assignment - 3

by Aparajit .

FILE	ASSIGNMENT_-3.PDF (155.6K)		
TIME SUBMITTED	04-FEB-2017 09:24PM	WORD COUNT	3019
SUBMISSION ID	766612010	CHARACTER COUNT	10367

Assignment- 3

Aparajit

2/4/2017

```
#Answer 1
library(Hmisc)
```

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

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'Hmisc'
```

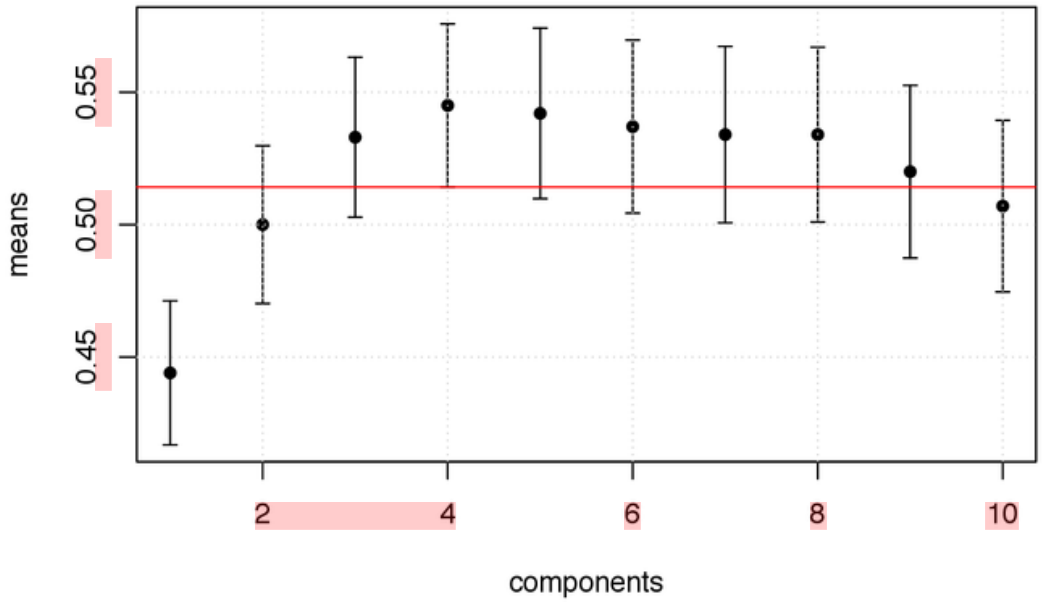
```
## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units
```

```
library(AppliedPredictiveModeling)
library(caret)
```

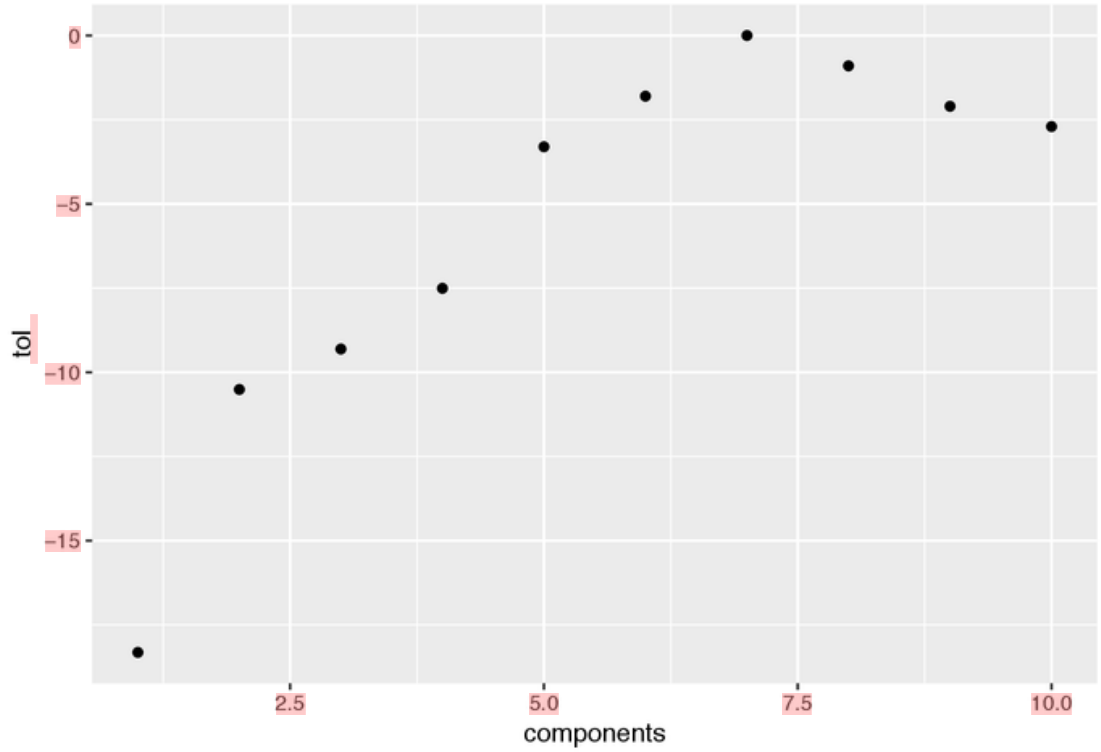
```
##
## Attaching package: 'caret'
```

```
## The following object is masked from 'package:survival':
##
##     cluster
```

```
data(ChemicalManufacturingProcess)
components = 1:10
means = c( 0.444, 0.500, 0.533, 0.545, 0.542, 0.537, 0.534, 0.534, 0.520, 0.507 )
std_errors = c( 0.0272, 0.0298, 0.0302, 0.0308, 0.0322, 0.0327, 0.0333, 0.0330, 0.0326, 0.0324 )
data = data.frame( components, means, std_errors )
errbar( components, means, means+std_errors, means-std_errors )
grid()
max_index = which.max( means )
abline( h=means[max_index] - std_errors[max_index], col='red' )
```




```
best <- subset(data,components==which.max(data$std_errors))
data$tol <- (data$std_errors-best$std_errors)/best$std_errors*100
qplot(components,tol,data = data)
```




```
## 1(a) Four components should be used to develope the most parsimonious model
```

```
## 1(b)for 10 % loss in R^2 values a two component model would work

##1(c) Basis the graph,we see that random forest has the highest R^2 so we will pick that

## 1(d) We would choose SVm because it has a reasonably fast processing time and has the second highest
```

```
library(caret)
data(oil)
str(oilType)
```

```
##  Factor w/ 7 levels "A","B","C","D",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```
table(oilType)
```

```
## oilType
##  A  B  C  D  E  F  G
## 37 26  3  7 11 10  2
```

```
samp <- floor(length(oilType)*.6) + 1
set.seed(629)
Split <- vector(mode = "list", length = 20)
for(i in seq(along = Split)) Split[[i]] <- table(sample(oilType, size = samp))
head(Split, 3)
```

```
## [[1]]
##
##  A  B  C  D  E  F  G
## 24 15  3  4  7  4  1
##
## [[2]]
##
##  A  B  C  D  E  F  G
## 25 14  2  4  6  6  1
##
## [[3]]
##
##  A  B  C  D  E  F  G
## 24 16  2  5  5  4  2
```

```
Split <- do.call("rbind", Split)
head(Split, 3)
```

```
##       A  B C D E F G
## [1,] 24 15 3 4 7 4 1
## [2,] 25 14 2 4 6 6 1
## [3,] 24 16 2 5 5 4 2
```

```
summary(Split/samp)
```

```
##        A                B                C                D
##  Min.   :0.3103   Min.   :0.2241   Min.   :0.00000   Min.   :0.03448
##  1st Qu.:0.3448   1st Qu.:0.2586   1st Qu.:0.01724   1st Qu.:0.06897
##  Median :0.3793   Median :0.2759   Median :0.03448   Median :0.08621
##  Mean   :0.3810   Mean   :0.2759   Mean   :0.03362   Mean   :0.08448
##  3rd Qu.:0.4138   3rd Qu.:0.2931   3rd Qu.:0.05172   3rd Qu.:0.10345
##  Max.   :0.4483   Max.   :0.3276   Max.   :0.05172   Max.   :0.12069
##        E                F                G
```

```
##  Min.   :0.05172   Min.   :0.03448   Min.   :0.00000
##  1st Qu.:0.10345   1st Qu.:0.08621   1st Qu.:0.01724
##  Median :0.10345   Median :0.10345   Median :0.01724
##  Mean   :0.10690   Mean   :0.09828   Mean   :0.01983
##  3rd Qu.:0.12069   3rd Qu.:0.10776   3rd Qu.:0.02155
##  Max.   :0.15517   Max.   :0.13793   Max.   :0.03448
```

```
set.seed(611)
Split2 <- createDataPartition(oilType, p = .60, times = 20)
Split2 <- lapply(Split2, function(x, y) table(y[x]), y = oilType)
head(Split2, 3)
```

```
## $Resample01
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
##
## $Resample02
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
##
## $Resample03
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
```

```
Split2 <- do.call("rbind",Split2)
summary(Split2)
```

```
##        A           B           C           D           E           F
##  Min.   :23   Min.   :16   Min.   :2   Min.   :5   Min.   :7   Min.   :6
##  1st Qu.:23   1st Qu.:16   1st Qu.:2   1st Qu.:5   1st Qu.:7   1st Qu.:6
##  Median :23   Median :16   Median :2   Median :5   Median :7   Median :6
##  Mean   :23   Mean   :16   Mean   :2   Mean   :5   Mean   :7   Mean   :6
##  3rd Qu.:23   3rd Qu.:16   3rd Qu.:2   3rd Qu.:5   3rd Qu.:7   3rd Qu.:6
##  Max.   :23   Max.   :16   Max.   :2   Max.   :5   Max.   :7   Max.   :6
##        G
##  Min.   :2
##  1st Qu.:2
##  Median :2
##  Mean   :2
##  3rd Qu.:2
##  Max.   :2
```

```
# a) We can see the summary using output command
# b) Stratified random sample has less variation and each partition has atleast one sample of each clas
# c)  For such a small sample size we should use K-fold cross validation technique to sample the data.
#     We will not use a test data in this case due to small data size
```

```
## 3.1
library(mlbench)
library(corrplot)
library(e1071)
```

```
##
## Attaching package: 'e1071'
```

```
## The following object is masked from 'package:Hmisc':
##
##     impute
```

```
library(caret)
library(AppliedPredictiveModeling)
data(Glass)
Glass_data <- subset(Glass,)
library(caTools)
set.seed(101)
sample <- sample.split(Glass_data$Si,SplitRatio = .75)
train_random <- subset(Glass_data,sample==TRUE)
test_random <- subset(Glass_data,sample==FALSE)
head(train_random)
```

```
##        RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 1 1.52101 13.64 4.49 1.10 71.78 0.06 8.75  0 0.00    1
## 2 1.51761 13.89 3.60 1.36 72.73 0.48 7.83  0 0.00    1
## 3 1.51618 13.53 3.55 1.54 72.99 0.39 7.78  0 0.00    1
## 4 1.51766 13.21 3.69 1.29 72.61 0.57 8.22  0 0.00    1
## 5 1.51742 13.27 3.62 1.24 73.08 0.55 8.07  0 0.00    1
## 6 1.51596 12.79 3.61 1.62 72.97 0.64 8.07  0 0.26    1
```

```
head(test_random)
```

```
##         RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 11 1.51571 12.72 3.46 1.56 73.20 0.67 8.09  0 0.24    1
## 14 1.51748 12.86 3.56 1.27 73.21 0.54 8.38  0 0.17    1
## 17 1.51784 12.68 3.67 1.16 73.11 0.61 8.70  0 0.00    1
## 22 1.51966 14.77 3.75 0.29 72.02 0.03 9.00  0 0.00    1
## 25 1.51720 13.38 3.50 1.15 72.85 0.50 8.43  0 0.00    1
## 26 1.51764 12.98 3.54 1.21 73.00 0.65 8.53  0 0.00    1
```

```
##3.2
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
```

```
##
## Attaching package: 'DAAG'
```

```
## The following object is masked from 'package:survival':
##
##     lung
```

```
set.seed(1)
cvsplit <- createFolds(Glass_data,k=3,returnTrain = TRUE)
str(cvsplit)
```

```
## List of 3
##  $ Fold1: int [1:7] 1 2 3 5 6 9 10
##  $ Fold2: int [1:6] 1 2 4 7 8 10
##  $ Fold3: int [1:7] 3 4 5 6 7 8 9
```

```r
fold1 <- cvsplit[[1]]
first <- Glass_data[,(cvsplit$Fold1)]
second <-Glass_data[,(cvsplit$Fold2)]
third <- Glass_data[,(cvsplit$Fold3)]
split <- lapply(cvsplit,function(ind,dat) dat[ind,],dat=Glass_data)
dim(Glass_data)
```

```
## [1] 214  10
```

```r
unlist(lapply(split,nrow))
```

```
## Fold1 Fold2 Fold3
##     7     6     7
```

```r
##3.3
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
bootstrap <- createResample(Glass_data,times=214,list=FALSE)
head.matrix(bootstrap)
```

```
##      Resample001 Resample002 Resample003 Resample004 Resample005
## [1,]           2           1           2           1           1
## [2,]           3           2           2           5           1
## [3,]           4           3           5           6           3
## [4,]           4           3           5           6           4
## [5,]           5           4           5           7           5
## [6,]           7           4           6           7           5
##      Resample006 Resample007 Resample008 Resample009 Resample010
## [1,]           1           4           2           1           2
## [2,]           3           4           2           3           3
## [3,]           3           4           3           5           3
## [4,]           4           4           3           5           4
## [5,]           5           5           4           7           5
## [6,]           5           8           4           7           6
##      Resample011 Resample012 Resample013 Resample014 Resample015
## [1,]           1           2           1           2           1
## [2,]           2           3           1           2           2
## [3,]           2           3           6           3           3
## [4,]           4           5           6           3           3
## [5,]           5           5           6           5           4
## [6,]           5           5           6           7           5
##      Resample016 Resample017 Resample018 Resample019 Resample020
## [1,]           1           4           2           2           2
## [2,]           3           4           2           2           2
## [3,]           4           7           3           3           3
## [4,]           4           7           6           4           3
## [5,]           5           7           6           4           3
## [6,]           5           8           8           6           3
##      Resample021 Resample022 Resample023 Resample024 Resample025
## [1,]           2           1           2           2           2
## [2,]           4           2           4           2           3
```

```
## [3,]         4         3         4         3         4
## [4,]         7         3         4         4         5
## [5,]         8         4         5         5         5
## [6,]         9         5         5         5         5
##      Resample026 Resample027 Resample028 Resample029 Resample030
## [1,]           1           1           1           1           1
## [2,]           2           2           1           1           1
## [3,]           2           2           2           2           2
## [4,]           3           2           2           2           4
## [5,]           4           2           2           3           5
## [6,]           6           3           3           4           5
##      Resample031 Resample032 Resample033 Resample034 Resample035
## [1,]           1           3           1           2           1
## [2,]           2           4           2           2           1
## [3,]           2           5           3           3           3
## [4,]           3           7           4           4           5
## [5,]           4           8           4           4           5
## [6,]           5           8           5           5           5
##      Resample036 Resample037 Resample038 Resample039 Resample040
## [1,]           1           2           2           2           1
## [2,]           2           3           3           2           2
## [3,]           3           3           3           2           4
## [4,]           4           4           5           2           7
## [5,]           5           5           6           4           8
## [6,]           6           5           7           4           8
##      Resample041 Resample042 Resample043 Resample044 Resample045
## [1,]           4           1           1           1           1
## [2,]           4           2           1           1           1
## [3,]           5           4           2           2           2
## [4,]           5           5           4           3           4
## [5,]           7           5           5           3           5
## [6,]           8           6           6           4           6
##      Resample046 Resample047 Resample048 Resample049 Resample050
## [1,]           1           2           1           1           5
## [2,]           2           3           2           1           5
## [3,]           2           4           2           2           6
## [4,]           3           5           2           3           6
## [5,]           5           5           6           4           7
## [6,]           5           5           6           5           7
##      Resample051 Resample052 Resample053 Resample054 Resample055
## [1,]           1           2           1           2           1
## [2,]           2           4           1           2           1
## [3,]           4           4           2           4           1
## [4,]           5           8           2           4           4
## [5,]           6           8           3           4           6
## [6,]           7           9           4           6           8
##      Resample056 Resample057 Resample058 Resample059 Resample060
## [1,]           2           1           1           3           2
## [2,]           3           2           1           4           2
## [3,]           3           3           4           5           4
## [4,]           4           5           5           5           4
## [5,]           7           5           6           5           8
## [6,]           7           6           6           6           9
##      Resample061 Resample062 Resample063 Resample064 Resample065
```

