Machine Learning Project

Hui ZENG

2020/5/24

Installing Packages and Getting Data

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 (http://rmarkdown.rstudio.com).

```
library(tibble)
library(bitops)
library(rpart)
library(rattle)
```

```
## Rattle: A free graphical interface for data science with R.
## XXXX 5.4.0 Copyright (c) 2006-2020 Togaware Pty Ltd.
## 键入'rattle()'去轻摇、晃动、翻滚你的数据。
```

```
library(e1071)
library(ggplot2)
library(lattice)
library(caret)
library(kernlab)
```

```
##
## Attaching package: 'kernlab'
```

```
## The following object is masked from 'package:ggplot2':
##
##     alpha
```

```
library(randomForest)
```

```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:ggplot2':
##
##     margin
```

```
## The following object is masked from 'package:rattle':
##
##     importance
```

```
library(gbm)
```

```
## Loaded gbm 2.1.5
```

Getting and Cleaning Data

Read data from computer and check the dimension of two datasets. Data from testing dataset is for testing and data from training dataset is for training

```
traindata <- read.csv('~/Downloads/pml-training.csv', header = TR
UE, na.strings=c("NA","#DIV/0!",""))
validdata <- read.csv('~/Downloads/pml-testing.csv', header = TRU
E, na.strings=c("NA","#DIV/0!",""))
dim(traindata);dim(validdata)
```

```
## [1] 19622   160
```

```
## [1]  20 160
```

Getting Training Data Set and Testing Data Set

training datasets divided into two parts. 70% data is in training set and 30% data is in testing set

```
set.seed(2332)
inTrain <- createDataPartition(traindata$classe, p = 0.7, list =
FALSE)
trainset <- traindata[inTrain, ]
testset <- traindata[-inTrain, ]
dim(trainset);dim(testset)
```

```
## [1] 13737   160
```

```
## [1] 5885  160
```

Continue to Clean Data

remove variance of some variables which equal to zero and keep the columns of training set and testing set same

```
nzv <- nearZeroVar(trainset, saveMetrics=TRUE)
trainset <- trainset[,nzv$nzv==FALSE]
nzv<- nearZeroVar(testset,saveMetrics=TRUE)
testset <- testset[,nzv$nzv==FALSE]

trainset <- trainset[c(-1)]

training <- trainset
for(i in 1:length(trainset)) {
  if( sum( is.na( trainset[, i] ) ) /nrow(trainset) >= .7) {
    for(j in 1:length(training)) {
      if( length( grep(names(trainset[i]), names(training)[j]) )
== 1)  {
        training <- training[ , -j]
      }
    }
  }
}
trainset <- training
rm(training)

clean1 <- colnames(trainset)
clean2 <- colnames(trainset[, -58])
testset<- testset[clean1]
dim(testset)
```

```
## [1] 5885   58
```

```
for (i in 1:length(testset) ) {
  for(j in 1:length(trainset)) {
    if( length( grep(names(trainset[i]), names(testset)[j]) ) ==
1)  {
      class(testset[j]) <- class(trainset[i])
    }
  }
}
testset <- rbind(trainset[2, ] , testset)
testset <- testset[-1,]
```

Building Models

in order to choose the best algorithm to predict valid data, I will train three machine learning algorithm. They are classification trees, random forests and generalized boosted regression

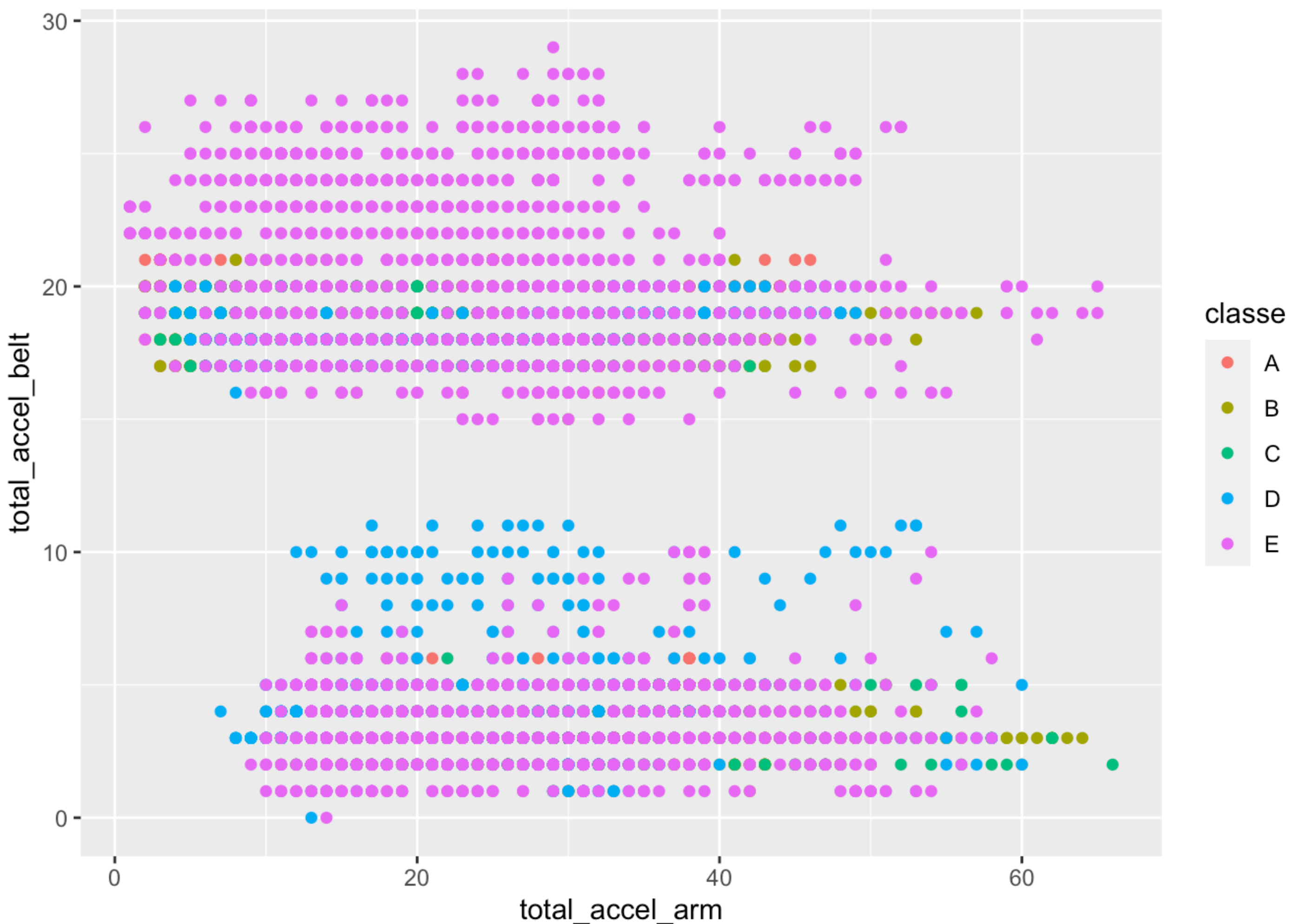
Predicting with trees

list the number of observations of five classes and plot total accelerated speed of belt to total accelerated speed of arm. Then use classification trees to train data in training set.

```
table(trainset$classe)
```

```
##
##    A    B    C    D    E
## 3906 2658 2396 2252 2525
```

```
qplot(total_accel_arm, total_accel_belt, colour = classe, data =
trainset)
```

