

# Week6\_3.R

*vinodkumar*

*Tue Dec 19 19:39:42 2017*

```
library(mlr)
```

```
## Loading required package: ParamHelpers
```

```
## Warning: replacing previous import 'BBmisc::isFALSE' by  
## 'backports::isFALSE' when loading 'mlr'
```

```
library(e1071)
```

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

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

```
library(MLmetrics)
```

```
##  
## Attaching package: 'MLmetrics'
```

```
## The following object is masked from 'package:base':  
##  
##      Recall
```

```
library(DMwR)
```

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

```
## Loading required package: grid
```

```
library(ggplot2)
setwd("G:/DATA SCIENCE/Internship/Mid TERM")
actual_data<-read.table("house-votes-84.data.txt",sep=',')
dim(actual_data)
```

```
## [1] 435 17
```

```
dimnames(actual_data)
```

```

## [[1]]
## [1] "1" "2" "3" "4" "5" "6" "7" "8" "9" "10" "11"
## [12] "12" "13" "14" "15" "16" "17" "18" "19" "20" "21" "22"
## [23] "23" "24" "25" "26" "27" "28" "29" "30" "31" "32" "33"
## [34] "34" "35" "36" "37" "38" "39" "40" "41" "42" "43" "44"
## [45] "45" "46" "47" "48" "49" "50" "51" "52" "53" "54" "55"
## [56] "56" "57" "58" "59" "60" "61" "62" "63" "64" "65" "66"
## [67] "67" "68" "69" "70" "71" "72" "73" "74" "75" "76" "77"
## [78] "78" "79" "80" "81" "82" "83" "84" "85" "86" "87" "88"
## [89] "89" "90" "91" "92" "93" "94" "95" "96" "97" "98" "99"
## [100] "100" "101" "102" "103" "104" "105" "106" "107" "108" "109" "110"
## [111] "111" "112" "113" "114" "115" "116" "117" "118" "119" "120" "121"
## [122] "122" "123" "124" "125" "126" "127" "128" "129" "130" "131" "132"
## [133] "133" "134" "135" "136" "137" "138" "139" "140" "141" "142" "143"
## [144] "144" "145" "146" "147" "148" "149" "150" "151" "152" "153" "154"
## [155] "155" "156" "157" "158" "159" "160" "161" "162" "163" "164" "165"
## [166] "166" "167" "168" "169" "170" "171" "172" "173" "174" "175" "176"
## [177] "177" "178" "179" "180" "181" "182" "183" "184" "185" "186" "187"
## [188] "188" "189" "190" "191" "192" "193" "194" "195" "196" "197" "198"
## [199] "199" "200" "201" "202" "203" "204" "205" "206" "207" "208" "209"
## [210] "210" "211" "212" "213" "214" "215" "216" "217" "218" "219" "220"
## [221] "221" "222" "223" "224" "225" "226" "227" "228" "229" "230" "231"
## [232] "232" "233" "234" "235" "236" "237" "238" "239" "240" "241" "242"
## [243] "243" "244" "245" "246" "247" "248" "249" "250" "251" "252" "253"
## [254] "254" "255" "256" "257" "258" "259" "260" "261" "262" "263" "264"
## [265] "265" "266" "267" "268" "269" "270" "271" "272" "273" "274" "275"
## [276] "276" "277" "278" "279" "280" "281" "282" "283" "284" "285" "286"
## [287] "287" "288" "289" "290" "291" "292" "293" "294" "295" "296" "297"
## [298] "298" "299" "300" "301" "302" "303" "304" "305" "306" "307" "308"
## [309] "309" "310" "311" "312" "313" "314" "315" "316" "317" "318" "319"
## [320] "320" "321" "322" "323" "324" "325" "326" "327" "328" "329" "330"
## [331] "331" "332" "333" "334" "335" "336" "337" "338" "339" "340" "341"
## [342] "342" "343" "344" "345" "346" "347" "348" "349" "350" "351" "352"
## [353] "353" "354" "355" "356" "357" "358" "359" "360" "361" "362" "363"
## [364] "364" "365" "366" "367" "368" "369" "370" "371" "372" "373" "374"
## [375] "375" "376" "377" "378" "379" "380" "381" "382" "383" "384" "385"
## [386] "386" "387" "388" "389" "390" "391" "392" "393" "394" "395" "396"
## [397] "397" "398" "399" "400" "401" "402" "403" "404" "405" "406" "407"
## [408] "408" "409" "410" "411" "412" "413" "414" "415" "416" "417" "418"
## [419] "419" "420" "421" "422" "423" "424" "425" "426" "427" "428" "429"
## [430] "430" "431" "432" "433" "434" "435"
##
## [[2]]
## [1] "V1" "V2" "V3" "V4" "V5" "V6" "V7" "V8" "V9" "V10" "V11"
## [12] "V12" "V13" "V14" "V15" "V16" "V17"

