min. :0.000
## 1st Qu.:0.000000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.000
## Median :0.000000 Median :0.0000 Median :0.0000 Median :0.000
## Mean :0.006663 Mean :0.1708 Mean :0.4034 Mean :0.272
## 3rd Qu.:0.000000 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:1.000
## Max. :1.000000 Max. :1.0000 Max. :1.0000 Max. :1.000
## bmi5 bmi6 lvedpm anemiapre

```

| | | | | | | | | |
|----|----------------|----------|------------|----------|---------------|---------|---------------|-----------|
| ## | Min. | :0.00000 | Min. | :0.00000 | Min. | :0.0000 | Min. | :0.0000 |
| ## | 1st Qu. | :0.00000 | 1st Qu. | :0.00000 | 1st Qu. | :0.0000 | 1st Qu. | :0.0000 |
| ## | Median | :0.00000 | Median | :0.00000 | Median | :0.0000 | Median | :0.0000 |
| ## | Mean | :0.09812 | Mean | :0.04906 | Mean | :0.3337 | Mean | :0.3882 |
| ## | 3rd Qu. | :0.00000 | 3rd Qu. | :0.00000 | 3rd Qu. | :1.0000 | 3rd Qu. | :1.0000 |
| ## | Max. | :1.00000 | Max. | :1.00000 | Max. | :1.0000 | Max. | :1.0000 |
| ## | | | | | | | | |
| ## | iabpintra | | lof1 | | ptimenumban | | ctimenumban | |
| ## | Min. | :0.00000 | Min. | :0.00000 | Min. | : 0.00 | Min. | : 0.00 |
| ## | 1st Qu. | :0.00000 | 1st Qu. | :0.00000 | 1st Qu. | : 23.20 | 1st Qu. | : 14.33 |
| ## | Median | :0.00000 | Median | :0.00000 | Median | : 27.60 | Median | : 17.50 |
| ## | Mean | :0.01817 | Mean | :0.03513 | Mean | : 30.21 | Mean | : 19.34 |
| ## | 3rd Qu. | :0.00000 | 3rd Qu. | :0.00000 | 3rd Qu. | : 33.50 | 3rd Qu. | : 22.00 |
| ## | Max. | :1.00000 | Max. | :1.00000 | Max. | :274.00 | Max. | :230.00 |
| ## | | | | | NA's | :18 | NA's | :18 |
| ## | cardtimenumban | | cardblood | | cardcold | | hotshot | |
| ## | Min. | : 0.000 | Min. | :0.0000 | Min. | :0.000 | Min. | :0.000 |
| ## | 1st Qu. | : 3.333 | 1st Qu. | :1.0000 | 1st Qu. | :0.000 | 1st Qu. | :0.000 |
| ## | Median | : 5.000 | Median | :1.0000 | Median | :0.000 | Median | :1.000 |
| ## | Mean | : 8.418 | Mean | :0.8904 | Mean | :0.341 | Mean | :0.699 |
| ## | 3rd Qu. | : 7.000 | 3rd Qu. | :1.0000 | 3rd Qu. | :1.000 | 3rd Qu. | :1.000 |
| ## | Max. | :499.500 | Max. | :1.0000 | Max. | :1.000 | Max. | :1.000 |
| ## | NA's | :18 | | | | | | |
| ## | aoxcon | | ultrafilyn | | cabg | | valve | |
| ## | Min. | :0.0000 | Min. | :0.0000 | Min. | :0.0000 | Min. | :0.000000 |
| ## | 1st Qu. | :0.0000 | 1st Qu. | :0.0000 | 1st Qu. | :1.0000 | 1st Qu. | :0.000000 |
| ## | Median | :0.0000 | Median | :0.0000 | Median | :1.0000 | Median | :0.000000 |
| ## | Mean | :0.3028 | Mean | :0.1114 | Mean | :0.9522 | Mean | :0.007874 |
| ## | 3rd Qu. | :1.0000 | 3rd Qu. | :0.0000 | 3rd Qu. | :1.0000 | 3rd Qu. | :0.000000 |
| ## | Max. | :1.0000 | Max. | :1.0000 | Max. | :1.0000 | Max. | :1.000000 |
| ## | | | | | | | | |
| ## | cabgvalve | | gfr60pre | | male | | notcoldcard | |
| ## | Min. | :0.00000 | Min. | :0.0000 | Min. | :0.0000 | Min. | :0.000 |
| ## | 1st Qu. | :0.00000 | 1st Qu. | :0.0000 | 1st Qu. | :1.0000 | 1st Qu. | :0.000 |
| ## | Median | :0.00000 | Median | :0.0000 | Median | :1.0000 | Median | :1.000 |
| ## | Mean | :0.03998 | Mean | :0.2459 | Mean | :0.7735 | Mean | :0.659 |
| ## | 3rd Qu. | :0.00000 | 3rd Qu. | :0.0000 | 3rd Qu. | :1.0000 | 3rd Qu. | :1.000 |
| ## | Max. | :1.00000 | Max. | :1.0000 | Max. | :1.0000 | Max. | :1.000 |
| ## | | | | | | | | |
| ## | fluidprel | | ptime120 | | heptotl | | heptot5 | |
| ## | Min. | :0.000 | Min. | :0.0000 | Min. | : 1.10 | Min. | :0.0000 |
| ## | 1st Qu. | :0.600 | 1st Qu. | :0.0000 | 1st Qu. | : 35.00 | 1st Qu. | :0.0000 |
| ## | Median | :1.000 | Median | :0.0000 | Median | : 43.20 | Median | :0.0000 |
| ## | Mean | :1.036 | Mean | :0.2641 | Mean | : 42.62 | Mean | :0.2477 |
| ## | 3rd Qu. | :1.500 | 3rd Qu. | :1.0000 | 3rd Qu. | : 49.80 | 3rd Qu. | :0.0000 |
| ## | Max. | :5.000 | Max. | :1.0000 | Max. | :110.00 | Max. | :1.0000 |
| ## | | | | | | | | |
| ## | tcys0 | | cyspre3cat | | I_cyspre3cat1 | | I_cyspre3cat2 | |
| ## | Min. | :1.000 | Min. | :1.000 | Min. | :0.0000 | Min. | :0.0000 |
| ## | 1st Qu. | :1.000 | 1st Qu. | :1.000 | 1st Qu. | :0.0000 | 1st Qu. | :0.0000 |
| ## | Median | :2.000 | Median | :2.000 | Median | :0.0000 | Median | :0.0000 |
| ## | Mean | :1.999 | Mean | :1.988 | Mean | :0.3377 | Mean | :0.3364 |
| ## | 3rd Qu. | :3.000 | 3rd Qu. | :3.000 | 3rd Qu. | :1.0000 | 3rd Qu. | :1.0000 |
| ## | Max. | :3.000 | Max. | :3.000 | Max. | :1.0000 | Max. | :1.0000 |

| | | | |
|---------------------|-------------------|------------------|------------------|
| ## NA's :123 | NA's :123 | NA's :123 | NA's :123 |
| ## I_cyspre3cat3 | logcys0 | tcys1 | cyspost3cat |
| ## Min. :0.0000 | Min. :-1.47810 | Min. :1 | Min. :1.000 |
| ## 1st Qu.:0.0000 | 1st Qu.: -0.38193 | 1st Qu.:1 | 1st Qu.:1.000 |
| ## Median :0.0000 | Median :-0.22575 | Median :2 | Median :2.000 |
| ## Mean :0.3259 | Mean :-0.17270 | Mean :2 | Mean :1.984 |
| ## 3rd Qu.:1.0000 | 3rd Qu.: -0.03344 | 3rd Qu.:3 | 3rd Qu.:3.000 |
| ## Max. :1.0000 | Max. : 1.98594 | Max. :3 | Max. :3.000 |
| ## NA's :123 | NA's :123 | NA's :394 | NA's :394 |
| ## I_cyspost3cat1 | I_cyspost3cat2 | I_cyspost3cat3 | logcys1 |
| ## Min. :0.0000 | Min. :0.0000 | Min. :0.000 | Min. :-0.9364 |
| ## 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.:0.000 | 1st Qu.: -0.3401 |
| ## Median :0.0000 | Median :0.0000 | Median :0.000 | Median :-0.1460 |
| ## Mean :0.3389 | Mean :0.3381 | Mean :0.323 | Mean :-0.0713 |
| ## 3rd Qu.:1.0000 | 3rd Qu.:1.0000 | 3rd Qu.:1.000 | 3rd Qu.: 0.1318 |
| ## Max. :1.0000 | Max. :1.0000 | Max. :1.000 | Max. : 1.7710 |
| ## NA's :394 | NA's :394 | NA's :394 | NA's :394 |
| ## cysdiff | tcysdiff | cysdiff3cat | I_cysdiff3cat1 |
| ## Min. :-2.0040 | Min. :1 | Min. :1.00 | Min. :0.0000 |
| ## 1st Qu.: -0.0396 | 1st Qu.:1 | 1st Qu.:1.00 | 1st Qu.:0.0000 |
| ## Median : 0.0715 | Median :2 | Median :2.00 | Median :0.0000 |
| ## Mean : 0.1216 | Mean :2 | Mean :1.99 | Mean :0.3377 |
| ## 3rd Qu.: 0.2143 | 3rd Qu.:3 | 3rd Qu.:3.00 | 3rd Qu.:1.0000 |
| ## Max. : 1.9013 | Max. :3 | Max. :3.00 | Max. :1.0000 |
| ## NA's :499 | NA's :499 | NA's :499 | NA's :499 |
| ## I_cysdiff3cat2 | I_cysdiff3cat3 | logcysdiff | til10_0 |
| ## Min. :0.0000 | Min. :0.0000 | Min. :-9.2829 | Min. :1.000 |
| ## 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.: -6.9078 | 1st Qu.:1.000 |
| ## Median :0.0000 | Median :0.0000 | Median :-2.6377 | Median :2.000 |
| ## Mean :0.3351 | Mean :0.3273 | Mean :-3.6729 | Mean :1.999 |
| ## 3rd Qu.:1.0000 | 3rd Qu.:1.0000 | 3rd Qu.: -1.5405 | 3rd Qu.:3.000 |
| ## Max. :1.0000 | Max. :1.0000 | Max. : 0.6426 | Max. :3.000 |
| ## NA's :499 | NA's :499 | NA's :499 | NA's :129 |
| ## il10pre3cat | I_il10pre3cat1 | I_il10pre3cat2 | I_il10pre3cat3 |
| ## Min. :1.000 | Min. :0.0000 | Min. :0.0000 | Min. :0.0000 |
| ## 1st Qu.:1.000 | 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.:0.0000 |
| ## Median :2.000 | Median :0.0000 | Median :0.0000 | Median :0.0000 |
| ## Mean :1.992 | Mean :0.3384 | Mean :0.3311 | Mean :0.3305 |
| ## 3rd Qu.:3.000 | 3rd Qu.:1.0000 | 3rd Qu.:1.0000 | 3rd Qu.:1.0000 |
| ## Max. :3.000 | Max. :1.0000 | Max. :1.0000 | Max. :1.0000 |
| ## NA's :129 | NA's :129 | NA's :129 | NA's :129 |
| ## logil100 | til10_1 | il10post3cat | I_il10post3cat1 |
| ## Min. :-3.0922 | Min. :1.000 | Min. :1.000 | Min. :0.0000 |
| ## 1st Qu.: -1.4397 | 1st Qu.:1.000 | 1st Qu.:1.000 | 1st Qu.:0.0000 |
| ## Median :-1.0966 | Median :2.000 | Median :2.000 | Median :0.0000 |
| ## Mean :-1.0214 | Mean :1.997 | Mean :1.987 | Mean :0.3386 |
| ## 3rd Qu.: -0.6912 | 3rd Qu.:3.000 | 3rd Qu.:3.000 | 3rd Qu.:1.0000 |
| ## Max. : 3.3286 | Max. :3.000 | Max. :3.000 | Max. :1.0000 |
| ## NA's :129 | NA's :393 | NA's :393 | NA's :393 |
| ## I_il10post3cat2 | I_il10post3cat3 | logil101 | il10diff |
| ## Min. :0.0000 | Min. :0.0000 | Min. :-2.8806 | Min. : -7.0350 |
| ## 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.: -0.5293 | 1st Qu.: 0.2255 |
| ## Median :0.0000 | Median :0.0000 | Median :-0.0910 | Median : 0.5290 |
| ## Mean :0.3355 | Mean :0.3259 | Mean :-0.0051 | Mean : 1.0419 |

| | | | |
|--------------------|------------------|-----------------|-------------------|
| ## 3rd Qu.:1.0000 | 3rd Qu.:1.0000 | 3rd Qu.: 0.4637 | 3rd Qu.: 1.0888 |
| ## Max. :1.0000 | Max. :1.0000 | Max. : 4.9416 | Max. :112.1000 |
| ## NA's :393 | NA's :393 | NA's :393 | NA's :501 |
| ## til10diff | il10diff3cat | I_il10diff3cat1 | I_il10diff3cat2 |
| ## Min. :1.000 | Min. :1.000 | Min. :0.0000 | Min. :0.0000 |
| ## 1st Qu.:1.000 | 1st Qu.:1.000 | 1st Qu.:0.0000 | 1st Qu.:0.0000 |
| ## Median :2.000 | Median :2.000 | Median :0.0000 | Median :0.0000 |
| ## Mean :1.999 | Mean :1.993 | Mean :0.3348 | Mean :0.3374 |
| ## 3rd Qu.:3.000 | 3rd Qu.:3.000 | 3rd Qu.:1.0000 | 3rd Qu.:1.0000 |
| ## Max. :3.000 | Max. :3.000 | Max. :1.0000 | Max. :1.0000 |
| ## NA's :501 | NA's :501 | NA's :501 | NA's :501 |
| ## I_il10diff3cat3 | logil10diff | til6_0 | il6pre3cat |
| ## Min. :0.0000 | Min. :-6.9078 | Min. :1.000 | Min. :1.000 |
| ## 1st Qu.:0.0000 | 1st Qu.: -1.4894 | 1st Qu.:1.000 | 1st Qu.:1.000 |
| ## Median :0.0000 | Median :-0.6368 | Median :2.000 | Median :2.000 |
| ## Mean :0.3278 | Mean :-1.0920 | Mean :1.995 | Mean :1.993 |
| ## 3rd Qu.:1.0000 | 3rd Qu.: 0.0850 | 3rd Qu.:3.000 | 3rd Qu.:3.000 |
| ## Max. :1.0000 | Max. : 4.7194 | Max. :3.000 | Max. :3.000 |
| ## NA's :501 | NA's :501 | NA's :123 | NA's :123 |
| ## I_il6pre3cat1 | I_il6pre3cat2 | I_il6pre3cat3 | logil60 |
| ## Min. :0.0000 | Min. :0.0000 | Min. :0.0000 | Min. :-1.67131 |
| ## 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.: -0.09569 |
| ## Median :0.0000 | Median :0.0000 | Median :0.0000 | Median : 0.44788 |
| ## Mean :0.3377 | Mean :0.3318 | Mean :0.3305 | Mean : 0.59375 |
| ## 3rd Qu.:1.0000 | 3rd Qu.:1.0000 | 3rd Qu.:1.0000 | 3rd Qu.: 1.08265 |
| ## Max. :1.0000 | Max. :1.0000 | Max. :1.0000 | Max. : 5.18739 |
| ## NA's :123 | NA's :123 | NA's :123 | NA's :123 |
| ## til6_1 | il6post3cat | I_il6post3cat1 | I_il6post3cat2 |
| ## Min. :1.000 | Min. :1.000 | Min. :0.0000 | Min. :0.0000 |
| ## 1st Qu.:1.000 | 1st Qu.:1.000 | 1st Qu.:0.0000 | 1st Qu.:0.0000 |
| ## Median :2.000 | Median :2.000 | Median :0.0000 | Median :0.0000 |
| ## Mean :1.998 | Mean :1.983 | Mean :0.3394 | Mean :0.3378 |
| ## 3rd Qu.:3.000 | 3rd Qu.:3.000 | 3rd Qu.:1.0000 | 3rd Qu.:1.0000 |
| ## Max. :3.000 | Max. :3.000 | Max. :1.0000 | Max. :1.0000 |
| ## NA's :393 | NA's :393 | NA's :393 | NA's :393 |
| ## I_il6post3cat3 | logil61 | il6diff | til6diff |
| ## Min. :0.0000 | Min. :-3.247 | Min. : -72.10 | Min. :1.000 |
| ## 1st Qu.:0.0000 | 1st Qu.: 2.815 | 1st Qu.: 14.57 | 1st Qu.:1.000 |
| ## Median :0.0000 | Median : 3.330 | Median : 25.59 | Median :2.000 |
| ## Mean :0.3227 | Mean : 3.334 | Mean : 36.49 | Mean :1.999 |
| ## 3rd Qu.:1.0000 | 3rd Qu.: 3.869 | 3rd Qu.: 44.64 | 3rd Qu.:3.000 |
| ## Max. :1.0000 | Max. : 7.345 | Max. :1445.00 | Max. :3.000 |
| ## NA's :393 | NA's :393 | NA's :498 | NA's :498 |
| ## il6diff3cat | I_il6diff3cat1 | I_il6diff3cat2 | I_il6diff3cat3 |
| ## Min. :1.000 | Min. :0.0000 | Min. :0.0000 | Min. :0.0000 |
| ## 1st Qu.:1.000 | 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.:0.0000 |
| ## Median :2.000 | Median :0.0000 | Median :0.0000 | Median :0.0000 |
| ## Mean :1.989 | Mean :0.3374 | Mean :0.3365 | Mean :0.3261 |
| ## 3rd Qu.:3.000 | 3rd Qu.:1.0000 | 3rd Qu.:1.0000 | 3rd Qu.:1.0000 |
| ## Max. :3.000 | Max. :1.0000 | Max. :1.0000 | Max. :1.0000 |
| ## NA's :498 | NA's :498 | NA's :498 | NA's :498 |
| ## logil6diff | gal30_adj | tgall3_0 | gal3pre3cat |
| ## Min. :-6.908 | Min. : 1375 | Min. :1.000 | Min. :1.000 |
| ## 1st Qu.: 2.679 | 1st Qu.: 6954 | 1st Qu.:1.000 | 1st Qu.:1.000 |