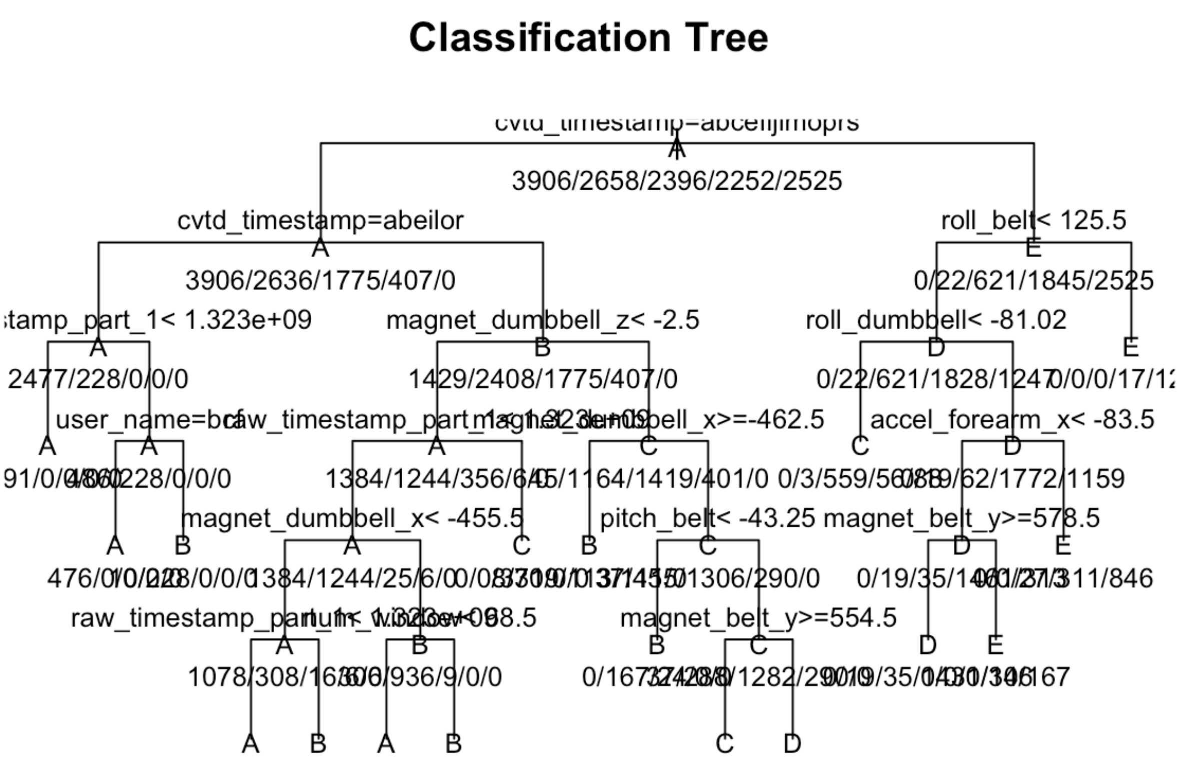


```
set.seed(11111)
modelfit <- rpart(classe ~ ., data = trainset, method = 'class')
```

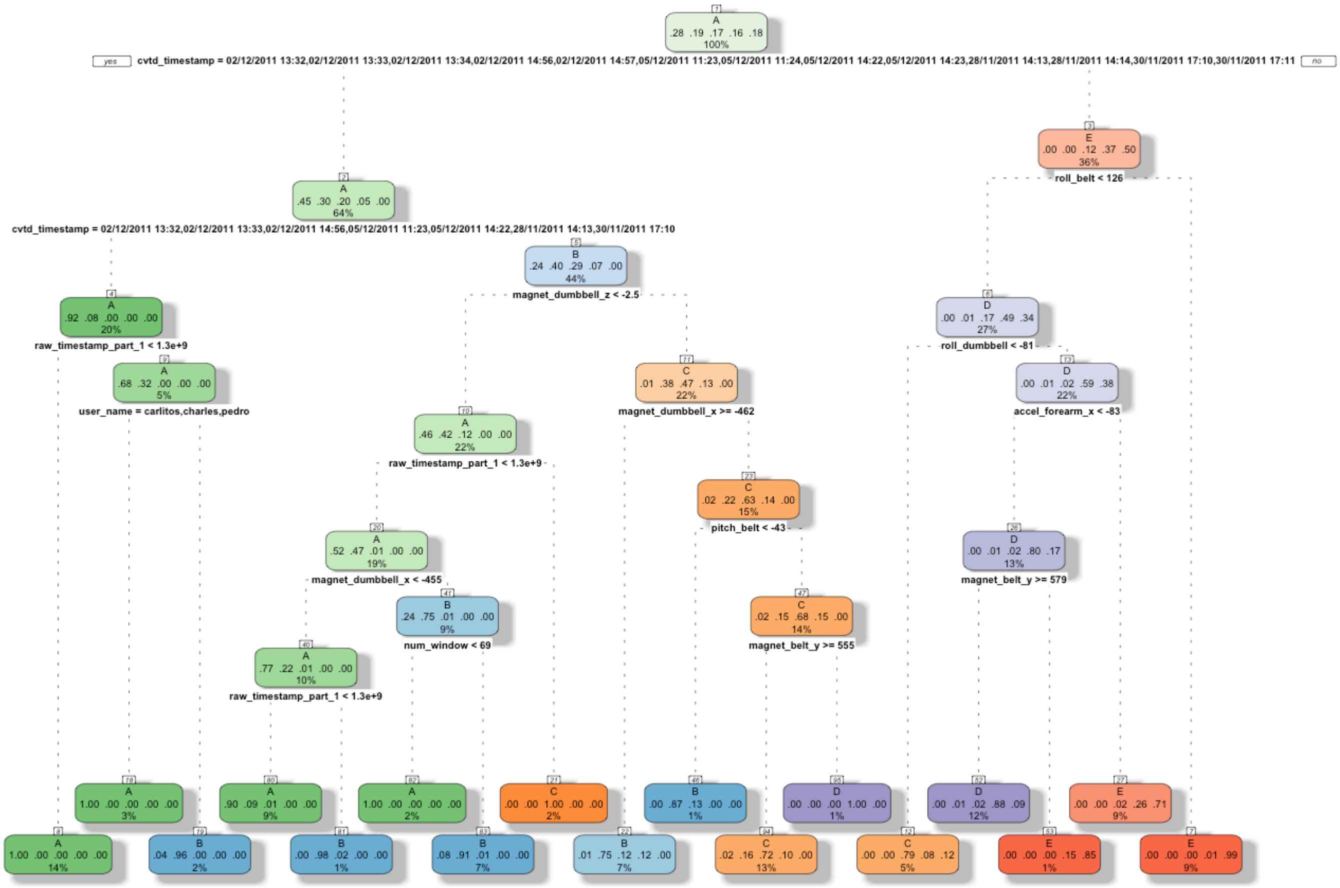
Ploting trees

plot normal classification trees and fancy classification trees. Then use trained model to predict data in testing set and check the result. The accuracy of this model is 0.8743

```
plot(modelfit, uniform = TRUE, main = 'Classification Tree')
text(modelfit, use.n = TRUE, all = TRUE, cex = 0.8)
```




```
fancyRpartPlot(modelfit)
```