```

```
str(actual_data)
```

```
## 'data.frame':    435 obs. of  17 variables:
## $ V1 : Factor w/ 2 levels "democrat","republican": 2 2 1 1 1 1 1 2 2 1 ...
## $ V2 : Factor w/ 3 levels "?","n","y": 2 2 1 2 3 2 2 2 2 3 ...
## $ V3 : Factor w/ 3 levels "?","n","y": 3 3 3 3 3 3 3 3 3 3 ...
## $ V4 : Factor w/ 3 levels "?","n","y": 2 2 3 3 3 3 2 2 2 3 ...
## $ V5 : Factor w/ 3 levels "?","n","y": 3 3 1 2 2 2 3 3 3 2 ...
## $ V6 : Factor w/ 3 levels "?","n","y": 3 3 3 1 3 3 3 3 3 2 ...
## $ V7 : Factor w/ 3 levels "?","n","y": 3 3 3 3 3 3 3 3 3 2 ...
## $ V8 : Factor w/ 3 levels "?","n","y": 2 2 2 2 2 2 2 2 2 3 ...
## $ V9 : Factor w/ 3 levels "?","n","y": 2 2 2 2 2 2 2 2 2 3 ...
## $ V10: Factor w/ 3 levels "?","n","y": 2 2 2 2 2 2 2 2 2 3 ...
## $ V11: Factor w/ 3 levels "?","n","y": 3 2 2 2 2 2 2 2 2 2 ...
## $ V12: Factor w/ 3 levels "?","n","y": 1 2 3 3 3 2 2 2 2 2 ...
## $ V13: Factor w/ 3 levels "?","n","y": 3 3 2 2 1 2 2 2 3 2 ...
## $ V14: Factor w/ 3 levels "?","n","y": 3 3 3 3 3 3 1 3 3 2 ...
## $ V15: Factor w/ 3 levels "?","n","y": 3 3 3 2 3 3 3 3 3 2 ...
## $ V16: Factor w/ 3 levels "?","n","y": 2 2 2 2 3 3 3 1 2 1 ...
## $ V17: Factor w/ 3 levels "?","n","y": 3 1 2 3 3 3 3 3 3 1 ...
```

```
summarizeColumns(actual_data)
```

```
##   name  type na mean      disp median mad min max nlevs
## 1  V1 factor  0  NA 0.3862069      NA  NA 168 267     2
## 2  V2 factor  0  NA 0.4574713      NA  NA  12 236     3
## 3  V3 factor  0  NA 0.5517241      NA  NA  48 195     3
## 4  V4 factor  0  NA 0.4183908      NA  NA  11 253     3
## 5  V5 factor  0  NA 0.4321839      NA  NA  11 247     3
## 6  V6 factor  0  NA 0.5126437      NA  NA  15 212     3
## 7  V7 factor  0  NA 0.3747126      NA  NA  11 272     3
## 8  V8 factor  0  NA 0.4505747      NA  NA  14 239     3
## 9  V9 factor  0  NA 0.4436782      NA  NA  15 242     3
## 10 V10 factor  0  NA 0.5241379      NA  NA  22 207     3
## 11 V11 factor  0  NA 0.5034483      NA  NA   7 216     3
## 12 V12 factor  0  NA 0.3931034      NA  NA  21 264     3
## 13 V13 factor  0  NA 0.4643678      NA  NA  31 233     3
## 14 V14 factor  0  NA 0.5195402      NA  NA  25 209     3
## 15 V15 factor  0  NA 0.4298851      NA  NA  17 248     3
## 16 V16 factor  0  NA 0.4643678      NA  NA  28 233     3
## 17 V17 factor  0  NA 0.3816092      NA  NA  62 269     3
```

```
summarizeLevels(actual_data)
```

```
## $V1
##
##   democrat republican
##      267      168
##
## $V2
##
##   ?   n   y
##  12 236 187
##
## $V3
##
##   ?   n   y
##  48 192 195
##
## $V4
##
##   ?   n   y
##  11 171 253
##
## $V5
##
##   ?   n   y
##  11 247 177
##
## $V6
##
##   ?   n   y
##  15 208 212
##
## $V7
##
##   ?   n   y
##  11 152 272
##
## $V8
##
##   ?   n   y
##  14 182 239
##
## $V9
##
##   ?   n   y
##  15 178 242
##
## $V10
##
##   ?   n   y
```

```
## 22 206 207
##
## $V11
##
## ? n y
## 7 212 216
##
## $V12
##
## ? n y
## 21 264 150
##
## $V13
##
## ? n y
## 31 233 171
##
## $V14
##
## ? n y
## 25 201 209
##
## $V15
##
## ? n y
## 17 170 248
##
## $V16
##
## ? n y
## 28 233 174
##
## $V17
##
## ? n y
## 104 62 269
```

```
sum(is.na(actual_data))
```

```
## [1] 0
```

```
class(actual_data)
```

```
## [1] "data.frame"
```

```
head(actual_data)
```

```
##           V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17
## 1 republican n  y n  y y  y n  n  n  y  ?  y  y  y  n  y
## 2 republican n  y n  y y  y n  n  n  n  n  y  y  y  n  ?
## 3  democrat ?  y y  ? y  y n  n  n  n  y  n  y  y  n  n
## 4  democrat n  y y  n ?  y n  n  n  n  y  n  y  n  n  y
## 5  democrat y  y y  n y  y n  n  n  n  y  ?  y  y  y  y
## 6  democrat n  y y  n y  y n  n  n  n  n  n  y  y  y  y
```

```
tail(actual_data)
```

```
##           V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17
## 430 democrat y  n y  n ?  n y  y  y  y  n  y  n  ?  y  y
## 431 republican n  n y  y y  y n  n  y  y  n  y  y  y  n  y
## 432  democrat n  n y  n n  n y  y  y  y  n  n  n  n  n  y
## 433 republican n  ? n  y y  y n  n  n  n  y  y  y  y  n  y
## 434 republican n  n n  y y  y ?  ?  ?  ?  n  y  y  y  n  y
## 435 republican n  y n  y y  y n  n  n  y  n  y  y  y  ?  n
```

```
plot(actual_data$V1)
summary(actual_data$V1)
```

```
##  democrat republican
##           267           168
```

```
actual_data[actual_data=="?"]<- NA
head(actual_data) #now ? is raplaced with NA
```

```
##           V1  V2 V3 V4  V5  V6 V7 V8 V9 V10 V11  V12  V13 V14 V15 V16
## 1 republican  n  y n  y  y  y n  n  n  y <NA>  y  y  y  n
## 2 republican  n  y n  y  y  y n  n  n  n  n  y  y  y  n
## 3  democrat <NA> y  y <NA>  y  y n  n  n  n  y  n  y  y  n
## 4  democrat  n  y y  n <NA> y  n n  n  n  y  n  y  n  n
## 5  democrat  y  y y  n  y  y n  n  n  n  y <NA>  y  y  y
## 6  democrat  n  y y  n  y  y n  n  n  n  n  n  y  y  y
##           V17
## 1  y
## 2 <NA>
## 3  n
## 4  y
## 5  y
## 6  y
```

```
str(actual_data)
```

```