```
## [1,]           1           2           1           1           1
## [2,]           2           3           1           3           2
## [3,]           4           3           4           3           4
## [4,]           4           4           4           4           4
## [5,]           5           5           4           7           6
## [6,]           6           6           5           8           6
##      Resample066 Resample067 Resample068 Resample069 Resample070
## [1,]           1           2           1           1           1
## [2,]           2           3           2           2           5
## [3,]           4           3           4           5           6
## [4,]           4           4           5           5           7
## [5,]           4           5           6           6           7
## [6,]           6           6           6           6           8
##      Resample071 Resample072 Resample073 Resample074 Resample075
## [1,]           1           1           1           2           2
## [2,]           2           1           1           3           2
## [3,]           2           3           1           3           3
## [4,]           5           4           3           4           5
## [5,]           5           5           3           4           5
## [6,]           5           6           6           5           5
##      Resample076 Resample077 Resample078 Resample079 Resample080
## [1,]           2           1           1           1           1
## [2,]           2           1           2           2           1
## [3,]           2           3           3           3           3
## [4,]           3           5           5           4           4
## [5,]           6           5           5           4           4
## [6,]           8           7           7           6           7
##      Resample081 Resample082 Resample083 Resample084 Resample085
## [1,]           3           1           1           1           1
## [2,]           4           3           5           2           1
## [3,]           4           4           6           3           3
## [4,]           4           4           6           4           3
## [5,]           4           4           7           5           3
## [6,]           6           6           8           5           3
##      Resample086 Resample087 Resample088 Resample089 Resample090
## [1,]           1           1           4           1           1
## [2,]           1           3           4           1           1
## [3,]           1           3           5           1           1
## [4,]           1           5           6           1           1
## [5,]           3           5           7           1           2
## [6,]           4           5           8           1           2
##      Resample091 Resample092 Resample093 Resample094 Resample095
## [1,]           1           1           3           1           1
## [2,]           1           3           3           3           1
## [3,]           2           4           4           5           2
## [4,]           5           5           5           5           2
## [5,]           6           6           7           5           3
## [6,]           7           8           8           7           5
##      Resample096 Resample097 Resample098 Resample099 Resample10[1]
## [1,]           2           1           2           2           1
## [2,]           3           1           2           2           2
## [3,]           4           2           2           2           4
## [4,]           4           5           4           3           6
## [5,]           4           6           4           3           6
```

```
## [6,]    5           7           5           3           6
##       Resample101 Resample102 Resample103 Resample104 Resample105
## [1,]           1           1           2           2           1
## [2,]           2           2           2           2           1
## [3,]           3           2           4           3           2
## [4,]           3           3           4           4           5
## [5,]           5           4           5           5           5
## [6,]           6           6           6           5           6
##       Resample106 Resample107 Resample108 Resample109 Resample110
## [1,]           2           1           1           1           2
## [2,]           2           1           2           3           2
## [3,]           2           3           4           3           3
## [4,]           4           4           4           5           4
## [5,]           5           4           5           6           4
## [6,]           6           5           6           7           5
##       Resample111 Resample112 Resample113 Resample114 Resample115
## [1,]           1           2           2           2           1
## [2,]           3           3           2           2           1
## [3,]           4           3           3           2           2
## [4,]           4           4           5           3           3
## [5,]           6           5           5           3           3
## [6,]           7           5           5           4           4
##       Resample116 Resample117 Resample118 Resample119 Resample120
## [1,]           1           1           1           1           4
## [2,]           2           1           1           3           5
## [3,]           3           3           2           3           7
## [4,]           3           5           2           4           7
## [5,]           4           5           4           7           8
## [6,]           4           7           4           8           8
##       Resample121 Resample122 Resample123 Resample124 Resample125
## [1,]           1           1           1           1           1
## [2,]           1           1           1           1           2
## [3,]           1           2           4           3           2
## [4,]           2           4           4           4           2
## [5,]           3           4           4           6           4
## [6,]           5           6           5           8           5
##       Resample126 Resample127 Resample128 Resample129 Resample130
## [1,]           1           2           1           1           1
## [2,]           2           3           3           1           2
## [3,]           3           4           3           1           3
## [4,]           4           7           4           4           5
## [5,]           4           7           4           4           6
## [6,]           4           8           5           5           6
##       Resample131 Resample132 Resample133 Resample134 Resample135
## [1,]           1           1           1           1           1
## [2,]           2           1           3           2           1
## [3,]           2           2           4           2           2
## [4,]           5           2           5           2           2
## [5,]           5           2           6           2           2
## [6,]           6           4           7           3           3
##       Resample136 Resample137 Resample138 Resample139 Resample140
## [1,]           1           2           1           1           1
## [2,]           2           2           3           3           2
## [3,]           2           3           3           3           2
```

```
## [4,]           4           4           4           4           3
## [5,]           4           4           5           6           4
## [6,]           6           6           6           8           5
##      Resample141 Resample142 Resample143 Resample144 Resample145
## [1,]           1           1           1           2           1
## [2,]           1           2           3           3           1
## [3,]           4           3           4           3           2
## [4,]           4           4           4           4           3
## [5,]           4           5           5           5           4
## [6,]           5           6           6           5           6
##      Resample146 Resample147 Resample148 Resample149 Resample150
## [1,]           2           2           1           1           1
## [2,]           2           4           1           2           1
## [3,]           2           5           1           3           2
## [4,]           4           5           2           5           2
## [5,]           4           5           4           7           4
## [6,]           4           7           5           7           6
##      Resample151 Resample152 Resample153 Resample154 Resample155
## [1,]           1           2           1           1           1
## [2,]           3           5           2           1           2
## [3,]           4           5           2           4           3
## [4,]           5           6           3           5           4
## [5,]           5           8           6           5           6
## [6,]           7           9           6           5           8
##      Resample156 Resample157 Resample158 Resample159 Resample160
## [1,]           1           1           5           1           1
## [2,]           2           2           5           1           1
## [3,]           3           3           5           2           2
## [4,]           4           5           6           2           2
## [5,]           7           6           6           2           2
## [6,]           8           7           7           2           3
##      Resample161 Resample162 Resample163 Resample164 Resample165
## [1,]           1           1           1           3           1
## [2,]           4           1           1           4           1
## [3,]           4           5           1           4           2
## [4,]           5           6           1           5           3
## [5,]           7           7           1           6           4
## [6,]           8           8           4           6           5
##      Resample166 Resample167 Resample168 Resample169 Resample170
## [1,]           3           2           1           1           1
## [2,]           3           2           4           1           2
## [3,]           3           2           4           2           2
## [4,]           5           2           7           3           2
## [5,]           5           7           9           6           3
## [6,]           6           8           9           6           4
##      Resample171 Resample172 Resample173 Resample174 Resample175
## [1,]           1           3           1           1           1
## [2,]           2           4           1           3           3
## [3,]           2           4           1           4           4
## [4,]           3           5           5           4           5
## [5,]           3           7           5           5           6
## [6,]           3           8           6           5           6
##      Resample176 Resample177 Resample178 Resample179 Resample180
## [1,]           2           1           2           1           1
```

```
## [2,]           4           1           2           2           2
## [3,]           5           2           3           2           4
## [4,]           5           2           3           3           4
## [5,]           6           3           4           3           5
## [6,] 1         6           3           4           5           6
##      Resample181 Resample182 Resample183 Resample184 Resample185
## [1,]           1           1           1           1           2
## [2,]           2           4           3           2           2
## [3,]           2           5           5           2           2
## [4,]           2           6           5           5           5
## [5,]           3           6           6           7           7
## [6,]           6           6           6           7           7
##      Resample186 Resample187 Resample188 Resample189 Resample190
## [1,]           1           1           2           1           1
## [2,]           1           2           2           2           4
## [3,]           3           3           3           4           4
## [4,]           5           3           3           4           5
## [5,]           5           4           4           5           5
## [6,]           6           6           5           7           6
##      Resample191 Resample192 Resample193 Resample194 Resample195
## [1,]           1           1           1           1           1
## [2,]           1           1           1           1           2
## [3,]           1           3           2           1           2
## [4,]           2           3           4           2           3
## [5,]           4           5           5           3           3
## [6,]           5           7           5           3           5
##      Resample196 Resample197 Resample198 Resample199 Resample200
## [1,]           2           1           3           1           1
## [2,]           3           3           4           2           1
## [3,]           4           3           4           2           2
## [4,]           6           5           5           3           5
## [5,]           7           5           6           3           5
## [6,]           7           6           6           4           6
##      Resample201 Resample202 Resample203 Resample204 Resample205
## [1,]           1           1           1           1           1
## [2,]           1           1           1           2           2
## [3,]           1           1           2           2           2
## [4,]           2           1           3           4           2
## [5,]           4           3           3           4           3
## [6,]           4           5           3           7           3
##      Resample206 Resample207 Resample208 Resample209 Resample210
## [1,]           1           1           2           1           1
## [2,]           3           3           2           2           2
## [3,]           3           5           4           6           3
## [4,]           4           5           6           6           6
## [5,]           5           5           6           6           6
## [6,]           6           7           6           7           7
##      Resample211 Resample212 Resample213 Resample214
## [1,]           2           1           1           2
## [2,]           3           1           1           2
## [3,]           4           1           3           2
## [4,]           4           1           4           3
## [5,]           5           3           4           4
## [6,]           9           3           5           5
```

assignment - 3

ORIGINALITY REPORT

%92	%13	%4	%92
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to UT, Dallas Student Paper	%92

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	OFF		

homework3

by Jasmine Sethi

FILE	HOMEWORK3.PDF (546.83K)		
TIME SUBMITTED	04-FEB-2017 10:08PM	WORD COUNT	3289
SUBMISSION ID	766621079	CHARACTER COUNT	11560

homework3

jasmine sethi

February 4, 2017

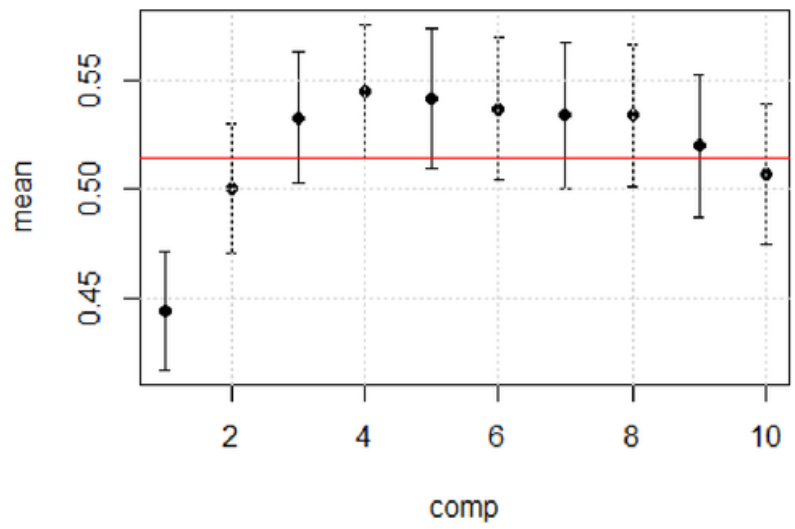
R Markdown

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.

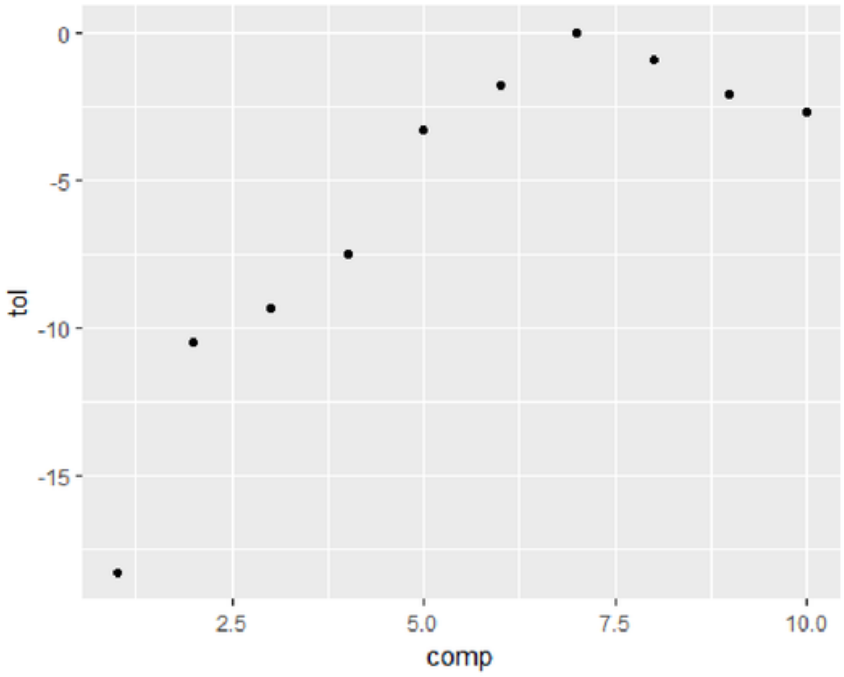
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
#Answer 1
library(caret)

## Loading required package: lattice

## Loading required package: ggplot2

library(Hmisc)

## Loading required package: survival

##
## Attaching package: 'survival'

## The following object is masked from 'package:caret':
##
##     cluster

## Loading required package: Formula

##
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units

library(AppliedPredictiveModeling)
data(ChemicalManufacturingProcess)
comp = 1:10
mean = c( 0.444, 0.500, 0.533, 0.545, 0.542, 0.537, 0.534, 0.534, 0.520, 0.50
7 )
std_err = c( 0.0272, 0.0298, 0.0302, 0.0308, 0.0322, 0.0327, 0.0333, 0.0330,
0.0326, 0.0324 )
```

```
x = data.frame( comp, mean, std_err )
errbar( comp, mean, mean+std_err, mean-std_err )
grid()
max_index = which.max( mean )
abline( h=mean[max_index] - std_err[max_index], col='red' )
```




```
y <- subset(x,comp==which.max(x$std_err))
x$tol <- (x$std_err-y$std_err)/y$std_err*100
qplot(comp,tol,data = x)
```




```
## (a) Using the"one-standard error"method, what number of PLS components pro
vides the most parsimonious model?
##1(a) 4 components

##(b) Compute the tolerance values for this example. If a 10% loss in R2 is a
cceptable, then what is the optimal number of PLS components?
## 1(b)if we consider 10% loss in R2 is acceptable, then two components are o
ptimal.

##(c) Several other models (discussed in Part II) with varying degrees of com
plexity were trained and tuned and the results are presented in Fig.4.13. If
the goal is to select the model that optimizes R2, then which model(s) would
you choose, and why?
##1(c) Random forest will be chosen as it has the highest R2

##(d) Prediction time, as well as model complexity (Sect.4.8) are other facto
rs to consider when selecting the optimal model(s). Given each model's predic
tion time, model complexity, and R2 estimates, which model(s) would you choos
e, and why?
## 1(d) We should choose SVm model as seen it has one of the highest r2 value
as well as good processing time

#Answer 2
##2(a) Use the sample function in base R to create a completely random sample
of 60 oils. How closely do the frequencies of the random sample match the ori
ginal samples? Repeat this procedure several times of understand the variatio
```

n in the sampling process.(b) Use the caret package function createDataPartit
ion to create a strati???ed random sample. How does this compare to the compl
etely random samples? 92 4 Over-Fitting and Model Tuning(c) With such a small
samples size, what are the options for determining performance of the model?
Should a test set be used?