```

## Median : 3.242      Median : 10312      Median :2.000      Median :3.000
## Mean   : 3.037      Mean   : 11848      Mean   :1.999      Mean   :2.194
## 3rd Qu.: 3.799      3rd Qu.: 14703      3rd Qu.:3.000      3rd Qu.:3.000
## Max.   : 7.276      Max.   :123355      Max.   :3.000      Max.   :3.000
## NA's   :498         NA's   :123         NA's   :123         NA's   :123
## I_gal3pre3cat1      I_gal3pre3cat2      I_gal3pre3cat3      loggal30
## Min.   :0.0000      Min.   :0.0000      Min.   :0.0000      Min.   : 7.226
## 1st Qu.:0.0000      1st Qu.:0.0000      1st Qu.:0.0000      1st Qu.: 8.847
## Median :0.0000      Median :0.0000      Median :1.0000      Median : 9.241
## Mean   :0.3357      Mean   :0.1348      Mean   :0.5294      Mean   : 9.206
## 3rd Qu.:1.0000      3rd Qu.:0.0000      3rd Qu.:1.0000      3rd Qu.: 9.596
## Max.   :1.0000      Max.   :1.0000      Max.   :1.0000      Max.   :11.723
## NA's   :123         NA's   :123         NA's   :123         NA's   :123
## gal31_adj           tgal3_1           gal3post3cat      I_gal3post3cat1
## Min.   : 464.9      Min.   :1.000      Min.   :1.00      Min.   :0.0000
## 1st Qu.: 7147.3      1st Qu.:1.000      1st Qu.:1.00      1st Qu.:0.0000
## Median : 10351.5      Median :2.000      Median :2.00      Median :0.0000
## Mean   : 12887.9      Mean   :1.999      Mean   :1.99      Mean   :0.3386
## 3rd Qu.: 15112.9      3rd Qu.:3.000      3rd Qu.:3.00      3rd Qu.:1.0000
## Max.   :390166.0      Max.   :3.000      Max.   :3.00      Max.   :1.0000
## NA's   :393         NA's   :393         NA's   :393         NA's   :393
## I_gal3post3cat2      I_gal3post3cat3      loggal31           gal3diff
## Min.   :0.0000      Min.   :0.0000      Min.   : 6.142      Min.   : -61540
## 1st Qu.:0.0000      1st Qu.:0.0000      1st Qu.: 8.874      1st Qu.: -1544
## Median :0.0000      Median :0.0000      Median : 9.245      Median : 214
## Mean   :0.3331      Mean   :0.3283      Mean   : 9.239      Mean   : 1389
## 3rd Qu.:1.0000      3rd Qu.:1.0000      3rd Qu.: 9.623      3rd Qu.: 2206
## Max.   :1.0000      Max.   :1.0000      Max.   :12.874      Max.   :375885
## NA's   :393         NA's   :393         NA's   :393         NA's   :498
## tgal3diff           gal3diff3cat      I_gal3diff3cat1      I_gal3diff3cat2
## Min.   :1.000      Min.   :1.000      Min.   :0.0000      Min.   :0.0000
## 1st Qu.:1.000      1st Qu.:1.000      1st Qu.:0.0000      1st Qu.:0.0000
## Median :2.000      Median :2.000      Median :0.0000      Median :0.0000
## Mean   :1.999      Mean   :1.991      Mean   :0.3356      Mean   :0.3374
## 3rd Qu.:3.000      3rd Qu.:3.000      3rd Qu.:1.0000      3rd Qu.:1.0000
## Max.   :3.000      Max.   :3.000      Max.   :1.0000      Max.   :1.0000
## NA's   :498         NA's   :498         NA's   :498         NA's   :498
## I_gal3diff3cat3      loggal3diff           ntpro0_adj           tntbnp_0
## Min.   :0.000      Min.   : -6.9078      Min.   : 121      Min.   :1.000
## 1st Qu.:0.000      1st Qu.: -6.9078      1st Qu.: 1161      1st Qu.:1.000
## Median :0.000      Median : 5.3658      Median : 2592      Median :2.000
## Mean   :0.327      Mean   : 0.8795      Mean   : 9002      Mean   :1.999
## 3rd Qu.:1.000      3rd Qu.: 7.6989      3rd Qu.: 6598      3rd Qu.:3.000
## Max.   :1.000      Max.   :12.8370      Max.   :568901      Max.   :3.000
## NA's   :498         NA's   :498         NA's   :123         NA's   :123
## ntbnppre3cat      I_ntbnppre3cat1      I_ntbnppre3cat2      I_ntbnppre3cat3
## Min.   :1.000      Min.   :0.000      Min.   :0.0000      Min.   :0.0000
## 1st Qu.:1.000      1st Qu.:0.000      1st Qu.:0.0000      1st Qu.:0.0000
## Median :2.000      Median :0.000      Median :0.0000      Median :0.0000
## Mean   :1.986      Mean   :0.339      Mean   :0.3357      Mean   :0.3253
## 3rd Qu.:3.000      3rd Qu.:1.000      3rd Qu.:1.0000      3rd Qu.:1.0000
## Max.   :3.000      Max.   :1.000      Max.   :1.0000      Max.   :1.0000
## NA's   :123         NA's   :123         NA's   :123         NA's   :123
## logntp0           ntpro1_adj           tntbnp_1           ntbnppost3cat

```

