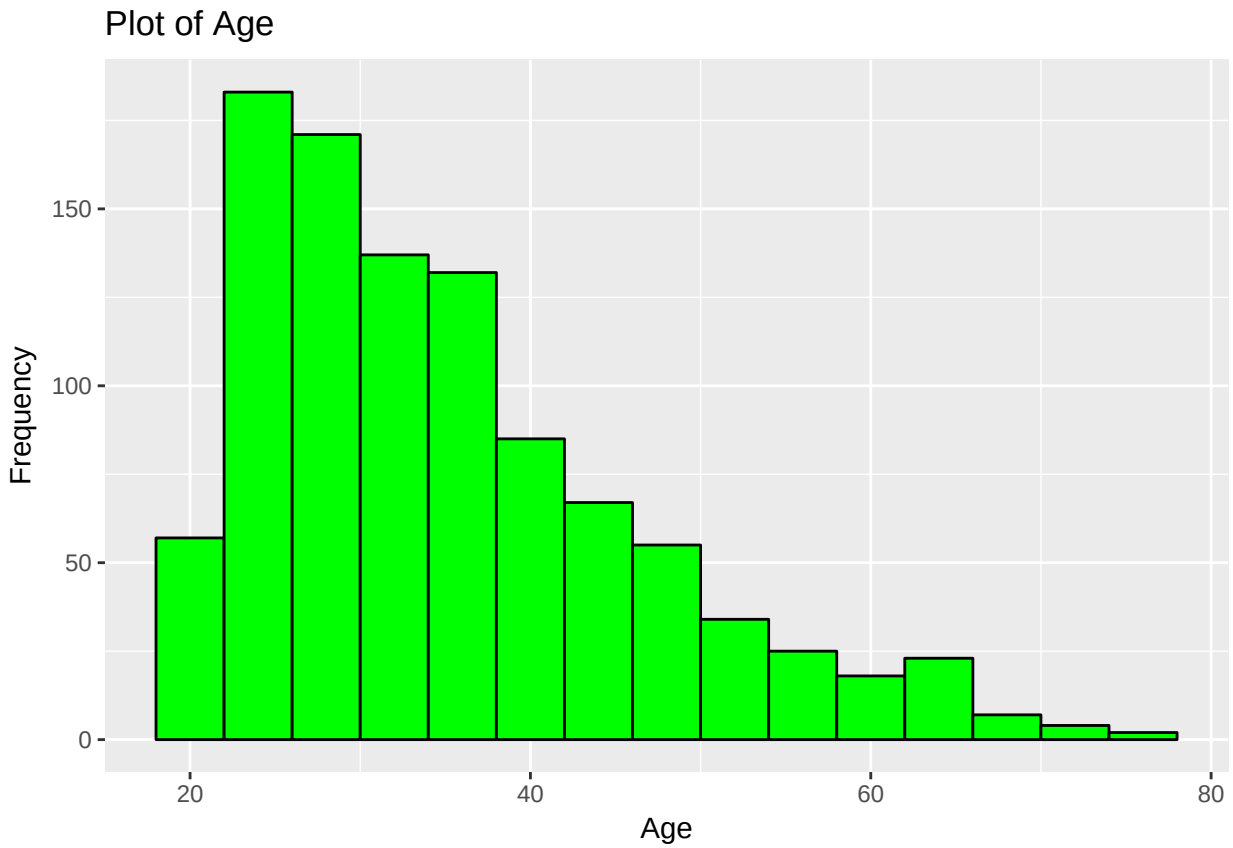
DTCreditApprovalC5.R

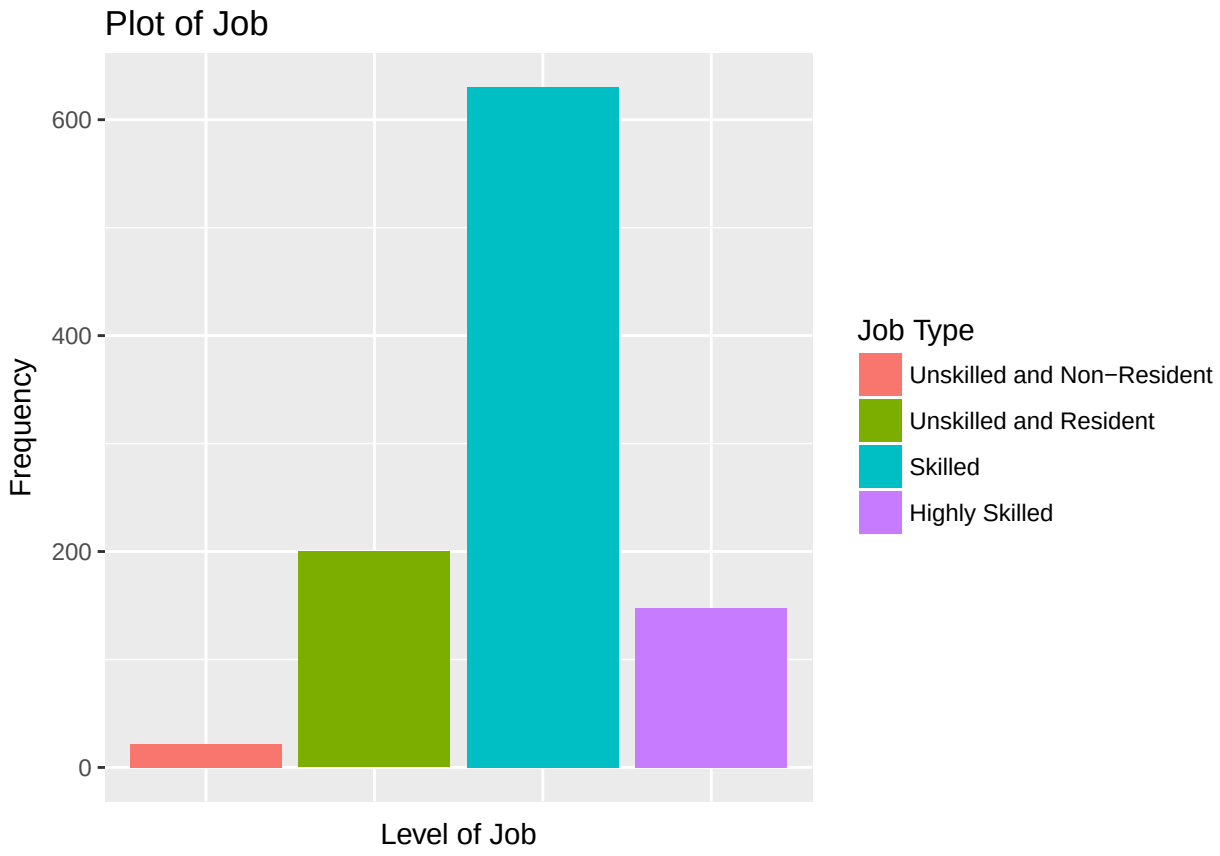
ai

Mon Jun 5 22:18:33 2017

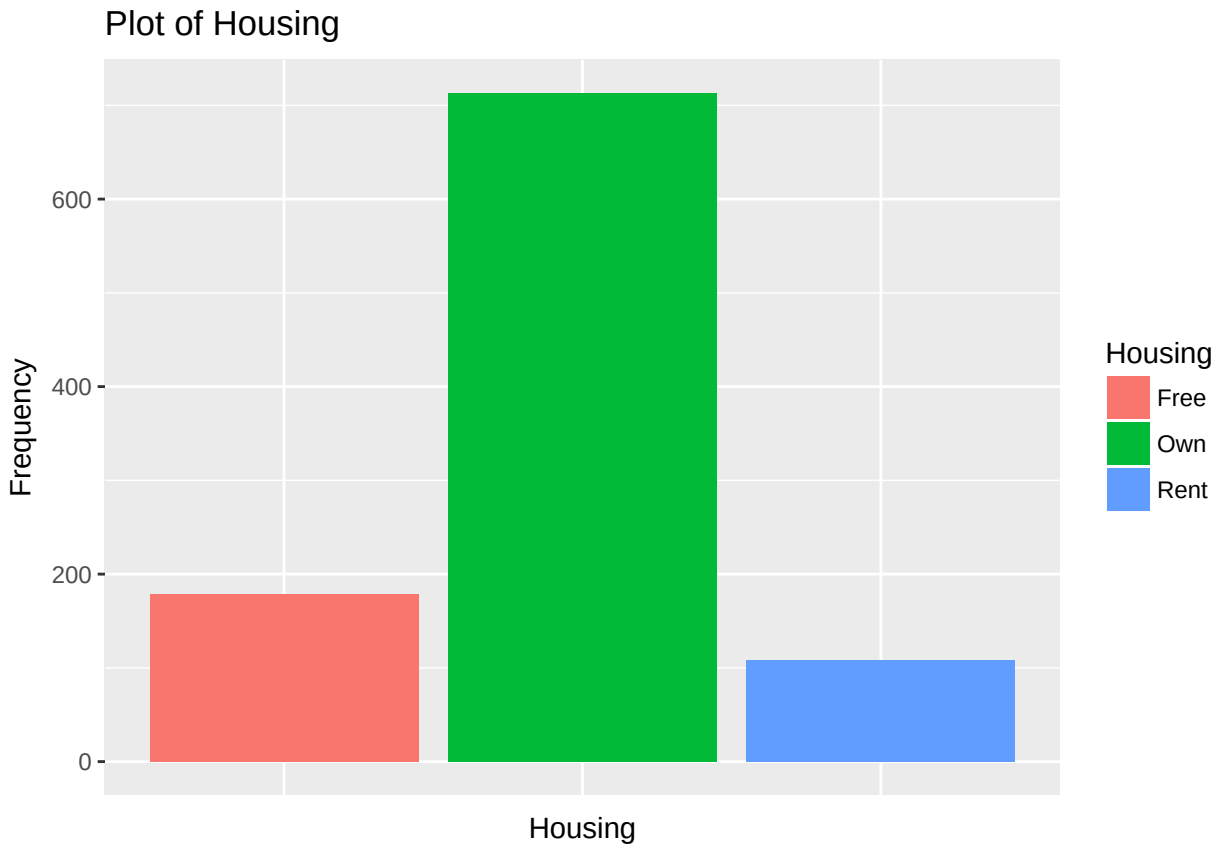
```
# Reference for data source (
# @misc{Lichman:2013 ,
# author = "M. Lichman",
# year = "2013",
# title = "{UCI} Machine Learning Repository",
# url = "http://archive.ics.uci.edu/ml",
# institution = "University of California, Irvine, School of Information and Computer Sciences" })

# Decision Trees
# Source of Data Set:- UCI Repository - German Credit Data(https://archive.ics.uci.edu/ml/machine-learn

# Step 01: Collecting data
# Download data from UCI repository
CreditDataUrl <- "https://archive.ics.uci.edu/ml/machine-learning-databases/statlog/german/german.data"