##a) Summary can be seen using the output
*# b) Among different sampling techniques stratified random sample has the lea
st variability and each partition will have a sample of each class*

1

```
library(caret)
data(oil)
str(oilType)
```

```
##  Factor w/ 7 levels "A","B","C","D",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```
table(oilType)
```

```
## oilType
##  A  B  C  D  E  F  G
## 37 26  3  7 11 10  2
```

```
x <- floor(length(oilType)*.6) + 1
set.seed(629)
y <- vector(mode = "list", length = 20)
for(i in seq(along = y)) y[[i]] <- table(sample(oilType, size = x))
head(y, 3)
```

```
## [[1]]
##
##  A  B  C  D  E  F  G
## 24 15  3  4  7  4  1
##
## [[2]]
##
##  A  B  C  D  E  F  G
## 25 14  2  4  6  6  1
##
## [[3]]
##
##  A  B  C  D  E  F  G
## 24 16  2  5  5  4  2
```

```
y <- do.call("rbind", y)
head(y, 3)
```

5

```
##       A  B C D E F G
## [1,] 24 15 3 4 7 4 1
```

```
## [2,] 25 14 2 4 6 6 1
## [3,] 24 16 2 5 5 4 2

summary(y/x)

##        A                B                C                D         
##  Min.   :0.3103   Min.   :0.2241   Min.   :0.00000   Min.   :0.03448  
##  1st Qu.:0.3448   1st Qu.:0.2586   1st Qu.:0.01724   1st Qu.:0.06897  
##  Median :0.3793   Median :0.2759   Median :0.03448   Median :0.08621  
##  Mean   :0.3810   Mean   :0.2759   Mean   :0.03362   Mean   :0.08448  
##  3rd Qu.:0.4138   3rd Qu.:0.2931   3rd Qu.:0.05172   3rd Qu.:0.10345  
##  Max.   :0.4483   Max.   :0.3276   Max.   :0.05172   Max.   :0.12069  
##        E                F                G         
##  Min.   :0.05172   Min.   :0.03448   Min.   :0.00000  
##  1st Qu.:0.10345   1st Qu.:0.08621   1st Qu.:0.01724  
##  Median :0.10345   Median :0.10345   Median :0.01724  
##  Mean   :0.10690   Mean   :0.09828   Mean   :0.01983  
##  3rd Qu.:0.12069   3rd Qu.:0.10776   3rd Qu.:0.02155  
##  Max.   :0.15517   Max.   :0.13793   Max.   :0.03448  

set.seed(629)
y2 <- createDataPartition(oilType, p = .60, times = 20)
y2 <- lapply(y2, function(x, y) table(y[x]), y = oilType)
head(y2, 3)

## $Resample01
##
##  A  B  C  D  E  F  G 
## 23 16  2  5  7  6  2 
##
## $Resample02
##
##  A  B  C  D  E  F  G 
## 23 16  2  5  7  6  2 
##
## $Resample03
##
##  A  B  C  D  E  F  G 
## 23 16  2  5  7  6  2 

y2 <- do.call("rbind",y2)
summary(y2)

##        A            B            C            D            E            F       
##  Min.   :23   Min.   :16   Min.   :2   Min.   :5   Min.   :7   Min.   :6  
##  1st Qu.:23   1st Qu.:16   1st Qu.:2   1st Qu.:5   1st Qu.:7   1st Qu.:6  
##  Median :23   Median :16   Median :2   Median :5   Median :7   Median :6  
##  Mean   :23   Mean   :16   Mean   :2   Mean   :5   Mean   :7   Mean   :6  
##  3rd Qu.:23   3rd Qu.:16   3rd Qu.:2   3rd Qu.:5   3rd Qu.:7   3rd Qu.:6  
##  Max.   :23   Max.   :16   Max.   :2   Max.   :5   Max.   :7   Max.   :6  
##        G
```

Assignment3

by Shivam Tiwari

FILE	ASSIGNMENT3.PDF (403.16K)		
TIME SUBMITTED	04-FEB-2017 10:04PM	WORD COUNT	3010
SUBMISSION ID	766620267	CHARACTER COUNT	10259

Assignment- 3

Shivam

2/4/2017

1

```
#Answer 1
library(Hmisc)
```

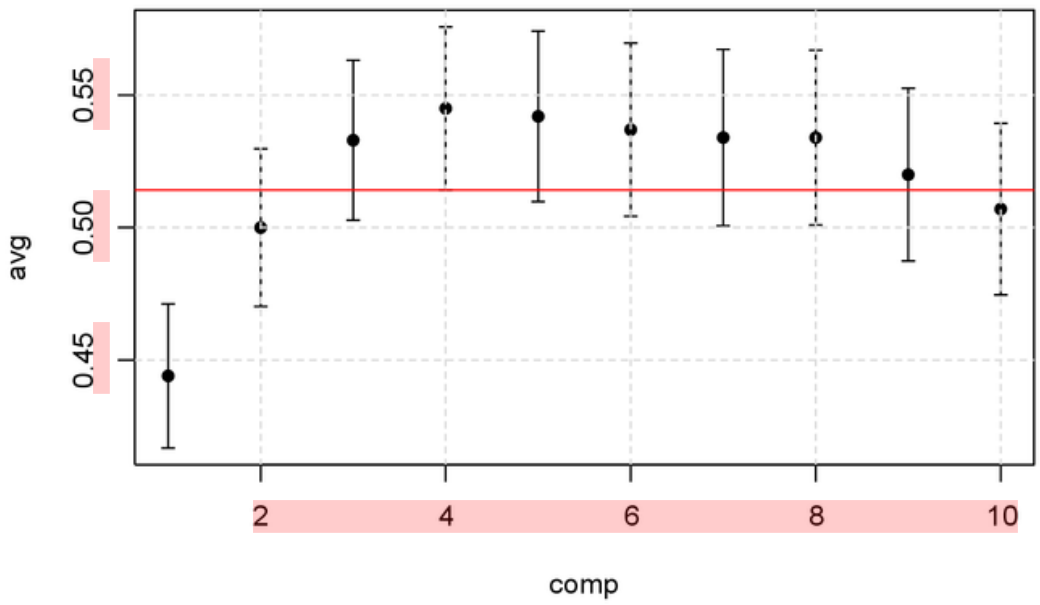
```
## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
## Loading required package: ggplot2
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units
```

```
library(AppliedPredictiveModeling)
library(caret)
```

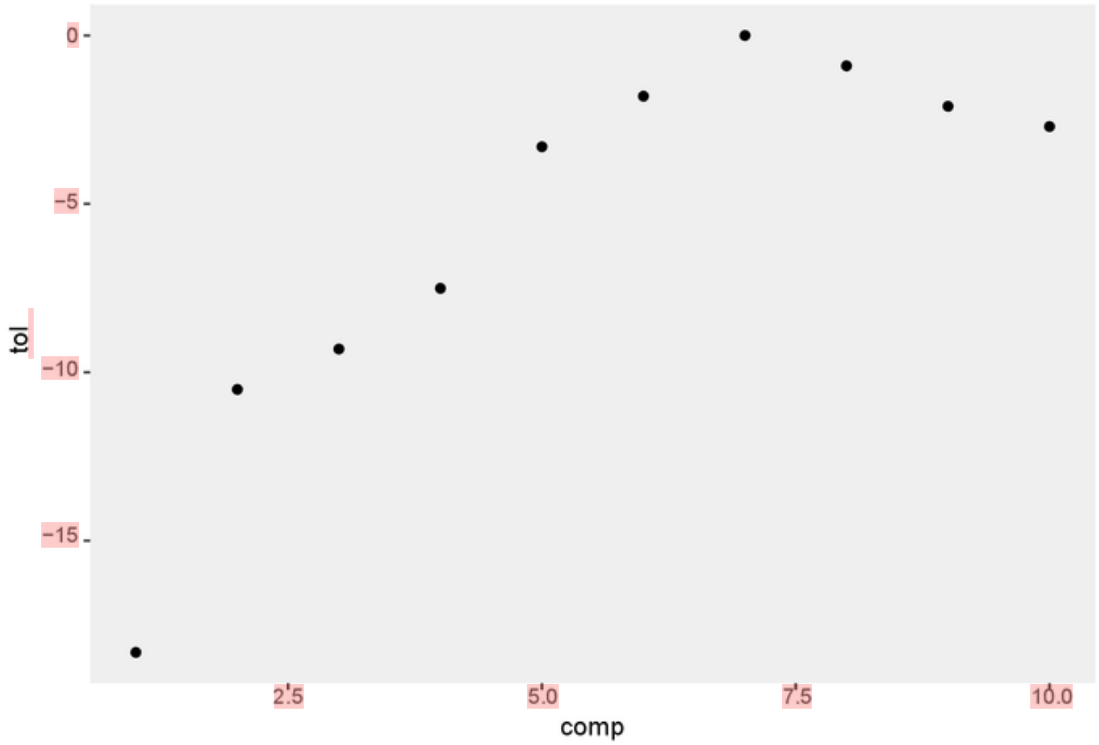
```
##
## Attaching package: 'caret'
## The following object is masked from 'package:survival':
##
##     cluster
```

```
data(ChemicalManufacturingProcess)
comp = 1:10
avg = c( 0.444, 0.500, 0.533, 0.545, 0.542, 0.537, 0.534, 0.534, 0.520, 0.507 )
std_error = c( 0.0272, 0.0298, 0.0302, 0.0308, 0.0322, 0.0327, 0.0333, 0.0330, 0.0326, 0.0324 )
data = data.frame( comp, avg, std_error )
errbar( comp, avg, avg+std_error, avg-std_error )
grid()
max1 = which.max( avg )
abline( h=avg[max1] - std_error[max1], col='red' )
```