## 'data.frame':    435 obs. of  17 variables:
## $ V1 : Factor w/ 2 levels "democrat","republican": 2 2 1 1 1 1 1 2 2 1 ...
## $ V2 : Factor w/ 3 levels "?","n","y": 2 2 NA 2 3 2 2 2 3 3 ...
## $ V3 : Factor w/ 3 levels "?","n","y": 3 3 3 3 3 3 3 3 3 3 ...
## $ V4 : Factor w/ 3 levels "?","n","y": 2 2 3 3 3 3 2 2 2 3 ...
## $ V5 : Factor w/ 3 levels "?","n","y": 3 3 NA 2 2 2 3 3 3 2 ...
## $ V6 : Factor w/ 3 levels "?","n","y": 3 3 3 NA 3 3 3 3 3 2 ...
## $ V7 : Factor w/ 3 levels "?","n","y": 3 3 3 3 3 3 3 3 3 2 ...
## $ V8 : Factor w/ 3 levels "?","n","y": 2 2 2 2 2 2 2 2 2 3 ...
## $ V9 : Factor w/ 3 levels "?","n","y": 2 2 2 2 2 2 2 2 2 3 ...
## $ V10: Factor w/ 3 levels "?","n","y": 2 2 2 2 2 2 2 2 2 3 ...
## $ V11: Factor w/ 3 levels "?","n","y": 3 2 2 2 2 2 2 2 2 2 ...
## $ V12: Factor w/ 3 levels "?","n","y": NA 2 3 3 3 2 2 2 2 2 ...
## $ V13: Factor w/ 3 levels "?","n","y": 3 3 2 2 NA 2 2 2 3 2 ...
## $ V14: Factor w/ 3 levels "?","n","y": 3 3 3 3 3 3 NA 3 3 2 ...
## $ V15: Factor w/ 3 levels "?","n","y": 3 3 3 2 3 3 3 3 3 2 ...
## $ V16: Factor w/ 3 levels "?","n","y": 2 2 2 2 3 3 3 NA 2 NA ...
## $ V17: Factor w/ 3 levels "?","n","y": 3 NA 2 3 3 3 3 3 3 NA ...
```

```
droplevels(actual_data)->actual_data
sum(is.na(actual_data)) #392
```

```
## [1] 392
```

```
sum(is.na(actual_data))/(nrow(actual_data)*ncol(actual_data))*100 #I found 5.300879%
of data missing
```

```
## [1] 5.300879
```

```
number_miss_val<-colSums(is.na(actual_data))
number_miss_val
```

```
## V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17
##  0 12 48 11 11 15 11 14 15 22  7 21 31 25 17 28 104
```

```
per_miss_val<-(colSums(is.na(actual_data))/nrow(actual_data))*100
per_miss_val
```



```
##          V1          V2          V3          V4          V5          V6          V7
## 0.000000  2.758621 11.034483  2.528736  2.528736  3.448276  2.528736
##          V8          V9          V10         V11         V12         V13         V14
## 3.218391  3.448276  5.057471  1.609195  4.827586  7.126437  5.747126
##          V15         V16         V17
## 3.908046  6.436782 23.908046
```

```
rbind(number_miss_val,per_miss_val)
```

```
##          V1          V2          V3          V4          V5          V6
## number_miss_val  0 12.000000 48.00000 11.000000 11.000000 15.000000
## per_miss_val    0  2.758621 11.03448  2.528736  2.528736  3.448276
##          V7          V8          V9          V10         V11         V12
## number_miss_val 11.000000 14.000000 15.000000 22.000000  7.000000 21.000000
## per_miss_val    2.528736  3.218391  3.448276  5.057471  1.609195  4.827586
##          V13         V14         V15         V16         V17
## number_miss_val 31.000000 25.000000 17.000000 28.000000 104.00000
## per_miss_val    7.126437  5.747126  3.908046  6.436782  23.90805
```

```
lables = c("classname","handicapped_infants","water_project_cost_sharing",
           "adoption_of_the_budget_resolution","physician_fee_freeze",
           "el_salvador_aid","religious_groups_in_schools",
           "anti_satellite_test_ban","aid_to_nicaraguan_contras",
           "mx_missile","immigration","synfuels_corporation_cutback",
           "education_spending","superfund_right_to_sue","crime","duty_free_exports",
           "export_administration_act_south_africa")
names(actual_data)[1:17]<-lables
names(actual_data)
```

```
## [1] "classname"
## [2] "handicapped_infants"
## [3] "water_project_cost_sharing"
## [4] "adoption_of_the_budget_resolution"
## [5] "physician_fee_freeze"
## [6] "el_salvador_aid"
## [7] "religious_groups_in_schools"
## [8] "anti_satellite_test_ban"
## [9] "aid_to_nicaraguan_contras"
## [10] "mx_missile"
## [11] "immigration"
## [12] "synfuels_corporation_cutback"
## [13] "education_spending"
## [14] "superfund_right_to_sue"
## [15] "crime"
## [16] "duty_free_exports"
## [17] "export_administration_act_south_africa"
```

```
actual_data_colnames<- c(colnames(actual_data))
for (i in actual_data_colnames){
  chi_sq_result<-chisq.test(actual_data$classname,actual_data[,i])
  print(i)
  print(chi_sq_result)
}
```

```

## [1] "classname"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 430.79, df = 1, p-value < 2.2e-16
##
## [1] "handicapped_infants"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 69.195, df = 1, p-value < 2.2e-16
##
## [1] "water_project_cost_sharing"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 0, df = 1, p-value = 1
##
## [1] "adoption_of_the_budget_resolution"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 234.65, df = 1, p-value < 2.2e-16
##
## [1] "physician_fee_freeze"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 357.59, df = 1, p-value < 2.2e-16
##
## [1] "el_salvador_aid"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 214.05, df = 1, p-value < 2.2e-16
##
## [1] "religious_groups_in_schools"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 75.972, df = 1, p-value < 2.2e-16

