Finance Data

forecast

4/9/2020

R Markdown

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```
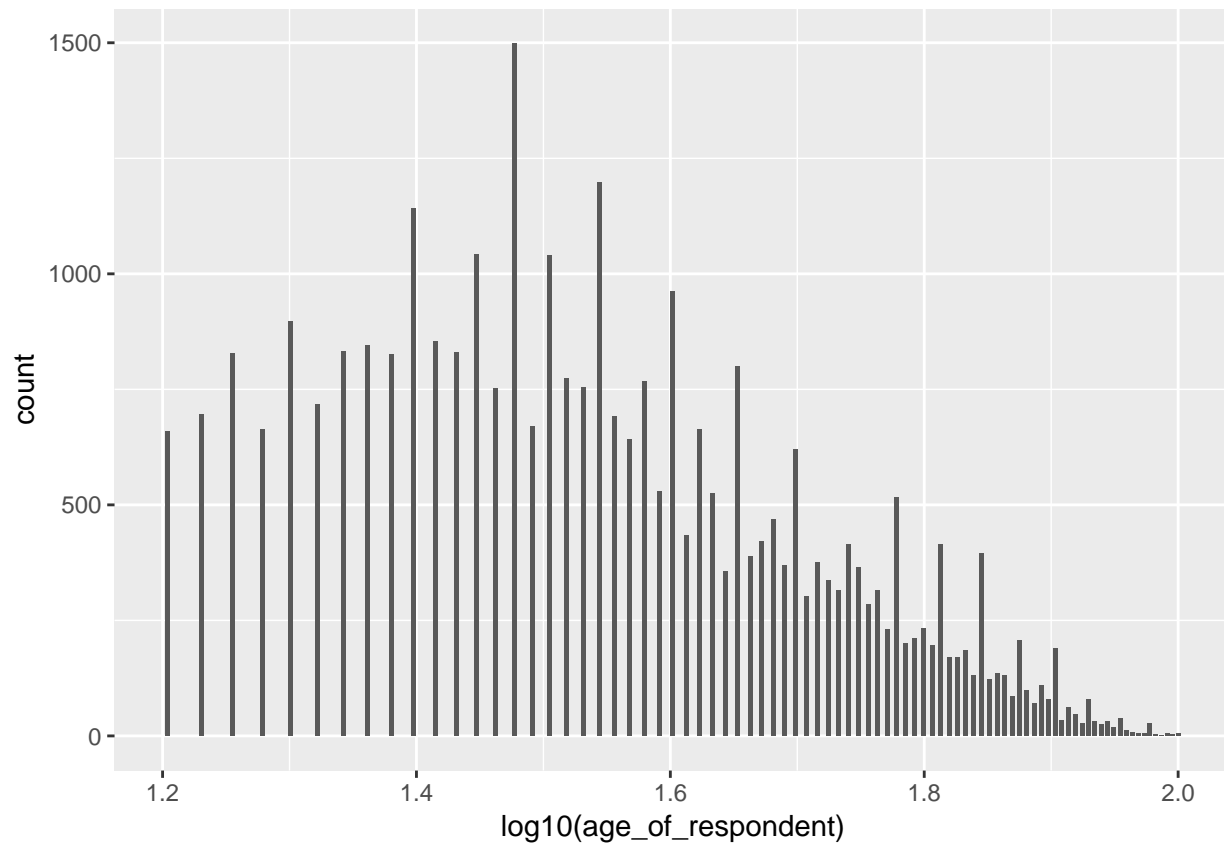
```r
library(stringr)
library(ggplot2)
fintrain <- read.csv('Train_v2.csv')
fintest <- read.csv('Test_v2.csv')
fintest['bank_account'] <- NA #rbind only works with equal ncols() and same colnames()
finfull <- rbind(fintrain,fintest) #fulldataset for manipulation if needed
```

```r
finfull %>% ggplot(aes(x = log10( age_of_respondent))) + geom_bar()
```

```
#right skew of age, need log transformation to make it symmetric or we can categorize them for our mode

groupages <- c(15,65,101)

finfull['age_groups'] <- cut(finfull$age_of_respondent,
                        breaks = groupages,
                        labels = c('Adult', 'Senior')) #age now binned into ordinal category

finfull$age_groups <- factor(finfull$age_groups, ordered = TRUE)

finfull <- finfull[-c(8)]

#household size can be binned into small, medium, big, #kind of subjective but worth trying

finfull['Household'] <- cut(finfull$household_size,4,
                        labels = c('S', 'M', 'L', 'XL'))

finfull <- finfull[-c(7)]

finfull$Household <- factor(finfull$Household, ordered = TRUE)

finfull$uniqueid <- as.character(finfull$uniqueid)

finfull$year <- factor(finfull$year, ordered = TRUE)
#back to train/test
```

```
finfull$relationship_with_head <- factor(finfull$relationship_with_head,
                                         ordered = TRUE)