# Read the url html file into a data frame titled CreditData.
CreditData <- read.table(CreditDataUrl)

# Assging attribute information
# The target function column name is class
colnames(CreditData) <- c("chk_status", "mth_duration", "credit_history", "purpose", "credit_amount", "

# Write a CSV file from CreditData
Credit_Data <- write.csv(CreditData, file = "CreditData.csv", row.names = FALSE)

# Exploring and preparing the data
# Step 2: Exploring and preparing the data
# Read the csv file into a data frame titled CreditData.
CreditData <- read.csv("CreditData.csv", header=TRUE)



# Data Visualization
# plot histogram

library(ggplot2)

#Plotting the data
ggplot(CreditData, aes(age)) + geom_histogram(binwidth=4, colour="black", fill="green") +
  labs(x= "Age",y= "Frequency" , title = "Plot of Age")
```




```
ggplot(CreditData, aes(job) ) + geom_bar(aes(fill = as.factor(CreditData$job))) +
  scale_fill_discrete(name="Job Type",
                      labels=c( "Unskilled and Non-Resident","Unskilled and Resident", "Skilled", "High
  theme(axis.text.x=element_blank(),
        axis.ticks.x=element_blank()) +
  labs(x= "Level of Job",y= "Frequency" , title = "Plot of Job")
```

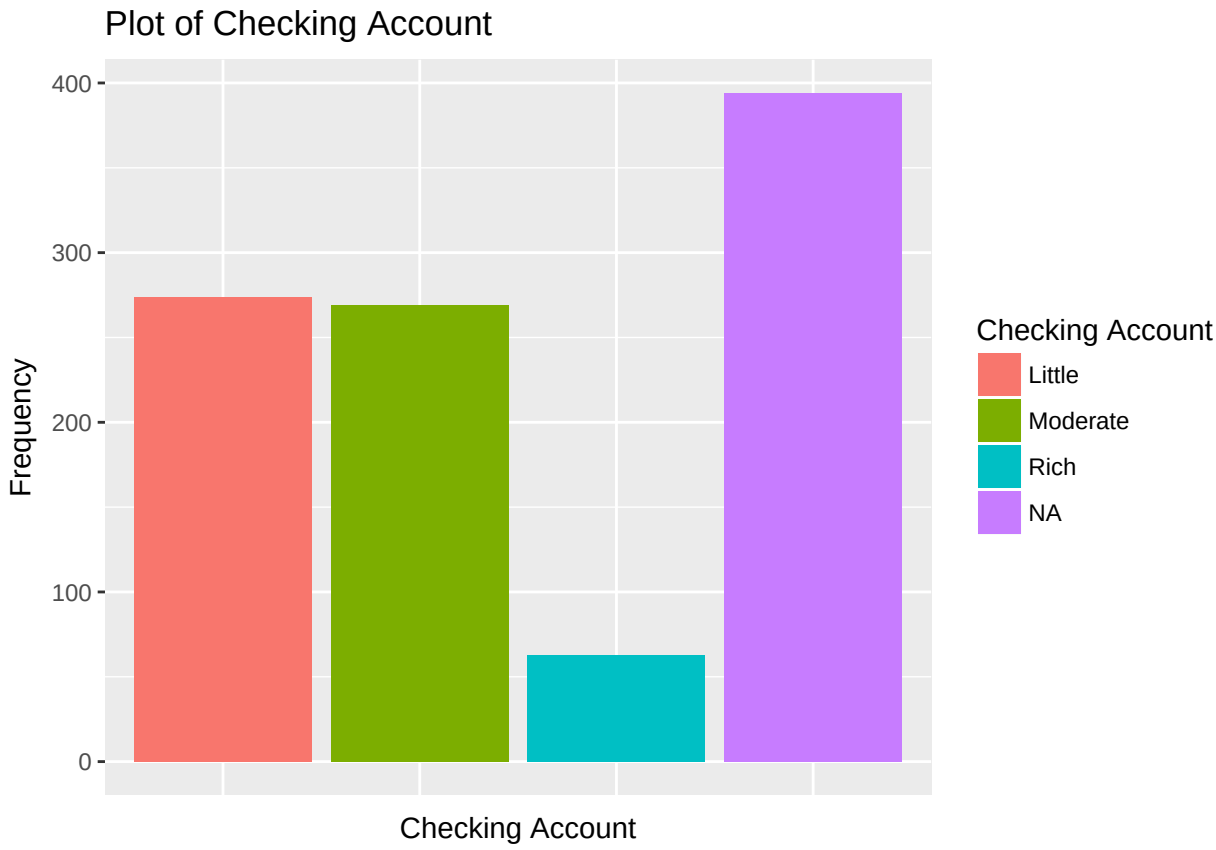



```
ggplot(CreditData, aes(housing) ) + geom_bar(aes(fill = as.factor(CreditData$housing))) +
  scale_fill_discrete(name="Housing",
                      labels=c( "Free","Own", "Rent")) +
  theme(axis.text.x=element_blank(),
        axis.ticks.x=element_blank()) +
  labs(x= "Housing",y= "Frequency" , title = "Plot of Housing")
```

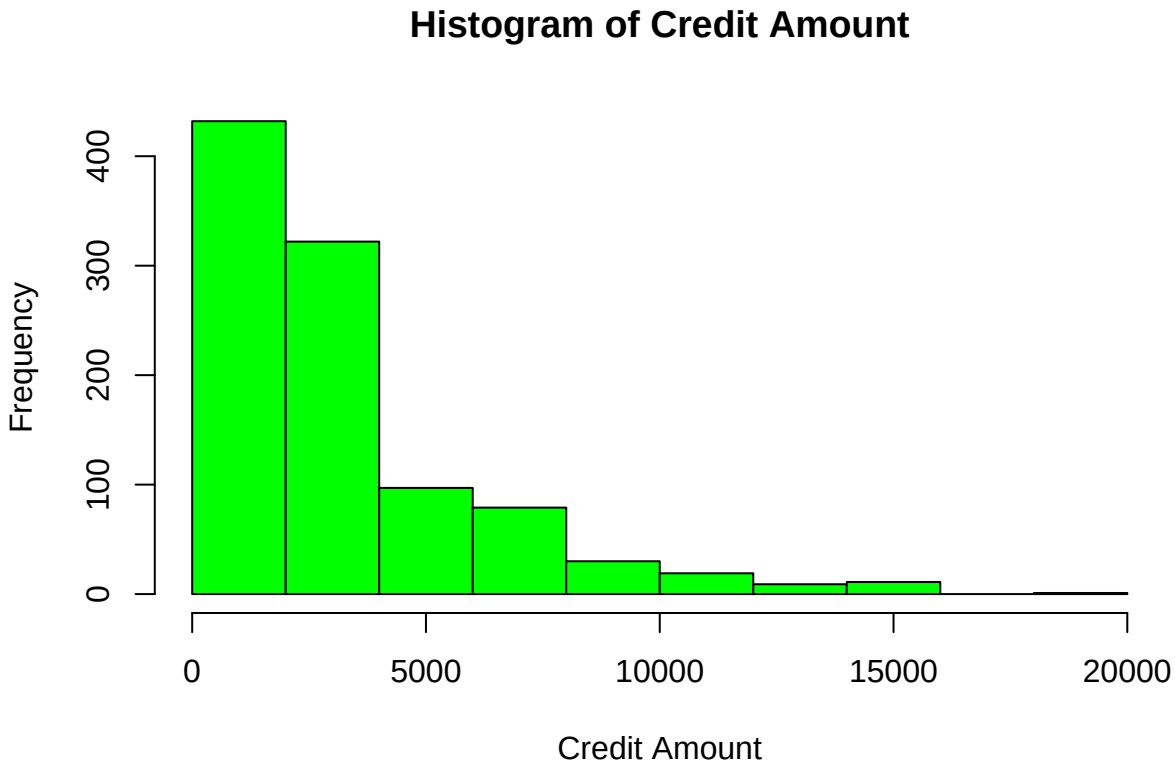



```
ggplot(CreditData, aes(saving) ) + geom_bar(aes(fill = as.factor(CreditData$saving)))  +
  scale_fill_discrete(name="Saving Accounts",
                      labels=c( "Little","Moderate", "Quite Rich", "Rich", "NA"))  +
  labs(x= "Saving Accounts",y= "Frequency" , title = "Plot of Saving Accounts")
```

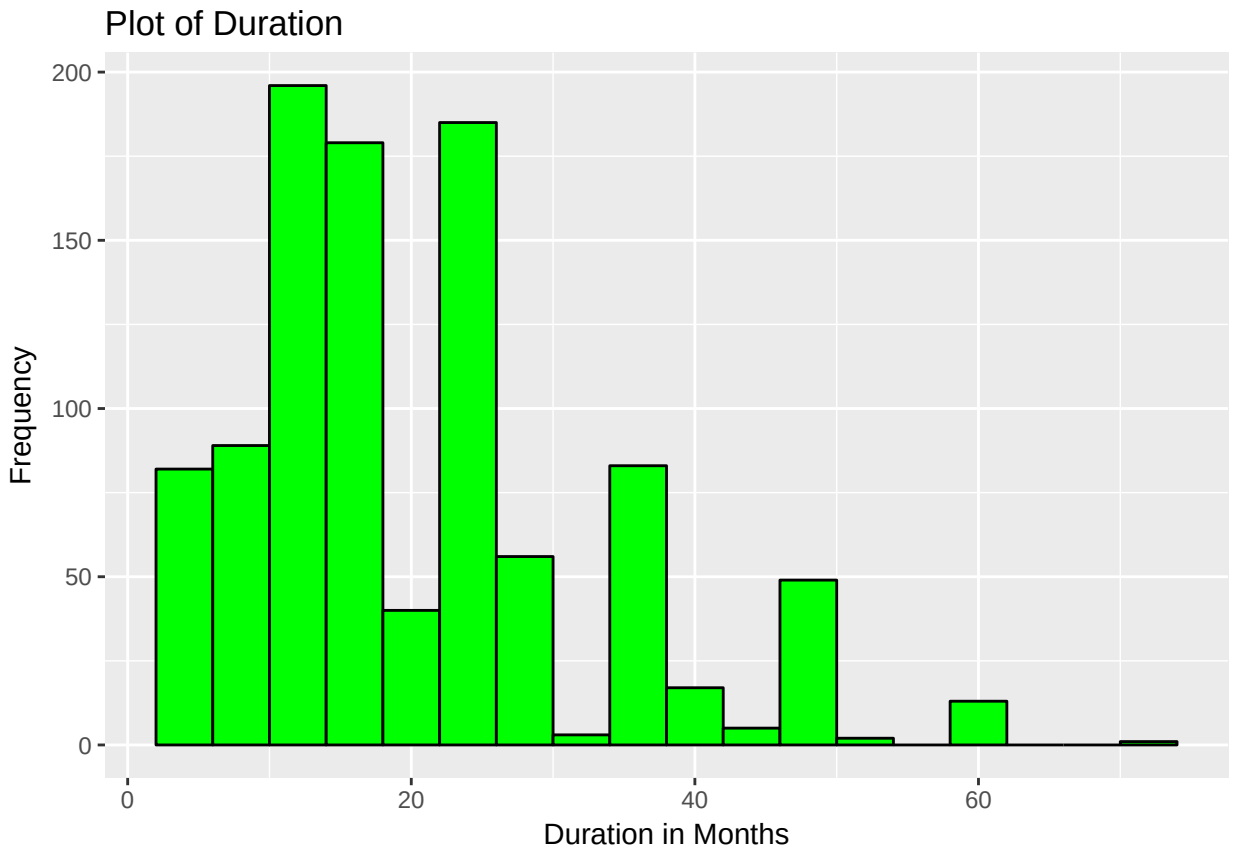

```
ggplot(CreditData, aes(chk_status) ) + geom_bar(aes(fill = as.factor(CreditData$chk_status))) +
  scale_fill_discrete(name="Checking Account",
                      labels=c( "Little","Moderate", "Rich")) +
  theme(axis.text.x=element_blank(),
        axis.ticks.x=element_blank()) +
  labs(x= "Checking Account",y= "Frequency" , title = "Plot of Checking Account")
```



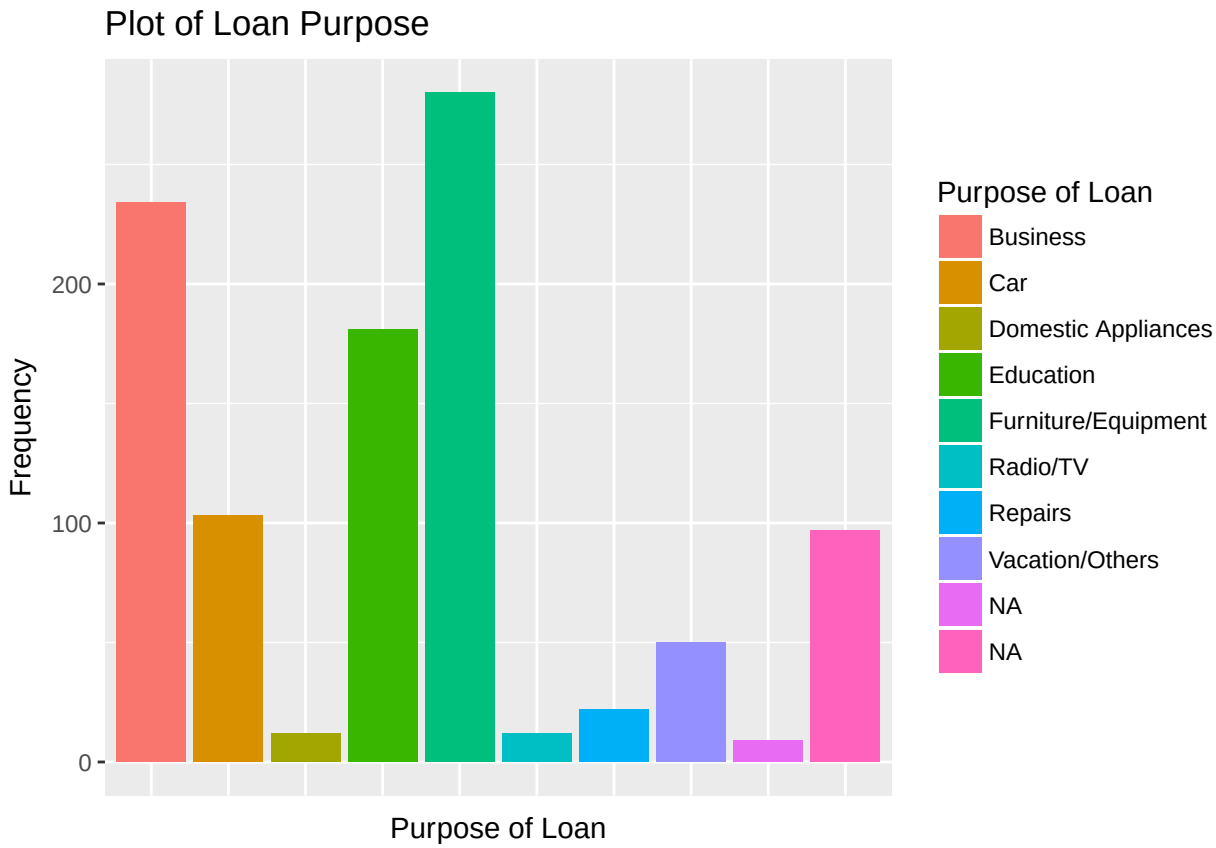

```
hist(CreditData$credit_amount, main = "Histogram of Credit Amount", xlab = "Credit Amount", ylab = "Fre
```



```
ggplot(CreditData, aes(mth_duration)) + geom_histogram(binwidth=4, colour="black", fill="green") +
  labs(x= "Duration in Months",y= "Frequency" , title = "Plot of Duration")
```




```
ggplot(CreditData, aes(purpose) ) + geom_bar(aes(fill = as.factor(CreditData$purpose))) +
  scale_fill_discrete(name="Purpose of Loan",
                      labels=c( "Business","Car", "Domestic Appliances","Education","Furniture/Equipmen
  theme(axis.text.x=element_blank(),
        axis.ticks.x=element_blank()) +
  labs(x= "Purpose of Loan",y= "Frequency" , title = "Plot of Loan Purpose")
```




```
# Class columns convert into facator
CreditData$class <- ifelse(CreditData$class==1, "good","bad")
CreditData$class = as.factor(CreditData$class)

# Displays description of each variable

head(CreditData)
```

```
##   chk_status mth_duration credit_history purpose credit_amount saving
## 1        A11            6            A34     A43          1169    A65
## 2        A12           48            A32     A43          5951    A61
## 3        A14           12            A34     A46          2096    A61
## 4        A11           42            A32     A42          7882    A61
## 5        A11           24            A33     A40          4870    A61
## 6        A14           36            A32     A46          9055    A65
##   employ_time pct_dpi status_gender other_debts residency_time property
## 1         A75       4           A93        A101              4     A121
## 2         A73       2           A92        A101              2     A121
## 3         A74       2           A93        A101              3     A121
## 4         A74       2           A93        A103              4     A122
## 5         A73       3           A93        A101              4     A124
## 6         A73       2           A93        A101              4     A124
##   age other_installments housing existing_credits  job dependents_num
## 1  67               A143    A152                2 A173             1
## 2  22               A143    A152                1 A173             1
## 3  49               A143    A152                1 A172             2
## 4  45               A143    A153                1 A173             2
```

```