```

## Min. : 4.796 Min. : 231.4 Min. :1.000 Min. :1.000
## 1st Qu.: 7.057 1st Qu.: 8799.8 1st Qu.:1.000 1st Qu.:1.000
## Median : 7.860 Median : 15424.3 Median :2.000 Median :2.000
## Mean : 7.994 Mean : 25538.6 Mean :1.999 Mean :1.984
## 3rd Qu.: 8.794 3rd Qu.: 28453.8 3rd Qu.:3.000 3rd Qu.:3.000
## Max. :13.251 Max. :444419.6 Max. :3.000 Max. :3.000
## NA's :123 NA's :393 NA's :393 NA's :393
## I_ntbnppost3cat1 I_ntbnppost3cat2 I_ntbnppost3cat3 logntp1
## Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. : 5.444
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.: 9.082
## Median :0.0000 Median :0.0000 Median :0.0000 Median : 9.644
## Mean :0.3402 Mean :0.3355 Mean :0.3243 Mean : 9.716
## 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:10.256
## Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :13.005
## NA's :393 NA's :393 NA's :393 NA's :393
## ntbnpdiff tntbnpdiff ntbnpdiff3cat I_ntbnpdiff3cat1
## Min. :-198617 Min. :1.000 Min. :1.00 Min. :0.0000
## 1st Qu.: 6342 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:0.0000
## Median : 11373 Median :2.000 Median :2.00 Median :0.0000
## Mean : 16903 Mean :1.999 Mean :1.99 Mean :0.3365
## 3rd Qu.: 20574 3rd Qu.:3.000 3rd Qu.:3.00 3rd Qu.:1.0000
## Max. : 218680 Max. :3.000 Max. :3.00 Max. :1.0000
## NA's :498 NA's :498 NA's :498 NA's :498
## I_ntbnpdiff3cat2 I_ntbnpdiff3cat3 logntbnpdiff st20_adj
## Min. :0.0000 Min. :0.000 Min. :-6.908 Min. : 858
## 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.: 8.755 1st Qu.: 2871
## Median :0.0000 Median :0.000 Median : 9.339 Median : 3794
## Mean :0.3365 Mean :0.327 Mean : 8.831 Mean : 5252
## 3rd Qu.:1.0000 3rd Qu.:1.000 3rd Qu.: 9.932 3rd Qu.: 5199
## Max. :1.0000 Max. :1.000 Max. :12.295 Max. :108253
## NA's :498 NA's :498 NA's :498 NA's :123
## tst2_0 st2pre3cat I_st2pre3cat1 I_st2pre3cat2
## Min. :1.000 Min. :1.000 Min. :0.000 Min. :0.0000
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:0.000 1st Qu.:0.0000
## Median :2.000 Median :2.000 Median :0.000 Median :0.0000
## Mean :1.999 Mean :1.991 Mean :0.337 Mean :0.3351
## 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:1.000 3rd Qu.:1.0000
## Max. :3.000 Max. :3.000 Max. :1.000 Max. :1.0000
## NA's :123 NA's :123 NA's :123 NA's :123
## I_st2pre3cat3 logst20 st21_adj tst2_1
## Min. :0.0000 Min. : 6.755 Min. : 156.2 Min. :1.000
## 1st Qu.:0.0000 1st Qu.: 7.962 1st Qu.: 24686.0 1st Qu.:1.000
## Median :0.0000 Median : 8.241 Median : 43944.5 Median :2.000
## Mean :0.3279 Mean : 8.311 Mean : 57786.5 Mean :1.999
## 3rd Qu.:1.0000 3rd Qu.: 8.556 3rd Qu.: 74095.2 3rd Qu.:3.000
## Max. :1.0000 Max. :11.592 Max. :545665.0 Max. :3.000
## NA's :123 NA's :123 NA's :393 NA's :393
## st2post3cat I_st2post3cat1 I_st2post3cat2 I_st2post3cat3
## Min. :1.000 Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:1.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :2.000 Median :0.0000 Median :0.0000 Median :0.0000
## Mean :1.985 Mean :0.3378 Mean :0.3394 Mean :0.3227
## 3rd Qu.:3.000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000
## Max. :3.000 Max. :1.0000 Max. :1.0000 Max. :1.0000

```

| ## | NA's :393 | NA's :393 | NA's :393 | NA's :393 |
|----|-----------------|-----------------|-----------------|-----------------|
| ## | logst21 | st2diff | tst2diff | st2diff3cat |
| ## | Min. : 5.051 | Min. : -70392 | Min. : 1.000 | Min. : 1.000 |
| ## | 1st Qu.: 10.114 | 1st Qu.: 20288 | 1st Qu.: 1.000 | 1st Qu.: 1.000 |
| ## | Median : 10.691 | Median : 40217 | Median : 2.000 | Median : 2.000 |
| ## | Mean : 10.640 | Mean : 53696 | Mean : 1.999 | Mean : 1.988 |
| ## | 3rd Qu.: 11.213 | 3rd Qu.: 69806 | 3rd Qu.: 3.000 | 3rd Qu.: 3.000 |
| ## | Max. : 13.210 | Max. : 542834 | Max. : 3.000 | Max. : 3.000 |
| ## | NA's :393 | NA's :498 | NA's :498 | NA's :498 |
| ## | I_st2diff3cat1 | I_st2diff3cat2 | I_st2diff3cat3 | logst2diff |
| ## | Min. : 0.0000 | Min. : 0.0000 | Min. : 0.0000 | Min. : -6.908 |
| ## | 1st Qu.: 0.0000 | 1st Qu.: 0.0000 | 1st Qu.: 0.0000 | 1st Qu.: 9.918 |
| ## | Median : 0.0000 | Median : 0.0000 | Median : 0.0000 | Median : 10.602 |
| ## | Mean : 0.3365 | Mean : 0.3391 | Mean : 0.3244 | Mean : 10.310 |
| ## | 3rd Qu.: 1.0000 | 3rd Qu.: 1.0000 | 3rd Qu.: 1.0000 | 3rd Qu.: 11.153 |
| ## | Max. : 1.0000 | Max. : 1.0000 | Max. : 1.0000 | Max. : 13.205 |
| ## | NA's :498 | NA's :498 | NA's :498 | NA's :498 |

```
str(data)
```

```
## tibble [1,651 x 260] (S3: tbl_df/tbl/data.frame)
## $ readmit30d_yn_state: num [1:1651] 1 1 0 0 0 0 1 0 0 0 ...
## $ dead                : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ obleed              : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ ocva                : num [1:1651] 0 0 0 0 0 0 0 0 1 0 ...
## $ oleginf              : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ otia                 : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ orf                  : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ opneu                : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ oafib2               : num [1:1651] 1 1 0 0 1 0 1 0 1 1 ...
## $ return               : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ akin                 : num [1:1651] 1 0 0 0 2 0 1 0 0 0 ...
## $ lm2cat               : num [1:1651] 1 1 0 1 0 0 0 0 0 1 ...
## $ rfcr_2               : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ iabppre_2            : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ atfibyn_2            : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ hyperyn_2            : num [1:1651] 1 0 1 0 1 1 1 0 1 0 ...
## $ ef50_neg             : num [1:1651] 0 13 0 NA 0 0 10 0 0 0 ...
## $ ef50_neg_d           : num [1:1651] 1 3 1 NA 1 1 2 1 1 1 ...
## $ priormi_21           : num [1:1651] 0 1 0 0 1 0 0 0 1 0 ...
## $ priormi_22           : num [1:1651] 0 1 0 0 1 0 0 0 1 0 ...
## $ priormi_23           : num [1:1651] 0 1 0 0 1 0 0 0 1 0 ...
## $ priormi_24           : num [1:1651] 0 1 0 0 1 0 0 0 0 0 ...
## $ age_d                : num [1:1651] 8 7 7 8 8 7 7 6 6 8 ...
## $ ua_nmi7              : num [1:1651] 1 1 0 0 1 1 1 1 0 0 ...
## $ creatpre_2           : num [1:1651] 1.5 1.2 1.2 0.9 1.1 ...
## $ rf_1                 : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ rf_2                 : num [1:1651] 0.5 0.2 0.2 0 0.1 ...
## $ rf_3                 : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ priority_21          : num [1:1651] 1 0 0 1 0 0 1 0 1 1 ...
## $ priority_22          : num [1:1651] 1 0 0 0 0 0 0 0 0 0 ...
## $ dm3cat               : num [1:1651] 1 1 2 2 1 1 1 1 1 1 ...
## $ dm3cat_21            : num [1:1651] 0 0 1 1 0 0 0 0 0 0 ...
## $ dm3cat_22            : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ dm3cat_21_2          : num [1:1651] 0 0 1 1 0 0 0 0 0 0 ...
```

```

## $ dm3cat_22_2      : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ pci_ta           : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ efvalue_r        : num [1:1651] 60 37 52 60 60 63 40 50 77 65 ...
## $ ef50_neg_r       : num [1:1651] 0 13 0 0 0 0 10 0 0 0 ...
## $ ef50_neg_d_r     : num [1:1651] 1 3 1 1 1 1 2 1 1 1 ...
## $ atfibyn_2_r      : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ priormi_r        : num [1:1651] 0 4 0 0 4 0 0 0 3 0 ...
## $ priormi_21_r     : num [1:1651] 0 1 0 0 1 0 0 0 1 0 ...
## $ priormi_22_r     : num [1:1651] 0 1 0 0 1 0 0 0 1 0 ...
## $ priormi_23_r     : num [1:1651] 0 1 0 0 1 0 0 0 1 0 ...
## $ priormi_24_r     : num [1:1651] 0 1 0 0 1 0 0 0 0 0 ...
## $ age_r            : num [1:1651] 74.1 69 64.7 73.5 77.3 ...
## $ age_d_r          : num [1:1651] 8 7 7 8 8 7 7 6 6 8 ...
## $ ua_r             : num [1:1651] 1 1 0 0 1 1 1 1 0 0 ...
## $ ua_nmi7_r        : num [1:1651] 1 1 0 0 1 1 1 1 0 0 ...
## $ chf_r            : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ chf2cat_r        : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ rf_r             : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ creatpre_2_r     : num [1:1651] 1.5 1.2 1.2 0.9 1.1 ...
## $ rf_1_r           : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ rf_2_r           : num [1:1651] 0.5 0.2 0.2 0 0.1 ...
## $ rf_3_r           : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ priority_r       : num [1:1651] 1 3 3 2 3 3 2 3 2 2 ...
## $ priority_21_r    : num [1:1651] 1 0 0 1 0 0 1 0 1 1 ...
## $ priority_22_r    : num [1:1651] 1 0 0 0 0 0 0 0 0 0 ...
## $ sex_r            : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ prcabg_r         : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ copd_r           : num [1:1651] 0 1 0 0 0 0 0 0 0 0 ...
## $ anydm_r          : num [1:1651] 0 0 1 1 0 0 0 0 0 0 ...
## $ dmtx_r           : num [1:1651] NA NA 2 2 NA NA NA NA NA NA ...
## $ dm3cat_r         : num [1:1651] 1 1 2 2 1 1 1 1 1 1 ...
## $ dm3cat_21_r      : num [1:1651] 0 0 1 1 0 0 0 0 0 0 ...
## $ dm3cat_22_r      : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ iabppre_r        : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ anyvad_r         : num [1:1651] 0 0 0 1 0 0 0 0 0 0 ...
## $ vad_r            : num [1:1651] 0 0 0 1 0 0 0 0 0 0 ...
## $ hyper_r          : num [1:1651] 3 0 1 0 3 1 3 0 1 0 ...
## $ hyperyn_2_r      : num [1:1651] 1 0 1 0 1 1 1 0 1 0 ...
## $ prptca6_r        : num [1:1651] 0 2 0 0 0 0 0 0 0 0 ...
## $ pci_ta_r         : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ lm3cat_r         : num [1:1651] 2 2 1 2 1 1 1 1 1 2 ...
## $ lm2cat_r         : num [1:1651] 1 1 0 1 0 0 0 0 0 1 ...
## $ aortic_insuff    : num [1:1651] 0 NA NA NA 0 NA 0 NA NA NA ...
## $ aortic_insuff_r  : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ aortic_sten      : num [1:1651] 0 NA NA NA 0 NA 0 NA NA NA ...
## $ aortic_sten_r    : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ chf_nyha_iv      : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ chf_nyha_ltiv    : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ chf_nyha_iv_r    : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ chf_nyha_ltiv_r  : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ smoker_r         : num [1:1651] 0 0 0 0 0 0 0 0 1 0 ...
## $ cvd              : num [1:1651] 0 0 0 1 0 0 0 0 0 0 ...
## $ cvd_r            : num [1:1651] 0 0 0 1 0 0 0 0 0 0 ...
## $ htcn_d           : num [1:1651] 17 16 17 17 19 18 18 17 17 17 ...

```

