R Notebook

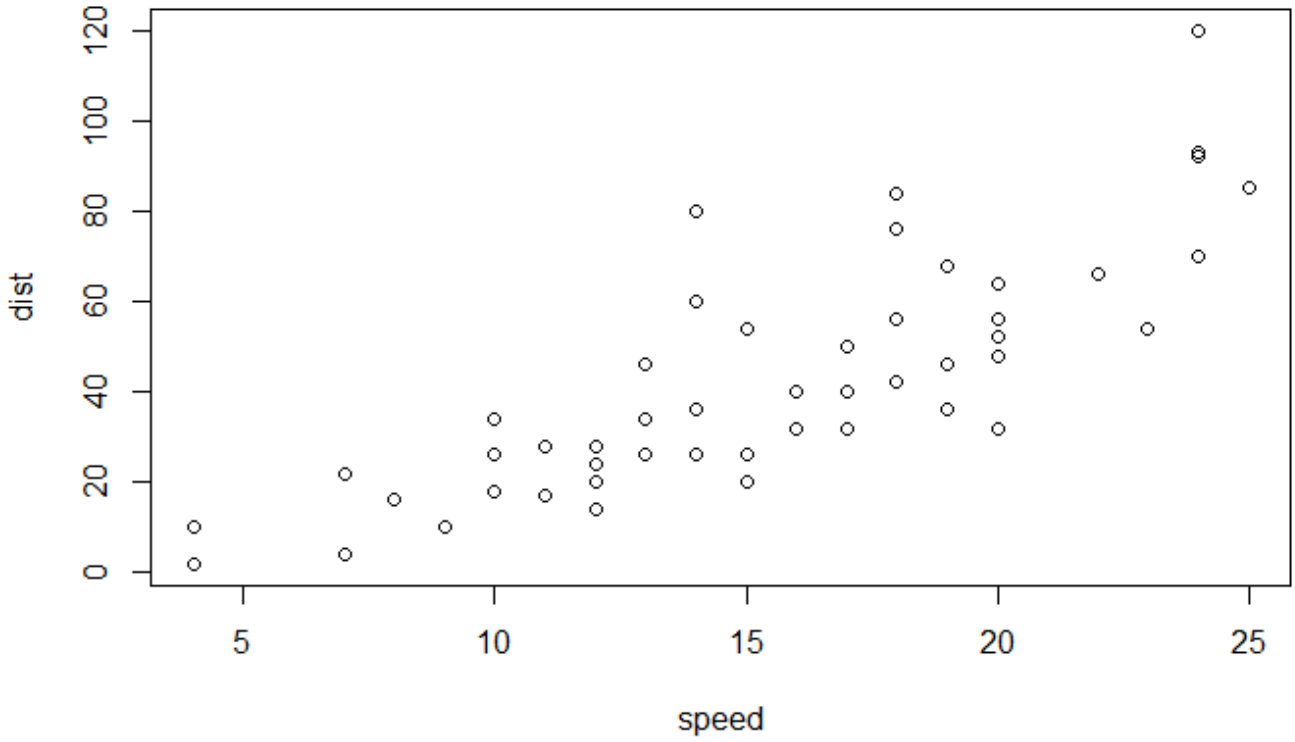
Code

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

Hide

```
plot(cars)
```



Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl+Alt+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview the HTML file).

Hide

```
###I downloaded the CSV file from Kaggle "HR-Analytics". data contains the informat
ion about employees which are best and leaving the company. Company wants to find o
ut why good employees are leaving the company and whom should they retain.
HR<-read.csv(file="C:/Users/CK/Documents/Intro to Data Mining and Machine Learning/
Final Project/HR_comma_sep.csv", header = TRUE)
HR
```