## 5  53               A143    A153                2 A173              2
## 6  35               A143    A153                1 A172              2
##   phone foreign class
## 1  A192    A201  good
## 2  A191    A201   bad
## 3  A191    A201  good
## 4  A191    A201  good
## 5  A191    A201   bad
## 6  A192    A201  good
```

```
# str(CreditData)

# Use table() function to look output for a couple of loan features
# Attribute 1: (qualitative)
# Status of existing checking account
# A11 : ... < 0 DM
# A12 : 0 <= ... < 200 DM
# A13 : ... >= 200 DM / salary assignments for at least 1 year
# A14 : no checking account
table(CreditData$chk_status)
```

```
##
## A11 A12 A13 A14
## 274 269  63 394
```

```
# Attribute 3: (qualitative)
# Credit history
# A30 : no credits taken/ all credits paid back duly
# A31 : all credits at this bank paid back duly
# A32 : existing credits paid back duly till now
# A33 : delay in paying off in the past
# A34 : critical account/ other credits existing (not at this bank)
table(CreditData$credit_history)
```

```
##
## A30 A31 A32 A33 A34
##  40  49 530  88 293
```

```
# Attribute 4: (qualitative)
# Purpose
# A40 : car (new)
# A41 : car (used)
# A42 : furniture/equipment
# A43 : radio/television
# A44 : domestic appliances
# A45 : repairs
# A46 : education
# A47 : (vacation - does not exist?)
# A48 : retraining
# A49 : business
# A410 : others
table(CreditData$purpose)
```

```
##
##  A40  A41 A410  A42  A43  A44  A45  A46  A48  A49
##  234  103   12  181  280   12   22   50    9   97
```

```
# Attibute 6: (qualitative)
# Savings account/bonds
# A61 : ... < 100 DM
# A62 : 100 <= ... < 500 DM
# A63 : 500 <= ... < 1000 DM
# A64 : .. >= 1000 DM
# A65 : unknown/ no savings account
table(CreditData$saving)
```

```
##
## A61 A62 A63 A64 A65
## 603 103  63  48 183
```

```
# Attribute 7: (qualitative)
# Present employment since
# A71 : unemployed
# A72 : ... < 1 year
# A73 : 1 <= ... < 4 years
# A74 : 4 <= ... < 7 years
# A75 : .. >= 7 years
table(CreditData$employ_time)
```

```
##
## A71 A72 A73 A74 A75
##  62 172 339 174 253
```

```
# Attribute 9: (qualitative)
# Personal status and sex
# A91 : male : divorced/separated
# A92 : female : divorced/separated/married
# A93 : male : single
# A94 : male : married/widowed
# A95 : female : single
table(CreditData$status_gender)
```

```
##
## A91 A92 A93 A94
##  50 310 548  92
```

```
# Attribute 10: (qualitative)
# Other debtors / guarantors
# A101 : none
# A102 : co-applicant
# A103 : guarantor
table(CreditData$other_debts)
```

```
##
## A101 A102 A103
##  907   41   52
```

```
# Attribute 12: (qualitative)
# Property
# A121 : real estate
# A122 : if not A121 : building society savings agreement/ life insurance
# A123 : if not A121/A122 : car or other, not in attribute 6
# A124 : unknown / no property
```

```r
table(CreditData$property)
```

```
##
## A121 A122 A123 A124
##  282  232  332  154
```

```r
# Attribute 14: (qualitative)
# Other installment plans
# A141 : bank
# A142 : stores
# A143 : none
table(CreditData$other_installments)
```

```
##
## A141 A142 A143
##  139   47  814
```

```r
# Attribute 15: (qualitative)
# Housing
# A151 : rent
# A152 : own
# A153 : for free
table(CreditData$housing)
```

```
##
## A151 A152 A153
##  179  713  108
```

```r
# Attribute 17: (qualitative)
# Job
# A171 : unemployed/ unskilled - non-resident
# A172 : unskilled - resident
# A173 : skilled employee / official
# A174 : management/ self-employed/
# highly qualified employee/ officer
table(CreditData$job)
```

```
##
## A171 A172 A173 A174
##   22  200  630  148
```

```r
# Attribute 19: (qualitative)
# Telephone
# A191 : none
# A192 : yes, registered under the customers name
table(CreditData$phone)
```

```
##
## A191 A192
##  596  404
```

```r
# Attribute 20: (qualitative)
# foreign worker
# A201 : yes
# A202 : no
table(CreditData$foreign)
```

```
##
## A201 A202
##  963   37
```

```
# Attribute 21:
table(CreditData$class)
```

```
##
##  bad good
##  300  700
```

```
# Check credit amount
summary(CreditData$credit_amount)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     250    1366    2320    3271    3972   18420
```

```
# Check duration
summary(CreditData$mth_duration)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     4.0    12.0    18.0    20.9    24.0    72.0
```

```
# Data preparation - creating random training and test datasets
# Create random sample
# Divide the data into a training set and a test set randomly with ratio 90:1

set.seed(123)
train_sample <- sample(nrow(CreditData), 0.9 * nrow(CreditData))
CreditData_train <- CreditData[train_sample, ]
CreditData_test <- CreditData[-train_sample, ]

# Check whether data set fairly even split
prop.table(table(CreditData_train$class))
```

```
##
##       bad      good
## 0.2966667 0.7033333
```

```
prop.table(table(CreditData_test$class))
```

```
##
##  bad good
## 0.33 0.67
```