finfull$education_level <- factor(finfull$education_level,
                                  ordered = TRUE)

findata <- finfull[1:23524, ]
fintest <- finfull[23525:33610, ]
```

```
#visualize country

findata %>% group_by(country, bank_account) %>%
  summarise(n = n()) %>%
  mutate(freq = n/ sum(n)) %>%
  arrange() %>%
  filter(bank_account == 'No')
```

```
## # A tibble: 4 x 4
## # Groups:   country [4]
##   country  bank_account     n  freq
##   <fct>    <fct>        <int> <dbl>
## 1 Kenya    No            4547 0.749
## 2 Rwanda   No            7732 0.885
## 3 Tanzania No            6013 0.908
## 4 Uganda   No            1920 0.914
```

```
#Kenya has the least proportion of people without bank account

findata %>% group_by(country, bank_account) %>%
  summarise(n = n()) %>%
  mutate(freq = n/ sum(n)) %>% arrange() %>%
  filter(bank_account == 'Yes')
```

```
## # A tibble: 4 x 4
## # Groups:   country [4]
##   country  bank_account     n   freq
##   <fct>    <fct>        <int>  <dbl>
## 1 Kenya    Yes           1521 0.251
## 2 Rwanda   Yes           1003 0.115
## 3 Tanzania Yes            607 0.0917
## 4 Uganda   Yes            181 0.0861
```
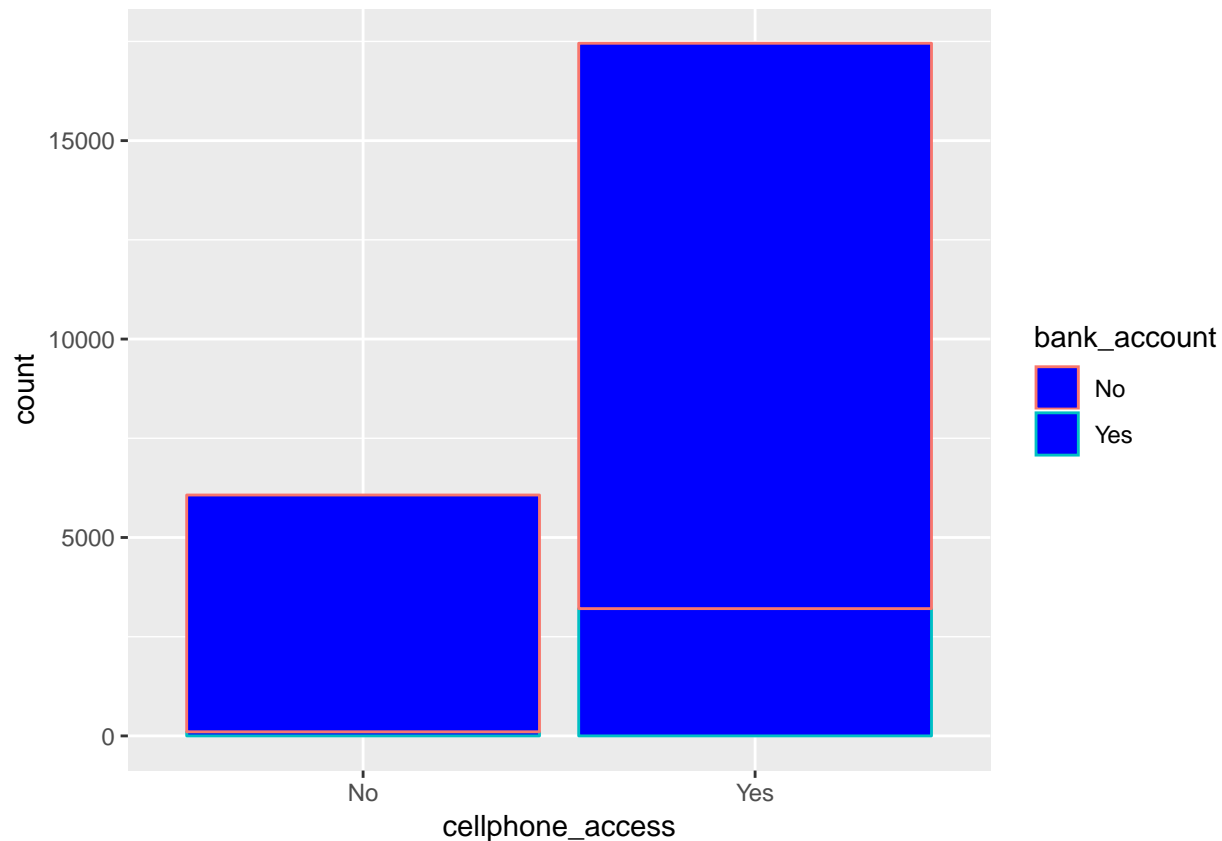
```
#Kenya has the biggest proportion of people with people with a bank account
```

```
# people with cellphone access have higher probability of opening a bank account?
findata %>% ggplot(aes(x = cellphone_access, color = bank_account)) +
  geom_bar(fill = 'blue')
```

```r
#household size doesnt seem to have a difference!