Hide

```
head(HR)
```

Hide

```
str(HR)
```

```
'data.frame':   14999 obs. of  10 variables:
 $ satisfaction_level   : num  0.38 0.8 0.11 0.72 0.37 0.41 0.1 0.92 0.89 0.42 ...
 $ last_evaluation      : num  0.53 0.86 0.88 0.87 0.52 0.5 0.77 0.85 1 0.53 ...
 $ number_project       : int  2 5 7 5 2 2 6 5 5 2 ...
 $ average_montly_hours : int  157 262 272 223 159 153 247 259 224 142 ...
 $ time_spend_company   : int  3 6 4 5 3 3 4 5 5 3 ...
 $ Work_accident        : int  0 0 0 0 0 0 0 0 0 0 ...
 $ left                 : int  1 1 1 1 1 1 1 1 1 1 ...
 $ promotion_last_5years: int  0 0 0 0 0 0 0 0 0 0 ...
 $ sales                : Factor w/ 10 levels "accounting","hr",..: 8 8 8 8 8 8 8
8 8 8 ...
 $ salary               : Factor w/ 3 levels "high","low","medium": 2 3 3 2 2 2 2
2 2 2 ...
```

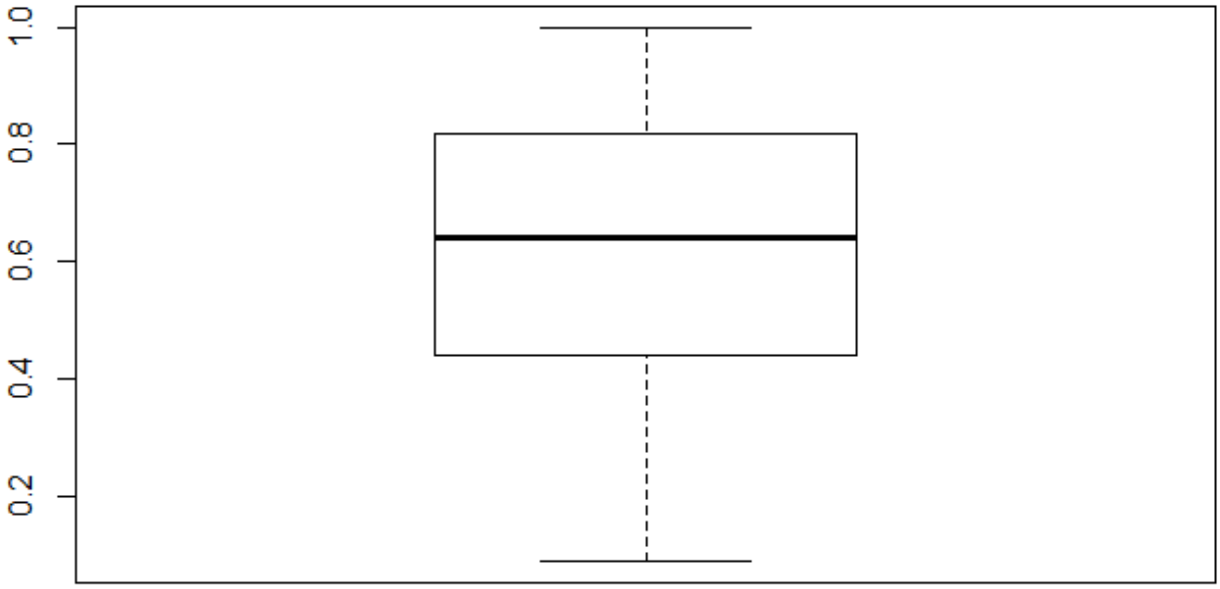
Hide

```
summary(HR)
```

```
 satisfaction_level last_evaluation  number_project  average_montly_hours time_spen
d_company Work_accident         left
 Min.   :0.0900     Min.   :0.3600   Min.   :2.000   Min.   : 96.0        Min.   :
2.000     Min.   :0.0000   Min.   :0.0000
 1st Qu.:0.4400     1st Qu.:0.5600   1st Qu.:3.000   1st Qu.:156.0        1st Qu.:
3.000     1st Qu.:0.0000   1st Qu.:0.0000
 Median :0.6400     Median :0.7200   Median :4.000   Median :200.0        Median :
3.000     Median :0.0000   Median :0.0000
 Mean   :0.6128     Mean   :0.7161   Mean   :3.803   Mean   :201.1        Mean   :
3.498     Mean   :0.1446   Mean   :0.2381
 3rd Qu.:0.8200     3rd Qu.:0.8700   3rd Qu.:5.000   3rd Qu.:245.0        3rd Qu.:
4.000     3rd Qu.:0.0000   3rd Qu.:0.0000
 Max.   :1.0000     Max.   :1.0000   Max.   :7.000   Max.   :310.0        Max.   :
10.000     Max.   :1.0000   Max.   :1.0000


 promotion_last_5years         sales         salary
 Min.   :0.00000       sales      :4140   high  :1237
 1st Qu.:0.00000       technical  :2720   low   :7316
 Median :0.00000       support    :2229   medium:6446
 Mean   :0.02127       IT         :1227
 3rd Qu.:0.00000       product_mng: 902
 Max.   :1.00000       marketing  : 858
                       (Other)    :2923
```