1




```
best1 <- subset(data,comp==which.max(data$std_error))
data$tol <- (data$std_error-best$std_error)/best1$std_error*100
qplot(comp,tol,data = data)
```




```
## 1(a) Four components should be used to develop the most parsimonious model
```

```
## 1(b)for 10 % loss in R^2 values a two component model would work

##1(c) Basis the graph,we see that random forest has the highest R^2 so we will pick that

## 1(d) We would choose SVm because it has a reasonably fast processing time and has the second highest
```

```
library(caret)
data(oil)
str(oilType)
```

```
##  Factor w/ 7 levels "A","B","C","D",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```
table(oilType)
```

```
## oilType
##  A  B  C  D  E  F  G
## 37 26  3  7 11 10  2
```

```
sample <- floor(length(oilType)*.6)  + 1
set.seed(629)
Splits <- vector(mode = "list", length = 20)
for(x in seq(along = Splits)) Splits[[x]] <- table(sample(oilType, size = sample))
head(Splits, 3)
```

```
## [[1]]
##
##  A  B  C  D  E  F  G
## 24 15  3  4  7  4  1
##
## [[2]]
##
##  A  B  C  D  E  F  G
## 25 14  2  4  6  6  1
##
## [[3]]
##
##  A  B  C  D  E  F  G
## 24 16  2  5  5  4  2
```

```
Splits <- do.call("rbind", Splits)
head(Splits, 3)
```

```
##      A  B C D E F G
## [1,] 24 15 3 4 7 4 1
## [2,] 25 14 2 4 6 6 1
## [3,] 24 16 2 5 5 4 2
```

```
summary(Splits/sample)
```

```
##        A                B                C                D
##  Min.   :0.3103   Min.   :0.2241   Min.   :0.00000   Min.   :0.03448
##  1st Qu.:0.3448   1st Qu.:0.2586   1st Qu.:0.01724   1st Qu.:0.06897
##  Median :0.3793   Median :0.2759   Median :0.03448   Median :0.08621
##  Mean   :0.3810   Mean   :0.2759   Mean   :0.03362   Mean   :0.08448
##  3rd Qu.:0.4138   3rd Qu.:0.2931   3rd Qu.:0.05172   3rd Qu.:0.10345
##  Max.   :0.4483   Max.   :0.3276   Max.   :0.05172   Max.   :0.12069
##        E                F                G
```

```
##  Min.   :0.05172   Min.   :0.03448   Min.   :0.00000
##  1st Qu.:0.10345   1st Qu.:0.08621   1st Qu.:0.01724
##  Median :0.10345   Median :0.10345   Median :0.01724
##  Mean   :0.10690   Mean   :0.09828   Mean   :0.01983
##  3rd Qu.:0.12069   3rd Qu.:0.10776   3rd Qu.:0.02155
##  Max.   :0.15517   Max.   :0.13793   Max.   :0.03448
```

```
set.seed(62)
Splits2 <- createDataPartition(oilType,  p = .60,  times = 20)
Splits2 <- lapply(Splits2,  function(i,  y) table(y[i]),  y = oilType)
head(Splits2,  3)
```

```
## $Resample01
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
##
## $Resample02
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
##
## $Resample03
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
```

```
Splits2 <- do.call("rbind",Splits2)
summary(Splits2)
```

```
##        A           B           C          D          E          F
##  Min.   :23   Min.   :16   Min.   :2   Min.   :5   Min.   :7   Min.   :6
##  1st Qu.:23   1st Qu.:16   1st Qu.:2   1st Qu.:5   1st Qu.:7   1st Qu.:6
##  Median :23   Median :16   Median :2   Median :5   Median :7   Median :6
##  Mean   :23   Mean   :16   Mean   :2   Mean   :5   Mean   :7   Mean   :6
##  3rd Qu.:23   3rd Qu.:16   3rd Qu.:2   3rd Qu.:5   3rd Qu.:7   3rd Qu.:6
##  Max.   :23   Max.   :16   Max.   :2   Max.   :5   Max.   :7   Max.   :6
##        G
##  Min.   :2
##  1st Qu.:2
##  Median :2
##  Mean   :2
##  3rd Qu.:2
##  Max.   :2
```

```
# a) We can see the summary using output command
# b) Stratified random sample has less variation and each partition has atleast one sample of each clas
# c)  For such a small sample size we should use K-fold cross validation technique to sample the data.
#     We will not use a test data in this case due to small data size
```

```
## 3.1
library(mlbench)
library(corrplot)
library(e1071)
```

```
##
## Attaching package: 'e1071'
```

```
## The following object is masked from 'package:Hmisc':
##
##     impute
```

```
library(caret)
library(AppliedPredictiveModeling)
data(Glass)
Glass_data <- subset(Glass,)
library(caTools)
set.seed(101)
sample <- sample.split(Glass_data$Si,SplitRatio = .75)
train <- subset(Glass_data,sample==TRUE)
test <- subset(Glass_data,sample==FALSE)
head(train)
```

```
##        RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 1 1.52101 13.64 4.49 1.10 71.78 0.06 8.75  0 0.00    1
## 2 1.51761 13.89 3.60 1.36 72.73 0.48 7.83  0 0.00    1
## 3 1.51618 13.53 3.55 1.54 72.99 0.39 7.78  0 0.00    1
## 4 1.51766 13.21 3.69 1.29 72.61 0.57 8.22  0 0.00    1
## 5 1.51742 13.27 3.62 1.24 73.08 0.55 8.07  0 0.00    1
## 6 1.51596 12.79 3.61 1.62 72.97 0.64 8.07  0 0.26    1
```

```
head(test)
```

```
##         RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 11 1.51571 12.72 3.46 1.56 73.20 0.67 8.09  0 0.24    1
## 14 1.51748 12.86 3.56 1.27 73.21 0.54 8.38  0 0.17    1
## 17 1.51784 12.68 3.67 1.16 73.11 0.61 8.70  0 0.00    1
## 22 1.51966 14.77 3.75 0.29 72.02 0.03 9.00  0 0.00    1
## 25 1.51720 13.38 3.50 1.15 72.85 0.50 8.43  0 0.00    1
## 26 1.51764 12.98 3.54 1.21 73.00 0.65 8.53  0 0.00    1
```

```
##3.2
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
```

```
##
## Attaching package: 'DAAG'
```

```
## The following object is masked from 'package:survival':
##
##     lung
```

```
set.seed(1)
splitx <- createFolds(Glass_data,k=3,returnTrain = TRUE)
str(splitx)
```

```
## List of 3
##  $ Fold1: int [1:7] 1 2 3 5 6 9 10
##  $ Fold2: int [1:6] 1 2 4 7 8 10
##  $ Fold3: int [1:7] 3 4 5 6 7 8 9
```

```
fold1 <- splitx[[1]]
first  <-  Glass_data[,(splitx$Fold1)]
second   <-Glass_data[,(splitx$Fold2)]
third <- Glass_data[,(splitx$Fold3)]
split <- lapply(splitx,function(ind,dat) dat[ind,],dat=Glass_data)
dim(Glass_data)
```

```
## [1] 214  10
```

```
unlist(lapply(split,nrow))
```

```
## Fold1 Fold2 Fold3
##     7     6     7
```

```
##3.3
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
boot <- createResample(Glass_data,times=214,list=FALSE)
head.matrix(boot)
```

```
##      Resample001 Resample002 Resample003 Resample004 Resample005
## [1,]           2           1           2           1           1
## [2,]           3           2           2           5           1
## [3,]           4           3           5           6           3
## [4,]           4           3           5           6           4
## [5,]           5           4           5           7           5
## [6,]           7           4           6           7           5
##      Resample006 Resample007 Resample008 Resample009 Resample010
## [1,]           1           4           2           1           2
## [2,]           3           4           2           3           3
## [3,]           3           4           3           5           3
## [4,]           4           4           3           5           4
## [5,]           5           5           4           7           5
## [6,]           5           8           4           7           6
##      Resample011 Resample012 Resample013 Resample014 Resample015
## [1,]           1           2           1           2           1
## [2,]           2           3           1           2           2
## [3,]           2           3           6           3           3
## [4,]           4           5           6           3           3
## [5,]           5           5           6           5           4
## [6,]           5           5           6           7           5
##      Resample016 Resample017 Resample018 Resample019 Resample020
## [1,]           1           4           2           2           2
## [2,]           3           4           2           2           2
## [3,]           4           7           3           3           3
## [4,]           4           7           6           4           3
## [5,]           5           7           6           4           3
## [6,]           5           8           8           6           3
##      Resample021 Resample022 Resample023 Resample024 Resample025
## [1,]           2           1           2           2           2
## [2,]           4           2           4           2           3
```

```
## [3,]        4         3         4         3         4
## [4,]        7         3         4         4         5
## [5,]        8         4         5         5         5
## [6,]        9         5         5         5         5
##      Resample026 Resample027 Resample028 Resample029 Resample030
## [1,]           1           1           1           1           1
## [2,]           2           2           1           1           1
## [3,]           2           2           2           2           2
## [4,]           3           2           2           2           4
## [5,]           4           2           2           3           5
## [6,]           6           3           3           4           5
##      Resample031 Resample032 Resample033 Resample034 Resample035
## [1,]           1           3           1           2           1
## [2,]           2           4           2           2           1
## [3,]           2           5           3           3           3
## [4,]           3           7           4           4           5
## [5,]           4           8           4           4           5
## [6,]           5           8           5           5           5
##      Resample036 Resample037 Resample038 Resample039 Resample040
## [1,]           1           2           2           2           1
## [2,]           2           3           3           2           2
## [3,]           3           3           3           2           4
## [4,]           4           4           5           2           7
## [5,]           5           5           6           4           8
## [6,]           6           5           7           4           8
##      Resample041 Resample042 Resample043 Resample044 Resample045
## [1,]           4           1           1           1           1
## [2,]           4           2           1           1           1
## [3,]           5           4           2           2           2
## [4,]           5           5           4           3           4
## [5,]           7           5           5           3           5
## [6,]           8           6           6           4           6
##      Resample046 Resample047 Resample048 Resample049 Resample050
## [1,]           1           2           1           1           5
## [2,]           2           3           2           1           5
## [3,]           2           4           2           2           6
## [4,]           3           5           2           3           6
## [5,]           5           5           6           4           7
## [6,]           5           5           6           5           7
##      Resample051 Resample052 Resample053 Resample054 Resample055
## [1,]           1           2           1           2           1
## [2,]           2           4           1           2           1
## [3,]           4           4           2           4           1
## [4,]           5           8           2           4           4
## [5,]           6           8           3           4           6
## [6,]           7           9           4           6           8
##      Resample056 Resample057 Resample058 Resample059 Resample060
## [1,]           2           1           1           3           2
## [2,]           3           2           1           4           2
## [3,]           3           3           4           5           4
## [4,]           4           5           5           5           4
## [5,]           7           5           6           5           8
## [6,]           7           6           6           6           9
##      Resample061 Resample062 Resample063 Resample064 Resample065
```

```
## [1,]        1        2        1        1        1
## [2,]        2        3        1        3        2
## [3,]        4        3        4        3        4
## [4,]        4        4        4        4        4
## [5,]        5        5        4        7        6
## [6,]        6        6        5        8        6
##      Resample066 Resample067 Resample068 Resample069 Resample070
## [1,]        1        2        1        1        1
## [2,]        2        3        2        2        5
## [3,]        4        3        4        5        6
## [4,]        4        4        5        5        7
## [5,]        4        5        6        6        7
## [6,]        6        6        6        6        8
##      Resample071 Resample072 Resample073 Resample074 Resample075
## [1,]        1        1        1        2        2
## [2,]        2        1        1        3        2
## [3,]        2        3        1        3        3
## [4,]        5        4        3        4        5
## [5,]        5        5        3        4        5
## [6,]        5        6        6        5        5
##      Resample076 Resample077 Resample078 Resample079 Resample080
## [1,]        2        1        1        1        1
## [2,]        2        1        2        2        1
## [3,]        2        3        3        3        3
## [4,]        3        5        5        4        4
## [5,]        6        5        5        4        4
## [6,]        8        7        7        6        7
##      Resample081 Resample082 Resample083 Resample084 Resample085
## [1,]        3        1        1        1        1
## [2,]        4        3        5        2        1
## [3,]        4        4        6        3        3
## [4,]        4        4        6        4        3
## [5,]        4        4        7        5        3
## [6,]        6        6        8        5        3
##      Resample086 Resample087 Resample088 Resample089 Resample090
## [1,]        1        1        4        1        1
## [2,]        1        3        4        1        1
## [3,]        1        3        5        1        1
## [4,]        1        5        6        1        1
## [5,]        3        5        7        1        2
## [6,]        4        5        8        1        2
##      Resample091 Resample092 Resample093 Resample094 Resample095
## [1,]        1        1        3        1        1
## [2,]        1        3        3        3        1
## [3,]        2        4        4        5        2
## [4,]        5        5        5        5        2
## [5,]        6        6        7        5        3
## [6,]        7        8        8        7        5
##      Resample096 Resample097 Resample098 Resample099 Resample100
## [1,]        2        1        2        2        1
## [2,]        3        1        2        2        2
## [3,]        4        2        2        2        4
## [4,]        4        5        4        3        6
## [5,]        4        6        4        3        6
```

```
## [6,] 1        5          7          5          3          6
##      Resample101 Resample102 Resample103 Resample104 Resample105
## [1,]           1          1          2          2          1
## [2,]           2          2          2          2          1
## [3,]           3          2          4          3          2
## [4,]           3          3          4          4          5
## [5,]           5          4          5          5          5
## [6,]           6          6          6          5          6
##      Resample106 Resample107 Resample108 Resample109 Resample110
## [1,]           2          1          1          1          2
## [2,]           2          1          2          3          2
## [3,]           2          3          4          3          3
## [4,]           4          4          4          5          4
## [5,]           5          4          5          6          4
## [6,]           6          5          6          7          5
##      Resample111 Resample112 Resample113 Resample114 Resample115
## [1,]           1          2          2          2          1
## [2,]           3          3          2          2          1
## [3,]           4          3          3          2          2
## [4,]           4          4          5          3          3
## [5,]           6          5          5          3          3
## [6,]           7          5          5          4          4
##      Resample116 Resample117 Resample118 Resample119 Resample120
## [1,]           1          1          1          1          4
## [2,]           2          1          1          3          5
## [3,]           3          3          2          3          7
## [4,]           3          5          2          4          7
## [5,]           4          5          4          7          8
## [6,]           4          7          4          8          8
##      Resample121 Resample122 Resample123 Resample124 Resample125
## [1,]           1          1          1          1          1
## [2,]           1          1          1          1          2
## [3,]           1          2          4          3          2
## [4,]           2          4          4          4          2
## [5,]           3          4          4          6          4
## [6,]           5          6          5          8          5
##      Resample126 Resample127 Resample128 Resample129 Resample130
## [1,]           1          2          1          1          1
## [2,]           2          3          3          1          2
## [3,]           3          4          3          1          3
## [4,]           4          7          4          4          5
## [5,]           4          7          4          4          6
## [6,]           4          8          5          5          6
##      Resample131 Resample132 Resample133 Resample134 Resample135
## [1,]           1          1          1          1          1
## [2,]           2          1          3          2          1
## [3,]           2          2          4          2          2
## [4,]           5          2          5          2          2
## [5,]           5          2          6          2          2
## [6,]           6          4          7          3          3
##      Resample136 Resample137 Resample138 Resample139 Resample140
## [1,]           1          2          1          1          1
## [2,] 1         2          2          3          3          2
## [3,]           2          3          3          3          2
```

```
## [4,]          4          4          4          4          3
## [5,]          4          4          5          6          4
## [6,]          6          6          6          8          5
##      Resample141 Resample142 Resample143 Resample144 Resample145
## [1,]          1          1          1          2          1
## [2,]          1          2          3          3          1
## [3,]          4          3          4          3          2
## [4,]          4          4          4          4          3
## [5,]          4          5          5          5          4
## [6,]          5          6          6          5          6
##      Resample146 Resample147 Resample148 Resample149 Resample150
## [1,]          2          2          1          1          1
## [2,]          2          4          1          2          1
## [3,]          2          5          1          3          2
## [4,]          4          5          2          5          2
## [5,]          4          5          4          7          4
## [6,]          4          7          5          7          6
##      Resample151 Resample152 Resample153 Resample154 Resample155
## [1,]          1          2          1          1          1
## [2,]          3          5          2          1          2
## [3,]          4          5          2          4          3
## [4,]          5          6          3          5          4
## [5,]          5          8          6          5          6
## [6,]          7          9          6          5          8
##      Resample156 Resample157 Resample158 Resample159 Resample160
## [1,]          1          1          5          1          1
## [2,]          2          2          5          1          1
## [3,]          3          3          5          2          2
## [4,]          4          5          6          2          2
## [5,]          7          6          6          2          2
## [6,]          8          7          7          2          3
##      Resample161 Resample162 Resample163 Resample164 Resample165
## [1,]          1          1          1          3          1
## [2,]          4          1          1          4          1
## [3,]          4          5          1          4          2
## [4,]          5          6          1          5          3
## [5,]          7          7          1          6          4
## [6,]          8          8          4          6          5
##      Resample166 Resample167 Resample168 Resample169 Resample170
## [1,]          3          2          1          1          1
## [2,]          3          2          4          1          2
## [3,]          3          2          4          2          2
## [4,]          5          2          7          3          2
## [5,]          5          7          9          6          3
## [6,]          6          8          9          6          4
##      Resample171 Resample172 Resample173 Resample174 Resample175
## [1,]          1          3          1          1          1
## [2,]          2          4          1          3          3
## [3,]          2          4          1          4          4
## [4,]          3          5          5          4          5
## [5,]          3          7          5          5          6
## [6,]          3          8          6          5          6
##      Resample176 Resample177 Resample178 Resample179 Resample180
## [1,]          2          1          2          1          1
```

```
## [2,]        4         1         2         2         2
## [3,]        5         2         3         2         4
## [4,]        5         2         3         3         4
## [5,]        6         3         4         3         5
## [6,]        6         3         4         5         6
##      Resample181 Resample182 Resample183 Resample184 Resample185
## [1,]           1           1           1           1           2
## [2,]           2           4           3           2           2
## [3,]           2           5           5           2           2
## [4,]           2           6           5           5           5
## [5,]           3           6           6           7           7
## [6,]           6           6           6           7           7
##      Resample186 Resample187 Resample188 Resample189 Resample190
## [1,]           1           1           2           1           1
## [2,]           1           2           2           2           4
## [3,]           3           3           3           4           4
## [4,]           5           3           3           4           5
## [5,]           5           4           4           5           5
## [6,]           6           6           5           7           6
##      Resample191 Resample192 Resample193 Resample194 Resample195
## [1,]           1           1           1           1           1
## [2,]           1           1           1           1           2
## [3,]           1           3           2           1           2
## [4,]           2           3           4           2           3
## [5,]           4           5           5           3           3
## [6,]           5           7           5           3           5
##      Resample196 Resample197 Resample198 Resample199 Resample200
## [1,]           2           1           3           1           1
## [2,]           3           3           4           2           1
## [3,]           4           3           4           2           2
## [4,]           6           5           5           3           5
## [5,]           7           5           6           3           5
## [6,]           7           6           6           4           6
##      Resample201 Resample202 Resample203 Resample204 Resample205
## [1,]           1           1           1           1           1
## [2,]           1           1           1           2           2
## [3,]           1           1           2           2           2
## [4,]           2           1           3           4           2
## [5,]           4           3           3           4           3
## [6,]           4           5           3           7           3
##      Resample206 Resample207 Resample208 Resample209 Resample210
## [1,]           1           1           2           1           1
## [2,]           3           3           2           2           2
## [3,]           3           5           4           6           3
## [4,]           4           5           6           6           6
## [5,]           5           5           6           6           6
## [6,]           6           7           6           7           7
##      Resample211 Resample212 Resample213 Resample214
## [1,]           2           1           1           2
## [2,]           3           1           1           2
## [3,]           4           1           3           2
## [4,]           4           1           4           3
## [5,]           5           3           4           4
## [6,]           9           3           5           5
```

Assignment3

ORIGINALITY REPORT

%89	%12	%4	%89
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to UT, Dallas Student Paper	%89