```
## $ htc_m_r      : num [1:1651] 175 168 173 175 193 183 185 173 175 171 ...
## $ htc_d_r      : num [1:1651] 17 16 17 17 19 18 18 17 17 17 ...
## $ mitral_insuff : num [1:1651] 1 NA NA NA 0 NA 0 NA NA NA ...
## $ mitral_insuff_r : num [1:1651] 1 0 0 0 0 0 0 0 0 0 ...
## $ carotid_sten  : num [1:1651] NA 0 0 0 NA 0 NA 0 0 0 ...
## $ carotid_sten_r : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ pvd           : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ pvd_r         : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ tricuspid_insuff : num [1:1651] 0 NA NA NA 0 NA 0 NA NA NA ...
## $ tricuspid_insuff_r : num [1:1651] 0 0 0 0 0 0 0 0 0 0 ...
## $ bmi_squared   : num [1:1651] 826 1868 904 717 904 ...
## [list output truncated]
```

Machine learning

```
head(data)
```

```
## # A tibble: 6 x 260
##   readmit30d_yn_state  dead  obleed  ocva  oleginf  otia  orf  opneu  oafib2  return
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1         1      0      0      0      0      0      0      0      1      0
## 2         1      0      0      0      0      0      0      0      1      0
## 3         0      0      0      0      0      0      0      0      0      0
## 4         0      0      0      0      0      0      0      0      0      0
## 5         0      0      0      0      0      0      0      0      1      0
## 6         0      0      0      0      0      0      0      0      0      0
## # ... with 250 more variables: akin <dbl>, lm2cat <dbl>, rfcr_2 <dbl>,
## #   iabppre_2 <dbl>, atfibyn_2 <dbl>, hyperyn_2 <dbl>, ef50_neg <dbl>,
## #   ef50_neg_d <dbl>, priormi_21 <dbl>, priormi_22 <dbl>, priormi_23 <dbl>,
## #   priormi_24 <dbl>, age_d <dbl>, ua_nmi7 <dbl>, creatpre_2 <dbl>, rf_1 <dbl>,
## #   rf_2 <dbl>, rf_3 <dbl>, priority_21 <dbl>, priority_22 <dbl>, dm3cat <dbl>,
## #   dm3cat_21 <dbl>, dm3cat_22 <dbl>, dm3cat_21_2 <dbl>, dm3cat_22_2 <dbl>,
## #   pci_ta <dbl>, efvalue_r <dbl>, ef50_neg_r <dbl>, ef50_neg_d_r <dbl>, ...
```

```
names(data)
```

```
##   [1] "readmit30d_yn_state" "dead"           "obleed"
##   [4] "ocva"                "oleginf"        "otia"
##   [7] "orf"                 "opneu"          "oafib2"
##  [10] "return"              "akin"           "lm2cat"
##  [13] "rfcr_2"              "iabppre_2"      "atfibyn_2"
##  [16] "hyperyn_2"           "ef50_neg"       "ef50_neg_d"
##  [19] "priormi_21"          "priormi_22"     "priormi_23"
##  [22] "priormi_24"          "age_d"          "ua_nmi7"
##  [25] "creatpre_2"          "rf_1"           "rf_2"
##  [28] "rf_3"                "priority_21"    "priority_22"
##  [31] "dm3cat"              "dm3cat_21"      "dm3cat_22"
##  [34] "dm3cat_21_2"         "dm3cat_22_2"    "pci_ta"
##  [37] "efvalue_r"           "ef50_neg_r"     "ef50_neg_d_r"
##  [40] "atfibyn_2_r"         "priormi_r"      "priormi_21_r"
##  [43] "priormi_22_r"        "priormi_23_r"   "priormi_24_r"
##  [46] "age_r"               "age_d_r"        "ua_r"
##  [49] "ua_nmi7_r"           "chf_r"          "chf2cat_r"
##  [52] "rf_r"                "creatpre_2_r"   "rf_1_r"
##  [55] "rf_2_r"              "rf_3_r"         "priority_r"
##  [58] "priority_21_r"       "priority_22_r"  "sex_r"
```

| | | | |
|----------|-----------------------|----------------------|-------------------|
| ## [61] | "prcabg_r" | "copd_r" | "anydm_r" |
| ## [64] | "dmtx_r" | "dm3cat_r" | "dm3cat_21_r" |
| ## [67] | "dm3cat_22_r" | "iabppre_r" | "anyvad_r" |
| ## [70] | "vad_r" | "hyper_r" | "hyperyn_2_r" |
| ## [73] | "prptca6_r" | "pci_ta_r" | "lm3cat_r" |
| ## [76] | "lm2cat_r" | "aortic_insuff" | "aortic_insuff_r" |
| ## [79] | "aortic_sten" | "aortic_sten_r" | "chf_nyha_iv" |
| ## [82] | "chf_nyha_ltiv" | "chf_nyha_iv_r" | "chf_nyha_ltiv_r" |
| ## [85] | "smoker_r" | "cvd" | "cvd_r" |
| ## [88] | "htcm_d" | "htcm_r" | "htcm_d_r" |
| ## [91] | "mitral_insuff" | "mitral_insuff_r" | "carotid_sten" |
| ## [94] | "carotid_sten_r" | "pvd" | "pvd_r" |
| ## [97] | "tricuspid_insuff" | "tricuspid_insuff_r" | "bmi_squared" |
| ## [100] | "bmi_r" | "bmi_squared_r" | "novsl_r" |
| ## [103] | "readmit_1y_yn_state" | "anyakin" | "creatcat" |
| ## [106] | "lm50" | "anymssd" | "lowoutput" |
| ## [109] | "emerg" | "urg" | "elec" |
| ## [112] | "bmicat" | "bmi1" | "bmi2" |
| ## [115] | "bmi3" | "bmi4" | "bmi5" |
| ## [118] | "bmi6" | "lvedpm" | "anemiapre" |
| ## [121] | "iabpintra" | "lof1" | "ptimenumban" |
| ## [124] | "ctimenumban" | "cardtimenumban" | "cardblood" |
| ## [127] | "cardcold" | "hotshot" | "aoxcon" |
| ## [130] | "ultrafilyn" | "cabg" | "valve" |
| ## [133] | "cabgvalve" | "gfr60pre" | "male" |
| ## [136] | "notcoldcard" | "fluidprel" | "ptime120" |
| ## [139] | "heptot1" | "heptot5" | "tcys0" |
| ## [142] | "cyspre3cat" | "I_cyspre3cat1" | "I_cyspre3cat2" |
| ## [145] | "I_cyspre3cat3" | "logcys0" | "tcys1" |
| ## [148] | "cyspost3cat" | "I_cyspost3cat1" | "I_cyspost3cat2" |
| ## [151] | "I_cyspost3cat3" | "logcys1" | "cysdiff" |
| ## [154] | "tcysdiff" | "cysdiff3cat" | "I_cysdiff3cat1" |
| ## [157] | "I_cysdiff3cat2" | "I_cysdiff3cat3" | "logcysdiff" |
| ## [160] | "til10_0" | "il10pre3cat" | "I_il10pre3cat1" |
| ## [163] | "I_il10pre3cat2" | "I_il10pre3cat3" | "logil100" |
| ## [166] | "til10_1" | "il10post3cat" | "I_il10post3cat1" |
| ## [169] | "I_il10post3cat2" | "I_il10post3cat3" | "logil101" |
| ## [172] | "il10diff" | "til10diff" | "il10diff3cat" |
| ## [175] | "I_il10diff3cat1" | "I_il10diff3cat2" | "I_il10diff3cat3" |
| ## [178] | "logil10diff" | "til6_0" | "il6pre3cat" |
| ## [181] | "I_il6pre3cat1" | "I_il6pre3cat2" | "I_il6pre3cat3" |
| ## [184] | "logil60" | "til6_1" | "il6post3cat" |
| ## [187] | "I_il6post3cat1" | "I_il6post3cat2" | "I_il6post3cat3" |
| ## [190] | "logil61" | "il6diff" | "til6diff" |
| ## [193] | "il6diff3cat" | "I_il6diff3cat1" | "I_il6diff3cat2" |
| ## [196] | "I_il6diff3cat3" | "logil6diff" | "gal30_adj" |
| ## [199] | "tgal3_0" | "gal3pre3cat" | "I_gal3pre3cat1" |
| ## [202] | "I_gal3pre3cat2" | "I_gal3pre3cat3" | "loggal30" |
| ## [205] | "gal31_adj" | "tgal3_1" | "gal3post3cat" |
| ## [208] | "I_gal3post3cat1" | "I_gal3post3cat2" | "I_gal3post3cat3" |
| ## [211] | "loggal31" | "gal3diff" | "tgal3diff" |
| ## [214] | "gal3diff3cat" | "I_gal3diff3cat1" | "I_gal3diff3cat2" |
| ## [217] | "I_gal3diff3cat3" | "loggal3diff" | "ntpro0_adj" |
| ## [220] | "tntbnp_0" | "ntbnp3cat" | "I_ntbnp3cat1" |

```
## [223] "I_ntbnppre3cat2"      "I_ntbnppre3cat3"      "logntp0"
## [226] "ntpro1_adj"           "tntbnp_1"             "ntbnppost3cat"
## [229] "I_ntbnppost3cat1"     "I_ntbnppost3cat2"     "I_ntbnppost3cat3"
## [232] "logntp1"             "ntbnpdiff"            "tntbnpdiff"
## [235] "ntbnpdiff3cat"        "I_ntbnpdiff3cat1"     "I_ntbnpdiff3cat2"
## [238] "I_ntbnpdiff3cat3"     "logntbnpdiff"         "st20_adj"
## [241] "tst2_0"              "st2pre3cat"           "I_st2pre3cat1"
## [244] "I_st2pre3cat2"        "I_st2pre3cat3"        "logst20"
## [247] "st21_adj"            "tst2_1"               "st2post3cat"
## [250] "I_st2post3cat1"       "I_st2post3cat2"       "I_st2post3cat3"
## [253] "logst21"             "st2diff"              "tst2diff"
## [256] "st2diff3cat"         "I_st2diff3cat1"       "I_st2diff3cat2"
## [259] "I_st2diff3cat3"      "logst2diff"
```

```
quick_models <- machine_learn(data, outcome = dead)
```

```
## Training new data prep recipe...
```

```
## Removing the following 25 near-zero variance column(s). If you don't want to remove them, call prep_recipe()
##   obleed, ocva, oleginf, otia, orf, opneu, rf_1, rf_3, priority_22, rf_r, rf_1_r, rf_3_r, priority_22_r
```

```
##
```

```
## dead looks categorical, so training classification algorithms.
```

```
##
```

```
## After data processing, models are being trained on 234 features with 1,651 observations.
```

```
## Based on n_folds = 5 and hyperparameter settings, the following number of models will be trained: 50
```

```
## Training with cross validation: Random Forest
```

```
## You may, or may not, see messages about progress in growing trees. The estimates are very rough, and
```

```
## Training with cross validation: eXtreme Gradient Boosting
```

```
## Training with cross validation: glmnet
```

```
##
```

```
## *** Models successfully trained. The model object contains the training data minus ignored ID columns.
```

```
## *** If there was PHI in training data, normal PHI protocols apply to the model object. ***
```