findata %>% group_by(Household, bank_account) %>%
  summarise(total = n()) %>%
  mutate(freq = total/ sum(total))
```

```
## # A tibble: 8 x 4
## # Groups:   Household [4]
##   Household bank_account total  freq
##   <ord>     <fct>        <int> <dbl>
## 1 S         No           17662 0.856
## 2 S         Yes           2967 0.144
## 3 M         No            2442 0.881
## 4 M         Yes            331 0.119
## 5 L         No             105 0.890
## 6 L         Yes             13 0.110
## 7 XL        No               3 0.75
## 8 XL        Yes              1 0.25
```

RANDOM FOREST

```r
#XGBOOST
library(xgboost)
```

```
##
## Attaching package: 'xgboost'
```

```
## The following object is masked from 'package:dplyr':
##
##     slice

findata <- data.frame(findata)

input <- data.matrix(findata[-c(3,4)]) #xgboost only works with a matrix

findata$bank_account <- ifelse(findata$bank_account == 'Yes',
                               1, 0)
```

```
classifier <- xgboost(data = input ,
              label = findata$bank_account,
              nrounds = 81,
              params = list(
                objective = 'binary:logistic',
                eta = 0.4,
                max_depth = 150,
                gamma = 0,
                colsample_bytree = 0.6,
                min_child_weight = 1,
                subsample = 1

              ))
```

```
## [1]   train-error:0.121620
## [2]   train-error:0.111886
## [3]   train-error:0.113161
## [4]   train-error:0.112098
## [5]   train-error:0.113586
## [6]   train-error:0.109633
## [7]   train-error:0.108527
## [8]   train-error:0.108017
## [9]   train-error:0.107635
## [10]  train-error:0.107422
## [11]  train-error:0.106657
## [12]  train-error:0.106062
## [13]  train-error:0.105212
## [14]  train-error:0.104149
## [15]  train-error:0.104021
## [16]  train-error:0.103596
## [17]  train-error:0.103384
## [18]  train-error:0.102916
## [19]  train-error:0.102746
## [20]  train-error:0.102746
## [21]  train-error:0.102491
## [22]  train-error:0.102449
## [23]  train-error:0.102364
## [24]  train-error:0.102321
## [25]  train-error:0.101683
## [26]  train-error:0.101428
## [27]  train-error:0.101258
## [28]  train-error:0.101343
```

```
## [29] train-error:0.101343
## [30] train-error:0.101131
## [31] train-error:0.101088
## [32] train-error:0.101003
## [33] train-error:0.100748
## [34] train-error:0.100748
## [35] train-error:0.100323
## [36] train-error:0.100323
## [37] train-error:0.100323
## [38] train-error:0.100111
## [39] train-error:0.099218
## [40] train-error:0.099133
## [41] train-error:0.099260
## [42] train-error:0.099090
## [43] train-error:0.098750
## [44] train-error:0.098750
## [45] train-error:0.098793
## [46] train-error:0.098538
## [47] train-error:0.098623
## [48] train-error:0.098538
## [49] train-error:0.098453
## [50] train-error:0.098198
## [51] train-error:0.098155
## [52] train-error:0.098198
## [53] train-error:0.097560
## [54] train-error:0.097475
## [55] train-error:0.097517
## [56] train-error:0.097645
## [57] train-error:0.097645
## [58] train-error:0.097602
## [59] train-error:0.097220
## [60] train-error:0.097220
## [61] train-error:0.097177
## [62] train-error:0.096540
## [63] train-error:0.096200
## [64] train-error:0.096115
## [65] train-error:0.095775
## [66] train-error:0.095690
## [67] train-error:0.095604
## [68] train-error:0.095562
## [69] train-error:0.095519
## [70] train-error:0.095519
## [71] train-error:0.095094
## [72] train-error:0.095052
## [73] train-error:0.094924
## [74] train-error:0.094669
## [75] train-error:0.094542
## [76] train-error:0.094457
## [77] train-error:0.094329
## [78] train-error:0.094457
## [79] train-error:0.094499
## [80] train-error:0.094372
## [81] train-error:0.094414
```

```r
#25 rounds best submission
```

```r
test <- fintest


test <- test[-c(3,4)]

test <- data.frame(test)

test <- data.matrix(test)



y_pred <- predict(classifier, test)
fintest['pred'] = y_pred

fintest <- fintest[-c(4)]

#pred <- predict(classifier,  newdata = fintest[-c(2,3)])
fintest$pred <- ifelse(fintest$pred < 0.5,
                       0,
                       1)
```

```r
#Applying K-Fold validation:
library(caret)
```

```r
## Loading required package: lattice
```

```r
folds = createFolds(findata$bank_account, k = 10)
#folds list
#applying function to list with lapply


cv = lapply(folds, function(x) {
  training_fold = input #training set of features
  test_fold = input
  classifier <- xgboost(data = training_fold,
             label = findata$bank_account,
             nrounds = 81,
             params = list(
               objective = 'binary:logistic',
               eta = 0.4,
               max_depth = 150,
               gamma = 0,
               colsample_bytree = 0.6,
               min_child_weight = 1,
               subsample = 1



           ))

y_pred2 = predict(classifier, newdata = input)
```

```r
y_pred2 = ifelse(y_pred2 < 0.5, 0, 1)