```

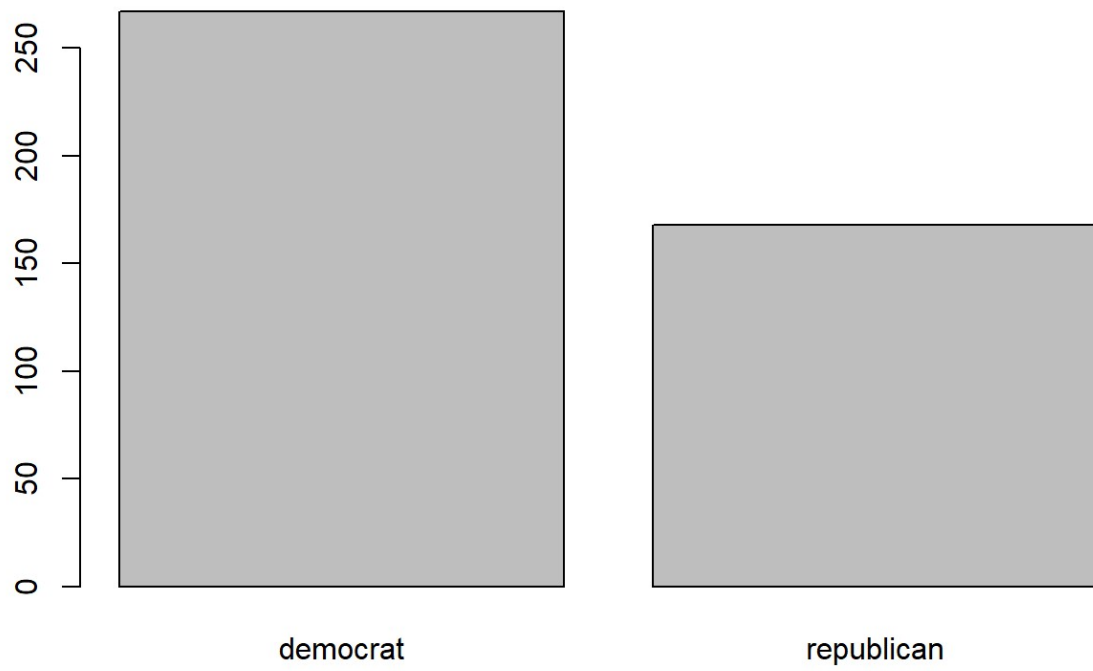
```

##
## [1] "anti_satellite_test_ban"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 112.55, df = 1, p-value < 2.2e-16
##
## [1] "aid_to_nicaraguan_contras"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 181.24, df = 1, p-value < 2.2e-16
##
## [1] "mx_missile"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 161.25, df = 1, p-value < 2.2e-16
##
## [1] "immigration"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 2.6717, df = 1, p-value = 0.1021
##
## [1] "synfuels_corporation_cutback"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 57.623, df = 1, p-value = 3.176e-14
##
## [1] "education_spending"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 203.53, df = 1, p-value < 2.2e-16
##
## [1] "superfund_right_to_sue"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 124.46, df = 1, p-value < 2.2e-16

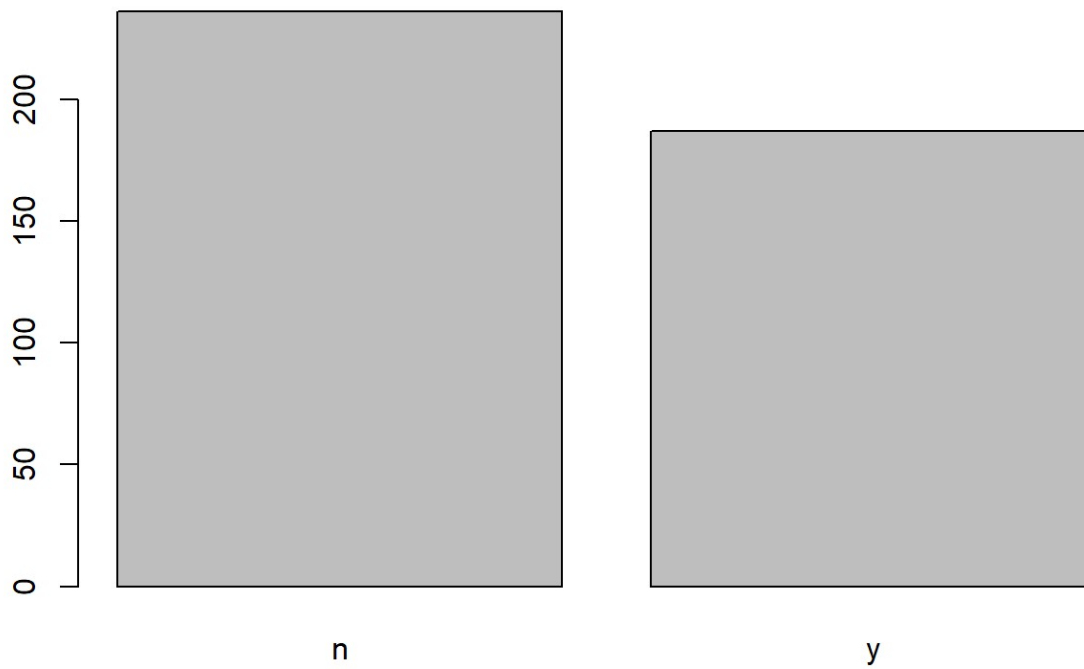
```

```
##
## [1] "crime"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 160.82, df = 1, p-value < 2.2e-16
##
## [1] "duty_free_exports"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 115.69, df = 1, p-value < 2.2e-16
##
## [1] "export_administration_act_south_africa"
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  actual_data$classname and actual_data[, i]
## X-squared = 39.506, df = 1, p-value = 3.27e-10
```

```
#to check the skewness of data
plot(actual_data$classname)
```



```
plot(actual_data$handicapped_infants)
```



```
d1<-c(actual_data$handicapped_infants,actual_data$water_project_cost_sharing)
republic <- actual_data[which(actual_data$classname == 'republican'),]
democrat <- actual_data[which(actual_data$classname == 'democrat'),]
sum(is.na(republic))  #131
```

```
## [1] 131
```

```
sum(is.na(democrat))  #261
```

```
## [1] 261
```

```
dim(republic)
```

```
## [1] 168  17
```

```
dim(democrat)
```

```
## [1] 267  17
```

```

republic <- centralImputation(republic)
democrat <- centralImputation(democrat)
summary(republic)