```
quick_models
```

```
## Algorithms Trained: Random Forest, eXtreme Gradient Boosting, and glmnet
```

```
## Model Name: dead
```

```
## Target: dead
```

```
## Class: Classification
```

```
## Performance Metric: AUROC
```

```
## Number of Observations: 1651
```

```
## Number of Features: 234
```

```
## Models Trained: 2021-09-05 10:15:27
```

```
##
```

```
## Models tuned via 5-fold cross validation over 10 combinations of hyperparameter values.
```

```
## Best model: glmnet
```

```
## AUPR = 0.38, AUROC = 0.78
```

```
## Optimal hyperparameter values:
```

```
##   alpha = 1
```

```
##   lambda = 0.0095
```

```
predictions <- predict(quick_models)
```

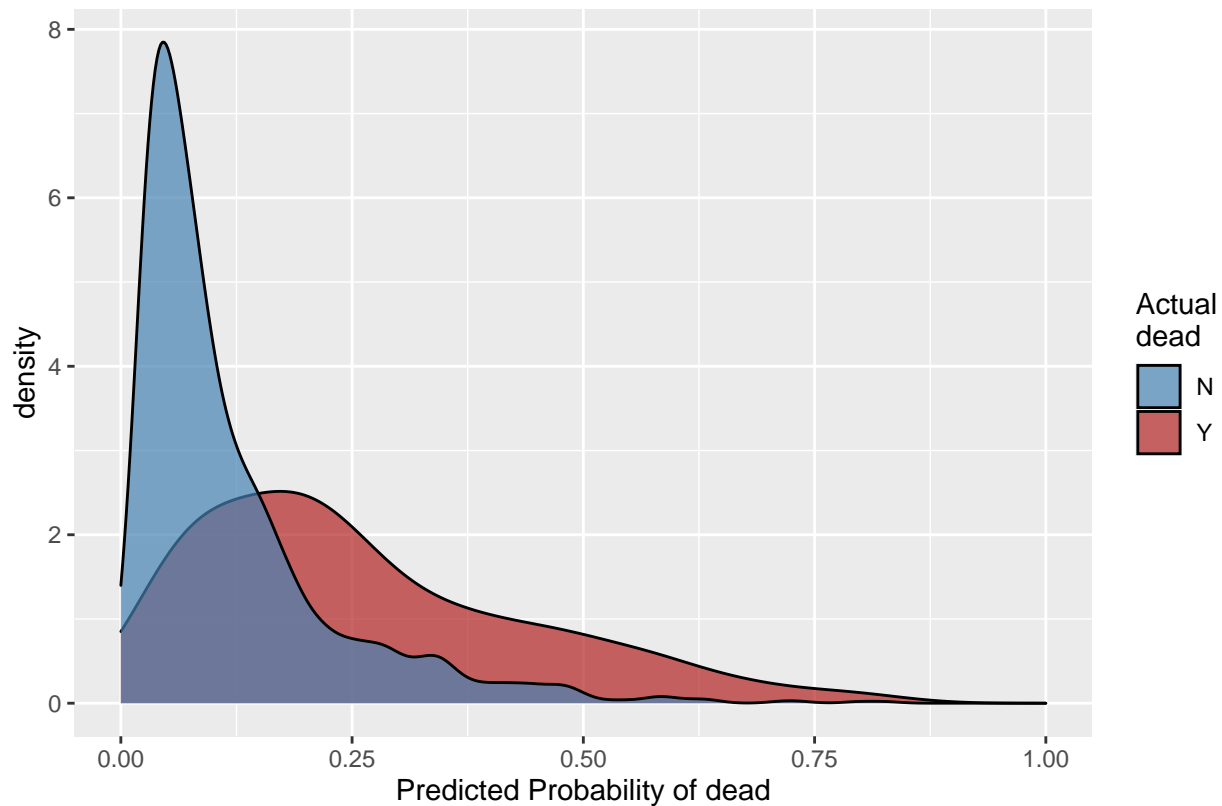
```
predictions
```



```
## "predicted_dead" predicted by glmnet last trained: 2021-09-05 10:15:27
## Performance in training: AUROC = 0.78
```

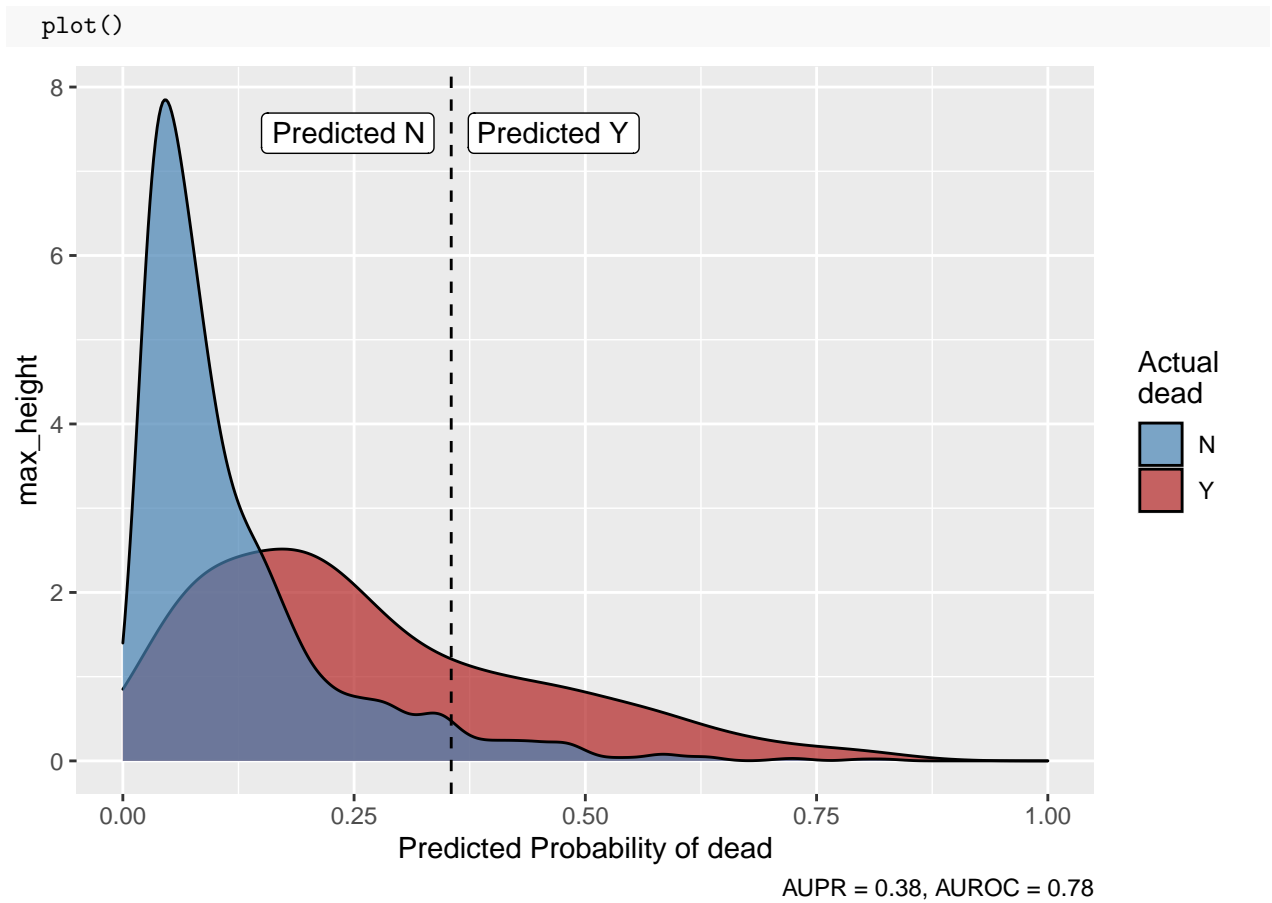
```
## # A tibble: 1,651 x 261
##   dead predicted_dead readmit30d_yn_st~ obleed ocva oleginf otia orf opneu
## * <fct>          <dbl>          <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 N              0.182              1     0     0     0     0     0     0
## 2 N              0.0728             1     0     0     0     0     0     0
## 3 N              0.0709             0     0     0     0     0     0     0
## 4 N              0.107              0     0     0     0     0     0     0
## 5 N              0.143              0     0     0     0     0     0     0
## 6 N              0.0285             0     0     0     0     0     0     0
## 7 N              0.116             1     0     0     0     0     0     0
## 8 N              0.0384             0     0     0     0     0     0     0
## 9 N              0.103             0     0     1     0     0     0     0
## 10 N             0.0720             0     0     0     0     0     0     0
## # ... with 1,641 more rows, and 252 more variables: oafib2 <dbl>, return <dbl>,
## #   akin <dbl>, lm2cat <dbl>, rfcr_2 <dbl>, iabppre_2 <dbl>, atfibyn_2 <dbl>,
## #   hyperyn_2 <dbl>, ef50_neg <dbl>, ef50_neg_d <dbl>, priormi_21 <dbl>,
## #   priormi_22 <dbl>, priormi_23 <dbl>, priormi_24 <dbl>, age_d <dbl>,
## #   ua_nmi7 <dbl>, creatpre_2 <dbl>, rf_1 <dbl>, rf_2 <dbl>, rf_3 <dbl>,
## #   priority_21 <dbl>, priority_22 <dbl>, dm3cat <dbl>, dm3cat_21 <dbl>,
## #   dm3cat_22 <dbl>, dm3cat_21_2 <dbl>, dm3cat_22_2 <dbl>, pci_ta <dbl>, ...
```

```
plot(predictions)
```



AUPR = 0.38, AUROC = 0.78

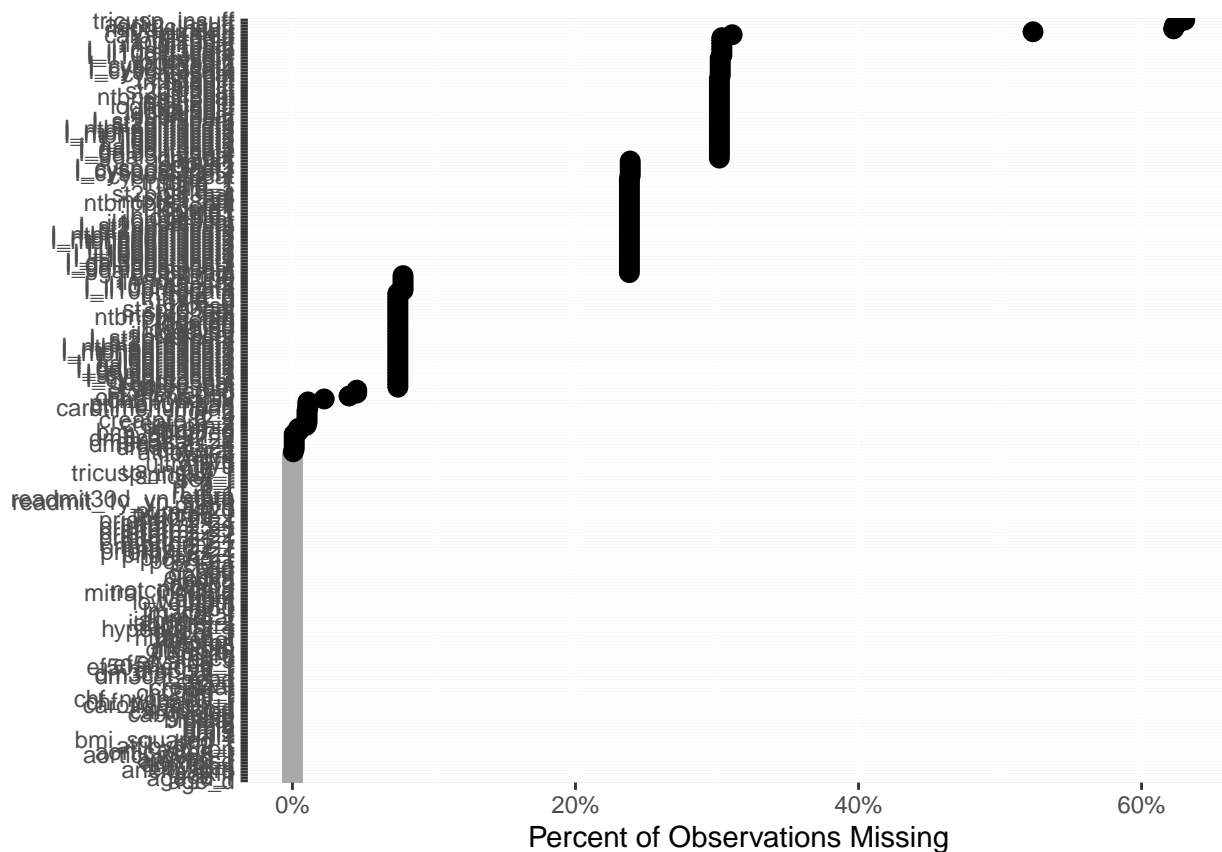
```
quick_models %>%
  predict(outcome_groups = 2) %>%
```



data profiling