Rattle 2020- 5-24 21:21:10 cenghui

```
predicttree <- predict(modelfit, testset, type = 'class')
cmtree <- confusionMatrix(predicttree, testset$classe)
cmtree
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1613   51    6    0    0
##          B   43  927   50   42    0
##          C   18  156  942   96   35
##          D    0    5   18  671   55
##          E    0    0   10  155  992
##
## Overall Statistics
##
##                Accuracy : 0.8743
##                  95% CI : (0.8655, 0.8826)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.8409
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class
: E
## Sensitivity            0.9636   0.8139   0.9181   0.6961   0.9
168
## Specificity            0.9865   0.9716   0.9372   0.9841   0.9
656
## Pos Pred Value         0.9659   0.8729   0.7554   0.8959   0.8
574
## Neg Pred Value         0.9855   0.9560   0.9819   0.9430   0.9
810
## Prevalence             0.2845   0.1935   0.1743   0.1638   0.1
839
## Detection Rate         0.2741   0.1575   0.1601   0.1140   0.1
686
## Detection Prevalence   0.2838   0.1805   0.2119   0.1273   0.1
966
## Balanced Accuracy      0.9750   0.8927   0.9277   0.8401   0.9
412
```

Random forests

use random forests to train data in training set

```
controlrf <- trainControl(method="cv", number=3, verboseIter=FALS
E)
modelfit2 <- train(classe~., data = trainset, method = 'rf', trCo
ntrol = controlrf)
modelfit2$finalModel
```

```
##
## Call:
##  randomForest(x = x, y = y, mtry = param$mtry)
##                Type of random forest: classification
##                      Number of trees: 500
## No. of variables tried at each split: 40
##
##         OOB estimate of  error rate: 0.07%
## Confusion matrix:
##      A    B    C    D    E  class.error
## A 3906    0    0    0    0 0.0000000000
## B    1 2656    1    0    0 0.0007524454
## C    0    3 2393    0    0 0.0012520868
## D    0    0    3 2249    0 0.0013321492
## E    0    0    0    1 2524 0.0003960396
```

```
getTree(modelfit2$finalModel, k=2)
```

```
##     left daughter right daughter split var   split point statu
s prediction
## 1               2              3        28  1.295000e+02
1         0
## 2               4              5        51  1.285000e+02
1         0
## 3               6              7        30  1.585000e+02
1         0
## 4               8              9        41 -5.655000e+01
1         0
## 5              10             11        30  1.705000e+02
1         0
## 6               0              0         0  0.000000e+00     -
1         5
## 7              12             13        51  3.045000e+02
1         0
## 8              14             15        66  7.200000e+01
```

```
1           0
## 9                 16                17           68 -3.205000e+01
1           0
## 10                18                19           23  5.000000e-01
1           0
## 11                20                21           34 -2.550000e-01
1           0
## 12                 0                 0            0  0.000000e+00      -
1           1
## 13                 0                 0            0  0.000000e+00      -
1           5
## 14                22                23           75  3.595000e+02
1           0
## 15                 0                 0            0  0.000000e+00      -
1           2
## 16                24                25           79  7.750000e+02
1           0
## 17                26                27           61  4.350000e+01
1           0
## 18                28                29           29 -4.255000e+01
1           0
## 19                30                31            6  1.322490e+09
1           0
## 20                32                33           29 -4.510000e+01
1           0
## 21                34                35            6  1.322833e+09
1           0
## 22                 0                 0            0  0.000000e+00      -
1           1
## 23                36                37           74  1.300000e+01
1           0
## 24                38                39            6  1.323084e+09
1           0
## 25                 0                 0            0  0.000000e+00      -
1           2
## 26                40                41           27  2.420000e+02
1           0
## 27                42                43            6  1.323084e+09
1           0
## 28                44                45            6  1.322833e+09
1           0
## 29                46                47           26  5.000000e-01
1           0
## 30                 0                 0            0  0.000000e+00      -
1           4
```

```
## 31          0          0     0  0.000000e+00     -
1         5
## 32          0          0     0  0.000000e+00     -
1         4
## 33         48         49    30  1.755000e+02
1         0
## 34          0          0     0  0.000000e+00     -
1         1
## 35         50         51    10  5.000000e-01
1         0
## 36          0          0     0  0.000000e+00     -
1         1
## 37          0          0     0  0.000000e+00     -
1         2
## 38          0          0     0  0.000000e+00     -
1         1
## 39          0          0     0  0.000000e+00     -
1         2
## 40         52         53    27  1.650000e+02
1         0
## 41         54         55    66 -2.150000e+01
1         0
## 42         56         57    23  5.000000e-01
1         0
## 43         58         59    28  6.421000e+01
1         0
## 44          0          0     0  0.000000e+00     -
1         2
## 45         60         61     6  1.322833e+09
1         0
## 46         62         63    27  4.750000e+01
1         0
## 47         64         65    74 -1.000000e+02
1         0
## 48          0          0     0  0.000000e+00     -
1         5
## 49          0          0     0  0.000000e+00     -
1         1
## 50          0          0     0  0.000000e+00     -
1         2
## 51          0          0     0  0.000000e+00     -
1         4
## 52         66         67     6  1.323095e+09
1         0
## 53          0          0     0  0.000000e+00     -
```

```
1          1
## 54              68             69         54  7.475947e+01
1          0
## 55              70             71         39  5.530000e+02
1          0
## 56              72             73         34 -3.650000e-01
1          0
## 57               0              0          0  0.000000e+00     -
1          5
## 58               0              0          0  0.000000e+00     -
1          5
## 59               0              0          0  0.000000e+00     -
1          2
## 60              74             75         29 -4.445000e+01
1          0
## 61              76             77         38  1.870000e+02
1          0
## 62              78             79          6  1.323084e+09
1          0
## 63              80             81          6  1.323095e+09
1          0
## 64              82             83          6  1.322673e+09
1          0
## 65              84             85          6  1.322673e+09
1          0
## 66              86             87         27  6.850000e+01
1          0
## 67              88             89         27  1.065000e+02
1          0
## 68              90             91         64 -4.815000e+02
1          0
## 69              92             93         72  4.000000e-02
1          0
## 70              94             95         76 -5.500000e+00
1          0
## 71              96             97          8  5.000000e-01
1          0
## 72              98             99         55  3.171861e+01
1          0
## 73             100            101         54  1.162102e+02
1          0
## 74               0              0          0  0.000000e+00     -
1          2
## 75               0              0          0  0.000000e+00     -
1          3
```

```
## 76           0          0         0  0.000000e+00       -
1          4
## 77           0          0         0  0.000000e+00       -
1          5
## 78         102        103        41  4.965000e+01
1          0
## 79           0          0         0  0.000000e+00       -
1          5
## 80         104        105        30 -2.700000e+00
1          0
## 81         106        107         6  1.323095e+09
1          0
## 82         108        109        27  4.785000e+02
1          0
## 83         110        111        47 -1.900000e-01
1          0
## 84           0          0         0  0.000000e+00       -
1          4
## 85           0          0         0  0.000000e+00       -
1          5
## 86         112        113        65 -2.185000e+02
1          0
## 87         114        115         5  5.000000e-01
1          0
## 88           0          0         0  0.000000e+00       -
1          3
## 89           0          0         0  0.000000e+00       -
1          4
## 90         116        117         6  1.323084e+09
1          0
## 91         118        119        67  1.380000e+02
1          0
## 92         120        121        49  1.680000e+02
1          0
## 93         122        123        29  6.650000e+00
1          0
## 94           0          0         0  0.000000e+00       -
1          5
## 95           0          0         0  0.000000e+00       -
1          4
## 96         124        125        59 -4.100000e-01
1          0
## 97         126        127         6  1.322833e+09
1          0
## 98           0          0         0  0.000000e+00       -
```

```
1          5
## 99             0              0         0  0.000000e+00     -
1          4
## 100          128            129        21  5.000000e-01
1          0
## 101            0              0         0  0.000000e+00     -
1          3
## 102            0              0         0  0.000000e+00     -
1          5
## 103            0              0         0  0.000000e+00     -
1          4
## 104          130            131         6  1.322490e+09
1          0
## 105          132            133        56  6.837196e+01
1          0