```

```

##      classname  handicapped_infants water_project_cost_sharing
## democrat   : 0   n:137                n:73
## republican:168  y: 31                y:95
## adoption_of_the_budget_resolution physician_fee_freeze el_salvador_aid
## n:146                n: 2                n: 8
## y: 22                y:166                y:160
## religious_groups_in_schools anti_satellite_test_ban
## n: 17                n:129
## y:151                y: 39
## aid_to_nicaraguan_contras mx_missile immigration
## n:144                n:149                n:73
## y: 24                y: 19                y:95
## synfuels_corporation_cutback education_spending superfund_right_to_sue
## n:147                n: 20                n: 22
## y: 21                y:148                y:146
## crime      duty_free_exports export_administration_act_south_africa
## n: 3   n:154                n: 50
## y:165  y: 14                y:118

```

```

summary(democrat)

```

```

##      classname  handicapped_infants water_project_cost_sharing
## democrat   :267  n:102                n:119
## republican: 0   y:165                y:148
## adoption_of_the_budget_resolution physician_fee_freeze el_salvador_aid
## n: 29                n:253                n:212
## y:238                y: 14                y: 55
## religious_groups_in_schools anti_satellite_test_ban
## n:144                n: 59
## y:123                y:208
## aid_to_nicaraguan_contras mx_missile immigration
## n: 45                n: 60                n:143
## y:222                y:207                y:124
## synfuels_corporation_cutback education_spending superfund_right_to_sue
## n:126                n:231                n:194
## y:141                y: 36                y: 73
## crime      duty_free_exports export_administration_act_south_africa
## n:177  n: 91                n: 12
## y: 90  y:176                y:255