  #remember cross validation is done on the training-set
  cm = table( c(y_pred2), findata$bank_account)
  accuracy = (cm[1,1] + cm[2,2]) / (cm[1,1] + cm[2,2] + cm[1,2] + cm[2,1])
  return(accuracy)

})
```

```
## [1]  train-error:0.111971
## [2]  train-error:0.113204
## [3]  train-error:0.110780
## [4]  train-error:0.111121
## [5]  train-error:0.108868
## [6]  train-error:0.108910
## [7]  train-error:0.108953
## [8]  train-error:0.107762
## [9]  train-error:0.106700
## [10] train-error:0.106104
## [11] train-error:0.106019
## [12] train-error:0.105849
## [13] train-error:0.104829
## [14] train-error:0.103299
## [15] train-error:0.102619
## [16] train-error:0.101981
## [17] train-error:0.101726
## [18] train-error:0.101513
## [19] train-error:0.101343
## [20] train-error:0.101258
## [21] train-error:0.101216
## [22] train-error:0.101131
## [23] train-error:0.101088
## [24] train-error:0.101046
## [25] train-error:0.100918
## [26] train-error:0.099728
## [27] train-error:0.099898
## [28] train-error:0.100026
## [29] train-error:0.099515
## [30] train-error:0.098878
## [31] train-error:0.098793
## [32] train-error:0.098623
## [33] train-error:0.098198
## [34] train-error:0.098198
## [35] train-error:0.098028
## [36] train-error:0.097730
## [37] train-error:0.097772
## [38] train-error:0.097262
## [39] train-error:0.097135
## [40] train-error:0.097305
## [41] train-error:0.097262
## [42] train-error:0.097135
## [43] train-error:0.097092
## [44] train-error:0.097050
```

```
## [45] train-error:0.096667
## [46] train-error:0.096582
## [47] train-error:0.096327
## [48] train-error:0.096115
## [49] train-error:0.096200
## [50] train-error:0.096200
## [51] train-error:0.095860
## [52] train-error:0.095945
## [53] train-error:0.095860
## [54] train-error:0.095434
## [55] train-error:0.095477
## [56] train-error:0.095349
## [57] train-error:0.095349
## [58] train-error:0.095349
## [59] train-error:0.095137
## [60] train-error:0.094797
## [61] train-error:0.094542
## [62] train-error:0.094159
## [63] train-error:0.094074
## [64] train-error:0.094117
## [65] train-error:0.094117
## [66] train-error:0.094032
## [67] train-error:0.093947
## [68] train-error:0.093904
## [69] train-error:0.093904
## [70] train-error:0.093947
## [71] train-error:0.093819
## [72] train-error:0.093436
## [73] train-error:0.093436
## [74] train-error:0.093436
## [75] train-error:0.093394
## [76] train-error:0.093309
## [77] train-error:0.093351
## [78] train-error:0.093266
## [79] train-error:0.093351
## [80] train-error:0.093309
## [81] train-error:0.093309
## [1]  train-error:0.109293
## [2]  train-error:0.109080
## [3]  train-error:0.110908
## [4]  train-error:0.110823
## [5]  train-error:0.108230
## [6]  train-error:0.106997
## [7]  train-error:0.106870
## [8]  train-error:0.105892
## [9]  train-error:0.105722
## [10] train-error:0.104574
## [11] train-error:0.104404
## [12] train-error:0.104191
## [13] train-error:0.103979
## [14] train-error:0.103979
## [15] train-error:0.103809
## [16] train-error:0.103001
## [17] train-error:0.102108
```

```
## [18] train-error:0.102066
## [19] train-error:0.101726
## [20] train-error:0.101258
## [21] train-error:0.101003
## [22] train-error:0.101003
## [23] train-error:0.100706
## [24] train-error:0.100748
## [25] train-error:0.100408
## [26] train-error:0.100366
## [27] train-error:0.099983
## [28] train-error:0.099728
## [29] train-error:0.099515
## [30] train-error:0.099473
## [31] train-error:0.099345
## [32] train-error:0.098750
## [33] train-error:0.098495
## [34] train-error:0.098708
## [35] train-error:0.098368
## [36] train-error:0.098113
## [37] train-error:0.098198
## [38] train-error:0.098113
## [39] train-error:0.098113
## [40] train-error:0.097985
## [41] train-error:0.097730
## [42] train-error:0.097220
## [43] train-error:0.097177
## [44] train-error:0.097007
## [45] train-error:0.096710
## [46] train-error:0.096795
## [47] train-error:0.096327
## [48] train-error:0.096370
## [49] train-error:0.096327
## [50] train-error:0.096285
## [51] train-error:0.096200
## [52] train-error:0.096157
## [53] train-error:0.095945
## [54] train-error:0.095690
## [55] train-error:0.095519
## [56] train-error:0.095817
## [57] train-error:0.095987
## [58] train-error:0.095775
## [59] train-error:0.095604
## [60] train-error:0.095477
## [61] train-error:0.095052
## [62] train-error:0.095094
## [63] train-error:0.095094
## [64] train-error:0.095052
## [65] train-error:0.095009
## [66] train-error:0.095009
## [67] train-error:0.094712
## [68] train-error:0.094712
## [69] train-error:0.094499
## [70] train-error:0.094499
## [71] train-error:0.094329
```

```