```
q<-missingness(data)
plot(q,guide = "none")
```

```
## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.
```



data preparation

prep_data

```
## function (d, ..., outcome, recipe = NULL, remove_near_zero_variance = TRUE,
##   convert_dates = TRUE, impute = TRUE, collapse_rare_factors = TRUE,
##   PCA = FALSE, center = FALSE, scale = FALSE, make_dummies = TRUE,
##   add_levels = TRUE, logical_to_numeric = TRUE, factor_outcome = TRUE,
##   no_prep = FALSE)
## {
##   if (!is.data.frame(d))
##     stop("\nd\" must be a data frame.")
##   orig_data <- d
##   new_recipe <- TRUE
##   if (!is.null(recipe)) {
##     new_recipe <- FALSE
##     recipe <- check_rec_obj(recipe)
##     no_prep <- attr(recipe, "no_prep")
##   }
##   if (no_prep)
##     remove_near_zero_variance <- convert_dates <- impute <- collapse_rare_factors <- center <- s
##   d_missing <- missingness(d, return_df = FALSE)
##   d_ods <- d[0, ]
##   d_levels <- get_factor_levels(d)
##   best_levels <- attr(d, "best_levels")
##   outcome <- rlang::enquo(outcome)
##   remove_outcome <- FALSE
##   ignore_columns <- rlang::quos(...)
```

```

## ignored <- purrr::map_chr(ignore_columns, rlang::quo_name)
## d_ignore <- NULL
## if (length(ignored)) {
##   present <- ignored %in% names(d)
##   if (any(!present))
##     stop(list_variables(ignored[!present]), " not found in d.")
##   if (length(ignored) >= ncol(d))
##     stop("You only have ignored columns. Try again.")
##   d_ignore <- dplyr::select(d, !!ignored)
##   d <- dplyr::select(d, -dplyr::one_of(ignored))
##   m <- missingness(d_ignore) %>% dplyr::filter(percent_missing >
##     0)
##   if (!purrr::is_empty(m$variable))
##     warning("These ignored variables have missingness: ",
##       list_variables(m$variable))
## }
## opt <- options("contrasts")[[1]][[1]]
## if (opt != "contr.treatment") {
##   w <- paste0("Your unordered-factor contrasts option is set to ",
##     opt, ". This may produce unexpected behavior, particularly in make_dummies in prep_data.
##     Consider resetting it by restarting R, or with: ",
##     "options(contrasts = c(\"contr.treatment\", \"contr.poly\"))")
##   warning(w)
## }
## if (!new_recipe) {
##   message("Prepping data based on provided recipe")
##   newvars <- setdiff(names(d), c(recipe$var_info$variable,
##     attr(recipe, "ignored_columns")))
##   if (length(newvars)) {
##     warning("These variables were not observed in training ",
##       "and will be ignored: ", list_variables(newvars))
##     ignored <- c(ignored, newvars)
##     d_ignore <- dplyr::bind_cols(d_ignore, dplyr::select(d,
##       !!newvars))
##   }
##   missing_vars <- setdiff(recipe$var_info$variable[recipe$var_info$role ==
##     "predictor"], names(d))
##   if (length(missing_vars))
##     warning("These variables were present in training but are missing or ignored here: ",
##       list_variables(missing_vars))
##   if (!is.null(recipe$steps)) {
##     if (attr(recipe$steps[[1]], "class")[1] == "step_nzv") {
##       if (length(recipe$steps[[1]]$removals)) {
##         missing_vars <- missing_vars[missing_vars ==
##           length(recipe$steps[[1]]$removals)]
##       }
##     }
##   }
##   if (length(missing_vars))
##     stop("These variables were present in training but are missing or ignored here: ",
##       list_variables(missing_vars))
##   newly_missing <- find_new_missingness(d, recipe)
##   if (length(newly_missing))
##     warning("The following variable(s) have missingness that was not present when recipe was

```

```

##           list_variables(newly_missing))
## outcome_var <- recipe$var_info$variable[recipe$var_info$role ==
##           "outcome"]
## if (length(outcome_var) && !outcome_var %in% names(d))
##   remove_outcome <- TRUE
## }
## else {
##   undeclared_ignores <- find_columns_to_ignore(d, c(rlang::quo_name(outcome),
##             ignored))
##   if (length(undeclared_ignores)) {
##     warning("The following variable(s) look a lot like identifiers: They are ",
##             "character-type and have a unique value on every row. They will ",
##             "be ignored: ", paste0(undeclared_ignores, collapse = ", "))
##     ignored <- c(ignored, undeclared_ignores)
##     d_ignore <- dplyr::bind_cols(d_ignore, d[, names(d) %in%
##               undeclared_ignores, drop = FALSE])
##     d <- d[, !names(d) %in% undeclared_ignores, drop = FALSE]
##   }
##   mes <- "Training new data prep recipe"
##   recipe <- recipes::recipe(d, ~.)
##   recipe$orig_data <- orig_data
##   if (!rlang::quo_is_missing(outcome)) {
##     outcome_name <- rlang::quo_name(outcome)
##     if (!outcome_name %in% names(d))
##       stop(paste(outcome_name, " not found in d."))
##     outcome_vec <- dplyr::pull(d, !!outcome)
##     if (is.logical(outcome_vec) || any(c("TRUE", "FALSE") %in%
##       outcome_vec))
##       stop("outcome looks logical. Please convert the outcome to character",
##           " with values other than TRUE and FALSE.")
##     if (any(is.na(outcome_vec)))
##       stop("Found NA values in the outcome column. Clean your data or ",
##           "remove these rows before training a model.")
##     suppressWarnings({
##       recipe <- recipes::update_role(recipe, !!outcome,
##         new_role = "outcome")
##     })
##     if (factor_outcome && all(outcome_vec %in% 0:1)) {
##       if (!is.numeric(outcome_vec))
##         stop("factor_outcome is TRUE, but ", outcome_name,
##             " is a character", "-type variable with 0s and 1s. Consider making it numeric wi
##             "`as.numeric(as.character())`")
##       recipe <- recipe %>% recipes::step_bin2factor(all_outcomes(),
##         levels = c("Y", "N"))
##     }
##     mes <- paste0(mes, "...\\n")
##   }
##   else {
##     mes <- paste0(mes, " with no outcome variable specified...\\n")
##   }
##   message(mes)
##   freq_cut <- 49
##   unique_cut <- 10
##   if (!is.logical(remove_near_zero_variance)) {

```

```

##         if (!is.numeric(remove_near_zero_variance))
##           stop("remove_near_zero_variance must be logical or numeric for step_nzv")
##         if (remove_near_zero_variance < 0 | remove_near_zero_variance >
##             1)
##           stop("remove_near_zero_variance must be numeric between 0 and 1")
##         freq_cut <- remove_near_zero_variance^-1
##         remove_near_zero_variance <- TRUE
##       }
##     if (remove_near_zero_variance) {
##       recipe <- recipe %>% recipes::step_nzv(all_predictors(),
##         freq_cut = freq_cut, unique_cut = unique_cut)
##     }
##     prep_check <- recipes::prep(recipe, training = d)
##     removing <- prep_check$steps[[1]]$removals
##     vi <- recipe$var_info
##     nom_preds <- vi$variable[vi$role == "predictor" & vi$type ==
##       "nominal"]
##     if (length(nom_preds) && all(nom_preds %in% removing))
##       stop("All your categorical columns will be removed because they have ",
##         "near-zero variance, which will break prep_data. ",
##         "Be less aggressive in removing near-zero variance columns by ",
##         "using a larger value of remove_near_zero_variance or setting it ",
##         "to FALSE.\n ", list_variables(removing))
##     if (!is.character(convert_dates)) {
##       if (!is.logical(convert_dates))
##         stop("convert_dates must be logical, \"none\", \"continuous\", or ",
##           "\"categories\"")
##       if (convert_dates)
##         convert_dates <- "continuous"
##       else convert_dates <- "none"
##     }
##     if (convert_dates %in% c("continuous", "categories")) {
##       cols <- find_date_cols(d)
##       if (!purrr::is_empty(cols)) {
##         recipe <- do.call(step_date_hcai, list(recipe = recipe,
##           cols, feature_type = convert_dates)) %>% recipes::step_rm(cols)
##       }
##     }
##     else if (convert_dates == "none") {
##       cols <- find_date_cols(d)
##       if (!purrr::is_empty(cols))
##         recipe <- recipes::step_rm(recipe, cols)
##     }
##     else {
##       stop("convert_dates must be logical, \"none\", \"continuous\", or ",
##         "\"categories\"")
##     }
##     if (isTRUE(impute)) {
##       recipe <- recipe %>% hcai_impute()
##     }
##     else if (is.list(impute)) {
##       ip <- list(numeric_method = "mean", nominal_method = "new_category",
##         numeric_params = NULL, nominal_params = NULL)
##       ip[names(ip) %in% names(impute)] <- impute[names(impute) %in%

```

```

##           names(ip)]
##     extras <- names(impute)[!(names(impute) %in% names(ip))]
##     if (length(extras > 0)) {
##       warning("You have extra imputation parameters that won't be used: ",
##             list_variables(extras), ". Available params are: ",
##             list_variables(names(ip)))
##     }
##     recipe <- recipe %>% hcai_impute(numeric_method = ip$numeric_method,
##           nominal_method = ip$nominal_method, numeric_params = ip$numeric_params,
##           nominal_params = ip$nominal_params)
##   }
##   else if (impute != FALSE) {
##     stop("impute must be boolean or list.")
##   }
##   if (!(is.numeric(PCA) || is.logical(PCA)))
##     stop("PCA must be logical or numeric")
##   if (as.logical(PCA)) {
##     if (!(as.logical(center) && as.logical(scale))) {
##       warning("\nd\" must be centered and scaled to perform PCA. Center and Scale are being
##       center <- as.logical(PCA)
##       scale <- as.logical(PCA)
##     }
##   }
## }
## var_info <- recipe$var_info
## if (any(var_info$type == "numeric" & var_info$role ==
##       "predictor")) {
##   if (isTRUE(as.logical(center))) {
##     recipe <- recipe %>% recipes::step_center(all_numeric(),
##           -all_outcomes())
##   }
##   if (isTRUE(as.logical(scale))) {
##     recipe <- recipe %>% recipes::step_scale(all_numeric(),
##           -all_outcomes())
##   }
## }
## if (any(var_info$type == "nominal" & var_info$role ==
##       "predictor")) {
##   if (add_levels)
##     recipe <- step_add_levels(recipe, all_nominal(),
##           -all_outcomes())
##   if (!is.logical(collapse_rare_factors)) {
##     if (!is.numeric(collapse_rare_factors))
##       stop("collapse_rare_factors must be logical or numeric")
##     if (collapse_rare_factors >= 1 || collapse_rare_factors <
##         0)
##       stop("If numeric, collapse_rare_factors should be between 0 and 1.")
##     fac_thresh <- collapse_rare_factors
##     collapse_rare_factors <- TRUE
##   }
##   if (collapse_rare_factors) {
##     if (!exists("fac_thresh"))
##       fac_thresh <- 0.03
##     recipe <- recipe %>% recipes::step_other(all_nominal(),
##           -all_outcomes(), threshold = fac_thresh)
##   }

```