```



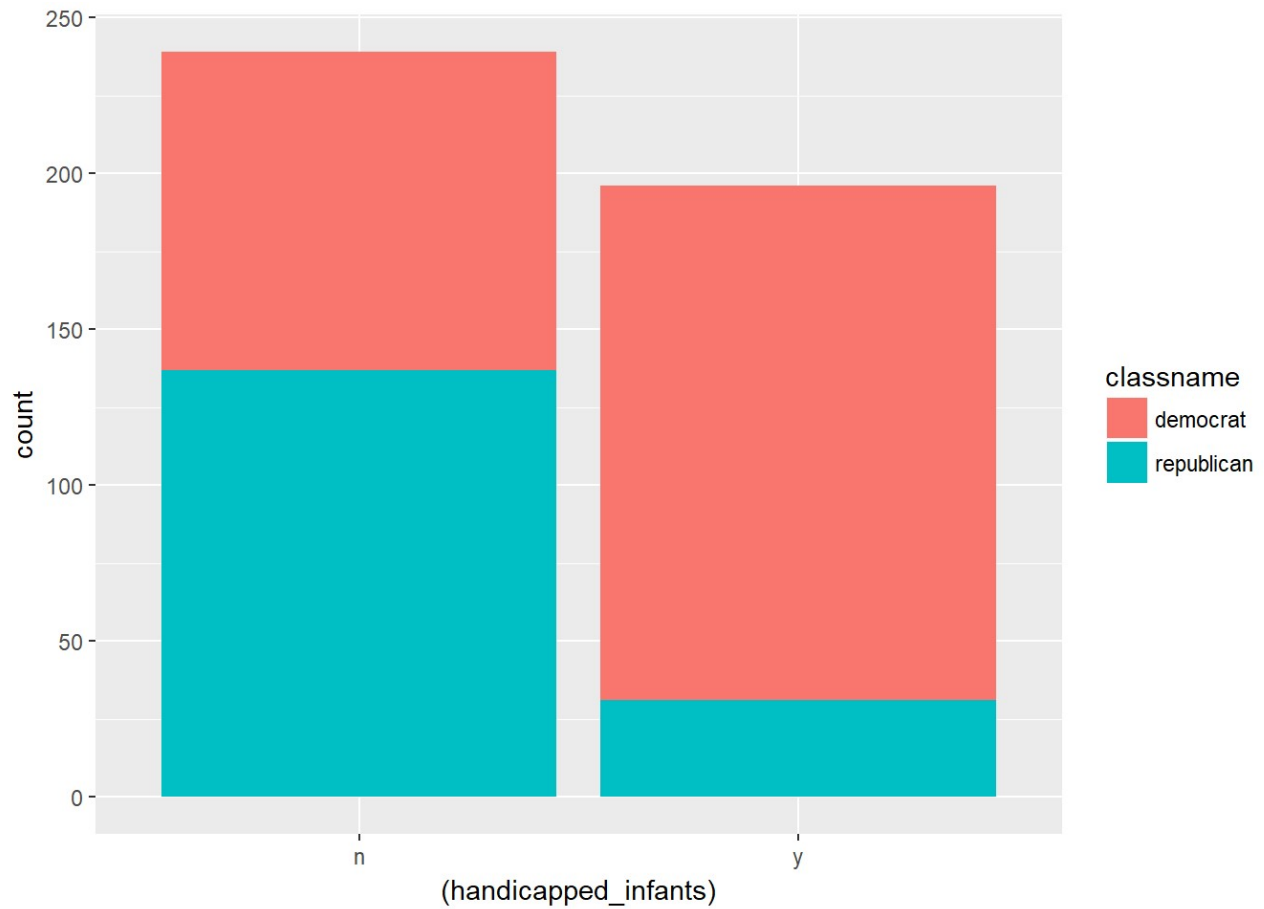
```
imp_data <- rbind(republic,democrat)
dim(imp_data)
```

```
## [1] 435  17
```