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	OFF		

```
##  Min.   :2
##  1st Qu.:2
##  Median :2
##  Mean   :2
##  3rd Qu.:2
##  Max.   :2

#c
#When sample size are small then K fold cross validation technique should be
used to  sample data.  As the data size is so small we will not consider test
data.

## Answer3.1
## 3.1
library(mlbench)
library(corrplot)
library(e1071)

##
## Attaching package: 'e1071'

## The following object is masked from 'package:Hmisc':
##
##     impute

library(caret)
library(AppliedPredictiveModeling)
data(Glass)
Glass_data <- subset(Glass,)
library(caTools)
set.seed(101)
sample <- sample.split(Glass_data$Si,SplitRatio = .75)
train_random <- subset(Glass_data,sample==TRUE)
test_random <- subset(Glass_data,sample==FALSE)
head(train_random)

##        RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 1 1.52101 13.64 4.49 1.10 71.78 0.06 8.75  0 0.00    1
## 2 1.51761 13.89 3.60 1.36 72.73 0.48 7.83  0 0.00    1
## 3 1.51618 13.53 3.55 1.54 72.99 0.39 7.78  0 0.00    1
## 4 1.51766 13.21 3.69 1.29 72.61 0.57 8.22  0 0.00    1
## 5 1.51742 13.27 3.62 1.24 73.08 0.55 8.07  0 0.00    1
## 6 1.51596 12.79 3.61 1.62 72.97 0.64 8.07  0 0.26    1

head(test_random)

##         RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 11 1.51571 12.72 3.46 1.56 73.20 0.67 8.09  0 0.24    1
## 14 1.51748 12.86 3.56 1.27 73.21 0.54 8.38  0 0.17    1
## 17 1.51784 12.68 3.67 1.16 73.11 0.61 8.70  0 0.00    1
## 22 1.51966 14.77 3.75 0.29 72.02 0.03 9.00  0 0.00    1
```

```
## 25 1.51720 13.38 3.50 1.15 72.85 0.50 8.43  0 0.00    1
## 26 1.51764 12.98 3.54 1.21 73.00 0.65 8.53  0 0.00    1
```

```
##3.2
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
```

```
##
## Attaching package: 'DAAG'
```

```
## The following object is masked from 'package:survival':
##
##     lung
```

```
set.seed(1)
cvsplit <- createFolds(Glass_data,k=3,returnTrain = TRUE)
str(cvsplit)
```

```
## List of 3
##  $ Fold1: int [1:7] 1 2 3 5 6 9 10
##  $ Fold2: int [1:6] 1 2 4 7 8 10
##  $ Fold3: int [1:7] 3 4 5 6 7 8 9
```

```
fold1 <- cvsplit[[1]]
first <- Glass_data[,(cvsplit$Fold1)]
second <-Glass_data[,(cvsplit$Fold2)]
third <- Glass_data[,(cvsplit$Fold3)]
split <- lapply(cvsplit,function(ind,dat) dat[ind,],dat=Glass_data)
dim(Glass_data)
```

```
## [1] 214  10
```

```
unlist(lapply(split,nrow))
```

```
## Fold1 Fold2 Fold3
##     7     6     7
```

```
##3.3
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
bootstrap <- createResample(Glass_data,times=214,list=FALSE)
head.matrix(bootstrap)
```

```
##      Resample001 Resample002 Resample003 Resample004 Resample005
## [1,]           2           1           2           1           1
## [2,]           3           2           2           5           1
## [3,]           4           3           5           6           3
## [4,]           4           3           5           6           4
## [5,]           5           4           5           7           5
## [6,]           7           4           6           7           5
##      Resample006 Resample007 Resample008 Resample009 Resample010
## [1,]           1           4           2           1           2
## [2,]           3           4           2           3           3
## [3,]           3           4           3           5           3
## [4,]           4           4           3           5           4
## [5,]           5           5           4           7           5
## [6,]           5           8           4           7           6
##      Resample011 Resample012 Resample013 Resample014 Resample015
## [1,]           1           2           1           2           1
## [2,]           2           3           1           2           2
## [3,]           2           3           6           3           3
## [4,]           4           5           6           3           3
## [5,]           5           5           6           5           4
## [6,]           5           5           6           7           5
##      Resample016 Resample017 Resample018 Resample019 Resample020
## [1,]           1           4           2           2           2
## [2,]           3           4           2           2           2
## [3,]           4           7           3           3           3
## [4,]           4           7           6           4           3
## [5,]           5           7           6           4           3
## [6,]           5           8           8           6           3
##      Resample021 Resample022 Resample023 Resample024 Resample025
## [1,]           2           1           2           2           2
## [2,]           4           2           4           2           3
## [3,]           4           3           4           3           4
## [4,]           7           3           4           4           5
## [5,]           8           4           5           5           5
## [6,]           9           5           5           5           5
##      Resample026 Resample027 Resample028 Resample029 Resample030
## [1,]           1           1           1           1           1
## [2,]           2           2           1           1           1
## [3,]           2           2           2           2           2
## [4,]           3           2           2           2           4
## [5,]           4           2           2           3           5
## [6,]           6           3           3           4           5
##      Resample031 Resample032 Resample033 Resample034 Resample035
## [1,]           1           3           1           2           1
## [2,]           2           4           2           2           1
## [3,]           2           5           3           3           3
## [4,]           3           7           4           4           5
## [5,]           4           8           4           4           5
## [6,]           5           8           5           5           5
##      Resample036 Resample037 Resample038 Resample039 Resample040
```

```
## [1,]            1            2            2            2            1
## [2,]            2            3            3            2            2
## [3,]            3            3            3            2            4
## [4,]            4            4            5            2            7
## [5,]            5            5            6            4            8
## [6,]            6            5            7            4            8
##      Resample041 Resample042 Resample043 Resample044 Resample045
## [1,]           4           1           1           1           1
## [2,]           4           2           1           1           1
## [3,]           5           4           2           2           2
## [4,]           5           5           4           3           4
## [5,]           7           5           5           3           5
## [6,]           8           6           6           4           6
##      Resample046 Resample047 Resample048 Resample049 Resample050
## [1,]           1           2           1           1           5
## [2,]           2           3           2           1           5
## [3,]           2           4           2           2           6
## [4,]           3           5           2           3           6
## [5,]           5           5           6           4           7
## [6,]           5           5           6           5           7
##      Resample051 Resample052 Resample053 Resample054 Resample055
## [1,]           1           2           1           2           1
## [2,]           2           4           1           2           1
## [3,]           4           4           2           4           1
## [4,]           5           8           2           4           4
## [5,]           6           8           3           4           6
## [6,]           7           9           4           6           8
##      Resample056 Resample057 Resample058 Resample059 Resample060
## [1,]           2           1           1           3           2
## [2,]           3           2           1           4           2
## [3,]           3           3           4           5           4
## [4,]           4           5           5           5           4
## [5,]           7           5           6           5           8
## [6,]           7           6           6           6           9
##      Resample061 Resample062 Resample063 Resample064 Resample065
## [1,]           1           2           1           1           1
## [2,]           2           3           1           3           2
## [3,]           4           3           4           3           4
## [4,]           4           4           4           4           4
## [5,]           5           5           4           7           6
## [6,]           6           6           5           8           6
##      Resample066 Resample067 Resample068 Resample069 Resample070
## [1,]           1           2           1           1           1
## [2,]           2           3           2           2           5
## [3,]           4           3           4           5           6
## [4,]           4           4           5           5           7
## [5,]           4           5           6           6           7
## [6,]           6           6           6           6           8
##      Resample071 Resample072 Resample073 Resample074 Resample075
## [1,]           1           1           1           2           2
```

```
## [2,]           2           1           1           3           2
## [3,]           2           3           1           3           3
## [4,]           5           4           3           4           5
## [5,]           5           5           3           4           5
## [6,]           5           6           6           5           5
##      Resample076 Resample077 Resample078 Resample079 Resample080
## [1,]           2           1           1           1           1
## [2,]           2           1           2           2           1
## [3,]           2           3           3           3           3
## [4,]           3           5           5           4           4
## [5,]           6           5           5           4           4
## [6,]           8           7           7           6           7
##      Resample081 Resample082 Resample083 Resample084 Resample085
## [1,]           3           1           1           1           1
## [2,]           4           3           5           2           1
## [3,]           4           4           6           3           3
## [4,]           4           4           6           4           3
## [5,]           4           4           7           5           3
## [6,]           6           6           8           5           3
##      Resample086 Resample087 Resample088 Resample089 Resample090
## [1,]           1           1           4           1           1
## [2,]           1           3           4           1           1
## [3,]           1           3           5           1           1
## [4,]           1           5           6           1           1
## [5,]           3           5           7           1           2
## [6,]           4           5           8           1           2
##      Resample091 Resample092 Resample093 Resample094 Resample095
## [1,]           1           1           3           1           1
## [2,]           1           3           3           3           1
## [3,]           2           4           4           5           2
## [4,]           5           5           5           5           2
## [5,]           6           6           7           5           3
## [6,]           7           8           8           7           5
##      Resample096 Resample097 Resample098 Resample099 Resample100
## [1,]           2           1           2           2           1
## [2,]           3           1           2           2           2
## [3,]           4           2           2           2           4
## [4,]           4           5           4           3           6
## [5,]           4           6           4           3           6
## [6,]           5           7           5           3           6
##      Resample101 Resample102 Resample103 Resample104 Resample105
## [1,]           1           1           2           2           1
## [2,]           2           2           2           2           1
## [3,]           3           2           4           3           2
## [4,]           3           3           4           4           5
## [5,]           5           4           5           5           5
## [6,]           6           6           6           5           6
##      Resample106 Resample107 Resample108 Resample109 Resample110
## [1,]           2           1           1           1           2
## [2,]           2           1           2           3           2
```

```
## [3,]           2           3           4           3           3
## [4,]           4           4           4           5           4
## [5,]           5           4           5           6           4
## [6,]           6           5           6           7           5
##      Resample111 Resample112 Resample113 Resample114 Resample115
## [1,]           1           2           2           2           1
## [2,]           3           3           2           2           1
## [3,]           4           3           3           2           2
## [4,]           4           4           5           3           3
## [5,]           6           5           5           3           3
## [6,]           7           5           5           4           4
##      Resample116 Resample117 Resample118 Resample119 Resample120
## [1,]           1           1           1           1           4
## [2,]           2           1           1           3           5
## [3,]           3           3           2           3           7
## [4,]           3           5           2           4           7
## [5,]           4           5           4           7           8
## [6,]           4           7           4           8           8
##      Resample121 Resample122 Resample123 Resample124 Resample125
## [1,]           1           1           1           1           1
## [2,]           1           1           1           1           2
## [3,]           1           2           4           3           2
## [4,]           2           4           4           4           2
## [5,]           3           4           4           6           4
## [6,]           5           6           5           8           5
##      Resample126 Resample127 Resample128 Resample129 Resample130
## [1,]           1           2           1           1           1
## [2,]           2           3           3           1           2
## [3,]           3           4           3           1           3
## [4,]           4           7           4           4           5
## [5,]           4           7           4           4           6
## [6,]           4           8           5           5           6
##      Resample131 Resample132 Resample133 Resample134 Resample135
## [1,]           1           1           1           1           1
## [2,]           2           1           3           2           1
## [3,]           2           2           4           2           2
## [4,]           5           2           5           2           2
## [5,]           5           2           6           2           2
## [6,]           6           4           7           3           3
##      Resample136 Resample137 Resample138 Resample139 Resample140
## [1,]           1           2           1           1           1
## [2,]           2           2           3           3           2
## [3,]           2           3           3           3           2
## [4,]           4           4           4           4           3
## [5,]           4           4           5           6           4
## [6,]           6           6           6           8           5
##      Resample141 Resample142 Resample143 Resample144 Resample145
## [1,]           1           1           1           2           1
## [2,]           1           2           3           3           1
## [3,]           4           3           4           3           2
```

```
## [4,]           4           4           4           4           3
## [5,]           4           5           5           5           4
## [6,]           5           6           6           5           6
##      Resample146 Resample147 Resample148 Resample149 Resample150
## [1,]           2           2           1           1           1
## [2,]           2           4           1           2           1
## [3,]           2           5           1           3           2
## [4,]           4           5           2           5           2
## [5,]           4           5           4           7           4
## [6,]           4           7           5           7           6
##      Resample151 Resample152 Resample153 Resample154 Resample155
## [1,]           1           2           1           1           1
## [2,]           3           5           2           1           2
## [3,]           4           5           2           4           3
## [4,]           5           6           3           5           4
## [5,]           5           8           6           5           6
## [6,]           7           9           6           5           8
##      Resample156 Resample157 Resample158 Resample159 Resample160
## [1,]           1           1           5           1           1
## [2,]           2           2           5           1           1
## [3,]           3           3           5           2           2
## [4,]           4           5           6           2           2
## [5,]           7           6           6           2           2
## [6,]           8           7           7           2           3
##      Resample161 Resample162 Resample163 Resample164 Resample165
## [1,]           1           1           1           3           1
## [2,]           4           1           1           4           1
## [3,]           4           5           1           4           2
## [4,]           5           6           1           5           3
## [5,]           7           7           1           6           4
## [6,]           8           8           4           6           5
##      Resample166 Resample167 Resample168 Resample169 Resample170
## [1,]           3           2           1           1           1
## [2,]           3           2           4           1           2
## [3,]           3           2           4           2           2
## [4,]           5           2           7           3           2
## [5,]           5           7           9           6           3
## [6,]           6           8           9           6           4
##      Resample171 Resample172 Resample173 Resample174 Resample175
## [1,]           1           3           1           1           1
## [2,]           2           4           1           3           3
## [3,]           2           4           1           4           4
## [4,]           3           5           5           4           5
## [5,]           3           7           5           5           6
## [6,]           3           8           6           5           6
##      Resample176 Resample177 Resample178 Resample179 Resample180
## [1,]           2           1           2           1           1
## [2,]           4           1           2           2           2
## [3,]           5           2           3           2           4
## [4,]           5           2           3           3           4
```

```
## [5,]           6           3           4           3           5
## [6,] 1         6           3           4           5           6
##      Resample181 Resample182 Resample183 Resample184 Resample185
## [1,]           1           1           1           1           2
## [2,]           2           4           3           2           2
## [3,]           2           5           5           2           2
## [4,]           2           6           5           5           5
## [5,]           3           6           6           7           7
## [6,]           6           6           6           7           7
##      Resample186 Resample187 Resample188 Resample189 Resample190
## [1,]           1           1           2           1           1
## [2,]           1           2           2           2           4
## [3,]           3           3           3           4           4
## [4,]           5           3           3           4           5
## [5,]           5           4           4           5           5
## [6,]           6           6           5           7           6
##      Resample191 Resample192 Resample193 Resample194 Resample195
## [1,]           1           1           1           1           1
## [2,]           1           1           1           1           2
## [3,]           1           3           2           1           2
## [4,]           2           3           4           2           3
## [5,]           4           5           5           3           3
## [6,]           5           7           5           3           5
##      Resample196 Resample197 Resample198 Resample199 Resample200
## [1,]           2           1           3           1           1
## [2,]           3           3           4           2           1
## [3,]           4           3           4           2           2
## [4,]           6           5           5           3           5
## [5,]           7           5           6           3           5
## [6,]           7           6           6           4           6
##      Resample201 Resample202 Resample203 Resample204 Resample205
## [1,]           1           1           1           1           1
## [2,]           1           1           1           2           2
## [3,]           1           1           2           2           2
## [4,]           2           1           3           4           2
## [5,]           4           3           3           4           3
## [6,]           4           5           3           7           3
##      Resample206 Resample207 Resample208 Resample209 Resample210
## [1,]           1           1           2           1           1
## [2,]           3           3           2           2           2
## [3,]           3           5           4           6           3
## [4,]           4           5           6           6           6
## [5,]           5           5           6           6           6
## [6,]           6           7           6           7           7
##      Resample211 Resample212 Resample213 Resample214
## [1,]           2           1           1           2
## [2,]           3           1           1           2
## [3,]           4           1           3           2
## [4,]           4           1           4           3
```

```
## [5,]          5          3          4          4
## [6,]          9          3          5          5
```

homework3

ORIGINALITY REPORT

%86	%13	%12	%85
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to UT, Dallas Student Paper	%78
2	Max Kuhn, Kjell Johnson. "Applied Predictive Modeling", Springer Nature, 2013 Publication	%4
3	rstudio-pubs-static.s3.amazonaws.com Internet Source	%2
4	Submitted to The University of Manchester Student Paper	%1
5	www.www.cottofederico.altervista.org Internet Source	<%1

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	OFF		

Assignment3

by Shivam Tiwari

FILE	ASSIGNMENT3.PDF (403.16K)		
TIME SUBMITTED	04-FEB-2017 10:04PM	WORD COUNT	3010
SUBMISSION ID	766620267	CHARACTER COUNT	10259

Assignment- 3

Shivam

2/4/2017

1

```
#Answer 1
library(Hmisc)
```

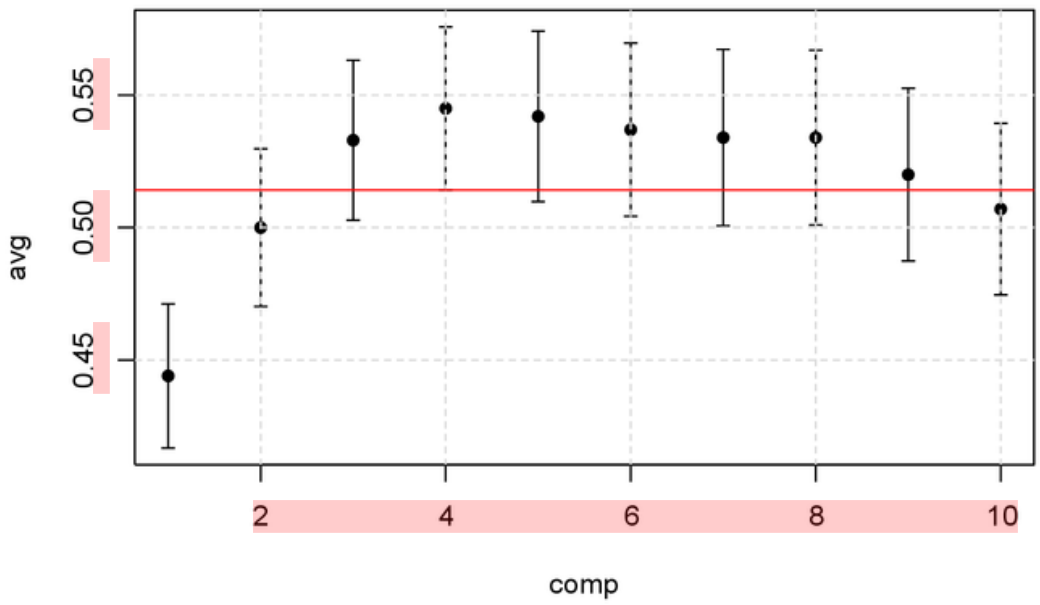
```
## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
## Loading required package: ggplot2
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units
```

```
library(AppliedPredictiveModeling)
library(caret)
```