## 106            0              0         0  0.000000e+00     -
1          4
## 107            0              0         0  0.000000e+00     -
1          5
## 108            0              0         0  0.000000e+00     -
1          3
## 109            0              0         0  0.000000e+00     -
1          4
## 110            0              0         0  0.000000e+00     -
1          4
## 111            0              0         0  0.000000e+00     -
1          5
## 112            0              0         0  0.000000e+00     -
1          1
## 113          134            135        44  2.250000e+01
1          0
## 114          136            137        54  4.250378e+00
1          0
## 115            0              0         0  0.000000e+00     -
1          2
## 116            0              0         0  0.000000e+00     -
1          1
## 117          138            139        17  5.000000e-01
1          0
## 118          140            141        27  4.365000e+02
1          0
## 119          142            143        27  8.355000e+02
1          0
## 120          144            145         6  1.322838e+09
1          0
```

```
## 121          0          0      0  0.000000e+00      -
1          1
## 122          0          0      0  0.000000e+00      -
1          2
## 123        146        147     68  1.690000e+01
1          0
## 124        148        149      6  1.322490e+09
1          0
## 125        150        151     29 -4.315000e+01
1          0
## 126        152        153     32  4.000000e-02
1          0
## 127          0          0      0  0.000000e+00      -
1          2
## 128        154        155     71  1.320000e+00
1          0
## 129          0          0      0  0.000000e+00      -
1          1
## 130          0          0      0  0.000000e+00      -
1          1
## 131        156        157     65  3.625000e+02
1          0
## 132        158        159     54  5.530236e+01
1          0
## 133        160        161     67  1.095000e+02
1          0
## 134        162        163     27  2.500000e+01
1          0
## 135          0          0      0  0.000000e+00      -
1          5
## 136          0          0      0  0.000000e+00      -
1          5
## 137          0          0      0  0.000000e+00      -
1          4
## 138          0          0      0  0.000000e+00      -
1          2
## 139        164        165     33  5.000000e-02
1          0
## 140          0          0      0  0.000000e+00      -
1          1
## 141        166        167     27  8.375000e+02
1          0
## 142          0          0      0  0.000000e+00      -
1          3
## 143          0          0      0  0.000000e+00      -
```

```
1          2
## 144             0              0        0  0.000000e+00     -
1          1
## 145           168            169       33 -1.000000e-02
1          0
## 146           170            171       70  4.500000e+01
1          0
## 147             0              0        0  0.000000e+00     -
1          4
## 148             0              0        0  0.000000e+00     -
1          1
## 149           172            173       32  9.000000e-02
1          0
## 150           174            175       78  6.080000e+02
1          0
## 151           176            177        6  1.323084e+09
1          0
## 152             0              0        0  0.000000e+00     -
1          2
## 153             0              0        0  0.000000e+00     -
1          1
## 154             0              0        0  0.000000e+00     -
1          2
## 155             0              0        0  0.000000e+00     -
1          3
## 156           178            179       36 -4.500000e+00
1          0
## 157           180            181       17  5.000000e-01
1          0
## 158           182            183        6  1.322838e+09
1          0
## 159           184            185       12  5.000000e-01
1          0
## 160           186            187       65  3.940000e+02
1          0
## 161           188            189       42  4.375000e+00
1          0
## 162             0              0        0  0.000000e+00     -
1          4
## 163             0              0        0  0.000000e+00     -
1          5
## 164             0              0        0  0.000000e+00     -
1          4
## 165             0              0        0  0.000000e+00     -
1          5
```

```
## 166          190          191          30 -1.810000e+00
1         0
## 167            0            0           0  0.000000e+00     -
1         1
## 168            0            0           0  0.000000e+00     -
1         4
## 169            0            0           0  0.000000e+00     -
1         2
## 170            0            0           0  0.000000e+00     -
1         1
## 171            0            0           0  0.000000e+00     -
1         2
## 172          192          193          47 -5.150000e-01
1         0
## 173          194          195          54  7.642837e+01
1         0
## 174          196          197          37 -1.575000e+02
1         0
## 175          198          199          62  2.160000e+02
1         0
## 176          200          201          26  5.000000e-01
1         0
## 177          202          203          37  6.000000e+00
1         0
## 178          204          205          22  5.000000e-01
1         0
## 179          206          207          65 -5.715000e+02
1         0
## 180          208          209          56 -6.883450e+01
1         0
## 181          210          211           6  1.323084e+09
1         0
## 182          212          213          40 -3.065000e+02
1         0
## 183          214          215          28  1.215000e+02
1         0
## 184          216          217          28  1.255000e+02
1         0
## 185            0            0           0  0.000000e+00     -
1         1
## 186          218          219          79  1.270000e+02
1         0
## 187          220          221          62  1.550000e+01
1         0
## 188          222          223          62 -3.250000e+01
```

```
1           0
## 189          224               225          29  2.575000e+01
1           0
## 190            0                 0           0  0.000000e+00      -
1           2
## 191          226               227          57  4.500000e+00
1           0
## 192          228               229          74  1.550000e+01
1           0
## 193          230               231          34  3.000000e-01
1           0
## 194          232               233          42  6.550000e+00
1           0
## 195          234               235          63 -7.200000e+01
1           0
## 196            0                 0           0  0.000000e+00      -
1           2
## 197          236               237          79  6.100000e+01
1           0
## 198            0                 0           0  0.000000e+00      -
1           3
## 199            0                 0           0  0.000000e+00      -
1           2
## 200          238               239          23  5.000000e-01
1           0
## 201          240               241           6  1.322673e+09
1           0
## 202            0                 0           0  0.000000e+00      -
1           5
## 203            0                 0           0  0.000000e+00      -
1           4
## 204            0                 0           0  0.000000e+00      -
1           5
## 205            0                 0           0  0.000000e+00      -
1           4
## 206          242               243          42  3.405000e+01
1           0
## 207          244               245          67  1.265000e+02
1           0
## 208          246               247           6  1.322673e+09
1           0
## 209          248               249          27  2.130000e+02
1           0
## 210            0                 0           0  0.000000e+00      -
1           4
```

```
## 211          0          0     0 0.000000e+00     -
1         5
## 212        250        251    12 5.000000e-01
1         0
## 213        252        253    28 1.255000e+02
1         0
## 214          0          0     0 0.000000e+00     -
1         4
## 215        254        255    75 4.450000e+01
1         0
## 216        256        257    69 2.240000e+01
1         0
## 217        258        259    71 6.000000e-01
1         0
## 218          0          0     0 0.000000e+00     -
1         2
## 219        260        261    65 3.795000e+02
1         0
## 220          0          0     0 0.000000e+00     -
1         4
## 221          0          0     0 0.000000e+00     -
1         2
## 222          0          0     0 0.000000e+00     -
1         1
## 223          0          0     0 0.000000e+00     -
1         2
## 224          0          0     0 0.000000e+00     -
1         3
## 225          0          0     0 0.000000e+00     -
1         1
## 226        262        263    61 0.000000e+00
1         0
## 227          0          0     0 0.000000e+00     -
1         3
## 228          0          0     0 0.000000e+00     -
1         4
## 229          0          0     0 0.000000e+00     -
1         3
## 230        264        265    76 2.315000e+02
1         0
## 231          0          0     0 0.000000e+00     -
1         5
## 232        266        267    50 2.000000e+01
1         0
## 233          0          0     0 0.000000e+00     -
```

```
1          3
## 234            0              0         0  0.000000e+00      -
1          3
## 235          268            269        45  7.600000e-01
1          0
## 236            0              0         0  0.000000e+00      -
1          2
## 237            0              0         0  0.000000e+00      -
1          3
## 238          270            271        58 -3.000000e-01
1          0
## 239          272            273        59  5.550000e-01
1          0
## 240          274            275        45 -2.810000e+00
1          0
## 241            0              0         0  0.000000e+00      -
1          5