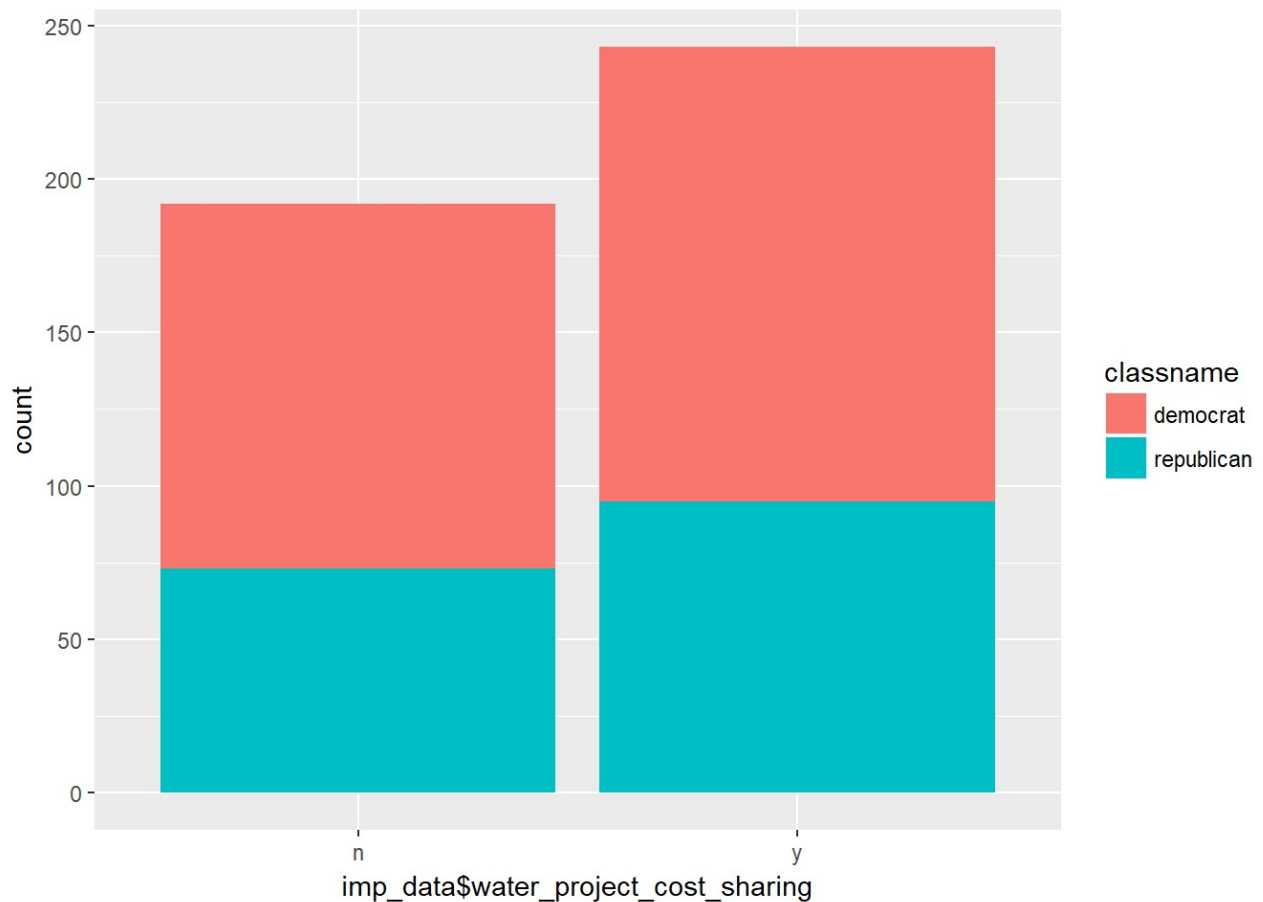
```
str(imp_data)
```

```
## 'data.frame':    435 obs. of  17 variables:
## $ classname                : Factor w/ 2 levels "democrat","republica
n": 2 2 2 2 2 2 2 2 2 2 ...
## $ handicapped_infants      : Factor w/ 2 levels "n","y": 1 1 1 1 1 1
1 1 1 2 ...
## $ water_project_cost_sharing : Factor w/ 2 levels "n","y": 2 2 2 2 2 2
2 2 2 1 ...
## $ adoption_of_the_budget_resolution : Factor w/ 2 levels "n","y": 1 1 1 1 1 1
1 1 1 1 ...
## $ physician_fee_freeze     : Factor w/ 2 levels "n","y": 2 2 2 2 2 2
2 2 2 2 ...
## $ el_salvador_aid          : Factor w/ 2 levels "n","y": 2 2 2 2 2 2
2 2 2 2 ...
## $ religious_groups_in_schools : Factor w/ 2 levels "n","y": 2 2 2 2 1 2
2 2 2 1 ...
## $ anti_satellite_test_ban  : Factor w/ 2 levels "n","y": 1 1 1 1 1 1
1 1 1 2 ...
## $ aid_to_nicaraguan_contras : Factor w/ 2 levels "n","y": 1 1 1 1 1 1
1 1 1 2 ...
## $ mx_missile               : Factor w/ 2 levels "n","y": 1 1 1 1 1 1
1 1 1 2 ...
## $ immigration              : Factor w/ 2 levels "n","y": 2 1 1 1 1 1
1 2 1 1 ...
## $ synfuels_corporation_cutback : Factor w/ 2 levels "n","y": 1 1 1 1 1 2
1 1 1 1 ...
## $ education_spending       : Factor w/ 2 levels "n","y": 2 2 1 2 2 2
2 2 2 2 ...
## $ superfund_right_to_sue    : Factor w/ 2 levels "n","y": 2 2 2 2 2 2
2 2 2 2 ...
## $ crime                    : Factor w/ 2 levels "n","y": 2 2 2 2 2 2
2 2 2 2 ...
## $ duty_free_exports         : Factor w/ 2 levels "n","y": 1 1 1 1 1 1
1 1 1 1 ...
## $ export_administration_act_south_africa: Factor w/ 2 levels "n","y": 2 2 2 2 1 2
2 2 1 2 ...
```

```
p1 = ggplot(data = imp_data, mapping = aes(x=(handicapped_infants), fill = classname))+  
  geom_bar()  
p2 = ggplot(data = imp_data, mapping = aes(x= imp_data$water_project_cost_sharing, fill  
  = classname))+geom_bar()  
p1
```



p2



```
set.seed(5000)
proportion_data <- sample(seq(1,2),size = nrow(imp_data),replace = TRUE, prob = c(.8,
.2))
train_imp_data <- imp_data[proportion_data == 1,]
test_imp_data <- imp_data[proportion_data == 2,]
svmfit<-svm(train_imp_data$classname~.,train_imp_data,kernal="linear",cost=10,scale =
FALSE)
svmfit
```

```
##
## Call:
## svm(formula = train_imp_data$classname ~ ., data = train_imp_data,
##     kernal = "linear", cost = 10, scale = FALSE)
##
##
## Parameters:
##   SVM-Type:  C-classification
##   SVM-Kernel: radial
##     cost:   10
##   gamma:  0.05882353
##
## Number of Support Vectors:  48
```

```
pd_svmfit<-predict(svmfit,train_imp_data)
pd_svmfit
```