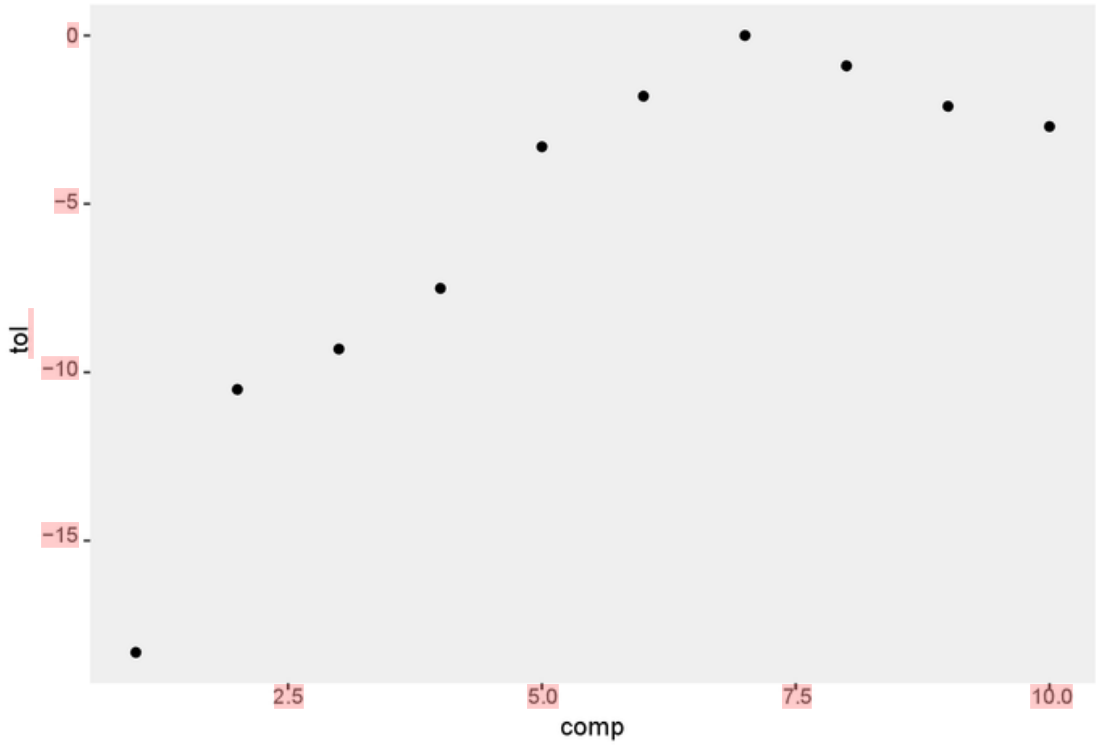
```
##
## Attaching package: 'caret'
## The following object is masked from 'package:survival':
##
##     cluster
```

```
data(ChemicalManufacturingProcess)
comp = 1:10
avg = c( 0.444, 0.500, 0.533, 0.545, 0.542, 0.537, 0.534, 0.534, 0.520, 0.507 )
std_error = c( 0.0272, 0.0298, 0.0302, 0.0308, 0.0322, 0.0327, 0.0333, 0.0330, 0.0326, 0.0324 )
data = data.frame( comp, avg, std_error )
errbar( comp, avg, avg+std_error, avg-std_error )
grid()
max1 = which.max( avg )
abline( h=avg[max1] - std_error[max1], col='red' )
```

1




```
best1 <- subset(data,comp==which.max(data$std_error))
data$tol <- (data$std_error-best$std_error)/best1$std_error*100
qplot(comp,tol,data = data)
```




```
## 1(a) Four components should be used to develop the most parsimonious model
```

```
## 1(b)for 10 % loss in R^2 values a two component model would work

##1(c) Basis the graph,we see that random forest has the highest R^2 so we will pick that

## 1(d) We would choose SVm because it has a reasonably fast processing time and has the second highest
```

```
library(caret)
data(oil)
str(oilType)
```

```
##  Factor w/ 7 levels "A","B","C","D",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```
table(oilType)
```

```
## oilType
##  A  B  C  D  E  F  G
## 37 26  3  7 11 10  2
```

```
sample <- floor(length(oilType)*.6)  + 1
set.seed(629)
Splits <- vector(mode = "list", length = 20)
for(x in seq(along = Splits)) Splits[[x]] <- table(sample(oilType, size = sample))
head(Splits, 3)
```

```
## [[1]]
##
##  A  B  C  D  E  F  G
## 24 15  3  4  7  4  1
##
## [[2]]
##
##  A  B  C  D  E  F  G
## 25 14  2  4  6  6  1
##
## [[3]]
##
##  A  B  C  D  E  F  G
## 24 16  2  5  5  4  2
```

```
Splits <- do.call("rbind", Splits)
head(Splits, 3)
```

```
##      A  B C D E F G
## [1,] 24 15 3 4 7 4 1
## [2,] 25 14 2 4 6 6 1
## [3,] 24 16 2 5 5 4 2
```

```
summary(Splits/sample)
```

```
##        A                B                C                D
##  Min.   :0.3103   Min.   :0.2241   Min.   :0.00000   Min.   :0.03448
##  1st Qu.:0.3448   1st Qu.:0.2586   1st Qu.:0.01724   1st Qu.:0.06897
##  Median :0.3793   Median :0.2759   Median :0.03448   Median :0.08621
##  Mean   :0.3810   Mean   :0.2759   Mean   :0.03362   Mean   :0.08448
##  3rd Qu.:0.4138   3rd Qu.:0.2931   3rd Qu.:0.05172   3rd Qu.:0.10345
##  Max.   :0.4483   Max.   :0.3276   Max.   :0.05172   Max.   :0.12069
##        E                F                G
```

```
##  Min.   :0.05172   Min.   :0.03448   Min.   :0.00000
##  1st Qu.:0.10345   1st Qu.:0.08621   1st Qu.:0.01724
##  Median :0.10345   Median :0.10345   Median :0.01724
##  Mean   :0.10690   Mean   :0.09828   Mean   :0.01983
##  3rd Qu.:0.12069   3rd Qu.:0.10776   3rd Qu.:0.02155
##  Max.   :0.15517   Max.   :0.13793   Max.   :0.03448
```

```
set.seed(62)
Splits2 <- createDataPartition(oilType,  p = .60,  times = 20)
Splits2 <- lapply(Splits2,  function(i,  y) table(y[i]),  y = oilType)
head(Splits2,  3)
```

```
## $Resample01
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
##
## $Resample02
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
##
## $Resample03
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
```

```
Splits2 <- do.call("rbind",Splits2)
summary(Splits2)
```

```
##        A           B           C          D          E          F
##  Min.   :23   Min.   :16   Min.   :2   Min.   :5   Min.   :7   Min.   :6
##  1st Qu.:23   1st Qu.:16   1st Qu.:2   1st Qu.:5   1st Qu.:7   1st Qu.:6
##  Median :23   Median :16   Median :2   Median :5   Median :7   Median :6
##  Mean   :23   Mean   :16   Mean   :2   Mean   :5   Mean   :7   Mean   :6
##  3rd Qu.:23   3rd Qu.:16   3rd Qu.:2   3rd Qu.:5   3rd Qu.:7   3rd Qu.:6
##  Max.   :23   Max.   :16   Max.   :2   Max.   :5   Max.   :7   Max.   :6
##        G
##  Min.   :2
##  1st Qu.:2
##  Median :2
##  Mean   :2
##  3rd Qu.:2
##  Max.   :2
```

```
# a) We can see the summary using output command
# b) Stratified random sample has less variation and each partition has atleast one sample of each clas
# c)  For such a small sample size we should use K-fold cross validation technique to sample the data.
#     We will not use a test data in this case due to small data size
```

```
## 3.1
library(mlbench)
library(corrplot)
library(e1071)
```

```
##
## Attaching package: 'e1071'
```

```
## The following object is masked from 'package:Hmisc':
##
##     impute
```

```
library(caret)
library(AppliedPredictiveModeling)
data(Glass)
Glass_data <- subset(Glass,)
library(caTools)
set.seed(101)
sample <- sample.split(Glass_data$Si,SplitRatio = .75)
train <- subset(Glass_data,sample==TRUE)
test <- subset(Glass_data,sample==FALSE)
head(train)
```

```
##        RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 1 1.52101 13.64 4.49 1.10 71.78 0.06 8.75  0 0.00    1
## 2 1.51761 13.89 3.60 1.36 72.73 0.48 7.83  0 0.00    1
## 3 1.51618 13.53 3.55 1.54 72.99 0.39 7.78  0 0.00    1
## 4 1.51766 13.21 3.69 1.29 72.61 0.57 8.22  0 0.00    1
## 5 1.51742 13.27 3.62 1.24 73.08 0.55 8.07  0 0.00    1
## 6 1.51596 12.79 3.61 1.62 72.97 0.64 8.07  0 0.26    1
```

```
head(test)
```

```
##         RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 11 1.51571 12.72 3.46 1.56 73.20 0.67 8.09  0 0.24    1
## 14 1.51748 12.86 3.56 1.27 73.21 0.54 8.38  0 0.17    1
## 17 1.51784 12.68 3.67 1.16 73.11 0.61 8.70  0 0.00    1
## 22 1.51966 14.77 3.75 0.29 72.02 0.03 9.00  0 0.00    1
## 25 1.51720 13.38 3.50 1.15 72.85 0.50 8.43  0 0.00    1
## 26 1.51764 12.98 3.54 1.21 73.00 0.65 8.53  0 0.00    1
```

```
##3.2
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
```

```
##
## Attaching package: 'DAAG'
```

```
## The following object is masked from 'package:survival':
##
##     lung
```

```
set.seed(1)
splitx <- createFolds(Glass_data,k=3,returnTrain = TRUE)
str(splitx)
```

```
## List of 3
##  $ Fold1: int [1:7] 1 2 3 5 6 9 10
##  $ Fold2: int [1:6] 1 2 4 7 8 10
##  $ Fold3: int [1:7] 3 4 5 6 7 8 9
```

```
fold1 <- splitx[[1]]
first  <-  Glass_data[,(splitx$Fold1)]
second   <-Glass_data[,(splitx$Fold2)]
third <- Glass_data[,(splitx$Fold3)]
split <- lapply(splitx,function(ind,dat) dat[ind,],dat=Glass_data)
dim(Glass_data)
```

```
## [1] 214  10
```

```
unlist(lapply(split,nrow))
```

```
## Fold1 Fold2 Fold3
##     7     6     7
```

```
##3.3
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
boot <- createResample(Glass_data,times=214,list=FALSE)
head.matrix(boot)
```

```
##      Resample001 Resample002 Resample003 Resample004 Resample005
## [1,]           2           1           2           1           1
## [2,]           3           2           2           5           1
## [3,]           4           3           5           6           3
## [4,]           4           3           5           6           4
## [5,]           5           4           5           7           5
## [6,]           7           4           6           7           5
##      Resample006 Resample007 Resample008 Resample009 Resample010
## [1,]           1           4           2           1           2
## [2,]           3           4           2           3           3
## [3,]           3           4           3           5           3
## [4,]           4           4           3           5           4
## [5,]           5           5           4           7           5
## [6,]           5           8           4           7           6
##      Resample011 Resample012 Resample013 Resample014 Resample015
## [1,]           1           2           1           2           1
## [2,]           2           3           1           2           2
## [3,]           2           3           6           3           3
## [4,]           4           5           6           3           3
## [5,]           5           5           6           5           4
## [6,]           5           5           6           7           5
##      Resample016 Resample017 Resample018 Resample019 Resample020
## [1,]           1           4           2           2           2
## [2,]           3           4           2           2           2
## [3,]           4           7           3           3           3
## [4,]           4           7           6           4           3
## [5,]           5           7           6           4           3
## [6,]           5           8           8           6           3
##      Resample021 Resample022 Resample023 Resample024 Resample025
## [1,]           2           1           2           2           2
## [2,]           4           2           4           2           3
```

```
## [3,]        4         3         4         3         4
## [4,]        7         3         4         4         5
## [5,]        8         4         5         5         5
## [6,]        9         5         5         5         5
##      Resample026 Resample027 Resample028 Resample029 Resample030
## [1,]           1           1           1           1           1
## [2,]           2           2           1           1           1
## [3,]           2           2           2           2           2
## [4,]           3           2           2           2           4
## [5,]           4           2           2           3           5
## [6,]           6           3           3           4           5
##      Resample031 Resample032 Resample033 Resample034 Resample035
## [1,]           1           3           1           2           1
## [2,]           2           4           2           2           1
## [3,]           2           5           3           3           3
## [4,]           3           7           4           4           5
## [5,]           4           8           4           4           5
## [6,]           5           8           5           5           5
##      Resample036 Resample037 Resample038 Resample039 Resample040
## [1,]           1           2           2           2           1
## [2,]           2           3           3           2           2
## [3,]           3           3           3           2           4
## [4,]           4           4           5           2           7
## [5,]           5           5           6           4           8
## [6,]           6           5           7           4           8
##      Resample041 Resample042 Resample043 Resample044 Resample045
## [1,]           4           1           1           1           1
## [2,]           4           2           1           1           1
## [3,]           5           4           2           2           2
## [4,]           5           5           4           3           4
## [5,]           7           5           5           3           5
## [6,]           8           6           6           4           6
##      Resample046 Resample047 Resample048 Resample049 Resample050
## [1,]           1           2           1           1           5
## [2,]           2           3           2           1           5
## [3,]           2           4           2           2           6
## [4,]           3           5           2           3           6
## [5,]           5           5           6           4           7
## [6,]           5           5           6           5           7
##      Resample051 Resample052 Resample053 Resample054 Resample055
## [1,]           1           2           1           2           1
## [2,]           2           4           1           2           1
## [3,]           4           4           2           4           1
## [4,]           5           8           2           4           4
## [5,]           6           8           3           4           6
## [6,]           7           9           4           6           8
##      Resample056 Resample057 Resample058 Resample059 Resample060
## [1,]           2           1           1           3           2
## [2,]           3           2           1           4           2
## [3,]           3           3           4           5           4
## [4,]           4           5           5           5           4
## [5,]           7           5           6           5           8
## [6,]           7           6           6           6           9
##      Resample061 Resample062 Resample063 Resample064 Resample065
```

```
## [1,]        1        2        1        1        1
## [2,]        2        3        1        3        2
## [3,]        4        3        4        3        4
## [4,]        4        4        4        4        4
## [5,]        5        5        4        7        6
## [6,]        6        6        5        8        6
##      Resample066 Resample067 Resample068 Resample069 Resample070
## [1,]        1        2        1        1        1
## [2,]        2        3        2        2        5
## [3,]        4        3        4        5        6
## [4,]        4        4        5        5        7
## [5,]        4        5        6        6        7
## [6,]        6        6        6        6        8
##      Resample071 Resample072 Resample073 Resample074 Resample075
## [1,]        1        1        1        2        2
## [2,]        2        1        1        3        2
## [3,]        2        3        1        3        3
## [4,]        5        4        3        4        5
## [5,]        5        5        3        4        5
## [6,]        5        6        6        5        5
##      Resample076 Resample077 Resample078 Resample079 Resample080
## [1,]        2        1        1        1        1
## [2,]        2        1        2        2        1
## [3,]        2        3        3        3        3
## [4,]        3        5        5        4        4
## [5,]        6        5        5        4        4
## [6,]        8        7        7        6        7
##      Resample081 Resample082 Resample083 Resample084 Resample085
## [1,]        3        1        1        1        1
## [2,]        4        3        5        2        1
## [3,]        4        4        6        3        3
## [4,]        4        4        6        4        3
## [5,]        4        4        7        5        3
## [6,]        6        6        8        5        3
##      Resample086 Resample087 Resample088 Resample089 Resample090
## [1,]        1        1        4        1        1
## [2,]        1        3        4        1        1
## [3,]        1        3        5        1        1
## [4,]        1        5        6        1        1
## [5,]        3        5        7        1        2
## [6,]        4        5        8        1        2
##      Resample091 Resample092 Resample093 Resample094 Resample095
## [1,]        1        1        3        1        1
## [2,]        1        3        3        3        1
## [3,]        2        4        4        5        2
## [4,]        5        5        5        5        2
## [5,]        6        6        7        5        3
## [6,]        7        8        8        7        5
##      Resample096 Resample097 Resample098 Resample099 Resample100
## [1,]        2        1        2        2        1
## [2,]        3        1        2        2        2
## [3,]        4        2        2        2        4
## [4,]        4        5        4        3        6
## [5,]        4        6        4        3        6
```

```
## [6,] 1        5          7          5          3          6
##      Resample101 Resample102 Resample103 Resample104 Resample105
## [1,]           1          1          2          2          1
## [2,]           2          2          2          2          1
## [3,]           3          2          4          3          2
## [4,]           3          3          4          4          5
## [5,]           5          4          5          5          5
## [6,]           6          6          6          5          6
##      Resample106 Resample107 Resample108 Resample109 Resample110
## [1,]           2          1          1          1          2
## [2,]           2          1          2          3          2
## [3,]           2          3          4          3          3
## [4,]           4          4          4          5          4
## [5,]           5          4          5          6          4
## [6,]           6          5          6          7          5
##      Resample111 Resample112 Resample113 Resample114 Resample115
## [1,]           1          2          2          2          1
## [2,]           3          3          2          2          1
## [3,]           4          3          3          2          2
## [4,]           4          4          5          3          3
## [5,]           6          5          5          3          3
## [6,]           7          5          5          4          4
##      Resample116 Resample117 Resample118 Resample119 Resample120
## [1,]           1          1          1          1          4
## [2,]           2          1          1          3          5
## [3,]           3          3          2          3          7
## [4,]           3          5          2          4          7
## [5,]           4          5          4          7          8
## [6,]           4          7          4          8          8
##      Resample121 Resample122 Resample123 Resample124 Resample125
## [1,]           1          1          1          1          1
## [2,]           1          1          1          1          2
## [3,]           1          2          4          3          2
## [4,]           2          4          4          4          2
## [5,]           3          4          4          6          4
## [6,]           5          6          5          8          5
##      Resample126 Resample127 Resample128 Resample129 Resample130
## [1,]           1          2          1          1          1
## [2,]           2          3          3          1          2
## [3,]           3          4          3          1          3
## [4,]           4          7          4          4          5
## [5,]           4          7          4          4          6
## [6,]           4          8          5          5          6
##      Resample131 Resample132 Resample133 Resample134 Resample135
## [1,]           1          1          1          1          1
## [2,]           2          1          3          2          1
## [3,]           2          2          4          2          2
## [4,]           5          2          5          2          2
## [5,]           5          2          6          2          2
## [6,]           6          4          7          3          3
##      Resample136 Resample137 Resample138 Resample139 Resample140
## [1,]           1          2          1          1          1
## [2,] 1         2          2          3          3          2
## [3,]           2          3          3          3          2
```

```
## [4,]          4          4          4          4          3
## [5,]          4          4          5          6          4
## [6,]          6          6          6          8          5
##      Resample141 Resample142 Resample143 Resample144 Resample145
## [1,]          1          1          1          2          1
## [2,]          1          2          3          3          1
## [3,]          4          3          4          3          2
## [4,]          4          4          4          4          3
## [5,]          4          5          5          5          4
## [6,]          5          6          6          5          6
##      Resample146 Resample147 Resample148 Resample149 Resample150
## [1,]          2          2          1          1          1
## [2,]          2          4          1          2          1
## [3,]          2          5          1          3          2
## [4,]          4          5          2          5          2
## [5,]          4          5          4          7          4
## [6,]          4          7          5          7          6
##      Resample151 Resample152 Resample153 Resample154 Resample155
## [1,]          1          2          1          1          1
## [2,]          3          5          2          1          2
## [3,]          4          5          2          4          3
## [4,]          5          6          3          5          4
## [5,]          5          8          6          5          6
## [6,]          7          9          6          5          8
##      Resample156 Resample157 Resample158 Resample159 Resample160
## [1,]          1          1          5          1          1
## [2,]          2          2          5          1          1
## [3,]          3          3          5          2          2
## [4,]          4          5          6          2          2
## [5,]          7          6          6          2          2
## [6,]          8          7          7          2          3
##      Resample161 Resample162 Resample163 Resample164 Resample165
## [1,]          1          1          1          3          1
## [2,]          4          1          1          4          1
## [3,]          4          5          1          4          2
## [4,]          5          6          1          5          3
## [5,]          7          7          1          6          4
## [6,]          8          8          4          6          5
##      Resample166 Resample167 Resample168 Resample169 Resample170
## [1,]          3          2          1          1          1
## [2,]          3          2          4          1          2
## [3,]          3          2          4          2          2
## [4,]          5          2          7          3          2
## [5,]          5          7          9          6          3
## [6,]          6          8          9          6          4
##      Resample171 Resample172 Resample173 Resample174 Resample175
## [1,]          1          3          1          1          1
## [2,]          2          4          1          3          3
## [3,]          2          4          1          4          4
## [4,]          3          5          5          4          5
## [5,]          3          7          5          5          6
## [6,]          3          8          6          5          6
##      Resample176 Resample177 Resample178 Resample179 Resample180
## [1,]          2          1          2          1          1
```

```
## [2,]        4         1         2         2         2
## [3,]        5         2         3         2         4
## [4,]        5         2         3         3         4
## [5,]        6         3         4         3         5
## [6,]        6         3         4         5         6
##      Resample181 Resample182 Resample183 Resample184 Resample185
## [1,]           1           1           1           1           2
## [2,]           2           4           3           2           2
## [3,]           2           5           5           2           2
## [4,]           2           6           5           5           5
## [5,]           3           6           6           7           7
## [6,]           6           6           6           7           7
##      Resample186 Resample187 Resample188 Resample189 Resample190
## [1,]           1           1           2           1           1
## [2,]           1           2           2           2           4
## [3,]           3           3           3           4           4
## [4,]           5           3           3           4           5
## [5,]           5           4           4           5           5
## [6,]           6           6           5           7           6
##      Resample191 Resample192 Resample193 Resample194 Resample195
## [1,]           1           1           1           1           1
## [2,]           1           1           1           1           2
## [3,]           1           3           2           1           2
## [4,]           2           3           4           2           3
## [5,]           4           5           5           3           3
## [6,]           5           7           5           3           5
##      Resample196 Resample197 Resample198 Resample199 Resample200
## [1,]           2           1           3           1           1
## [2,]           3           3           4           2           1
## [3,]           4           3           4           2           2
## [4,]           6           5           5           3           5
## [5,]           7           5           6           3           5
## [6,]           7           6           6           4           6
##      Resample201 Resample202 Resample203 Resample204 Resample205
## [1,]           1           1           1           1           1
## [2,]           1           1           1           2           2
## [3,]           1           1           2           2           2
## [4,]           2           1           3           4           2
## [5,]           4           3           3           4           3
## [6,]           4           5           3           7           3
##      Resample206 Resample207 Resample208 Resample209 Resample210
## [1,]           1           1           2           1           1
## [2,]           3           3           2           2           2
## [3,]           3           5           4           6           3
## [4,]           4           5           6           6           6
## [5,]           5           5           6           6           6
## [6,]           6           7           6           7           7
##      Resample211 Resample212 Resample213 Resample214
## [1,]           2           1           1           2
## [2,]           3           1           1           2
## [3,]           4           1           3           2
## [4,]           4           1           4           3
## [5,]           5           3           4           4
## [6,]           9           3           5           5
```