```
# Training a model on the data
# The C5.0 package can be installed via the install.packages("C50") and
# loaded with the library(C50) command.
library(C50)

# Train model
CreditData_model <- C5.0(CreditData_train[-21], CreditData_train$class)

CreditData_model
```

```
##
## Call:
## C5.0.default(x = CreditData_train[-21], y = CreditData_train$class)
##
```

```
## Classification Tree
## Number of samples: 900
## Number of predictors: 20
##
## Tree size: 55
##
## Non-standard options: attempt to group attributes
```

```
# See the tree's decisions
summary(CreditData_model)
```

```
##
## Call:
## C5.0.default(x = CreditData_train[-21], y = CreditData_train$class)
##
##
## C5.0 [Release 2.07 GPL Edition]      Mon Jun  5 22:18:35 2017
## -------------------------------
##
## Class specified by attribute `outcome'
##
## Read 900 cases (21 attributes) from undefined.data
##
## Decision tree:
##
## chk_status in {A13,A14}: good (412/50)
## chk_status in {A11,A12}:
## :...other_debts = A103:
##     :...mth_duration > 36: bad (4/1)
##     :   mth_duration <= 36:
##     :   :...other_installments in {A142,A143}: good (24)
##     :       other_installments = A141:
##     :       :...purpose = A40: bad (3)
##     :           purpose in {A41,A410,A42,A43,A44,A45,A46,A48,
##     :                       A49}: good (7/1)
##     other_debts in {A101,A102}:
##     :...credit_history = A30: bad (27/6)
##         credit_history = A34: good (102/30)
##         credit_history = A31:
##         :...other_debts = A101: bad (26/8)
##         :   other_debts = A102: good (2)
##         credit_history in {A32,A33}:
##         :...saving in {A63,A64}: good (19/3)
##             saving = A62:
##             :...other_debts = A102: bad (3)
##             :   other_debts = A101:
##             :   :...status_gender in {A91,A94}: bad (6/1)
##             :       status_gender = A92:
##             :       :...pct_dpi <= 3: good (4/1)
##             :       :   pct_dpi > 3: bad (4)
##             :       status_gender = A93:
##             :       :...age <= 41: good (15/2)
##             :           age > 41: bad (2)
##             saving = A65:
##             :...credit_history = A33: good (8)
```

```
##             :   credit_history = A32:
##             :   :...foreign = A202: good (2)
##             :       foreign = A201:
##             :       :...chk_status = A11:
##             :           :...phone = A191: bad (11/2)
##             :           :   phone = A192:
##             :           :   :...credit_amount <= 5045: good (5/1)
##             :           :       credit_amount > 5045: bad (2)
##             :           chk_status = A12:
##             :           :...residency_time > 3: good (9)
##             :               residency_time <= 3:
##             :               :...property in {A122,A124}: bad (4)
##             :                   property = A123: good (6)
##             :                   property = A121:
##             :                   :...job = A171: good (2)
##             :                       job in {A172,A173,A174}: bad (2)
##             saving = A61:
##             :...mth_duration > 39:
##                 :...residency_time <= 1: good (2)
##                 :   residency_time > 1: bad (19/1)
##                 mth_duration <= 39:
##                 :...purpose in {A40,A48}: bad (47/16)
##                     purpose in {A410,A44}: good (3)
##                     purpose = A41:
##                     :...credit_amount <= 8086: good (9/1)
##                     :   credit_amount > 8086: bad (5)
##                     purpose = A45:
##                     :...residency_time <= 3: bad (4/1)
##                     :   residency_time > 3: good (3)
##                     purpose = A46:
##                     :...chk_status = A11: bad (5)
##                     :   chk_status = A12: good (2)
##                     purpose = A49:
##                     :...credit_history = A33: bad (2)
##                     :   credit_history = A32:
##                     :   :...age <= 34: good (5)
##                     :       age > 34: bad (2)
##                     purpose = A43:
##                     :...employ_time in {A71,A72}: bad (14/5)
##                     :   employ_time = A74: good (3)
##                     :   employ_time = A75:
##                     :   :...credit_amount <= 932: bad (2)
##                     :   :   credit_amount > 932: good (7)
##                     :   employ_time = A73:
##                     :   :...mth_duration <= 15: good (6)
##                     :       mth_duration > 15:
##                     :       :...credit_amount <= 3275: bad (7)
##                     :           credit_amount > 3275: good (2)
##                     purpose = A42:
##                     :...residency_time <= 1: good (8/1)
##                         residency_time > 1:
##                         :...other_installments = A141: bad (2/1)
##                             other_installments = A142: good (1)
##                             other_installments = A143:
```

```
##                               :...phone = A192: bad (7/1)
##                                   phone = A191:
##                                   :...mth_duration > 27: bad (3)
##                                       mth_duration <= 27:
##                                       :...chk_status = A12: bad (5/2)
##                                           chk_status = A11:
##                                           :...property in {A121,A122,
##                                               :            A124}: good (8)
##                                               property = A123:
##                                               :...pct_dpi <= 1: good (2)
##                                                   pct_dpi > 1: bad (4)
##
##
## Evaluation on training data (900 cases):
##
##      Decision Tree
##    ----------------
##    Size      Errors
##
##      55  135(15.0%)   <<
##
##
##     (a)   (b)    <-classified as
##    ----  ----
##     177    90    (a): class bad
##      45   588    (b): class good
##
##
##  Attribute usage:
##
##  100.00% chk_status
##   54.22% other_debts
##   50.00% credit_history
##   32.56% saving
##   25.22% mth_duration
##   19.78% purpose
##   10.11% residency_time
##    7.33% other_installments
##    5.22% phone
##    4.78% foreign
##    4.56% employ_time
##    4.33% credit_amount
##    3.44% status_gender
##    3.11% property
##    2.67% age
##    1.56% pct_dpi
##    0.44% job
##
##
## Time: 0.0 secs
```

```r
# Evaluating model performance
CreditData_predict <- predict(CreditData_model, CreditData_test)
```

```r
# Various R Programming Tools for Model Fitting
library(gmodels)

# create a cross tabulation indicating the agreement between the two vectors.
# Specifying prop.chisq = FALSE will remove the unnecessary chi-square
# values from the output.
# Setting the prop.c and prop.r parameters to FALSE removes the column and row percentages
# from the table. The remaining percentage ( prop.t ) indicates the proportion of
# records in the cell out of the total number of records:
CrossTable(CreditData_test$class, CreditData_predict, prop.chisq = FALSE, prop.c= FALSE, prop.r = FALSE
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  100
##
##
##              | Predicted class
## Actual class |       bad |      good | Row Total |
## -------------|-----------|-----------|-----------|
##          bad |        14 |        19 |        33 |
##              |     0.140 |     0.190 |           |
## -------------|-----------|-----------|-----------|
##         good |         8 |        59 |        67 |
##              |     0.080 |     0.590 |           |
## -------------|-----------|-----------|-----------|
## Column Total |        22 |        78 |       100 |
## -------------|-----------|-----------|-----------|
##
##
```

```r
# Accuracy : Measures of performance
library(caret)
```

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

```r
confusionMatrix(CreditData_test$class, CreditData_predict, positive = "good")
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction bad good
##       bad   14   19
##       good   8   59
##
##                Accuracy : 0.73
##                  95% CI : (0.632, 0.8139)
##     No Information Rate : 0.78
##     P-Value [Acc > NIR] : 0.90547
```

```
##
##                   Kappa : 0.333
##  Mcnemar's Test P-Value : 0.05429
##
##             Sensitivity : 0.7564
##             Specificity : 0.6364
##          Pos Pred Value : 0.8806
##          Neg Pred Value : 0.4242
##              Prevalence : 0.7800
##          Detection Rate : 0.5900
##    Detection Prevalence : 0.6700
##       Balanced Accuracy : 0.6964
##
##        'Positive' Class : good
##
```