## [72] train-error:0.094287
## [73] train-error:0.094159
## [74] train-error:0.093989
## [75] train-error:0.093989
## [76] train-error:0.093734
## [77] train-error:0.093777
## [78] train-error:0.093607
## [79] train-error:0.093607
## [80] train-error:0.093351
## [81] train-error:0.093181
## [1]  train-error:0.116519
## [2]  train-error:0.125106
## [3]  train-error:0.118730
## [4]  train-error:0.118772
## [5]  train-error:0.110228
## [6]  train-error:0.107465
## [7]  train-error:0.105212
## [8]  train-error:0.104829
## [9]  train-error:0.104021
## [10] train-error:0.103596
## [11] train-error:0.103384
## [12] train-error:0.102491
## [13] train-error:0.102321
## [14] train-error:0.102279
## [15] train-error:0.101981
## [16] train-error:0.102236
## [17] train-error:0.102236
## [18] train-error:0.101853
## [19] train-error:0.101726
## [20] train-error:0.100876
## [21] train-error:0.100281
## [22] train-error:0.100323
## [23] train-error:0.100111
## [24] train-error:0.099855
## [25] train-error:0.099643
## [26] train-error:0.099558
## [27] train-error:0.099090
## [28] train-error:0.099430
## [29] train-error:0.098708
## [30] train-error:0.098538
## [31] train-error:0.098538
## [32] train-error:0.098028
## [33] train-error:0.097858
## [34] train-error:0.097517
## [35] train-error:0.097560
## [36] train-error:0.097347
## [37] train-error:0.097177
## [38] train-error:0.096752
## [39] train-error:0.096412
## [40] train-error:0.096412
## [41] train-error:0.096370
## [42] train-error:0.096200
## [43] train-error:0.096242
## [44] train-error:0.096242
```

```
## [45] train-error:0.096200
## [46] train-error:0.095860
## [47] train-error:0.095817
## [48] train-error:0.095817
## [49] train-error:0.095775
## [50] train-error:0.095775
## [51] train-error:0.095732
## [52] train-error:0.095519
## [53] train-error:0.095434
## [54] train-error:0.095179
## [55] train-error:0.095264
## [56] train-error:0.095179
## [57] train-error:0.095052
## [58] train-error:0.095222
## [59] train-error:0.095179
## [60] train-error:0.095137
## [61] train-error:0.095052
## [62] train-error:0.094924
## [63] train-error:0.094414
## [64] train-error:0.094202
## [65] train-error:0.094159
## [66] train-error:0.093904
## [67] train-error:0.093904
## [68] train-error:0.093862
## [69] train-error:0.093862
## [70] train-error:0.093862
## [71] train-error:0.093649
## [72] train-error:0.093734
## [73] train-error:0.093436
## [74] train-error:0.093436
## [75] train-error:0.093436
## [76] train-error:0.093479
## [77] train-error:0.093309
## [78] train-error:0.093351
## [79] train-error:0.093224
## [80] train-error:0.093224
## [81] train-error:0.093181
## [1]  train-error:0.140665
## [2]  train-error:0.125064
## [3]  train-error:0.127189
## [4]  train-error:0.119155
## [5]  train-error:0.112226
## [6]  train-error:0.110568
## [7]  train-error:0.110143
## [8]  train-error:0.108442
## [9]  train-error:0.107295
## [10] train-error:0.105679
## [11] train-error:0.105254
## [12] train-error:0.104914
## [13] train-error:0.104999
## [14] train-error:0.104829
## [15] train-error:0.103809
## [16] train-error:0.103809
## [17] train-error:0.103044
```

```
## [18] train-error:0.102576
## [19] train-error:0.102279
## [20] train-error:0.102279
## [21] train-error:0.102066
## [22] train-error:0.101046
## [23] train-error:0.100493
## [24] train-error:0.100323
## [25] train-error:0.100196
## [26] train-error:0.099473
## [27] train-error:0.099558
## [28] train-error:0.099090
## [29] train-error:0.098793
## [30] train-error:0.098580
## [31] train-error:0.098538
## [32] train-error:0.098155
## [33] train-error:0.098070
## [34] train-error:0.097772
## [35] train-error:0.097602
## [36] train-error:0.097347
## [37] train-error:0.097092
## [38] train-error:0.097177
## [39] train-error:0.097220
## [40] train-error:0.097092
## [41] train-error:0.097050
## [42] train-error:0.097050
## [43] train-error:0.096922
## [44] train-error:0.096880
## [45] train-error:0.096497
## [46] train-error:0.096540
## [47] train-error:0.096370
## [48] train-error:0.096370
## [49] train-error:0.096157
## [50] train-error:0.096030
## [51] train-error:0.095987
## [52] train-error:0.095987
## [53] train-error:0.095817
## [54] train-error:0.095987
## [55] train-error:0.095860
## [56] train-error:0.095647
## [57] train-error:0.095562
## [58] train-error:0.095477
## [59] train-error:0.095477
## [60] train-error:0.095392
## [61] train-error:0.095179
## [62] train-error:0.095179
## [63] train-error:0.095179
## [64] train-error:0.095137
## [65] train-error:0.095094
## [66] train-error:0.095094
## [67] train-error:0.095052
## [68] train-error:0.095094
## [69] train-error:0.095179
## [70] train-error:0.095264
## [71] train-error:0.095094
```

```