##	1	2	8	11	12	15
##	republican	republican	republican	republican	republican	republican
##	16	19	29	31	37	38
##	republican	republican	republican	republican	republican	republican
##	39	50	52	56	57	58
##	republican	republican	republican	republican	republican	republican
##	59	60	62	66	67	68
##	republican	republican	republican	republican	republican	republican
##	74	80	83	84	85	87
##	republican	republican	republican	republican	republican	republican
##	88	90	100	108	112	114
##	republican	republican	republican	republican	republican	republican
##	118	121	122	123	124	126
##	republican	republican	republican	republican	republican	republican
##	127	134	135	136	137	141
##	republican	republican	republican	republican	republican	republican
##	142	147	149	155	159	164
##	republican	republican	republican	republican	republican	republican
##	167	168	172	177	189	191
##	republican	republican	republican	republican	republican	republican
##	192	196	198	205	207	208
##	republican	republican	republican	republican	republican	republican
##	212	224	226	229	230	231
##	republican	republican	republican	republican	republican	republican
##	232	236	240	241	243	249
##	republican	republican	republican	republican	republican	republican
##	251	252	254	258	267	268
##	republican	republican	republican	republican	republican	democrat
##	274	275	277	279	280	282
##	republican	republican	republican	republican	republican	republican
##	283	284	296	301	304	305
##	republican	republican	republican	republican	republican	republican
##	306	307	309	311	315	316
##	republican	republican	republican	republican	republican	republican
##	325	328	331	336	340	341
##	republican	republican	republican	republican	republican	republican
##	344	346	347	350	352	357
##	republican	republican	republican	republican	republican	republican
##	358	360	364	370	378	379
##	republican	republican	republican	republican	republican	republican
##	380	394	400	401	402	403
##	republican	republican	republican	republican	republican	republican
##	404	406	410	413	414	421
##	republican	republican	republican	republican	republican	republican
##	428	431	433	434	435	4
##	republican	republican	republican	republican	republican	democrat
##	5	6	7	10	17	18
##	democrat	democrat	republican	democrat	democrat	democrat

##	22	23	24	25	26	27
##	democrat	democrat	democrat	democrat	democrat	democrat
##	28	32	33	40	42	43
##	democrat	democrat	democrat	democrat	democrat	democrat
##	44	45	47	48	49	51
##	democrat	democrat	democrat	democrat	democrat	democrat
##	53	55	64	65	69	71
##	democrat	democrat	democrat	democrat	democrat	democrat
##	73	75	76	77	78	79
##	democrat	democrat	democrat	democrat	democrat	democrat
##	82	86	89	91	92	93
##	democrat	democrat	democrat	democrat	democrat	democrat
##	95	96	97	98	99	101
##	democrat	democrat	democrat	democrat	democrat	democrat
##	102	104	105	106	109	111
##	democrat	democrat	democrat	democrat	democrat	democrat
##	113	115	116	119	125	128
##	democrat	democrat	democrat	democrat	democrat	democrat
##	130	131	132	138	140	145
##	democrat	democrat	democrat	democrat	democrat	democrat
##	146	148	152	153	154	158
##	democrat	democrat	democrat	democrat	democrat	democrat
##	161	162	163	165	169	170
##	democrat	republican	democrat	democrat	democrat	democrat
##	171	173	174	175	176	179
##	democrat	democrat	democrat	democrat	democrat	democrat
##	180	181	182	183	184	186
##	democrat	democrat	democrat	democrat	democrat	democrat
##	190	193	195	197	199	201
##	democrat	democrat	democrat	democrat	democrat	democrat
##	202	203	204	206	209	210
##	democrat	democrat	democrat	democrat	democrat	democrat
##	211	213	216	217	219	220
##	democrat	democrat	democrat	democrat	democrat	democrat
##	221	222	223	227	228	235
##	democrat	democrat	democrat	democrat	democrat	democrat
##	237	238	239	244	245	246
##	democrat	democrat	democrat	democrat	democrat	democrat
##	247	250	253	255	256	260
##	democrat	democrat	democrat	democrat	democrat	democrat
##	261	262	264	265	266	269
##	democrat	democrat	democrat	democrat	democrat	democrat
##	271	276	281	286	287	288
##	democrat	democrat	democrat	democrat	democrat	democrat
##	291	292	293	294	298	299
##	democrat	democrat	democrat	democrat	democrat	democrat
##	300	302	308	310	312	313
##	democrat	democrat	democrat	democrat	democrat	democrat
##	317	318	319	320	321	322

```
## democrat democrat democrat democrat democrat democrat
##      323      324      326      327      329      330
## democrat democrat democrat democrat democrat democrat
##      332      333      334      335      337      338
## democrat democrat democrat democrat democrat democrat
##      339      342      343      345      349      351
## democrat democrat democrat democrat democrat democrat
##      353      355      361      362      366      367
## democrat democrat democrat democrat democrat democrat
##      368      371      372      373      374      376
## democrat democrat democrat democrat democrat republican
##      381      382      383      384      386      387
## democrat democrat democrat democrat democrat democrat
##      388      389      390      391      392      395
## democrat republican democrat democrat democrat democrat
##      396      397      407      408      412      415
## democrat democrat democrat republican democrat democrat
##      416      418      419      420      422      423
## democrat democrat democrat democrat democrat democrat
##      425      426      427      429      432
## democrat democrat democrat democrat democrat
## Levels: democrat republican
```

```
ConfusionMatrix(pd_svmfit,train_imp_data$classname)
```

```
##           y_pred
## y_true    democrat republican
## democrat    211         5
## republican    1        136
```