```
# Improving model performance
# Boosting the accuracy of decision trees
# Add additional trials parameter indicating the number of
# separate decision trees to use in the boosted team.
CreditData_boost10 <- C5.0(CreditData_train[-21], CreditData_train$class, trials = 10)
CreditData_boost10
```

```
##
## Call:
## C5.0.default(x = CreditData_train[-21], y = CreditData_train$class,
##  trials = 10)
##
## Classification Tree
## Number of samples: 900
## Number of predictors: 20
##
## Number of boosting iterations: 10
## Average tree size: 46.3
##
## Non-standard options: attempt to group attributes
```

```
# See all 10 trees
summary(CreditData_boost10)
```

```
##
## Call:
## C5.0.default(x = CreditData_train[-21], y = CreditData_train$class,
##  trials = 10)
##
##
## C5.0 [Release 2.07 GPL Edition]      Mon Jun  5 22:18:36 2017
## -------------------------------
##
## Class specified by attribute `outcome'
##
## Read 900 cases (21 attributes) from undefined.data
##
## -----  Trial 0:  -----
##
## Decision tree:
```

```
##
## chk_status in {A13,A14}: good (412/50)
## chk_status in {A11,A12}:
## :...other_debts = A103:
##     :...mth_duration > 36: bad (4/1)
##     :   mth_duration <= 36:
##     :   :...other_installments in {A142,A143}: good (24)
##     :       other_installments = A141:
##     :       :...purpose = A40: bad (3)
##     :           purpose in {A41,A410,A42,A43,A44,A45,A46,A48,
##     :                       A49}: good (7/1)
##     other_debts in {A101,A102}:
##     :...credit_history = A30: bad (27/6)
##         credit_history = A34: good (102/30)
##         credit_history = A31:
##         :...other_debts = A101: bad (26/8)
##         :   other_debts = A102: good (2)
##         credit_history in {A32,A33}:
##         :...saving in {A63,A64}: good (19/3)
##             saving = A62:
##             :...other_debts = A102: bad (3)
##             :   other_debts = A101:
##             :   :...status_gender in {A91,A94}: bad (6/1)
##             :       status_gender = A92:
##             :       :...pct_dpi <= 3: good (4/1)
##             :       :   pct_dpi > 3: bad (4)
##             :       status_gender = A93:
##             :       :...age <= 41: good (15/2)
##             :           age > 41: bad (2)
##             saving = A65:
##             :...credit_history = A33: good (8)
##             :   credit_history = A32:
##             :   :...foreign = A202: good (2)
##             :       foreign = A201:
##             :       :...chk_status = A11:
##             :           :...phone = A191: bad (11/2)
##             :           :   phone = A192:
##             :           :   :...credit_amount <= 5045: good (5/1)
##             :           :       credit_amount > 5045: bad (2)
##             :           chk_status = A12:
##             :           :...residency_time > 3: good (9)
##             :               residency_time <= 3:
##             :               :...property in {A122,A124}: bad (4)
##             :                   property = A123: good (6)
##             :                   property = A121:
##             :                   :...job = A171: good (2)
##             :                       job in {A172,A173,A174}: bad (2)
##             saving = A61:
##             :...mth_duration > 39:
##                 :...residency_time <= 1: good (2)
##                 :   residency_time > 1: bad (19/1)
##                 mth_duration <= 39:
##                 :...purpose in {A40,A48}: bad (47/16)
##                     purpose in {A410,A44}: good (3)
```

```
##                     purpose = A41:
##                     :...credit_amount <= 8086: good (9/1)
##                     :   credit_amount > 8086: bad (5)
##                     purpose = A45:
##                     :...residency_time <= 3: bad (4/1)
##                     :   residency_time > 3: good (3)
##                     purpose = A46:
##                     :...chk_status = A11: bad (5)
##                     :   chk_status = A12: good (2)
##                     purpose = A49:
##                     :...credit_history = A33: bad (2)
##                     :   credit_history = A32:
##                     :   :...age <= 34: good (5)
##                     :       age > 34: bad (2)
##                     purpose = A43:
##                     :...employ_time in {A71,A72}: bad (14/5)
##                     :   employ_time = A74: good (3)
##                     :   employ_time = A75:
##                     :   :...credit_amount <= 932: bad (2)
##                     :   :   credit_amount > 932: good (7)
##                     :   employ_time = A73:
##                     :   :...mth_duration <= 15: good (6)
##                     :       mth_duration > 15:
##                     :       :...credit_amount <= 3275: bad (7)
##                     :           credit_amount > 3275: good (2)
##                     purpose = A42:
##                     :...residency_time <= 1: good (8/1)
##                         residency_time > 1:
##                         :...other_installments = A141: bad (2/1)
##                             other_installments = A142: good (1)
##                             other_installments = A143:
##                             :...phone = A192: bad (7/1)
##                                 phone = A191:
##                                 :...mth_duration > 27: bad (3)
##                                     mth_duration <= 27:
##                                     :...chk_status = A12: bad (5/2)
##                                         chk_status = A11:
##                                         :...property in {A121,A122,
##                                             :            A124}: good (8)
##                                             property = A123:
##                                             :...pct_dpi <= 1: good (2)
##                                                 pct_dpi > 1: bad (4)
##
## -----  Trial 1:  -----
##
## Decision tree:
##
## foreign = A202: good (28.4/2.4)
## foreign = A201:
## :...chk_status = A14:
##     :...other_installments in {A141,A142}:
##     :   :...other_debts in {A102,A103}: good (2.4)
##     :   :   other_debts = A101:
##     :   :   :...employ_time in {A71,A73}: bad (31/7.1)
```

```
##     :   :       employ_time in {A72,A74,A75}: good (32.3/10.8)
##     :   other_installments = A143:
##     :   :...credit_history in {A30,A31,A32,A34}: good (224.7/32.5)
##     :       credit_history = A33:
##     :       :...residency_time <= 1: bad (4.3)
##     :           residency_time > 1:
##     :           :...pct_dpi <= 3: good (11.9)
##     :               pct_dpi > 3: bad (14.2/5.6)
##     chk_status in {A11,A12,A13}:
##     :...other_debts = A102: bad (24.3/7.9)
##         other_debts = A103:
##         :...property in {A121,A122,A124}: good (27.6/4)
##         :   property = A123: bad (3)
##         other_debts = A101:
##         :...pct_dpi <= 2:
##             :...purpose in {A40,A41,A410,A43,A44,A48,A49}: good (125.5/34.3)
##             :   purpose in {A45,A46}: bad (13.6/4.8)
##             :   purpose = A42:
##             :   :...job = A171: good (0)
##             :       job = A174: bad (4.3)
##             :       job in {A172,A173}:
##             :       :...dependents_num > 1: bad (2.2)
##             :           dependents_num <= 1:
##             :           :...chk_status = A13: good (4)
##             :               chk_status in {A11,A12}:
##             :               :...mth_duration <= 22: good (23/6.1)
##             :                   mth_duration > 22: bad (12.1/2.4)
##             pct_dpi > 2:
##             :...residency_time <= 1:
##                 :...job = A171: bad (2.2)
##                 :   job in {A172,A173,A174}: good (36.8/6.4)
##                 residency_time > 1:
##                 :...credit_history = A30: bad (11.7)
##                     credit_history in {A31,A32,A33,A34}:
##                     :...mth_duration <= 11:
##                         :...purpose in {A40,A41,A410,A42,A43,A44,A45,A48,
##                         :   :           A49}: good (35.2/6.9)
##                         :   purpose = A46: bad (5.3/0.8)
##                         mth_duration > 11:
##                         :...saving = A63: bad (15.4/5.9)
##                             saving = A64: good (9.1/2.2)
##                             saving = A62:
##                             :...other_installments in {A141,
##                             :   :                      A142}: bad (8.3/0.8)
##                             :   other_installments = A143: good (16.2/4.5)
##                             saving = A65:
##                             :...chk_status in {A11,A13}: bad (20.8/5.6)
##                             :   chk_status = A12: good (12.7/1.6)
##                             saving = A61:
##                             :...other_installments in {A141,
##                                 :                      A142}: bad (25.3/3.2)
##                                 other_installments = A143:
##                                 :...dependents_num > 1: good (14.4/5.6)
##                                     dependents_num <= 1:
```

```
##                                     :...mth_duration > 42: bad (11.5)
##                                         mth_duration <= 42:
##                                         :...credit_history in {A31,
##                                             :                  A33}: bad (5.3)
##                                             credit_history = A32:
##                                             :...job in {A172,
##                                             :   :       A174}: good (23.2/8.7)
##                                             :   job in {A171,
##                                             :           A173}: bad (24.2/7.1)
##                                             credit_history = A34: [S1]
##
## SubTree [S1]
##
## existing_credits <= 1: good (6.9/2.2)
## existing_credits > 1:
## :...purpose in {A40,A410,A42,A44,A45,A46,A48,A49}: bad (22.7/3.2)
##     purpose in {A41,A43}: good (4)
##
## -----  Trial 2:  -----
##
## Decision tree:
##
## chk_status in {A11,A12}:
## :...mth_duration > 42:
## :   :...saving in {A61,A62,A64}: bad (41.6/6.1)
## :   :   saving in {A63,A65}: good (7.2)
## :   mth_duration <= 42:
## :   :...property = A124:
## :       :...pct_dpi <= 2: good (26.1/10.8)
## :       :   pct_dpi > 2: bad (48.4/12.6)
## :       property = A121:
## :       :...status_gender = A91: bad (7.9/1.3)
## :       :   status_gender in {A93,A94}: good (82.5/20.9)
## :       :   status_gender = A92:
## :       :   :...age <= 38: bad (28/8.9)
## :       :       age > 38: good (8.2)
## :       property = A122:
## :       :...other_debts in {A102,A103}: good (15.5/4)
## :       :   other_debts = A101:
## :       :   :...existing_credits <= 1: bad (66.4/27.1)
## :       :       existing_credits > 1:
## :       :       :...credit_history in {A30,A31}: bad (3.5)
## :       :           credit_history in {A32,A33,A34}:
## :       :           :...residency_time > 3: good (10.9)
## :       :               residency_time <= 3:
## :       :               :...age <= 54: good (18.1/4.8)
## :       :                   age > 54: bad (4.7)
## :       property = A123:
## :       :...credit_amount <= 1386: bad (39/6.3)
## :           credit_amount > 1386:
## :           :...saving in {A64,A65}: good (15.5/1.3)
## :               saving in {A61,A62,A63}:
## :               :...other_installments = A142: good (6.1/0.6)
## :                   other_installments in {A141,A143}:
```

```