## 242            0              0         0  0.000000e+00      -
1          2
## 243            0              0         0  0.000000e+00      -
1          3
## 244          276            277        79  1.125000e+02
1          0
## 245          278            279         6  1.322490e+09
1          0
## 246            0              0         0  0.000000e+00      -
1          1
## 247          280            281        69  9.045000e+01
1          0
## 248          282            283        68  2.606000e+01
1          0
## 249          284            285        74  1.595000e+02
1          0
## 250          286            287        30  1.625000e+02
1          0
## 251          288            289        42 -2.895000e+01
1          0
## 252          290            291         6  1.322833e+09
1          0
## 253          292            293        73  9.550000e-01
1          0
## 254            0              0         0  0.000000e+00      -
1          5
## 255            0              0         0  0.000000e+00      -
1          3
```

```
## 256          294          295          6  1.322833e+09
1         0
## 257            0            0          0  0.000000e+00     -
1         5
## 258            0            0          0  0.000000e+00     -
1         5
## 259          296          297         72 -7.200000e-01
1         0
## 260          298          299         73  1.010000e+00
1         0
## 261          300          301         67  2.730000e+01
1         0
## 262            0            0          0  0.000000e+00     -
1         1
## 263            0            0          0  0.000000e+00     -
1         4
## 264          302          303         50  9.400000e+01
1         0
## 265            0            0          0  0.000000e+00     -
1         3
## 266            0            0          0  0.000000e+00     -
1         2
## 267            0            0          0  0.000000e+00     -
1         5
## 268            0            0          0  0.000000e+00     -
1         5
## 269            0            0          0  0.000000e+00     -
1         2
## 270          304          305         30 -8.440000e+01
1         0
## 271          306          307          6  1.322490e+09
1         0
## 272          308          309         46 -1.210000e+00
1         0
## 273            0            0          0  0.000000e+00     -
1         4
## 274            0            0          0  0.000000e+00     -
1         3
## 275            0            0          0  0.000000e+00     -
1         4
## 276          310          311          6  1.322490e+09
1         0
## 277          312          313          6  1.323084e+09
1         0
## 278          314          315         48  1.130000e+02
```

```
1         0
## 279          316               317        68  2.575000e+01
1         0
## 280          318               319        48 -1.730000e+02
1         0
## 281          320               321        74  1.145000e+02
1         0
## 282            0                 0         0  0.000000e+00      -
1         1
## 283            0                 0         0  0.000000e+00      -
1         4
## 284          322               323        24  5.000000e-01
1         0
## 285          324               325         6  1.322673e+09
1         0
## 286          326               327         6  1.322838e+09
1         0
## 287          328               329        64 -5.695000e+02
1         0
## 288            0                 0         0  0.000000e+00      -
1         1
## 289            0                 0         0  0.000000e+00      -
1         2
## 290            0                 0         0  0.000000e+00      -
1         3
## 291          330               331        35  1.500000e+01
1         0
## 292          332               333        32  7.000000e-02
1         0
## 293            0                 0         0  0.000000e+00      -
1         2
## 294            0                 0         0  0.000000e+00      -
1         3
## 295          334               335        39  5.395000e+02
1         0
## 296            0                 0         0  0.000000e+00      -
1         4
## 297            0                 0         0  0.000000e+00      -
1         3
## 298            0                 0         0  0.000000e+00      -
1         1
## 299            0                 0         0  0.000000e+00      -
1         2
## 300            0                 0         0  0.000000e+00      -
1         1
```

```
## 301              0              0       0  0.000000e+00      -
1          2
## 302            336            337      46 -1.735000e+00
1          0
## 303              0              0       0  0.000000e+00      -
1          3
## 304              0              0       0  0.000000e+00      -
1          2
## 305            338            339      63 -1.710000e+02
1          0
## 306            340            341      50 -1.290000e+02
1          0
## 307            342            343      16  5.000000e-01
1          0
## 308              0              0       0  0.000000e+00      -
1          4
## 309              0              0       0  0.000000e+00      -
1          5
## 310              0              0       0  0.000000e+00      -
1          2
## 311            344            345      76 -2.650000e+01
1          0
## 312            346            347      35 -1.050000e+01
1          0
## 313            348            349      30 -5.105000e+00
1          0
## 314              0              0       0  0.000000e+00      -
1          2
## 315            350            351      63  9.800000e+01
1          0
## 316            352            353      30 -9.345000e+01
1          0
## 317            354            355      57  3.500000e+00
1          0
## 318              0              0       0  0.000000e+00      -
1          3
## 319              0              0       0  0.000000e+00      -
1          2
## 320              0              0       0  0.000000e+00      -
1          1
## 321              0              0       0  0.000000e+00      -
1          3
## 322            356            357      56 -6.768388e+01
1          0
## 323              0              0       0  0.000000e+00      -
```

```
1          1
## 324          358            359         68 -2.985000e+01
1          0
## 325          360            361         45 -3.145000e+00
1          0
## 326            0              0          0  0.000000e+00      -
1          3
## 327          362            363         78  7.545000e+02
1          0
## 328            0              0          0  0.000000e+00      -
1          3
## 329          364            365         64 -5.680000e+02
1          0
## 330          366            367         74 -3.325000e+02
1          0
## 331            0              0          0  0.000000e+00      -
1          4
## 332            0              0          0  0.000000e+00      -
1          4
## 333          368            369          6  1.322833e+09
1          0
## 334            0              0          0  0.000000e+00      -
1          5
## 335            0              0          0  0.000000e+00      -
1          4
## 336          370            371         55 -5.170101e+01
1          0
## 337            0              0          0  0.000000e+00      -
1          2
## 338            0              0          0  0.000000e+00      -
1          3
## 339            0              0          0  0.000000e+00      -
1          4
## 340            0              0          0  0.000000e+00      -
1          2
## 341            0              0          0  0.000000e+00      -
1          1
## 342          372            373         34 -4.900000e-01
1          0
## 343          374            375          6  1.323084e+09
1          0
## 344          376            377         48  3.095000e+02
1          0
## 345          378            379         29 -3.350000e-01
1          0
```

```
## 346          380          381          63  3.950000e+01
1          0
## 347          382          383          52 -2.800000e+01
1          0
## 348          384          385          17  5.000000e-01
1          0
## 349          386          387           6  1.323095e+09
1          0
## 350            0            0           0  0.000000e+00     -
1          1
## 351            0            0           0  0.000000e+00     -
1          2
## 352          388          389          29  8.150000e+00
1          0
## 353          390          391          29  3.000000e-01
1          0
## 354          392          393           6  1.323084e+09
1          0
## 355          394          395          68  5.935000e+01
1          0
## 356          396          397          27  4.335000e+02
1          0
## 357            0            0           0  0.000000e+00     -
1          2
## 358            0            0           0  0.000000e+00     -
1          1
## 359            0            0           0  0.000000e+00     -
1          2
## 360            0            0           0  0.000000e+00     -
1          2
## 361            0            0           0  0.000000e+00     -
1          3
## 362            0            0           0  0.000000e+00     -
1          3
## 363            0            0           0  0.000000e+00     -
1          4
## 364            0            0           0  0.000000e+00     -
1          2
## 365            0            0           0  0.000000e+00     -
1          4
## 366            0            0           0  0.000000e+00     -
1          2
## 367            0            0           0  0.000000e+00     -
1          3
## 368            0            0           0  0.000000e+00     -
```

```
1          3
## 369          0             0         0  0.000000e+00      -
1          4
## 370          0             0         0  0.000000e+00      -
1          2
## 371          0             0         0  0.000000e+00      -
1          3
## 372          0             0         0  0.000000e+00      -
1          4