Assignment3

ORIGINALITY REPORT

%89	%12	%4	%89
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to UT, Dallas Student Paper	%89

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	OFF		

Assignment 3

by Kshitij Yadav

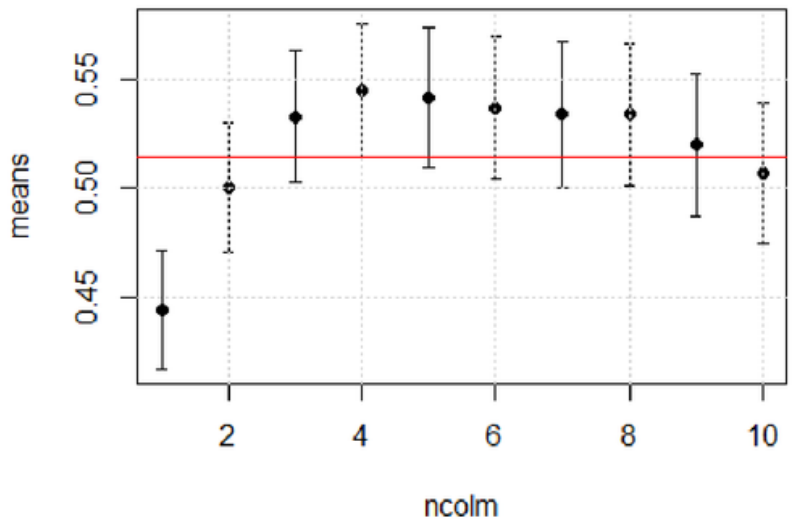
FILE	ASSINGMENT_3.PDF (465.54K)		
TIME SUBMITTED	04-FEB-2017 09:57PM	WORD COUNT	2897
SUBMISSION ID	766618754	CHARACTER COUNT	9703

Assingment 3

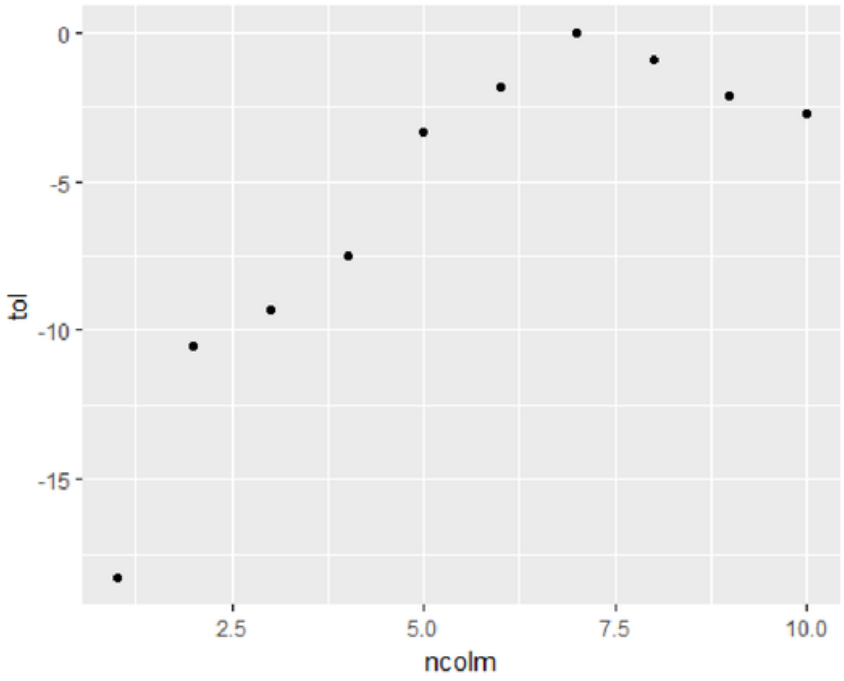
Kshitij

February 4, 2017

```
library(Hmisc)

library(AppliedPredictiveModeling)
data(ChemicalManufacturingProcess)
ncolm = 1:10
means = c( 0.444, 0.500, 0.533, 0.545, 0.542, 0.537, 0.534, 0.534, 0.520, 0.5
07 )
std_errors = c( 0.0272, 0.0298, 0.0302, 0.0308, 0.0322, 0.0327, 0.0333, 0.033
1, 0.0326, 0.0324 )
data = data.frame( ncolm, means, std_errors )
errbar( ncolm, means, means+std_errors, means-std_errors )
grid()
max_index = which.max( means )
abline( h=means[max_index] - std_errors[max_index], col='red' )
```




```
tol <- subset(data,ncolm==which.max(data$std_errors))
data$tol <- (data$std_errors-tol$std_errors)/tol$std_errors*100
qplot(ncolm,tol,data = data)
```




```
# a) We will use 4 PLS component with a lower bound of aprrox .51 .

# b) The lowest setting that does not exceed 10% tolerance is 2 component mod
el.

# c) Random Forest as it has the highest R Square.

# d) SVM as it has the second highest R square and faster running time.




#Q2

library(caret)

data(oil)
str(oilType)

##  Factor w/ 7 levels "A","B","C","D",..: 1 1 1 1 1 1 1 1 1 1 ...

table(oilType)

## oilType
##  A  B  C  D  E  F  G
## 37 26  3  7 11 10  2
```

```r
sample1 <- floor(length(oilType)*.6) + 1
set.seed(629)
Split1 <- vector(mode = "list", length = 20)
for(i in seq(along = Split1)) Split1[[i]] <- table(sample(oilType, size = sam
ple1))
head(Split1, 3)
```

```
## [[1]]
##
##  A  B  C  D  E  F  G
## 24 15  3  4  7  4  1
##
## [[2]]
##
##  A  B  C  D  E  F  G
## 25 14  2  4  6  6  1
##
## [[3]]
##
##  A  B  C  D  E  F  G
## 24 16  2  5  5  4  2
```

```r
Split1 <- do.call("rbind", Split1)
head(Split1, 3)
```

```
##       A  B C D E F G
## [1,] 24 15 3 4 7 4 1
## [2,] 25 14 2 4 6 6 1
## [3,] 24 16 2 5 5 4 2
```

```r
summary(Split1/sample1)
```

```
##        A                B                C                D
##  Min.   :0.3103   Min.   :0.2241   Min.   :0.00000   Min.   :0.03448
##  1st Qu.:0.3448   1st Qu.:0.2586   1st Qu.:0.01724   1st Qu.:0.06897
##  Median :0.3793   Median :0.2759   Median :0.03448   Median :0.08621
##  Mean   :0.3810   Mean   :0.2759   Mean   :0.03362   Mean   :0.08448
##  3rd Qu.:0.4138   3rd Qu.:0.2931   3rd Qu.:0.05172   3rd Qu.:0.10345
##  Max.   :0.4483   Max.   :0.3276   Max.   :0.05172   Max.   :0.12069
##        E                F                G
##  Min.   :0.05172   Min.   :0.03448   Min.   :0.00000
##  1st Qu.:0.10345   1st Qu.:0.08621   1st Qu.:0.01724
##  Median :0.10345   Median :0.10345   Median :0.01724
##  Mean   :0.10690   Mean   :0.09828   Mean   :0.01983
##  3rd Qu.:0.12069   3rd Qu.:0.10776   3rd Qu.:0.02155
##  Max.   :0.15517   Max.   :0.13793   Max.   :0.03448
```

```r
set.seed(629)
Splits2 <- createDataPartition(oilType, p = .60, times = 20)
Splits2 <- lapply(Splits2, function(x, y) table(y[x]), y = oilType)
head(Splits2, 3)
```

```
## $Resample01
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
##
## $Resample02
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
##
## $Resample03
##
##  A  B  C  D  E  F  G
## 23 16  2  5  7  6  2
```

```
Splits2 <- do.call("rbind",Splits2)
summary(Splits2)
```

```
##        A            B            C            D            E            F
##  Min.   :23   Min.   :16   Min.   :2   Min.   :5   Min.   :7   Min.   :6
##  1st Qu.:23   1st Qu.:16   1st Qu.:2   1st Qu.:5   1st Qu.:7   1st Qu.:6
##  Median :23   Median :16   Median :2   Median :5   Median :7   Median :6
##  Mean   :23   Mean   :16   Mean   :2   Mean   :5   Mean   :7   Mean   :6
##  3rd Qu.:23   3rd Qu.:16   3rd Qu.:2   3rd Qu.:5   3rd Qu.:7   3rd Qu.:6
##  Max.   :23   Max.   :16   Max.   :2   Max.   :5   Max.   :7   Max.   :6
##        G
##  Min.   :2
##  1st Qu.:2
##  Median :2
##  Mean   :2
##  3rd Qu.:2
##  Max.   :2
```

```
# a) We can see the summary using output
# b) The main difference between is that Stratified random sample has less va
riability and each partition has atleast one sample of each class
# c)  We will usw K-fold cross validation technique to sample the data becaus
e the data set is very small. Since the data is very small we will avoid usin
g test data cases.



#Q3

library(mlbench)

library(corrplot)

library(e1071)

library(caret)
library(AppliedPredictiveModeling)
```

```r
data(Glass)
Glass_data <- subset(Glass,)
library(caTools)
set.seed(101)
sample <- sample.split(Glass_data$Si,SplitRatio = .75)
train <- subset(Glass_data,sample==TRUE)
test <- subset(Glass_data,sample==FALSE)
head(train)
```

```
##        RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 1 1.52101 13.64 4.49 1.10 71.78 0.06 8.75  0 0.00    1
## 2 1.51761 13.89 3.60 1.36 72.73 0.48 7.83  0 0.00    1
## 3 1.51618 13.53 3.55 1.54 72.99 0.39 7.78  0 0.00    1
## 4 1.51766 13.21 3.69 1.29 72.61 0.57 8.22  0 0.00    1
## 5 1.51742 13.27 3.62 1.24 73.08 0.55 8.07  0 0.00    1
## 6 1.51596 12.79 3.61 1.62 72.97 0.64 8.07  0 0.26    1
```

```r
head(test)
```

```
##         RI    Na   Mg   Al    Si    K   Ca Ba   Fe Type
## 11 1.51571 12.72 3.46 1.56 73.20 0.67 8.09  0 0.24    1
## 14 1.51748 12.86 3.56 1.27 73.21 0.54 8.38  0 0.17    1
## 17 1.51784 12.68 3.67 1.16 73.11 0.61 8.70  0 0.00    1
## 22 1.51966 14.77 3.75 0.29 72.02 0.03 9.00  0 0.00    1
## 25 1.51720 13.38 3.50 1.15 72.85 0.50 8.43  0 0.00    1
## 26 1.51764 12.98 3.54 1.21 73.00 0.65 8.53  0 0.00    1
```

```r
##b)
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
```