```

##      }
##      if (add_levels)
##        recipe <- step_add_levels(recipe, all_nominal(),
##          -all_outcomes())
##      if (isTRUE(make_dummies)) {
##        make_dummies <- list()
##      }
##      if (is.list(make_dummies)) {
##        recipe <- recipe %>% step_dummy_hcai(all_nominal(),
##          -all_outcomes(), levels = make_dummies)
##      }
##      else if (!is.logical(make_dummies)) {
##        stop("step_dummies must be logical or list")
##      }
##    }
##    if (as.logical(PCA)) {
##      if (!impute && !is.list(impute))
##        stop("NAs present in \"d\". PCA not compatible when NAs are present.")
##      if (is.logical(PCA))
##        PCA <- 5
##      if (PCA > length(recipes::prep(recipe, training = d)$term_info$role ==
##        "predictor"))
##        stop("Can't have more components than columns in \"d\".")
##      recipe <- recipe %>% recipes::step_pca(all_numeric(),
##        -all_outcomes(), num_comp = as.integer(PCA))
##    }
##    recipe <- recipes::prep(recipe, training = d)
##    attr(recipe, "missingness") <- d_missing
##    attr(recipe, "factor_levels") <- d_levels
##  }
##  if (logical_to_numeric)
##    d <- dplyr::mutate_if(d, is.logical, as.numeric)
##  d <- recipes::bake(recipe, d)
##  steps <- map_chr(recipe$steps, ~attr(.x, "class")[1])
##  if ("step_nzv" %in% steps && length(nzv_removed <- recipe$steps[[which(steps ==
##    "step_nzv")]]$removals))
##    message("Removing the following ", length(nzv_removed),
##      " near-zero variance column(s). ", "If you don't want to remove them, call prep_data with
##      \"remove_near_zero_variance as a smaller numeric or FALSE.\n ",
##      list_variables(nzv_removed))
##  if (remove_outcome && outcome_var %in% names(d))
##    d <- select_not(d, outcome_var)
##  if (rlang::quo_name(outcome) %in% names(d_ods))
##    d_ods <- select_not(d_ods, outcome)
##  d <- dplyr::bind_cols(d_ignore, d)
##  if (new_recipe)
##    recipe$template <- dplyr::bind_cols(d_ignore, recipe$template)
##  attr(recipe, "ignored_columns") <- unname(ignored)
##  attr(recipe, "no_prep") <- no_prep
##  attr(d, "recipe") <- recipe
##  attr(d, "best_levels") <- best_levels
##  attr(d, "original_data_str") <- d_ods
##  d <- tibble::as_tibble(d)
##  class(d) <- c("prepped_df", class(d))

```



```

##   return(d)
## }
## <bytecode: 0x7fc5d863c8e0>
## <environment: namespace:healthcareai>

split_data <- split_train_test(d = data,
                              outcome = dead,
                              p = .9,
                              seed = 84105)
prepped_training_data <- prep_data(split_data$train, outcome = dead,
                                   center = TRUE, scale = TRUE,
                                   collapse_rare_factors = FALSE)

## Training new data prep recipe...

## Removing the following 23 near-zero variance column(s). If you don't want to remove them, call prep_recipe()
##   obleed, ocva, oleginf, otia, orf, opneu, rf_1, priority_22, rf_r, rf_1_r, priority_22_r, aortic_injury,
model training
models <- tune_models(d = prepped_training_data,
                     outcome = dead,
                     tune_depth = 25,
                     metric = "PR")

##
## dead looks categorical, so training classification algorithms.

##
## After data processing, models are being trained on 236 features with 1,486 observations.
## Based on n_folds = 5 and hyperparameter settings, the following number of models will be trained: 120
## Model training may take a few minutes.

## Training with cross validation: Random Forest

## You may, or may not, see messages about progress in growing trees. The estimates are very rough, and
## Training with cross validation: eXtreme Gradient Boosting

## Training with cross validation: glmnet

##
## *** Models successfully trained. The model object contains the training data minus ignored ID columns.
## *** If there was PHI in training data, normal PHI protocols apply to the model object. ***

evaluate(models, all_models = TRUE)

## # A tibble: 3 x 3
##   model                AUPR AUROC
##   <chr>                <dbl> <dbl>
## 1 glmnet                0.386 0.777
## 2 eXtreme Gradient Boosting 0.382 0.786
## 3 Random Forest          0.364 0.761

models["Random Forest"] %>%
  plot()

## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.

## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please

```

```
## use `guide = "none"` instead.
```

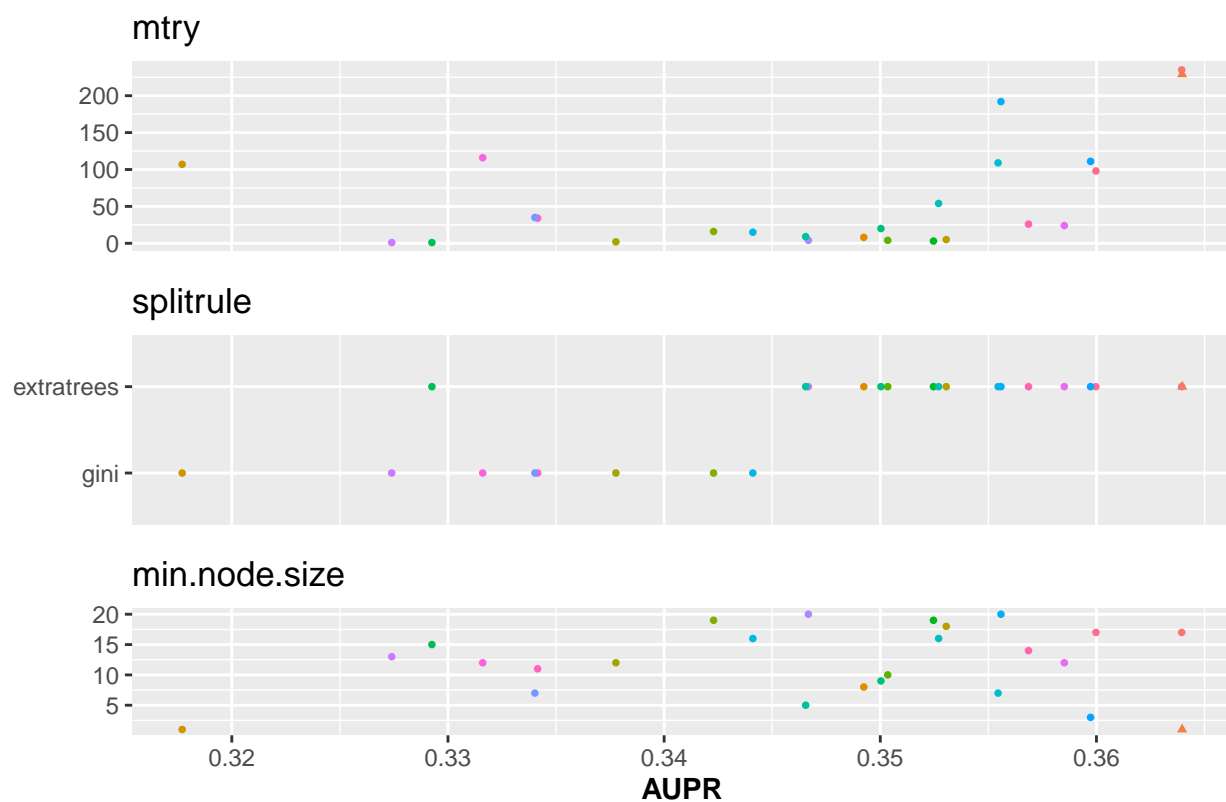
```
## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.
```

```
## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.
```

```
## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.
```

```
## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.
```

Random Forest



faster model training

```
untuned_rf <- flash_models(d = prepped_training_data,
                           outcome = dead,
                           models = "RF",
                           metric = "PR")
```

```
##
```

```
## dead looks categorical, so training classification algorithms.
```

```
##
```

```
## After data processing, models are being trained on 236 features with 1,486 observations.
```

```
## Based on n_folds = 5 and hyperparameter settings, the following number of models will be trained: 5
```

```
## Training at fixed values: Random Forest
```

```
## You may, or may not, see messages about progress in growing trees. The estimates are very rough, and
##
## *** Models successfully trained. The model object contains the training data minus ignored ID column.
## *** If there was PHI in training data, normal PHI protocols apply to the model object. ***
```

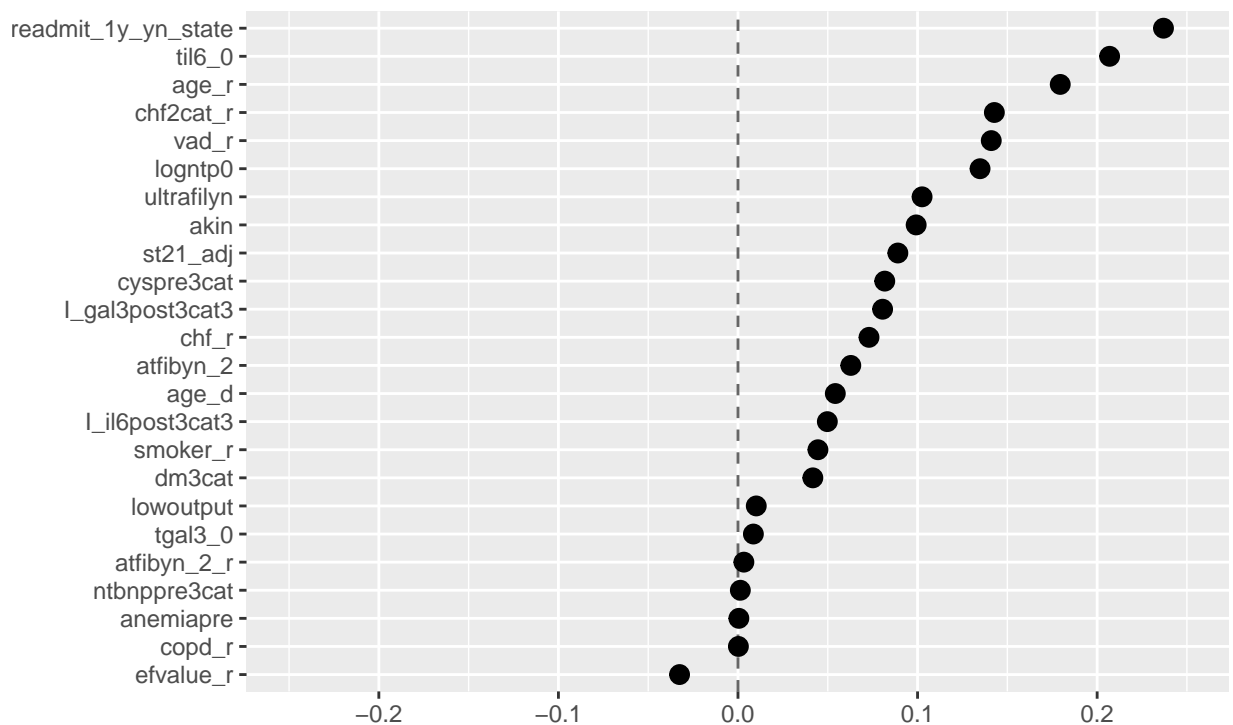
```
untuned_rf
```