## 373        398           399        62  1.555000e+02
1          0
## 374          0             0         0  0.000000e+00      -
1          2
## 375          0             0         0  0.000000e+00      -
1          3
## 376        400           401        72  3.115000e+00
1          0
## 377          0             0         0  0.000000e+00      -
1          1
## 378          0             0         0  0.000000e+00      -
1          4
## 379          0             0         0  0.000000e+00      -
1          3
## 380        402           403        67  1.195000e+02
1          0
## 381          0             0         0  0.000000e+00      -
1          2
## 382          0             0         0  0.000000e+00      -
1          2
## 383          0             0         0  0.000000e+00      -
1          3
## 384        404           405        20  5.000000e-01
1          0
## 385        406           407         6  1.323084e+09
1          0
## 386          0             0         0  0.000000e+00      -
1          1
## 387          0             0         0  0.000000e+00      -
1          2
## 388          0             0         0  0.000000e+00      -
1          4
## 389          0             0         0  0.000000e+00      -
1          3
## 390          0             0         0  0.000000e+00      -
1          4
```

```
## 391          408            409          68 -2.070000e+01
1         0
## 392            0              0           0  0.000000e+00      -
1         3
## 393          410            411          34  1.300000e-01
1         0
## 394          412            413          64 -5.175000e+02
1         0
## 395          414            415          73  1.465000e+00
1         0
## 396            0              0           0  0.000000e+00      -
1         1
## 397            0              0           0  0.000000e+00      -
1         2
## 398          416            417          30  1.635000e+02
1         0
## 399          418            419           6  1.322673e+09
1         0
## 400            0              0           0  0.000000e+00      -
1         2
## 401          420            421          27  3.660000e+02
1         0
## 402            0              0           0  0.000000e+00      -
1         1
## 403          422            423          78  4.080000e+02
1         0
## 404            0              0           0  0.000000e+00      -
1         2
## 405            0              0           0  0.000000e+00      -
1         4
## 406            0              0           0  0.000000e+00      -
1         3
## 407          424            425          30 -9.290000e+01
1         0
## 408            0              0           0  0.000000e+00      -
1         2
## 409          426            427          49 -1.135000e+02
1         0
## 410            0              0           0  0.000000e+00      -
1         4
## 411            0              0           0  0.000000e+00      -
1         5
## 412          428            429          71 -5.850000e-01
1         0
## 413          430            431           6  1.323095e+09
```

```
1          0
## 414              0              0         0  0.000000e+00      -
1          1
## 415              0              0         0  0.000000e+00      -
1          2
## 416            432            433        68 -2.485000e+01
1          0
## 417            434            435        77 -4.895000e+02
1          0
## 418            436            437        66  1.400000e+01
1          0
## 419              0              0         0  0.000000e+00      -
1          3
## 420              0              0         0  0.000000e+00      -
1          2
## 421              0              0         0  0.000000e+00      -
1          3
## 422              0              0         0  0.000000e+00      -
1          1
## 423              0              0         0  0.000000e+00      -
1          3
## 424              0              0         0  0.000000e+00      -
1          4
## 425              0              0         0  0.000000e+00      -
1          5
## 426              0              0         0  0.000000e+00      -
1          2
## 427            438            439        58 -3.150000e-01
1          0
## 428            440            441        48  6.500000e+00
1          0
## 429            442            443        66  8.500000e+00
1          0
## 430              0              0         0  0.000000e+00      -
1          2
## 431            444            445        65 -5.165000e+02
1          0
## 432              0              0         0  0.000000e+00      -
1          1
## 433            446            447        60 -1.360000e+00
1          0
## 434            448            449        65  2.455000e+02
1          0
## 435            450            451        41 -1.017500e+01
1          0
```

```
## 436            0           0         0  0.000000e+00      -
1          1
## 437            0           0         0  0.000000e+00      -
1          2
## 438          452         453        17  5.000000e-01
1          0
## 439          454         455        65  3.245000e+02
1          0
## 440            0           0         0  0.000000e+00      -
1          1
## 441            0           0         0  0.000000e+00      -
1          5
## 442          456         457        36  7.500000e+00
1          0
## 443            0           0         0  0.000000e+00      -
1          3
## 444            0           0         0  0.000000e+00      -
1          3
## 445            0           0         0  0.000000e+00      -
1          4
## 446          458         459        32  6.000000e-02
1          0
## 447          460         461        53  6.435000e+02
1          0
## 448          462         463        75  4.385000e+02
1          0
## 449          464         465        53  5.110000e+02
1          0
## 450            0           0         0  0.000000e+00      -
1          4
## 451            0           0         0  0.000000e+00      -
1          3
## 452            0           0         0  0.000000e+00      -
1          2
## 453            0           0         0  0.000000e+00      -
1          5
## 454          466         467        28 -6.500000e-02
1          0
## 455          468         469        38  3.550000e+01
1          0
## 456            0           0         0  0.000000e+00      -
1          5
## 457            0           0         0  0.000000e+00      -
1          4
## 458            0           0         0  0.000000e+00      -
```

```
1          1
## 459           0             0         0  0.000000e+00      -
1          2
## 460         470           471        76 -2.450000e+02
1          0
## 461         472           473        62 -4.800000e+01
1          0
## 462           0             0         0  0.000000e+00      -
1          3
## 463           0             0         0  0.000000e+00      -
1          4
## 464           0             0         0  0.000000e+00      -
1          3
## 465           0             0         0  0.000000e+00      -
1          4
## 466           0             0         0  0.000000e+00      -
1          4
## 467         474           475        70  5.200000e+01
1          0
## 468           0             0         0  0.000000e+00      -
1          2
## 469           0             0         0  0.000000e+00      -
1          3
## 470         476           477        62  4.950000e+01
1          0
## 471         478           479        29  1.465000e+01
1          0
## 472           0             0         0  0.000000e+00      -
1          3
## 473           0             0         0  0.000000e+00      -
1          4
## 474         480           481         6  1.323095e+09
1          0
## 475           0             0         0  0.000000e+00      -
1          2
## 476           0             0         0  0.000000e+00      -
1          3
## 477           0             0         0  0.000000e+00      -
1          2
## 478         482           483        68  4.990000e+01
1          0
## 479         484           485        71 -2.450000e-01
1          0
## 480         486           487        56 -8.884194e+01
1          0
```

```
## 481          488          489       45 -1.325000e+00
1          0
## 482          490          491       24  5.000000e-01
1          0
## 483          492          493       76 -7.700000e+01
1          0
## 484            0            0        0  0.000000e+00     -
1          4
## 485            0            0        0  0.000000e+00     -
1          3
## 486          494          495       48 -5.350000e+01
1          0
## 487          496          497       39  5.490000e+02
1          0
## 488            0            0        0  0.000000e+00     -
1          3
## 489            0            0        0  0.000000e+00     -
1          4
## 490          498          499       32  2.200000e-01
1          0
## 491            0            0        0  0.000000e+00     -
1          1
## 492            0            0        0  0.000000e+00     -
1          3
## 493            0            0        0  0.000000e+00     -
1          4
## 494            0            0        0  0.000000e+00     -
1          4
## 495            0            0        0  0.000000e+00     -
1          3
## 496            0            0        0  0.000000e+00     -
1          5
## 497          500          501       75  3.455000e+02
1          0
## 498            0            0        0  0.000000e+00     -
1          3
## 499            0            0        0  0.000000e+00     -
1          2
## 500            0            0        0  0.000000e+00     -
1          3
## 501          502          503       66  1.100000e+02
1          0
## 502            0            0        0  0.000000e+00     -
1          3
## 503            0            0        0  0.000000e+00     -
```

```
1             2
```