## :                   :...age <= 28: bad (48.5/16.8)
## :                       age > 28:
## :                       :...pct_dpi <= 1: good (6.6)
## :                           pct_dpi > 1:
## :                           :...employ_time in {A71,A72,A74}: good (19.2/1.9)
## :                               employ_time in {A73,A75}: bad (22.3/8.5)
## chk_status in {A13,A14}:
## :...foreign = A202: good (11)
##     foreign = A201:
##     :...other_installments = A141:
##         :...housing = A151: bad (5.3)
##         :   housing in {A152,A153}:
##         :   :...age > 44: good (11.3/1.8)
##         :       age <= 44:
##         :       :...job = A171: bad (0)
##         :           job = A174: good (3.2)
##         :           job in {A172,A173}:
##         :           :...mth_duration <= 9: good (2.8)
##         :               mth_duration > 9: bad (35.9/10.7)
##         other_installments in {A142,A143}:
##         :...purpose in {A41,A410,A44,A48}: good (37.3)
##             purpose in {A40,A42,A43,A45,A46,A49}:
##             :...employ_time in {A71,A72}:
##                 :...other_debts = A102: bad (8.8)
##                 :   other_debts = A103: good (0.6)
##                 :   other_debts = A101:
##                 :   :...credit_amount > 4594: bad (14.1/0.6)
##                 :       credit_amount <= 4594:
##                 :       :...chk_status = A13: bad (10.6/4.1)
##                 :           chk_status = A14:
##                 :           :...purpose in {A40,A42,A43,A46}: good (20.7)
##                 :               purpose in {A45,A49}: bad (4.7)
##                 employ_time in {A73,A74,A75}:
##                 :...job in {A171,A173}: good (140.7/24.6)
##                     job in {A172,A174}:
##                     :...credit_history in {A30,A31,A34}: good (18.6)
##                         credit_history in {A32,A33}:
##                         :...other_debts in {A102,A103}: bad (2.9)
##                             other_debts = A101:
##                             :...employ_time = A74: good (4.8)
##                                 employ_time in {A73,A75}:
##                                 :...saving = A65: good (4.8)
##                                     saving = A64: bad (2.9)
##                                     saving in {A61,A62,A63}:
##                                     :...dependents_num > 1: bad (7.1/0.6)
##                                         dependents_num <= 1:
##                                         :...pct_dpi <= 2: good (4.5)
##                                             pct_dpi > 2: bad (21.2/8.9)
##
## -----  Trial 3:  -----
##
## Decision tree:
##
## foreign = A202:
```

```
## :...job in {A171,A172,A173}: good (21.1/1.3)
## :   job = A174: bad (2.5)
## foreign = A201:
## :...mth_duration <= 7:
##     :...credit_amount <= 3380: good (49.2/4.1)
##     :   credit_amount > 3380: bad (10/2.6)
##     mth_duration > 7:
##     :...chk_status in {A11,A12,A13}:
##         :...purpose in {A410,A44,A48}: good (17/4.1)
##         :   purpose = A45: bad (16.6/7.1)
##         :   purpose = A46:
##         :   :...employ_time in {A71,A72,A73,A75}: bad (22.4/7.6)
##         :   :   employ_time = A74: good (5/0.5)
##         :   purpose = A43:
##         :   :...employ_time = A75: good (22.3/2.7)
##         :   :   employ_time in {A71,A72,A73,A74}:
##         :   :   :...credit_history in {A30,A32,A34}: bad (99.7/38.1)
##         :   :       credit_history in {A31,A33}: good (19.5/8.9)
##         :   purpose = A49:
##         :   :...mth_duration <= 18: good (14.5)
##         :   :   mth_duration > 18:
##         :   :   :...residency_time > 3: bad (17.8/2.8)
##         :   :       residency_time <= 3:
##         :   :       :...phone = A191: good (11.2/0.5)
##         :   :           phone = A192: bad (11.2/4.4)
##         :   purpose = A41:
##         :   :...credit_amount <= 3161: good (13.1)
##         :   :   credit_amount > 3161:
##         :   :   :...mth_duration <= 18: bad (6.3)
##         :   :       mth_duration > 18:
##         :   :       :...other_installments = A141: good (2)
##         :   :           other_installments = A142: bad (2)
##         :   :           other_installments = A143:
##         :   :           :...credit_amount <= 8086: good (18.9/3.4)
##         :   :               credit_amount > 8086: bad (7/1.4)
##         :   purpose = A42:
##         :   :...other_installments = A142: good (10.1)
##         :   :   other_installments in {A141,A143}:
##         :   :   :...other_debts = A103: good (4.9)
##         :   :       other_debts in {A101,A102}:
##         :   :       :...saving in {A62,A63}: bad (5.5/1.4)
##         :   :           saving in {A64,A65}: good (19.5/4.4)
##         :   :           saving = A61:
##         :   :           :...credit_amount > 4281: bad (10.2/0.5)
##         :   :               credit_amount <= 4281:
##         :   :               :...property = A121: bad (15.3/4.4)
##         :   :                   property in {A122,A123,A124}: good (53.8/15.5)
##         :   purpose = A40:
##         :   :...pct_dpi <= 2:
##         :       :...employ_time = A71: good (8.7)
##         :       :   employ_time = A75: bad (7.7/0.5)
##         :       :   employ_time in {A72,A73,A74}:
##         :       :   :...age <= 29: bad (14.5/3.6)
##         :       :       age > 29: good (20.3/2.7)
```

```
##         :       pct_dpi > 2:
##         :       :...other_installments in {A141,A142}: bad (13)
##         :           other_installments = A143:
##         :           :...chk_status = A11: bad (38.9/9.6)
##         :               chk_status = A13: good (5.9/2.5)
##         :               chk_status = A12:
##         :               :...other_debts = A103: bad (0)
##         :                   other_debts = A102: good (1.3)
##         :                   other_debts = A101:
##         :                   :...property = A121: good (5)
##         :                       property in {A122,A123,A124}:
##         :                       :...credit_amount <= 4057: good (19.1/7)
##         :                           credit_amount > 4057: bad (8.3)
##         chk_status = A14:
##         :...other_installments = A142: good (16.2/6.5)
##             other_installments = A141:
##             :...housing = A153: good (6.6/1.4)
##             :   housing = A151: bad (4.3)
##             :   housing = A152:
##             :   :...phone = A192: bad (11.3/2.6)
##             :       phone = A191:
##             :       :...purpose in {A40,A46}: bad (6.2)
##             :           purpose in {A41,A410,A42,A43,A44,A45,A48,
##             :                       A49}: good (19.8/3.8)
##             other_installments = A143:
##             :...other_debts in {A102,A103}: bad (14.8/5.3)
##                 other_debts = A101:
##                 :...age > 31: good (92.9/8.7)
##                     age <= 31:
##                     :...saving in {A64,A65}: good (12.1)
##                         saving in {A61,A62,A63}:
##                         :...credit_history = A31: bad (0)
##                             credit_history in {A30,A34}: good (16/6.1)
##                             credit_history in {A32,A33}:
##                             :...housing = A153: bad (0)
##                                 housing = A151: good (9.2/2.4)
##                                 housing = A152:
##                                 :...saving = A63: good (3)
##                                     saving in {A61,A62}:
##                                     :...job = A172: good (3.9)
##                                         job in {A171,A173,A174}: bad (32.9/10.5)
##
## -----  Trial 4:  -----
##
## Decision tree:
##
## chk_status in {A13,A14}:
## :...other_installments = A142: good (21.6/10.6)
## :   other_installments = A141:
## :   :...other_debts in {A102,A103}: good (2.4)
## :   :   other_debts = A101:
## :   :   :...chk_status = A13: good (8.9/2.3)
## :   :       chk_status = A14:
## :   :       :...status_gender = A94: good (0)
```

```
## :   :           status_gender in {A91,A92}: bad (15.7/6.1)
## :   :           status_gender = A93:
## :   :           :...residency_time <= 2: bad (15.3/3.9)
## :   :               residency_time > 2: good (15.3/1.9)
## :   other_installments = A143:
## :   :...employ_time in {A71,A72}:
## :       :...other_debts = A102: bad (5.8)
## :       :   other_debts = A103: good (0.4)
## :       :   other_debts = A101:
## :       :   :...credit_amount <= 4594: good (32.6/6.2)
## :       :       credit_amount > 4594: bad (13.2/2)
## :       employ_time in {A73,A74,A75}:
## :       :...existing_credits > 2: bad (9.7/3.4)
## :           existing_credits <= 2:
## :           :...age > 31: good (105/11.7)
## :               age <= 31:
## :               :...pct_dpi <= 3: good (46/4.9)
## :                   pct_dpi > 3:
## :                   :...other_debts = A102: good (2.2)
## :                       other_debts = A103: bad (3.2/1.2)
## :                       other_debts = A101:
## :                       :...property in {A122,A124}: good (5.4)
## :                           property in {A121,A123}:
## :                           :...credit_amount <= 1056: good (5.6)
## :                               credit_amount > 1056:
## :                               :...mth_duration <= 21: bad (16.3/1.7)
## :                                   mth_duration > 21: good (14.9/5.2)
## chk_status in {A11,A12}:
## :...credit_history in {A30,A31}:
##     :...age <= 22: good (5)
##     :   age > 22:
##     :   :...other_debts = A102: good (2.9)
##     :       other_debts in {A101,A103}:
##     :       :...housing in {A151,A153}: bad (26.5/1.2)
##     :           housing = A152:
##     :           :...residency_time <= 2: good (19.2/6.3)
##     :               residency_time > 2: bad (17.1/4.7)
##     credit_history in {A32,A33,A34}:
##     :...mth_duration <= 11: good (69.3/16.7)
##         mth_duration > 11:
##         :...other_debts = A102: bad (20.7/7.9)
##             other_debts = A103: good (26.6/7.7)
##             other_debts = A101:
##             :...other_installments = A142: bad (22.2/7.6)
##                 other_installments = A141:
##                 :...residency_time <= 1: good (4.1)
##                 :   residency_time > 1:
##                 :   :...age <= 26: bad (9.3)
##                 :       age > 26: good (19.3/6.9)
##                 other_installments = A143:
##                 :...saving in {A62,A64,A65}: good (86.4/29.3)
##                     saving = A63: bad (13.8/6.4)
##                     saving = A61:
##                     :...mth_duration > 39: bad (18.1/4.4)
```

```
##                           mth_duration <= 39:
##                           :...credit_amount > 7629: bad (15.3/2.5)
##                               credit_amount <= 7629:
##                               :...credit_amount > 5248: good (16.6)
##                                   credit_amount <= 5248:
##                                   :...dependents_num > 1:
##                                       :...phone = A191: bad (17.7/3)
##                                       :   phone = A192: good (6.8/1)
##                                       dependents_num <= 1:
##                                       :...purpose in {A41,A410,A43,A44,
##                                           :           A45}: good (47.7/12.3)
##                                           purpose in {A46,A48,
##                                           :           A49}: bad (6.6/1.6)
##                                           purpose = A40:
##                                           :...job = A171: good (1.1)
##                                           :   job in {A172,A174}: bad (13.7/2.1)
##                                           :   job = A173: [S1]
##                                           purpose = A42:
##                                           :...pct_dpi <= 1: good (4.3)
##                                               pct_dpi > 1:
##                                               :...job in {A171,
##                                                   :       A172}: bad (8.1)
##                                                   job in {A173,A174}: [S2]
##
## SubTree [S1]
##
## employ_time in {A71,A73}: good (10/0.4)
## employ_time in {A72,A74,A75}: bad (16.6/5.2)
##
## SubTree [S2]
##
## residency_time <= 1: good (6.7)
## residency_time > 1:
## :...property in {A121,A122,A124}: good (21.7/7.8)
##     property = A123: bad (7.2)
##
## -----  Trial 5:  -----
##
## Decision tree:
##
## chk_status = A14:
## :...purpose in {A41,A410,A43,A44,A48}: good (109.3/19)
## :   purpose in {A45,A46}: bad (28/11)
## :   purpose = A40:
## :   :...other_installments in {A141,A142}: bad (12.9/3.7)
## :   :   other_installments = A143: good (35.4/4.5)
## :   purpose = A42:
## :   :...status_gender in {A91,A93}: good (16.3/1)
## :   :   status_gender in {A92,A94}: bad (24.2/8.4)
## :   purpose = A49:
## :   :...residency_time > 3: good (6.9)
## :       residency_time <= 3:
## :       :...dependents_num <= 1: bad (20.3/8)
## :           dependents_num > 1: good (2)
```