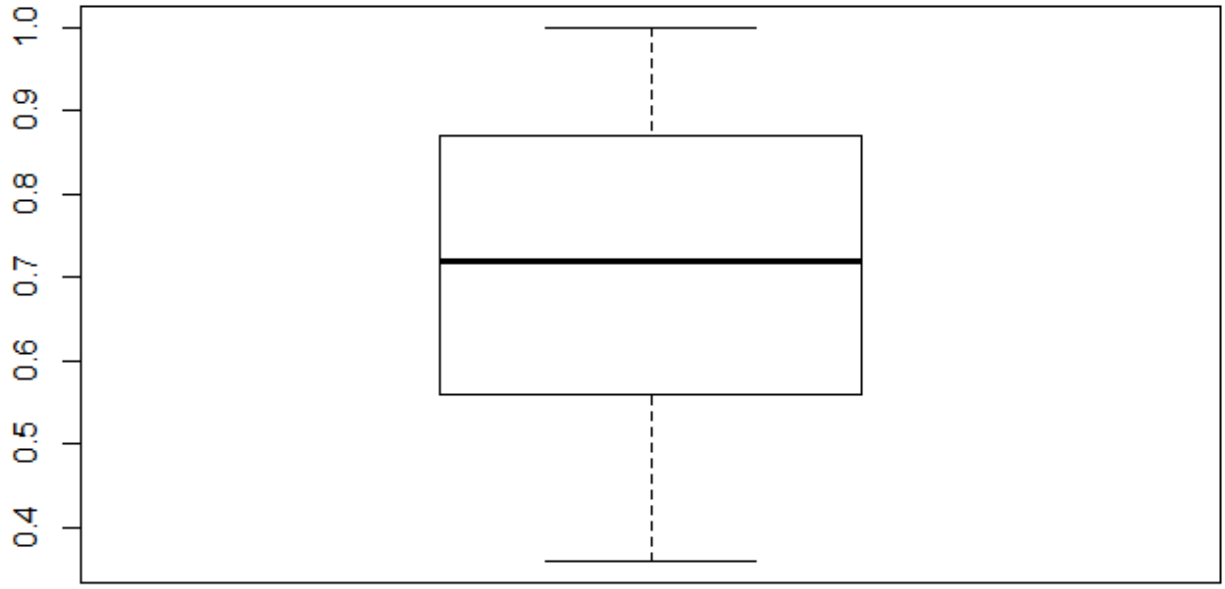
Hide

```
### through summary we can say that the satisfaction level around employees is arou
nd 62% and number of projects they are working is around 4, and evaluation of emplo
yees is at 71%
### corelation analysis
#install.packages("magrittr")
library(dplyr)
### with corrgram we can say that the blue color signifies the most corelated varia
ble and red color least significant variable
###
### Detection of outliers
boxplot(HR$satisfaction_level) # no outliers
```