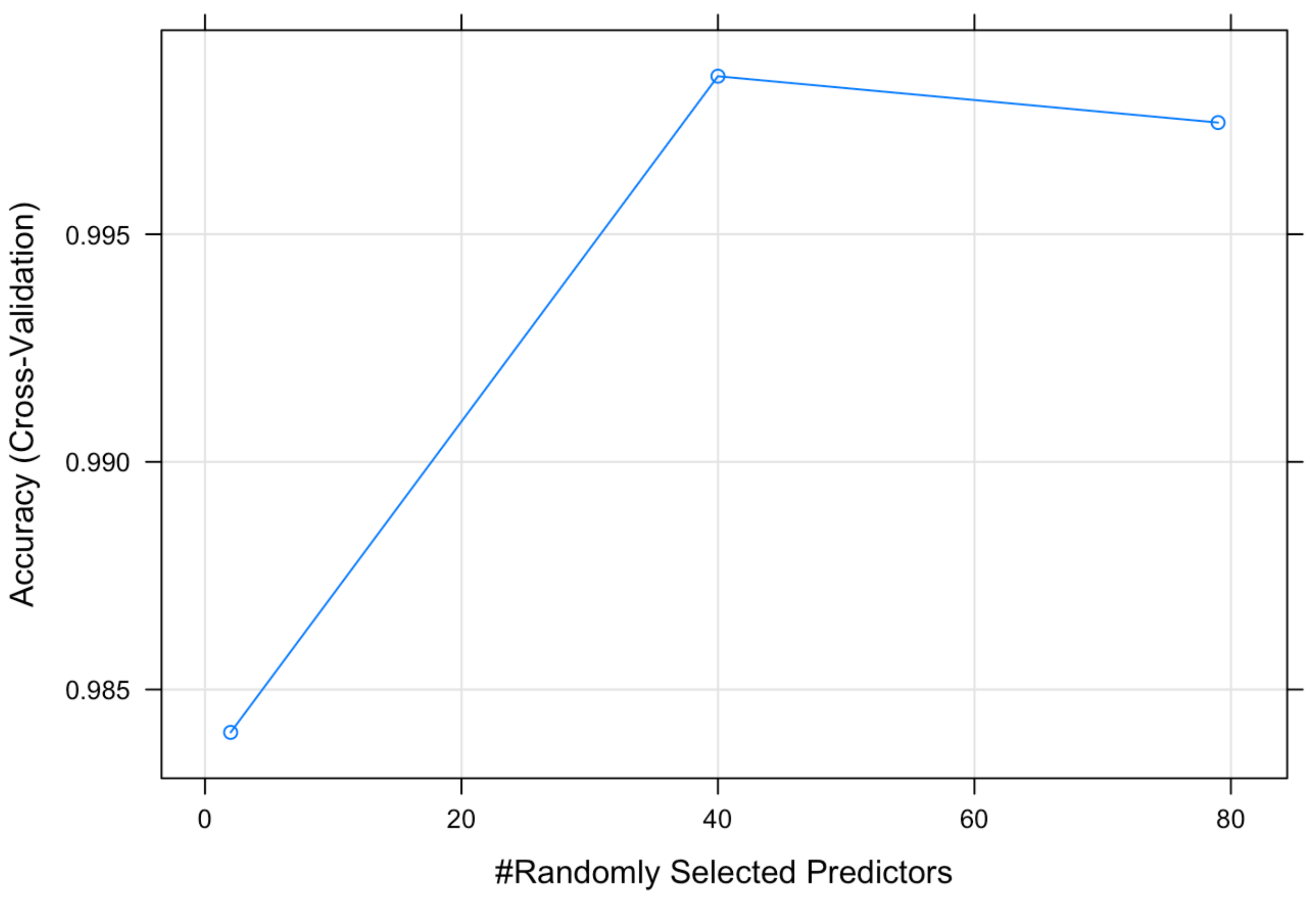
Predicting New Values

use trained model to predict data in testing set and check the result of prediction. The accuracy of this model is 0.9995

```
pred <- predict(modelfit2, testset); testset$predRight <- pred==t
estset$classe
table(pred, testset$classe)
```

```
##
## pred    A    B    C    D    E
##     A 1674    1    0    0    0
##     B    0 1138    0    0    0
##     C    0    0 1026    0    0
##     D    0    0    0  962    0
##     E    0    0    0    2 1082
```

```
plot(modelfit2)
```



```
predictrf <- predict(modelfit2, testset)
cmrf <- confusionMatrix(predictrf, testset$classe)
cmrf
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1674    1    0    0    0
##          B    0 1138    0    0    0
##          C    0    0 1026    0    0
##          D    0    0    0  962    0
##          E    0    0    0    2 1082
##
## Overall Statistics
##
##                Accuracy : 0.9995
##                  95% CI : (0.9985, 0.9999)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9994
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class
: E
## Sensitivity            1.0000   0.9991   1.0000   0.9979   1.0
000
## Specificity            0.9998   1.0000   1.0000   1.0000   0.9
996
## Pos Pred Value         0.9994   1.0000   1.0000   1.0000   0.9
982
## Neg Pred Value         1.0000   0.9998   1.0000   0.9996   1.0
000
## Prevalence             0.2845   0.1935   0.1743   0.1638   0.1
839
## Detection Rate         0.2845   0.1934   0.1743   0.1635   0.1
839
## Detection Prevalence   0.2846   0.1934   0.1743   0.1635   0.1
842
## Balanced Accuracy      0.9999   0.9996   1.0000   0.9990   0.9
998
```

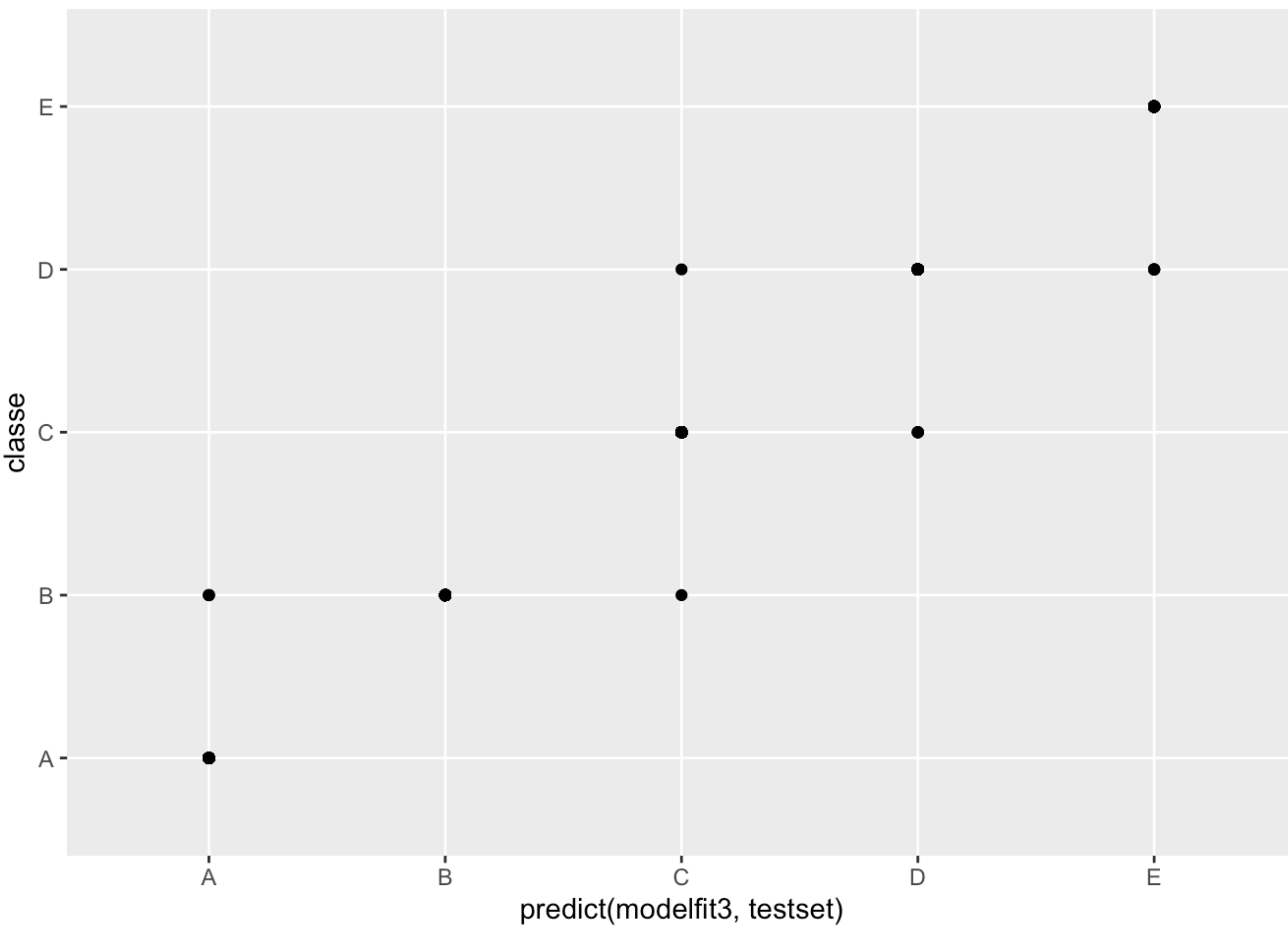
Boosting

use generalized boosted regression to train data in training set and to predict data in testing set. Plot the result of prediction in testing set and check the accuracy of this model(0.9978)

```