```
## chk_status in {A11,A12,A13}:
## :...purpose in {A410,A44}: bad (16.1/7.5)
##     purpose in {A45,A48}: good (24.9/9.1)
##     purpose = A41:
##     :...residency_time <= 3: good (13.1/1.7)
##     :   residency_time > 3:
##     :   :...mth_duration <= 18: bad (7/0.9)
##     :       mth_duration > 18: good (24.4/8.7)
##     purpose = A46:
##     :...other_installments in {A141,A142}: bad (4.3)
##     :   other_installments = A143:
##     :   :...employ_time in {A71,A72,A73,A75}: bad (26.3/9.9)
##     :       employ_time = A74: good (4.8)
##     purpose = A49:
##     :...age > 46: bad (5.5)
##     :   age <= 46:
##     :   :...credit_amount <= 10722: good (47.7/10.5)
##     :       credit_amount > 10722: bad (3.5)
##     purpose = A40:
##     :...other_installments = A142: bad (4/0.4)
##     :   other_installments = A141:
##     :   :...saving in {A61,A64,A65}: bad (24.9/2.3)
##     :   :   saving in {A62,A63}: good (4/0.4)
##     :   other_installments = A143:
##     :   :...property in {A122,A123,A124}: bad (91.7/35.3)
##     :       property = A121:
##     :       :...saving in {A61,A62,A64,A65}: good (38.1/9.5)
##     :           saving = A63: bad (2.6)
##     purpose = A42:
##     :...other_installments = A142: good (8.7)
##     :   other_installments in {A141,A143}:
##     :   :...other_debts in {A102,A103}: good (12.5/2.7)
##     :       other_debts = A101:
##     :       :...dependents_num > 1: bad (6.2/0.9)
##     :           dependents_num <= 1:
##     :           :...chk_status = A13: good (5.6)
##     :               chk_status in {A11,A12}:
##     :               :...residency_time <= 1: good (13.8/4.1)
##     :                   residency_time > 1: bad (82.4/31.2)
##     purpose = A43:
##     :...mth_duration > 36: bad (15.3/1.7)
##         mth_duration <= 36:
##         :...mth_duration <= 8: good (7.9)
##             mth_duration > 8:
##             :...employ_time in {A74,A75}:
##                 :...credit_amount <= 939: bad (6.1/1)
##                 :   credit_amount > 939: good (28.8/1.7)
##                 employ_time in {A71,A72,A73}:
##                 :...other_debts = A103: good (6.7)
##                     other_debts in {A101,A102}:
##                     :...saving in {A61,A62,A64,A65}: bad (94.1/34.2)
##                         saving = A63: good (13.6/3.4)
##
## -----  Trial 6:  -----
```

```
##
## Decision tree:
##
## chk_status in {A13,A14}:
## :...mth_duration <= 8: good (26.7/1.3)
## :   mth_duration > 8:
## :   :...other_installments = A142: bad (20.4/8.9)
## :       other_installments = A141:
## :       :...housing = A151: bad (3.7)
## :       :   housing in {A152,A153}:
## :       :   :...age <= 32: good (22.5/4.3)
## :       :       age > 32:
## :       :       :...residency_time <= 3: bad (14.7/1.3)
## :       :           residency_time > 3: good (12.6/2.3)
## :       other_installments = A143:
## :       :...credit_history in {A30,A31,A34}: good (72.1/12.9)
## :           credit_history = A33:
## :           :...pct_dpi <= 3: good (7.6/0.8)
## :           :   pct_dpi > 3: bad (19.1/5.5)
## :           credit_history = A32:
## :           :...existing_credits > 1: bad (19.3/7.3)
## :               existing_credits <= 1:
## :               :...phone = A192: good (41.3/2.3)
## :                   phone = A191:
## :                   :...job in {A171,A173,A174}: good (38.8/9.6)
## :                       job = A172:
## :                       :...mth_duration <= 21: bad (20.8/6.9)
## :                           mth_duration > 21: good (2.8)
## chk_status in {A11,A12}:
## :...credit_history = A30: bad (31/9.9)
##     credit_history = A31:
##     :...age <= 23: good (5.9)
##     :   age > 23:
##     :   :...other_debts in {A101,A103}: bad (28.3/6.9)
##     :       other_debts = A102: good (3.3)
##     credit_history = A33:
##     :...pct_dpi <= 1: good (6.5)
##     :   pct_dpi > 1:
##     :   :...other_debts = A103: bad (0)
##     :       other_debts = A102: good (2.3)
##     :       other_debts = A101:
##     :       :...status_gender in {A91,A94}: bad (5.8)
##     :           status_gender in {A92,A93}:
##     :           :...chk_status = A11: bad (11.7/2.4)
##     :               chk_status = A12: good (25.9/8.2)
##     credit_history = A34:
##     :...residency_time <= 1: good (7)
##     :   residency_time > 1:
##     :   :...existing_credits > 2: good (16.3/2.2)
##     :       existing_credits <= 2:
##     :       :...saving in {A62,A63,A64,A65}: good (22.8/6.2)
##     :           saving = A61:
##     :           :...age > 61: good (4.8)
##     :               age <= 61:
```

```
##     :                :...mth_duration <= 15:
##     :                    :...credit_amount <= 2181: good (21.4/3)
##     :                    :   credit_amount > 2181: bad (8.4/1.5)
##     :                    mth_duration > 15:
##     :                    :...property in {A121,A123,A124}: bad (31.6/3.7)
##     :                        property = A122: good (10.5/3.8)
##     credit_history = A32:
##     :...other_debts = A103: good (16.9/3)
##         other_debts in {A101,A102}:
##         :...saving in {A63,A64}: good (19.5/4.4)
##             saving = A62:
##             :...property = A121: good (2.2)
##             :   property = A122: bad (7.7)
##             :   property in {A123,A124}:
##             :   :...status_gender in {A91,A92,A94}: bad (10.5)
##             :       status_gender = A93: good (12.3/2.4)
##             saving = A65:
##             :...existing_credits > 1: good (3.2)
##             :   existing_credits <= 1:
##             :   :...foreign = A202: good (2.4)
##             :       foreign = A201:
##             :       :...age <= 56: bad (46.7/17.1)
##             :           age > 56: good (4.4)
##             saving = A61:
##             :...employ_time = A71: good (17.6/4.6)
##                 employ_time = A74:
##                 :...status_gender in {A91,A94}: bad (5)
##                 :   status_gender in {A92,A93}: good (24.9/6.6)
##                 employ_time = A75:
##                 :...age <= 34: good (7.2)
##                 :   age > 34: bad (25.3/4.8)
##                 employ_time = A72:
##                 :...foreign = A202: good (2.4)
##                 :   foreign = A201:
##                 :   :...other_installments in {A141,A142}: good (5.6/1.5)
##                 :       other_installments = A143:
##                 :       :...housing in {A151,A153}: bad (16.7/4.4)
##                 :           housing = A152:
##                 :           :...chk_status = A11: bad (15.3/5.4)
##                 :               chk_status = A12: good (16.2/4.6)
##                 employ_time = A73:
##                 :...foreign = A202: bad (2.3)
##                     foreign = A201:
##                     :...existing_credits > 1: bad (3.6)
##                         existing_credits <= 1:
##                         :...dependents_num > 1: bad (7.1/1.5)
##                             dependents_num <= 1:
##                             :...mth_duration > 40: bad (3.1)
##                                 mth_duration <= 40:
##                                 :...age <= 20: bad (3.4)
##                                     age > 20: good (53/15.1)
##
## -----  Trial 7:  -----
##
```

```
## Decision tree:
##
## foreign = A202:
## :...chk_status in {A11,A13,A14}: good (20.1/0.6)
## :   chk_status = A12: bad (3.6/0.5)
## foreign = A201:
## :...chk_status = A14:
##     :...employ_time in {A71,A72}:
##     :   :...credit_amount > 6681: bad (9.4)
##     :   :   credit_amount <= 6681:
##     :   :   :...other_debts in {A101,A103}: good (31.2/10.2)
##     :   :       other_debts = A102: bad (5.7/0.7)
##     :   employ_time in {A73,A74,A75}:
##     :   :...other_installments = A141:
##     :       :...age > 44: good (7.7)
##     :       :   age <= 44:
##     :       :   :...credit_amount <= 709: good (4.3)
##     :       :       credit_amount > 709: bad (26.6/9.5)
##     :       other_installments in {A142,A143}:
##     :       :...saving = A65: good (31.3/1.7)
##     :           saving in {A61,A62,A63,A64}:
##     :           :...pct_dpi <= 3: good (55/5.2)
##     :               pct_dpi > 3:
##     :               :...mth_duration > 33: bad (14.5/4.1)
##     :                   mth_duration <= 33:
##     :                   :...age <= 23: bad (4.8)
##     :                       age > 23: good (58.8/12.1)
##     chk_status in {A11,A12,A13}:
##     :...saving in {A64,A65}: good (99.6/34.7)
##         saving = A62:
##         :...existing_credits > 3: bad (3.1)
##         :   existing_credits <= 3:
##         :   :...property in {A121,A123}: good (28.9/8.9)
##         :       property in {A122,A124}: bad (30/8.2)
##         saving = A63:
##         :...other_installments in {A141,A142}: good (3.1)
##         :   other_installments = A143:
##         :   :...pct_dpi <= 3: good (17.4/4.5)
##         :       pct_dpi > 3: bad (10.4/1.3)
##         saving = A61:
##         :...mth_duration > 15:
##             :...mth_duration > 47: bad (24.4/3.2)
##             :   mth_duration <= 47:
##             :   :...credit_amount <= 2320: bad (74.1/19.2)
##             :       credit_amount > 2320:
##             :       :...other_installments = A142: good (11.3/1.9)
##             :           other_installments = A141:
##             :           :...mth_duration <= 16: bad (3)
##             :           :   mth_duration > 16: good (21.8/5.9)
##             :           other_installments = A143:
##             :           :...residency_time <= 3: bad (76.1/30.8)
##             :               residency_time > 3:
##             :               :...credit_history = A30: bad (5)
##             :                   credit_history in {A31,A32,A33,
```

```
##             :                                        A34}: good (41.6/11.6)
##             mth_duration <= 15:
##             :...other_debts = A103: good (7.9)
##                 other_debts in {A101,A102}:
##                 :...purpose in {A41,A45}: bad (9.8/3.8)
##                     purpose in {A410,A42,A44,A46,A48,A49}: good (59.6/16.4)
##                     purpose = A43:
##                     :...employ_time = A72: bad (12.8/1.8)
##                     :   employ_time in {A71,A73,A74,A75}: good (29.6/6.1)
##                     purpose = A40:
##                     :...residency_time <= 1: bad (6.6/0.6)
##                         residency_time > 1:
##                         :...dependents_num > 1: good (12.2/1.1)
##                             dependents_num <= 1:
##                             :...pct_dpi <= 2: good (15.9/3.5)
##                                 pct_dpi > 2:
##                                 :...status_gender in {A91,A92,
##                                     :                 A93}: bad (20.6/4.7)
##                                     status_gender = A94: good (2.6)
##
## -----  Trial 8:  -----
##
## Decision tree:
##
## chk_status in {A13,A14}:
## :...employ_time in {A71,A72}:
## :   :...credit_amount > 6681: bad (7.8)
## :   :   credit_amount <= 6681:
## :   :   :...property in {A121,A123}: bad (32.9/13.2)
## :   :       property in {A122,A124}: good (20.8/3.7)
## :   employ_time in {A73,A74,A75}:
## :   :...other_installments in {A141,A142}: good (68/23.5)
## :       other_installments = A143:
## :       :...age > 29: good (97.1/6.2)
## :           age <= 29:
## :           :...pct_dpi <= 3: good (26.8/2.6)
## :               pct_dpi > 3:
## :               :...existing_credits > 2: bad (2.5)
## :                   existing_credits <= 2:
## :                   :...credit_history in {A30,A31,A34}: good (7.1)
## :                       credit_history in {A32,A33}:
## :                       :...credit_amount <= 1056: good (4.4)
## :                           credit_amount > 1056: bad (25.7/8.2)
## chk_status in {A11,A12}:
## :...other_debts = A102:
##     :...credit_amount <= 2214: good (8.5)
##     :   credit_amount > 2214: bad (21.5/8.6)
##     other_debts = A103:
##     :...other_installments = A142: good (0)
##     :   other_installments = A141: bad (11.5/4.2)
##     :   other_installments = A143:
##     :   :...mth_duration <= 36: good (18)
##     :       mth_duration > 36: bad (4.5/0.5)
##     other_debts = A101:
```