## [72] train-error:0.095052
## [73] train-error:0.094542
## [74] train-error:0.094499
## [75] train-error:0.094542
## [76] train-error:0.094499
## [77] train-error:0.094457
## [78] train-error:0.094372
## [79] train-error:0.094329
## [80] train-error:0.094329
## [81] train-error:0.094287
## [1]  train-error:0.122556
## [2]  train-error:0.110143
## [3]  train-error:0.111673
## [4]  train-error:0.113969
## [5]  train-error:0.111886
## [6]  train-error:0.109208
## [7]  train-error:0.108187
## [8]  train-error:0.107720
## [9]  train-error:0.107720
## [10] train-error:0.107210
## [11] train-error:0.107040
## [12] train-error:0.106615
## [13] train-error:0.106232
## [14] train-error:0.104447
## [15] train-error:0.103809
## [16] train-error:0.103639
## [17] train-error:0.103384
## [18] train-error:0.102321
## [19] train-error:0.101811
## [20] train-error:0.101768
## [21] train-error:0.101641
## [22] train-error:0.101301
## [23] train-error:0.100748
## [24] train-error:0.100876
## [25] train-error:0.100621
## [26] train-error:0.100451
## [27] train-error:0.099855
## [28] train-error:0.099558
## [29] train-error:0.099303
## [30] train-error:0.099090
## [31] train-error:0.098750
## [32] train-error:0.098623
## [33] train-error:0.098623
## [34] train-error:0.098453
## [35] train-error:0.098453
## [36] train-error:0.098113
## [37] train-error:0.097432
## [38] train-error:0.097475
## [39] train-error:0.097560
## [40] train-error:0.097517
## [41] train-error:0.097347
## [42] train-error:0.097220
## [43] train-error:0.097177
## [44] train-error:0.097177
```

```
## [45] train-error:0.097220
## [46] train-error:0.097007
## [47] train-error:0.096922
## [48] train-error:0.096837
## [49] train-error:0.096880
## [50] train-error:0.096880
## [51] train-error:0.096455
## [52] train-error:0.096370
## [53] train-error:0.096157
## [54] train-error:0.096072
## [55] train-error:0.095987
## [56] train-error:0.095690
## [57] train-error:0.095775
## [58] train-error:0.095945
## [59] train-error:0.095477
## [60] train-error:0.095519
## [61] train-error:0.095264
## [62] train-error:0.095179
## [63] train-error:0.095179
## [64] train-error:0.095052
## [65] train-error:0.095179
## [66] train-error:0.094839
## [67] train-error:0.094669
## [68] train-error:0.094584
## [69] train-error:0.094457
## [70] train-error:0.094202
## [71] train-error:0.094244
## [72] train-error:0.094244
## [73] train-error:0.094117
## [74] train-error:0.094074
## [75] train-error:0.094032
## [76] train-error:0.093862
## [77] train-error:0.093734
## [78] train-error:0.093692
## [79] train-error:0.093607
## [80] train-error:0.093479
## [81] train-error:0.093309
## [1]  train-error:0.112863
## [2]  train-error:0.116477
## [3]  train-error:0.122471
## [4]  train-error:0.114734
## [5]  train-error:0.109208
## [6]  train-error:0.108570
## [7]  train-error:0.108570
## [8]  train-error:0.108145
## [9]  train-error:0.108060
## [10] train-error:0.108017
## [11] train-error:0.107550
## [12] train-error:0.107252
## [13] train-error:0.106955
## [14] train-error:0.105212
## [15] train-error:0.104744
## [16] train-error:0.104702
## [17] train-error:0.104574
```

```
## [18] train-error:0.104191
## [19] train-error:0.104064
## [20] train-error:0.103851
## [21] train-error:0.104021
## [22] train-error:0.102831
## [23] train-error:0.102661
## [24] train-error:0.102491
## [25] train-error:0.102151
## [26] train-error:0.101428
## [27] train-error:0.100621
## [28] train-error:0.100663
## [29] train-error:0.100153
## [30] train-error:0.100111
## [31] train-error:0.099430
## [32] train-error:0.099388
## [33] train-error:0.099643
## [34] train-error:0.099770
## [35] train-error:0.099600
## [36] train-error:0.099048
## [37] train-error:0.099005
## [38] train-error:0.098963
## [39] train-error:0.099048
## [40] train-error:0.098963
## [41] train-error:0.098963
## [42] train-error:0.098453
## [43] train-error:0.098495
## [44] train-error:0.097943
## [45] train-error:0.097815
## [46] train-error:0.097772
## [47] train-error:0.097815
## [48] train-error:0.097475
## [49] train-error:0.097645
## [50] train-error:0.097220
## [51] train-error:0.097177
## [52] train-error:0.097007
## [53] train-error:0.097050
## [54] train-error:0.097007
## [55] train-error:0.096880
## [56] train-error:0.097092
## [57] train-error:0.096710
## [58] train-error:0.096752
## [59] train-error:0.096710
## [60] train-error:0.096625
## [61] train-error:0.096497
## [62] train-error:0.096455
## [63] train-error:0.096455
## [64] train-error:0.096370
## [65] train-error:0.096200
## [66] train-error:0.096285
## [67] train-error:0.096157
## [68] train-error:0.096157
## [69] train-error:0.096200
## [70] train-error:0.096157
## [71] train-error:0.096072
```

```
## [72] train-error:0.096115
## [73] train-error:0.096072
## [74] train-error:0.095477
## [75] train-error:0.095137
## [76] train-error:0.095222
## [77] train-error:0.095222
## [78] train-error:0.094797
## [79] train-error:0.094797
## [80] train-error:0.094754
## [81] train-error:0.094669
## [1]  train-error:0.119112
## [2]  train-error:0.109590
## [3]  train-error:0.110951
## [4]  train-error:0.109123
## [5]  train-error:0.107295
## [6]  train-error:0.106912
## [7]  train-error:0.105509
## [8]  train-error:0.105339
## [9]  train-error:0.104106
## [10] train-error:0.103044
## [11] train-error:0.102534
## [12] train-error:0.102576
## [13] train-error:0.102364
## [14] train-error:0.101258
## [15] train-error:0.100111
## [16] train-error:0.099685