1

```r
set.seed(1)
cvsplit <- createFolds(Glass_data,k=3,returnTrain = TRUE)
str(cvsplit)
```

```
## List of 3
##  $ Fold1: int [1:7] 1 2 3 5 6 9 10
##  $ Fold2: int [1:6] 1 2 4 7 8 10
##  $ Fold3: int [1:7] 3 4 5 6 7 8 9
```

```r
fold1 <- cvsplit[[1]]
firstsplit <- Glass_data[,(cvsplit$Fold1)]
secondsplit <-Glass_data[,(cvsplit$Fold2)]
thirdsplit <- Glass_data[,(cvsplit$Fold3)]
split <- lapply(cvsplit,function(ind,dat) dat[ind,],dat=Glass_data)
dim(Glass_data)
```

```
## [1] 214  10
```

```r
unlist(lapply(split,nrow))
```

```
## Fold1 Fold2 Fold3
##     7     6     7
```

```r
##b)
library(mlbench)
library(corrplot)
library(e1071)
library(caret)
library(AppliedPredictiveModeling)
library(DAAG)
bootstrap <- createResample(Glass_data,times=214,list=FALSE)
head.matrix(bootstrap)
```

```
##      Resample001 Resample002 Resample003 Resample004 Resample005
## [1,]           2           1           2           1           1
## [2,]           3           2           2           5           1
## [3,]           4           3           5           6           3
## [4,]           4           3           5           6           4
## [5,]           5           4           5           7           5
## [6,]           7           4           6           7           5
##      Resample006 Resample007 Resample008 Resample009 Resample010
## [1,]           1           4           2           1           2
## [2,]           3           4           2           3           3
## [3,]           3           4           3           5           3
## [4,]           4           4           3           5           4
## [5,]           5           5           4           7           5
## [6,]           5           8           4           7           6
##      Resample011 Resample012 Resample013 Resample014 Resample015
## [1,]           1           2           1           2           1
## [2,]           2           3           1           2           2
## [3,]           2           3           6           3           3
## [4,]           4           5           6           3           3
## [5,]           5           5           6           5           4
## [6,]           5           5           6           7           5
##      Resample016 Resample017 Resample018 Resample019 Resample020
## [1,]           1           4           2           2           2
## [2,]           3           4           2           2           2
## [3,]           4           7           3           3           3
## [4,]           4           7           6           4           3
## [5,]           5           7           6           4           3
## [6,]           5           8           8           6           3
##      Resample021 Resample022 Resample023 Resample024 Resample025
## [1,]           2           1           2           2           2
## [2,]           4           2           4           2           3
## [3,]           4           3           4           3           4
## [4,]           7           3           4           4           5
## [5,]           8           4           5           5           5
## [6,]           9           5           5           5           5
```

```
##      Resample026 Resample027 Resample028 Resample029 Resample030
## [1,]           1           1           1           1           1
## [2,]           2           2           1           1           1
## [3,]           2           2           2           2           2
## [4,]           3           2           2           2           4
## [5,]           4           2           2           3           5
## [6,]           6           3           3           4           5
##      Resample031 Resample032 Resample033 Resample034 Resample035
## [1,]           1           3           1           2           1
## [2,]           2           4           2           2           1
## [3,]           2           5           3           3           3
## [4,]           3           7           4           4           5
## [5,]           4           8           4           4           5
## [6,]           5           8           5           5           5
##      Resample036 Resample037 Resample038 Resample039 Resample040
## [1,]           1           2           2           2           1
## [2,]           2           3           3           2           2
## [3,]           3           3           3           2           4
## [4,]           4           4           5           2           7
## [5,]           5           5           6           4           8
## [6,]           6           5           7           4           8
##      Resample041 Resample042 Resample043 Resample044 Resample045
## [1,]           4           1           1           1           1
## [2,]           4           2           1           1           1
## [3,]           5           4           2           2           2
## [4,]           5           5           4           3           4
## [5,]           7           5           5           3           5
## [6,]           8           6           6           4           6
##      Resample046 Resample047 Resample048 Resample049 Resample050
## [1,]           1           2           1           1           5
## [2,]           2           3           2           1           5
## [3,]           2           4           2           2           6
## [4,]           3           5           2           3           6
## [5,]           5           5           6           4           7
## [6,]           5           5           6           5           7
##      Resample051 Resample052 Resample053 Resample054 Resample055
## [1,]           1           2           1           2           1
## [2,]           2           4           1           2           1
## [3,]           4           4           2           4           1
## [4,]           5           8           2           4           4
## [5,]           6           8           3           4           6
## [6,]           7           9           4           6           8
##      Resample056 Resample057 Resample058 Resample059 Resample060
## [1,]           2           1           1           3           2
## [2,]           3           2           1           4           2
## [3,]           3           3           4           5           4
## [4,]           4           5           5           5           4
## [5,]           7           5           6           5           8
## [6,]           7           6           6           6           9
##      Resample061 Resample062 Resample063 Resample064 Resample065
```

```
## [1,]           1           2           1           1           1
## [2,]           2           3           1           3           2
## [3,]           4           3           4           3           4
## [4,]           4           4           4           4           4
## [5,]           5           5           4           7           6
## [6,]           6           6           5           8           6
##      Resample066 Resample067 Resample068 Resample069 Resample070
## [1,]           1           2           1           1           1
## [2,]           2           3           2           2           5
## [3,]           4           3           4           5           6
## [4,]           4           4           5           5           7
## [5,]           4           5           6           6           7
## [6,]           6           6           6           6           8
##      Resample071 Resample072 Resample073 Resample074 Resample075
## [1,]           1           1           1           2           2
## [2,]           2           1           1           3           2
## [3,]           2           3           1           3           3
## [4,]           5           4           3           4           5
## [5,]           5           5           3           4           5
## [6,]           5           6           6           5           5
##      Resample076 Resample077 Resample078 Resample079 Resample080
## [1,]           2           1           1           1           1
## [2,]           2           1           2           2           1
## [3,]           2           3           3           3           3
## [4,]           3           5           5           4           4
## [5,]           6           5           5           4           4
## [6,]           8           7           7           6           7
##      Resample081 Resample082 Resample083 Resample084 Resample085
## [1,]           3           1           1           1           1
## [2,]           4           3           5           2           1
## [3,]           4           4           6           3           3
## [4,]           4           4           6           4           3
## [5,]           4           4           7           5           3
## [6,]           6           6           8           5           3
##      Resample086 Resample087 Resample088 Resample089 Resample090
## [1,]           1           1           4           1           1
## [2,]           1           3           4           1           1
## [3,]           1           3           5           1           1
## [4,]           1           5           6           1           1
## [5,]           3           5           7           1           2
## [6,]           4           5           8           1           2
##      Resample091 Resample092 Resample093 Resample094 Resample095
## [1,]           1           1           3           1           1
## [2,]           1           3           3           3           1
## [3,]           2           4           4           5           2
## [4,]           5           5           5           5           2
## [5,]           6           6           7           5           3
## [6,]           7           8           8           7           5
##      Resample096 Resample097 Resample098 Resample099 Resample100
## [1,]           2           1           2           2           1
```

```
## [2,]           3           1           2           2           2
## [3,]           4           2           2           2           4
## [4,]           4           5           4           3           6
## [5,]           4           6           4           3           6
## [6,]           5           7           5           3           6
##      Resample101 Resample102 Resample103 Resample104 Resample105
## [1,]           1           1           2           2           1
## [2,]           2           2           2           2           1
## [3,]           3           2           4           3           2
## [4,]           3           3           4           4           5
## [5,]           5           4           5           5           5
## [6,]           6           6           6           5           6
##      Resample106 Resample107 Resample108 Resample109 Resample110
## [1,]           2           1           1           1           2
## [2,]           2           1           2           3           2
## [3,]           2           3           4           3           3
## [4,]           4           4           4           5           4
## [5,]           5           4           5           6           4
## [6,]           6           5           6           7           5
##      Resample111 Resample112 Resample113 Resample114 Resample115
## [1,]           1           2           2           2           1
## [2,]           3           3           2           2           1
## [3,]           4           3           3           2           2
## [4,]           4           4           5           3           3
## [5,]           6           5           5           3           3
## [6,]           7           5           5           4           4
##      Resample116 Resample117 Resample118 Resample119 Resample120
## [1,]           1           1           1           1           4
## [2,]           2           1           1           3           5
## [3,]           3           3           2           3           7
## [4,]           3           5           2           4           7
## [5,]           4           5           4           7           8
## [6,]           4           7           4           8           8
##      Resample121 Resample122 Resample123 Resample124 Resample125
## [1,]           1           1           1           1           1
## [2,]           1           1           1           1           2
## [3,]           1           2           4           3           2
## [4,]           2           4           4           4           2
## [5,]           3           4           4           6           4
## [6,]           5           6           5           8           5
##      Resample126 Resample127 Resample128 Resample129 Resample130
## [1,]           1           2           1           1           1
## [2,]           2           3           3           1           2
## [3,]           3           4           3           1           3
## [4,]           4           7           4           4           5
## [5,]           4           7           4           4           6
## [6,]           4           8           5           5           6
##      Resample131 Resample132 Resample133 Resample134 Resample135
## [1,]           1           1           1           1           1
## [2,]           2           1           3           2           1
```

```
## [3,]          2          2          4          2          2
## [4,]          5          2          5          2          2
## [5,]          5          2          6          2          2
## [6,] 1        6          4          7          3          3
##      Resample136 Resample137 Resample138 Resample139 Resample140
## [1,]           1           2           1           1           1
## [2,]           2           2           3           3           2
## [3,]           2           3           3           3           2
## [4,]           4           4           4           4           3
## [5,]           4           4           5           6           4
## [6,] 1         6           6           6           8           5
##      Resample141 Resample142 Resample143 Resample144 Resample145
## [1,]           1           1           1           2           1
## [2,]           1           2           3           3           1
## [3,]           4           3           4           3           2
## [4,]           4           4           4           4           3
## [5,]           4           5           5           5           4
## [6,]           5           6           6           5           6
##      Resample146 Resample147 Resample148 Resample149 Resample150
## [1,]           2           2           1           1           1
## [2,]           2           4           1           2           1
## [3,]           2           5           1           3           2
## [4,]           4           5           2           5           2
## [5,]           4           5           4           7           4
## [6,]           4           7           5           7           6
##      Resample151 Resample152 Resample153 Resample154 Resample155
## [1,]           1           2           1           1           1
## [2,]           3           5           2           1           2
## [3,]           4           5           2           4           3
## [4,]           5           6           3           5           4
## [5,]           5           8           6           5           6
## [6,]           7           9           6           5           8
##      Resample156 Resample157 Resample158 Resample159 Resample160
## [1,]           1           1           5           1           1
## [2,]           2           2           5           1           1
## [3,]           3           3           5           2           2
## [4,]           4           5           6           2           2
## [5,]           7           6           6           2           2
## [6,]           8           7           7           2           3
##      Resample161 Resample162 Resample163 Resample164 Resample165
## [1,]           1           1           1           3           1
## [2,]           4           1           1           4           1
## [3,]           4           5           1           4           2
## [4,]           5           6           1           5           3
## [5,]           7           7           1           6           4
## [6,]           8           8           4           6           5
##      Resample166 Resample167 Resample168 Resample169 Resample170
## [1,]           3           2           1           1           1
## [2,]           3           2           4           1           2
## [3,]           3           2           4           2           2
```

```
## [4,]          5          2          7          3          2
## [5,]          5          7          9          6          3
## [6,]          6          8          9          6          4
##      Resample171 Resample172 Resample173 Resample174 Resample175
## [1,]          1          3          1          1          1
## [2,]          2          4          1          3          3
## [3,]          2          4          1          4          4
## [4,]          3          5          5          4          5
## [5,]          3          7          5          5          6
## [6,]          3          8          6          5          6
##      Resample176 Resample177 Resample178 Resample179 Resample180
## [1,]          2          1          2          1          1
## [2,]          4          1          2          2          2
## [3,]          5          2          3          2          4
## [4,]          5          2          3          3          4
## [5,]          6          3          4          3          5
## [6,]          6          3          4          5          6
##      Resample181 Resample182 Resample183 Resample184 Resample185
## [1,]          1          1          1          1          2
## [2,]          2          4          3          2          2
## [3,]          2          5          5          2          2
## [4,]          2          6          5          5          5
## [5,]          3          6          6          7          7
## [6,]          6          6          6          7          7
##      Resample186 Resample187 Resample188 Resample189 Resample190
## [1,]          1          1          2          1          1
## [2,]          1          2          2          2          4
## [3,]          3          3          3          4          4
## [4,]          5          3          3          4          5
## [5,]          5          4          4          5          5
## [6,]          6          6          5          7          6
##      Resample191 Resample192 Resample193 Resample194 Resample195
## [1,]          1          1          1          1          1
## [2,]          1          1          1          1          2
## [3,]          1          3          2          1          2
## [4,]          2          3          4          2          3
## [5,]          4          5          5          3          3
## [6,]          5          7          5          3          5
##      Resample196 Resample197 Resample198 Resample199 Resample200
## [1,]          2          1          3          1          1
## [2,]          3          3          4          2          1
## [3,]          4          3          4          2          2
## [4,]          6          5          5          3          5
## [5,]          7          5          6          3          5
## [6,]          7          6          6          4          6
##      Resample201 Resample202 Resample203 Resample204 Resample205
## [1,]          1          1          1          1          1
## [2,]          1          1          1          2          2
## [3,]          1          1          2          2          2
## [4,]          2          1          3          4          2
```

```
## [5,]           4           3           3           4           3
## [6,]           4           5           3           7           3
##      Resample206 Resample207 Resample208 Resample209 Resample210
## [1,]           1           1           2           1           1
## [2,]           3           3           2           2           2
## [3,]           3           5           4           6           3
## [4,]           4           5           6           6           6
## [5,]           5           5           6           6           6
## [6,]           6           7           6           7           7
##      Resample211 Resample212 Resample213 Resample214
## [1,]           2           1           1           2
## [2,]           3           1           1           2
## [3,]           4           1           3           2
## [4,]           4           1           4           3
## [5,]           5           3           4           4
## [6,]           9           3           5           5
```

Assignment 3

ORIGINALITY REPORT

%90	%9	%4	%90
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to UT, Dallas Student Paper	%90

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	OFF		