```
##     :...residency_time <= 1:
##         :...employ_time = A71: bad (6.1)
##         :   employ_time in {A72,A73,A74,A75}: good (81.4/25.1)
##         residency_time > 1:
##         :...pct_dpi > 2:
##             :...credit_history = A30: bad (10.5)
##             :   credit_history in {A31,A32,A33,A34}:
##             :   :...status_gender = A92:
##             :       :...age <= 38: bad (66.4/10)
##             :       :   age > 38: good (13.2/3.4)
##             :       status_gender in {A91,A93,A94}:
##             :       :...purpose in {A41,A410,A44,A49}: good (42.9/13.3)
##             :           purpose in {A45,A46,A48}: bad (21.4/4.6)
##             :           purpose = A43:
##             :           :...mth_duration <= 11: good (11.3)
##             :           :   mth_duration > 11: bad (37.6/10.4)
##             :           purpose = A40:
##             :           :...property = A121: good (15.4/2.7)
##             :           :   property in {A122,A123,A124}:
##             :           :   :...saving in {A61,A64,A65}: bad (29.3/5)
##             :           :       saving in {A62,A63}: good (8/1.5)
##             :           purpose = A42:
##             :           :...housing = A151: bad (1.5)
##             :               housing = A153: good (5/1.5)
##             :               housing = A152:
##             :               :...phone = A191: bad (17.3/5.4)
##             :                   phone = A192: good (14.2/1.7)
##             pct_dpi <= 2:
##             :...credit_amount > 11760: bad (9.4)
##                 credit_amount <= 11760:
##                 :...status_gender in {A91,A94}: bad (27.2/12.1)
##                     status_gender = A93:
##                     :...other_installments = A141: good (10.5)
##                     :   other_installments = A142: bad (2.3/1)
##                     :   other_installments = A143:
##                     :   :...mth_duration <= 16: good (17.3/0.9)
##                     :       mth_duration > 16:
##                     :       :...credit_amount <= 3051: bad (4)
##                     :           credit_amount > 3051:
##                     :           :...age <= 35: good (14.2/0.9)
##                     :               age > 35: bad (9.3/2.8)
##                     status_gender = A92:
##                     :...job = A174: bad (8.6/0.9)
##                         job in {A171,A172,A173}:
##                         :...dependents_num > 1: bad (3.3)
##                             dependents_num <= 1:
##                             :...age > 29: good (12.6)
##                                 age <= 29:
##                                 :...other_installments = A142: good (0)
##                                     other_installments = A141: bad (3.6)
##                                     other_installments = A143:
##                                     :...employ_time in {A71,A74,
##                                         :               A75}: good (11.8)
##                                         employ_time in {A72,A73}: bad (21.5/8.7)
```

```
##
## -----  Trial 9:  -----
##
## Decision tree:
##
## foreign = A202: good (24.2/3.1)
## foreign = A201:
## :...chk_status = A14:
##     :...other_installments in {A141,A142}:
##     :   :...other_debts in {A102,A103}: good (3.1)
##     :   :   other_debts = A101:
##     :   :   :...employ_time in {A71,A73}: bad (33.2/11.3)
##     :   :       employ_time in {A72,A74,A75}: good (36.2/12.3)
##     :   other_installments = A143:
##     :   :...age > 32: good (62.8/1.7)
##     :       age <= 32:
##     :       :...saving in {A62,A64,A65}: good (24.1/2)
##     :           saving in {A61,A63}:
##     :           :...credit_amount > 3345: bad (15.4/4.6)
##     :               credit_amount <= 3345:
##     :               :...age <= 22: bad (7.4/1.4)
##     :                   age > 22: good (40.2/3.5)
##     chk_status in {A11,A12,A13}:
##     :...other_debts in {A102,A103}: good (63.6/22.3)
##         other_debts = A101:
##         :...property = A124:
##             :...job = A172: bad (9.1)
##             :   job in {A171,A173,A174}:
##             :   :...credit_amount > 12169: bad (9.6)
##             :       credit_amount <= 12169:
##             :       :...pct_dpi <= 2: good (26.5/7.6)
##             :           pct_dpi > 2:
##             :           :...chk_status = A11: bad (37.2/8.3)
##             :               chk_status in {A12,A13}: good (25.7/11.3)
##             property in {A121,A122,A123}:
##             :...purpose in {A410,A44,A46}: bad (29/11.9)
##                 purpose in {A45,A48}: good (19.2/6)
##                 purpose = A41:
##                 :...credit_amount <= 7253: good (16.8)
##                 :   credit_amount > 7253: bad (7.4/1.2)
##                 purpose = A49:
##                 :...age <= 46: good (40.6/13.7)
##                 :   age > 46: bad (5.8)
##                 purpose = A43:
##                 :...employ_time in {A71,A72}: bad (28.5/8.8)
##                 :   employ_time in {A74,A75}: good (40.5/8.7)
##                 :   employ_time = A73:
##                 :   :...credit_history in {A30,A34}: bad (9/3.2)
##                 :       credit_history in {A31,A33}: good (2.9/1.1)
##                 :       credit_history = A32:
##                 :       :...property = A121: bad (16.2/6.6)
##                 :           property = A122: good (2.3)
##                 :           property = A123:
##                 :           :...saving in {A61,A63,A64,A65}: good (19.2/3.7)
```

```
##                 :                saving = A62: bad (3.2)
##                 purpose = A40:
##                 :...other_installments in {A141,A142}: bad (16.2/3.3)
##                 :   other_installments = A143:
##                 :   :...credit_history in {A30,A31}: bad (6.7/1.5)
##                 :       credit_history in {A33,A34}: good (35.7/12.2)
##                 :       credit_history = A32:
##                 :       :...chk_status = A13: good (3.7)
##                 :           chk_status in {A11,A12}:
##                 :           :...dependents_num > 1: good (7.5/2.2)
##                 :               dependents_num <= 1:
##                 :               :...pct_dpi <= 2: good (21.7/9.3)
##                 :                   pct_dpi > 2: bad (30.8/7.5)
##                 purpose = A42:
##                 :...other_installments = A142: good (6.2)
##                     other_installments in {A141,A143}:
##                     :...saving in {A62,A63}: bad (3.4)
##                         saving in {A64,A65}: good (21.5/7.5)
##                         saving = A61:
##                         :...credit_amount > 4473: bad (7.4)
##                             credit_amount <= 4473:
##                             :...status_gender in {A91,A94}: good (10.4/3.4)
##                                 status_gender = A93:
##                                 :...credit_amount <= 3578: bad (20.1/5.5)
##                                 :   credit_amount > 3578: good (5)
##                                 status_gender = A92:
##                                 :...property = A121: bad (9.6/2.7)
##                                     property in {A122,A123}:
##                                     :...employ_time = A71: bad (2.7)
##                                         employ_time in {A72,A73,A74,
##                                                         A75}: good (22.4/3.4)
##
##
## Evaluation on training data (900 cases):
##
## Trial        Decision Tree
## -----      ----------------
##    Size      Errors
##
##    0     55  135(15.0%)
##    1     38  178(19.8%)
##    2     43  176(19.6%)
##    3     53  188(20.9%)
##    4     49  167(18.6%)
##    5     39  228(25.3%)
##    6     57  174(19.3%)
##    7     38  186(20.7%)
##    8     45  177(19.7%)
##    9     46  185(20.6%)
## boost             24( 2.7%)   <<
##
##
##     (a)   (b)    <-classified as
##    ----  ----
```

```
##     245    22    (a): class bad
##       2   631    (b): class good
##
##
##  Attribute usage:
##
##  100.00% chk_status
##  100.00% mth_duration
##  100.00% purpose
##  100.00% foreign
##   97.00% other_installments
##   95.67% other_debts
##   92.67% credit_history
##   91.44% saving
##   84.67% age
##   81.00% employ_time
##   78.67% pct_dpi
##   72.33% credit_amount
##   70.00% property
##   68.11% existing_credits
##   63.33% residency_time
##   61.89% job
##   56.89% status_gender
##   35.56% dependents_num
##   28.78% phone
##   23.11% housing
##
##
## Time: 0.1 secs
```

```
CreditData_boost10_predict <- predict(CreditData_boost10, CreditData_test)
CrossTable(CreditData_test$class, CreditData_boost10_predict, prop.chisq = FALSE, prop.c = FALSE, prop.
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  100
##
##
##              | Predicted Class
## Actual Class |       bad |      good | Row Total |
## -------------|-----------|-----------|-----------|
##          bad |        18 |        15 |        33 |
##              |     0.180 |     0.150 |           |
## -------------|-----------|-----------|-----------|
##         good |        11 |        56 |        67 |
##              |     0.110 |     0.560 |           |
## -------------|-----------|-----------|-----------|
## Column Total |        29 |        71 |       100 |
```

```
## -------------|-----------|-----------|-----------|
##
##
```

```
confusionMatrix(CreditData_test$class, CreditData_boost10_predict, positive = "good")
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction bad good
##       bad   18   15
##       good  11   56
##
##                Accuracy : 0.74
##                  95% CI : (0.6427, 0.8226)
##     No Information Rate : 0.71
##     P-Value [Acc > NIR] : 0.2947
##
##                   Kappa : 0.3934
##  Mcnemar's Test P-Value : 0.5563
##
##             Sensitivity : 0.7887
##             Specificity : 0.6207
##          Pos Pred Value : 0.8358
##          Neg Pred Value : 0.5455
##              Prevalence : 0.7100
##          Detection Rate : 0.5600
##    Detection Prevalence : 0.6700
##       Balanced Accuracy : 0.7047
##
##        'Positive' Class : good
##
```

```
# Making mistakes more costlier than others
Matrix_dimensions <- list(c("bad", "good"), c("bad", "good"))
names(Matrix_dimensions) <- c("Predicted", "Actual")

Matrix_dimensions
```

```
## $Predicted
## [1] "bad"  "good"
##
## $Actual
## [1] "bad"  "good"
```

```
error_cost <- matrix(c(0,1,4,0), nrow = 2, dimnames = Matrix_dimensions)

error_cost
```

```
##          Actual
## Predicted bad good
##      bad    0    4
##      good   1    0
```

```
CreditData_cost <- C5.0(CreditData_train[-21], CreditData_train$class, costs = error_cost)
CreditData_cost_predict <- predict(CreditData_cost, CreditData_test)
CrossTable(CreditData_test$class, CreditData_cost_predict, prop.chisq=FALSE, prop.c = FALSE, prop.r=FALS
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  100
##
##
##              | Predicted
##       Actual |       bad |      good | Row Total |
## -------------|-----------|-----------|-----------|
##          bad |         4 |        29 |        33 |
##              |     0.040 |     0.290 |           |
## -------------|-----------|-----------|-----------|
##         good |         2 |        65 |        67 |
##              |     0.020 |     0.650 |           |
## -------------|-----------|-----------|-----------|
## Column Total |         6 |        94 |       100 |
## -------------|-----------|-----------|-----------|
##
##
```