## [17] train-error:0.099940
## [18] train-error:0.099728
## [19] train-error:0.099218
## [20] train-error:0.098835
## [21] train-error:0.098538
## [22] train-error:0.098453
## [23] train-error:0.098240
## [24] train-error:0.098113
## [25] train-error:0.097815
## [26] train-error:0.097772
## [27] train-error:0.097560
## [28] train-error:0.097347
## [29] train-error:0.097390
## [30] train-error:0.097390
## [31] train-error:0.097177
## [32] train-error:0.097135
## [33] train-error:0.097050
## [34] train-error:0.096922
## [35] train-error:0.096880
## [36] train-error:0.096497
## [37] train-error:0.096285
## [38] train-error:0.096327
## [39] train-error:0.096157
## [40] train-error:0.096072
## [41] train-error:0.095945
## [42] train-error:0.096072
## [43] train-error:0.095392
## [44] train-error:0.095094
```

```
## [45]  train-error:0.095094
## [46]  train-error:0.094627
## [47]  train-error:0.094542
## [48]  train-error:0.094287
## [49]  train-error:0.094074
## [50]  train-error:0.094032
## [51]  train-error:0.093819
## [52]  train-error:0.093947
## [53]  train-error:0.093777
## [54]  train-error:0.093607
## [55]  train-error:0.093522
## [56]  train-error:0.093522
## [57]  train-error:0.093394
## [58]  train-error:0.093351
## [59]  train-error:0.093351
## [60]  train-error:0.093224
## [61]  train-error:0.093139
## [62]  train-error:0.093266
## [63]  train-error:0.093224
## [64]  train-error:0.093181
## [65]  train-error:0.093139
## [66]  train-error:0.092884
## [67]  train-error:0.092841
## [68]  train-error:0.092799
## [69]  train-error:0.092799
## [70]  train-error:0.092756
## [71]  train-error:0.092586
## [72]  train-error:0.092501
## [73]  train-error:0.092671
## [74]  train-error:0.092714
## [75]  train-error:0.092416
## [76]  train-error:0.092374
## [77]  train-error:0.092246
## [78]  train-error:0.092246
## [79]  train-error:0.092204
## [80]  train-error:0.092246
## [81]  train-error:0.092161
## [1]   train-error:0.140792
## [2]   train-error:0.122513
## [3]   train-error:0.117369
## [4]   train-error:0.115882
## [5]   train-error:0.115542
## [6]   train-error:0.111418
## [7]   train-error:0.108783
## [8]   train-error:0.108485
## [9]   train-error:0.106317
## [10]  train-error:0.106317
## [11]  train-error:0.104957
## [12]  train-error:0.104149
## [13]  train-error:0.104149
## [14]  train-error:0.103681
## [15]  train-error:0.102704
## [16]  train-error:0.102321
## [17]  train-error:0.102364
```

```
## [18] train-error:0.101896
## [19] train-error:0.102108
## [20] train-error:0.101598
## [21] train-error:0.101428
## [22] train-error:0.101386
## [23] train-error:0.101173
## [24] train-error:0.101131
## [25] train-error:0.101088
## [26] train-error:0.101046
## [27] train-error:0.100748
## [28] train-error:0.100961
## [29] train-error:0.100833
## [30] train-error:0.100706
## [31] train-error:0.100281
## [32] train-error:0.099558
## [33] train-error:0.099345
## [34] train-error:0.099303
## [35] train-error:0.099260
## [36] train-error:0.097985
## [37] train-error:0.097900
## [38] train-error:0.097687
## [39] train-error:0.097560
## [40] train-error:0.097602
## [41] train-error:0.097645
## [42] train-error:0.097390
## [43] train-error:0.097390
## [44] train-error:0.097135
## [45] train-error:0.097050
## [46] train-error:0.097007
## [47] train-error:0.096965
## [48] train-error:0.097092
## [49] train-error:0.097007
## [50] train-error:0.096922
## [51] train-error:0.096965
## [52] train-error:0.097177
## [53] train-error:0.097007
## [54] train-error:0.096922
## [55] train-error:0.096625
## [56] train-error:0.096370
## [57] train-error:0.096370
## [58] train-error:0.096072
## [59] train-error:0.095987
## [60] train-error:0.096157
## [61] train-error:0.095860
## [62] train-error:0.095902
## [63] train-error:0.095860
## [64] train-error:0.095945
## [65] train-error:0.095647
## [66] train-error:0.095094
## [67] train-error:0.094967
## [68] train-error:0.094414
## [69] train-error:0.094627
## [70] train-error:0.094074
## [71] train-error:0.094202
```

```
## [72] train-error:0.094074
## [73] train-error:0.094202
## [74] train-error:0.093989
## [75] train-error:0.094117
## [76] train-error:0.094032
## [77] train-error:0.093947
## [78] train-error:0.093904
## [79] train-error:0.093904
## [80] train-error:0.093862
## [81] train-error:0.093819
## [1]  train-error:0.121068
## [2]  train-error:0.111503
## [3]  train-error:0.113076
## [4]  train-error:0.112991
## [5]  train-error:0.113501
## [6]  train-error:0.113119
## [7]  train-error:0.112353
## [8]  train-error:0.111886
## [9]  train-error:0.111461
## [10] train-error:0.111121
## [11] train-error:0.110993
## [12] train-error:0.108145
## [13] train-error:0.106530
## [14] train-error:0.106742
## [15] train-error:0.106530
## [16] train-error:0.105934
## [17] train-error:0.105594
## [18] train-error:0.105764
## [19] train-error:0.104744
## [20] train-error:0.104744
## [21] train-error:0.103596
## [22] train-error:0.103426
## [23] train-error:0.102364
## [24] train-error:0.102279
## [25] train-error:0.102151
## [26] train-error:0.101896
## [27] train-error:0.101768
## [28] train-error:0.101683
## [29] train-error:0.101046
## [30] train-error:0.100876
## [31] train-error:0.100281
## [32] train-error:0.099728
## [33] train-error:0.099770
## [34] train-error:0.098963
## [35] train-error:0.098665
## [36] train-error:0.098155
## [37] train-error:0.098070
## [38] train-error:0.097858
## [39] train-error:0.097772
## [40] train-error:0.097858
## [41] train-error:0.097815
## [42] train-error:0.097943
## [43] train-error:0.097645
## [44] train-error:0.097517
```