```
## Algorithms Trained: Random Forest
## Model Name: dead
## Target: dead
## Class: Classification
## Performance Metric: AUPR
## Number of Observations: 1486
## Number of Features: 236
## Models Trained: 2021-09-05 10:24:14
##
## Models have not been tuned. Performance estimated via 5-fold cross validation at fixed hyperparameters.
## Best model: Random Forest
## AUPR = 0.33, AUROC = 0.75
## User-selected hyperparameter values:
##   mtry = 15
##   splitrule = extratrees
##   min.node.size = 1
```

model interpretation

```
interpret(models) %>%
  plot()
```

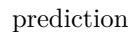
Coefficients for regularized classification model of dead



Hyperparameter values $\alpha = 1$ and $\lambda = 0.0183$

```
get_variable_importance(models) %>%  
  plot()
```

eXtreme Gradient Boosting variable importance



```
## "predicted_dead" predicted by glmnet last trained: 2021-09-05 10:24:09
## Performance in training: AUPR = 0.39

## # A tibble: 1,486 x 261
##   dead predicted_dead readmit30d_yn_st~ oblead ocva oleginf otia orf opneu
## * <fct>          <dbl>          <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 N              0.188              1      0      0      0      0      0      0
## 2 N              0.0960             1      0      0      0      0      0      0
## 3 N              0.0860             0      0      0      0      0      0      0
## 4 N              0.0935             0      0      0      0      0      0      0
## 5 N              0.156              0      0      0      0      0      0      0
## 6 N              0.0432             0      0      0      0      0      0      0
## 7 N              0.142              1      0      0      0      0      0      0
## 8 N              0.0530             0      0      0      0      0      0      0
## 9 N              0.117              0      0      1      0      0      0      0
```

```
## 10 N          0.110          0      0      0      0      0      0      0
## # ... with 1,476 more rows, and 252 more variables: oafib2 <dbl>, return <dbl>,
## #   akin <dbl>, lm2cat <dbl>, rfcr_2 <dbl>, iabppre_2 <dbl>, atfibyn_2 <dbl>,
## #   hyperyn_2 <dbl>, ef50_neg <dbl>, ef50_neg_d <dbl>, priormi_21 <dbl>,
## #   priormi_22 <dbl>, priormi_23 <dbl>, priormi_24 <dbl>, age_d <dbl>,
## #   ua_nmi7 <dbl>, creatpre_2 <dbl>, rf_1 <dbl>, rf_2 <dbl>, rf_3 <dbl>,
## #   priority_21 <dbl>, priority_22 <dbl>, dm3cat <dbl>, dm3cat_21 <dbl>,
## #   dm3cat_22 <dbl>, dm3cat_21_2 <dbl>, dm3cat_22_2 <dbl>, pci_ta <dbl>, ...
```

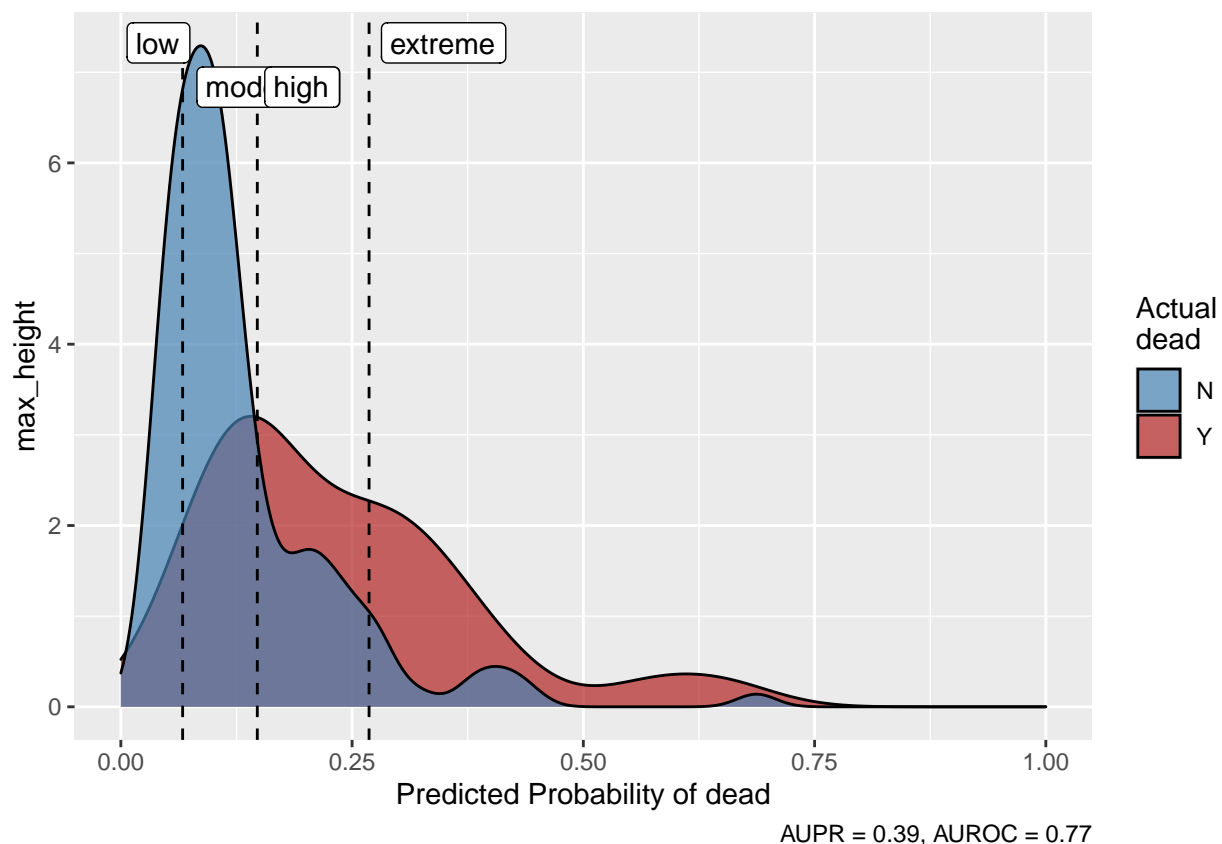
```
test_predictions <-
  predict(models,
    split_data$test,
    risk_groups = c(low = 30, moderate = 40, high = 20, extreme = 10)
  )
```

Prepping data based on provided recipe

Removing the following 23 near-zero variance column(s). If you don't want to remove them, call prep_
obleed, ocva, oleginf, otia, orf, opneu, rf_1, priority_22, rf_r, rf_1_r, priority_22_r, aortic_in

Prepping data based on provided recipe

```
plot(test_predictions)
```



Saving, Moving, and Loading Models

```
save_models(models, file = "my_models.RDS")
models <- load_models("my_models.RDS")
```

regression models

```

regression_models <- machine_learn(data, outcome = dead)

## Training new data prep recipe...

## Removing the following 25 near-zero variance column(s). If you don't want to remove them, call prep_recipe()
##   obleed, ocva, oleginf, otia, orf, opneu, rf_1, rf_3, priority_22, rf_r, rf_1_r, rf_3_r, priority_22_r
##
## dead looks categorical, so training classification algorithms.

##
## After data processing, models are being trained on 234 features with 1,651 observations.
## Based on n_folds = 5 and hyperparameter settings, the following number of models will be trained: 50

## Training with cross validation: Random Forest

## You may, or may not, see messages about progress in growing trees. The estimates are very rough, and
## Training with cross validation: eXtreme Gradient Boosting

## Training with cross validation: glmnet

##
## *** Models successfully trained. The model object contains the training data minus ignored ID columns. ***
## *** If there was PHI in training data, normal PHI protocols apply to the model object. ***

summary(regression_models)

## Models trained: 2021-09-05 10:27:10
##
## Models tuned via 5-fold cross validation over 10 combinations of hyperparameter values.
## Best performance: AUPR = 0.37, AUROC = 0.78
## By glmnet with hyperparameters:
##   alpha = 1
##   lambda = 0.011
##
## Out-of-fold performance of all trained models:
##
## $`Random Forest`
## # A tibble: 10 x 9
##   mtry splitrule min.node.size AUROC   Sens Spec ROCSD SensSD SpecSD
##   <int> <chr>      <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1    37 extratrees      11 0.770 0.0701 0.994 0.0291 0.0180 0.00293
## 2    51 extratrees      19 0.768 0.0702 0.994 0.0265 0.0286 0.00314
## 3     7 extratrees       3 0.764 0.0131 0.998 0.0302 0.0120 0.00192
## 4     1 extratrees       2 0.763 0      1      0.0390 0      0
## 5     2 gini            1 0.759 0.00435 0.999 0.0397 0.00972 0.00314
## 6    12 gini           12 0.754 0.0263 0.994 0.0301 0.0240 0.00675
## 7    18 gini           11 0.752 0.0350 0.993 0.0305 0.0249 0.00823
## 8   193 gini           12 0.740 0.0789 0.984 0.0344 0.0252 0.0128
## 9   231 gini           20 0.739 0.0833 0.986 0.0351 0.0180 0.00744
## 10  119 gini           15 0.739 0.0833 0.984 0.0374 0.0286 0.00845
##
## $`eXtreme Gradient Boosting`
## # A tibble: 10 x 13
##   eta max_depth gamma colsample_bytree min_child_weight subsample nrounds
##   <dbl> <int> <dbl> <dbl> <dbl> <dbl> <int>
## 1 0.213      1 4.83      0.784      2.30      0.901     284

```

```

## 2 0.316      8 9.84      0.687      6.45      0.776      60
## 3 0.320      3 6.59      0.505      9.91      0.878      982
## 4 0.123      5 3.54      0.596      1.63      0.623      407
## 5 0.167      8 4.36      0.510     12.3      0.899      463
## 6 0.485      2 9.78      0.769      6.59      0.573      66
## 7 0.357      7 5.54      0.611      2.71      0.812      52
## 8 0.268      4 4.96      0.835     12.5      0.646     141
## 9 0.364      4 8.21      0.526      2.48      0.561     147
## 10 0.465      4 6.05      0.557      0.162     0.499     672
## # ... with 6 more variables: AUROC <dbl>, Sens <dbl>, Spec <dbl>, ROCSD <dbl>,
## #   SensSD <dbl>, SpecSD <dbl>
##
## $glmnet
## # A tibble: 20 x 8
##   alpha lambda AUROC   Sens Spec ROCSD SensSD SpecSD
##   <dbl>   <dbl> <dbl>   <dbl> <dbl> <dbl>   <dbl>   <dbl>
## 1     1 0.0111 0.782 0.114 0.989 0.0329 0.0461 0.00627
## 2     1 0.0229 0.779 0.0352 0.995 0.0399 0.0199 0.00401
## 3     0 1.17    0.770 0.0175 0.997 0.0420 0.0182 0.00457
## 4     1 0.00695 0.770 0.132 0.985 0.0300 0.0449 0.00880
## 5     0 0.297   0.768 0.0924 0.992 0.0327 0.0335 0.00520
## 6     0 3.73    0.767 0      1    0.0458 0      0
## 7     1 0.0384 0.764 0      1    0.0453 0      0
## 8     1 0.00555 0.763 0.149 0.982 0.0294 0.0294 0.00859
## 9     0 0.102   0.756 0.136 0.980 0.0238 0.0339 0.0125
## 10    0 0.0384 0.738 0.162 0.968 0.0212 0.0339 0.00893
## 11    0 0.0229 0.731 0.18  0.963 0.0214 0.0401 0.00918
## 12    1 0.00187 0.726 0.171 0.960 0.0213 0.0580 0.0128
## 13    0 0.0111 0.720 0.193 0.962 0.0216 0.0622 0.0109
## 14    0 0.00695 0.715 0.189 0.960 0.0216 0.0682 0.00948
## 15    0 0.00555 0.715 0.189 0.960 0.0216 0.0682 0.00948
## 16    0 0.00187 0.715 0.189 0.960 0.0216 0.0682 0.00948
## 17    1 3.73    0.5  0      1    0      0      0
## 18    1 1.17    0.5  0      1    0      0      0
## 19    1 0.297   0.5  0      1    0      0      0
## 20    1 0.102   0.5  0      1    0      0      0

```