set.seed(11111)
controlgbm <- trainControl(method = "repeatedcv", number = 5, rep
eats = 1)
modelfit3 <- train(classe~., method = 'gbm', data = trainset, trC
ontrol = controlgbm, verbose = FALSE)
modelfit3$finalModel
```

```
## A gradient boosted model with multinomial loss function.
## 150 iterations were performed.
## There were 79 predictors of which 70 had non-zero influence.
```

```
qplot(predict(modelfit3, testset), classe, data = testset)
```




```
predictboost <- predict(modelfit3, testset)
cmboost <- confusionMatrix(predictboost, testset$classe)
cmboost
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1674    3    0    0    0
##          B    0 1135    0    0    0
##          C    0    1 1022    1    0
##          D    0    0    4  959    0
##          E    0    0    0    4 1082
##
## Overall Statistics
##
##                Accuracy : 0.9978
##                  95% CI : (0.9962, 0.9988)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9972
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class
: E
## Sensitivity            1.0000   0.9965   0.9961   0.9948   1.0
000
## Specificity            0.9993   1.0000   0.9996   0.9992   0.9
992
## Pos Pred Value         0.9982   1.0000   0.9980   0.9958   0.9
963
## Neg Pred Value         1.0000   0.9992   0.9992   0.9990   1.0
000
## Prevalence             0.2845   0.1935   0.1743   0.1638   0.1
839
## Detection Rate         0.2845   0.1929   0.1737   0.1630   0.1
839
## Detection Prevalence   0.2850   0.1929   0.1740   0.1636   0.1
845
## Balanced Accuracy      0.9996   0.9982   0.9978   0.9970   0.9
996
```

Getting Results from Valide Data

As the accuracy of three models above is 0.8743, 0.9995 and 0.9978 respectively. So use random forests algorithm to predict data in valid dataset

```
results <- predict(modelfit2, newdata = validdata)
results
```

```
##  [1] B A B A A E D B A A B C B A E E A B B B
## Levels: A B C D E
```