```
## [45] train-error:0.097517
## [46] train-error:0.097135
## [47] train-error:0.097220
## [48] train-error:0.097135
## [49] train-error:0.096710
## [50] train-error:0.096752
## [51] train-error:0.096242
## [52] train-error:0.095987
## [53] train-error:0.095860
## [54] train-error:0.095647
## [55] train-error:0.095562
## [56] train-error:0.095477
## [57] train-error:0.095052
## [58] train-error:0.095094
## [59] train-error:0.095052
## [60] train-error:0.094882
## [61] train-error:0.094669
## [62] train-error:0.094542
## [63] train-error:0.094584
## [64] train-error:0.094712
## [65] train-error:0.094499
## [66] train-error:0.094542
## [67] train-error:0.094542
## [68] train-error:0.094287
## [69] train-error:0.094202
## [70] train-error:0.093989
## [71] train-error:0.093862
## [72] train-error:0.093777
## [73] train-error:0.093819
## [74] train-error:0.093479
## [75] train-error:0.093436
## [76] train-error:0.093309
## [77] train-error:0.093351
## [78] train-error:0.093351
## [79] train-error:0.093266
## [80] train-error:0.093181
## [81] train-error:0.093181
## [1]  train-error:0.121961
## [2]  train-error:0.122896
## [3]  train-error:0.118007
## [4]  train-error:0.112141
## [5]  train-error:0.111886
## [6]  train-error:0.111673
## [7]  train-error:0.111291
## [8]  train-error:0.111333
## [9]  train-error:0.108187
## [10] train-error:0.107805
## [11] train-error:0.106700
## [12] train-error:0.106742
## [13] train-error:0.106997
## [14] train-error:0.104702
## [15] train-error:0.104404
## [16] train-error:0.104276
## [17] train-error:0.104234
```

```
## [18] train-error:0.103809
## [19] train-error:0.103766
## [20] train-error:0.103596
## [21] train-error:0.103384
## [22] train-error:0.103129
## [23] train-error:0.102789
## [24] train-error:0.101938
## [25] train-error:0.101428
## [26] train-error:0.100791
## [27] train-error:0.100833
## [28] train-error:0.100621
## [29] train-error:0.099898
## [30] train-error:0.100153
## [31] train-error:0.099558
## [32] train-error:0.098835
## [33] train-error:0.098708
## [34] train-error:0.098453
## [35] train-error:0.098538
## [36] train-error:0.098538
## [37] train-error:0.098495
## [38] train-error:0.098538
## [39] train-error:0.098070
## [40] train-error:0.097943
## [41] train-error:0.097262
## [42] train-error:0.097092
## [43] train-error:0.097092
## [44] train-error:0.096837
## [45] train-error:0.096412
## [46] train-error:0.096582
## [47] train-error:0.096667
## [48] train-error:0.096582
## [49] train-error:0.096540
## [50] train-error:0.096455
## [51] train-error:0.096497
## [52] train-error:0.096327
## [53] train-error:0.096157
## [54] train-error:0.096285
## [55] train-error:0.095264
## [56] train-error:0.095307
## [57] train-error:0.094967
## [58] train-error:0.095137
## [59] train-error:0.094754
## [60] train-error:0.094754
## [61] train-error:0.094839
## [62] train-error:0.094797
## [63] train-error:0.094627
## [64] train-error:0.094457
## [65] train-error:0.094457
## [66] train-error:0.094457
## [67] train-error:0.094202
## [68] train-error:0.094244
## [69] train-error:0.094244
## [70] train-error:0.093989
## [71] train-error:0.093904
```

```
## [72] train-error:0.093947
## [73] train-error:0.093777
## [74] train-error:0.093734
## [75] train-error:0.093819
## [76] train-error:0.093734
## [77] train-error:0.093692
## [78] train-error:0.093692
## [79] train-error:0.093564
## [80] train-error:0.093436
## [81] train-error:0.093479
```

```r
accuracy = mean(as.numeric(cv))
#average error
error = 1 - accuracy
0.0932 #error rate
```

```
## [1] 0.0932
```

```r
library(caret)

#gridsearch for best parameters
#trialclass <- train(input,factor(findata$bank_account),
                  #  method = 'xgbTree', metric = 'Accuracy')

#trialclass$bestTune
#caret suggested parameters for optimal accuracy

fintest <- fintest %>% select(uniqueid, country, pred)

colnames(fintest) <- c('uniqueid', 'country', 'bank_account')

fintest$uniqueid <- as.character(fintest$uniqueid)

library(dplyr)
library(tidyr)
fintest <- fintest %>% unite(uniqueid, uniqueid, country, sep = ' x ' )


write.csv(fintest, 'financialinclusion.csv', col.names = TRUE, row.names = FALSE )
```

```
## Warning in write.csv(fintest, "financialinclusion.csv", col.names = TRUE, :
## attempt to set 'col.names' ignored
```