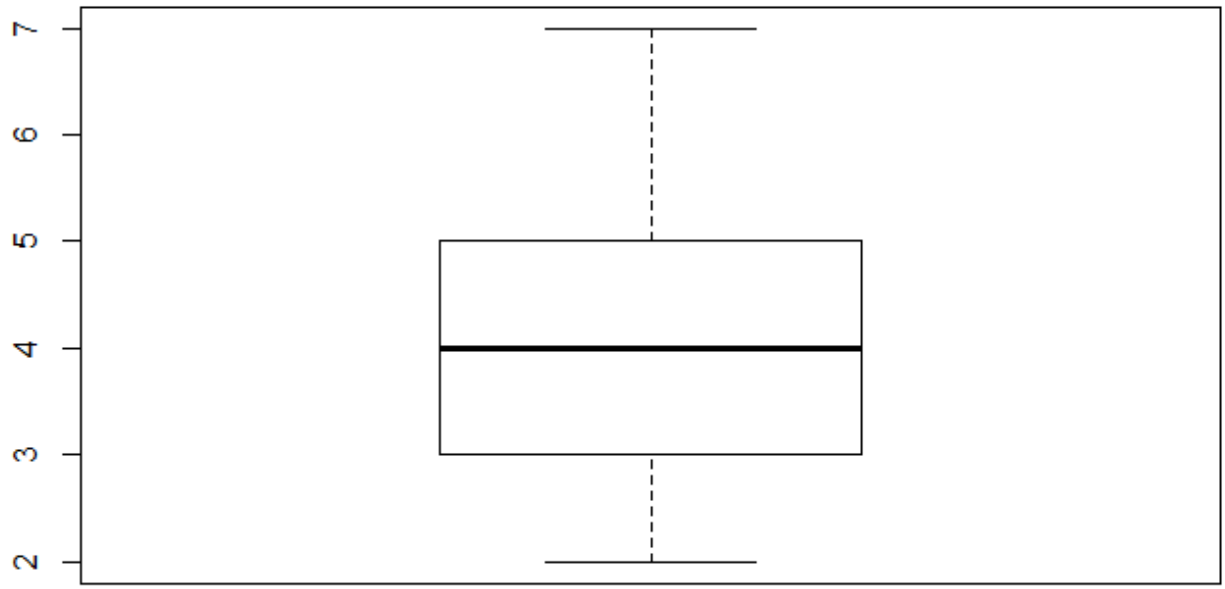
Hide

```
boxplot(HR$last_evaluation) # no outliers
```



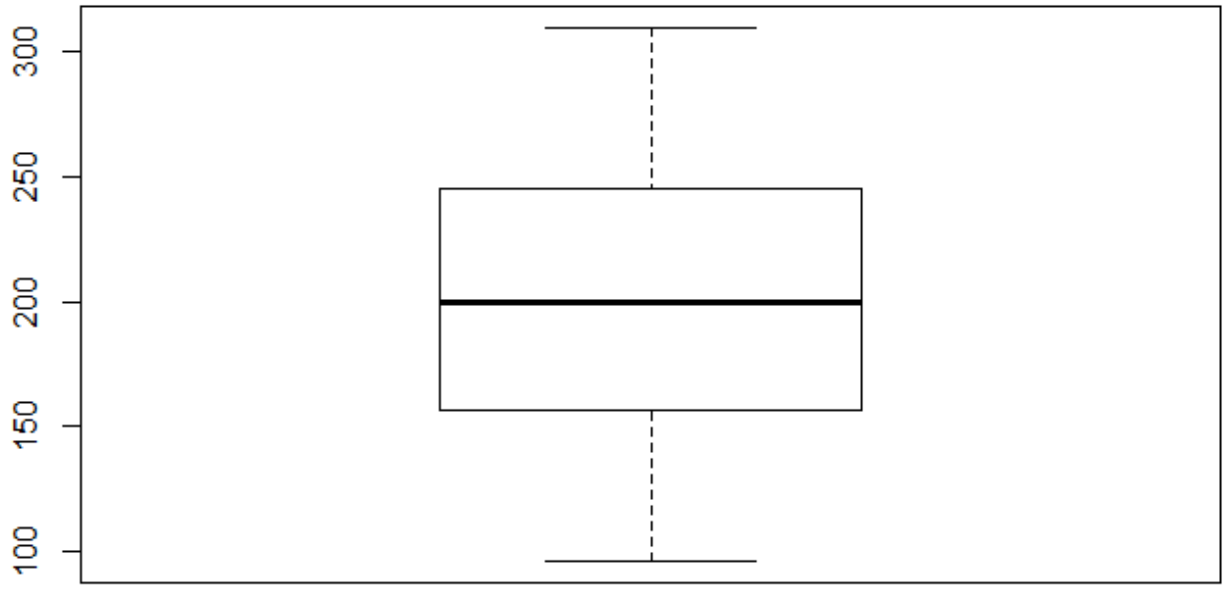
Hide

```
boxplot(HR$number_project) # no outliers
```



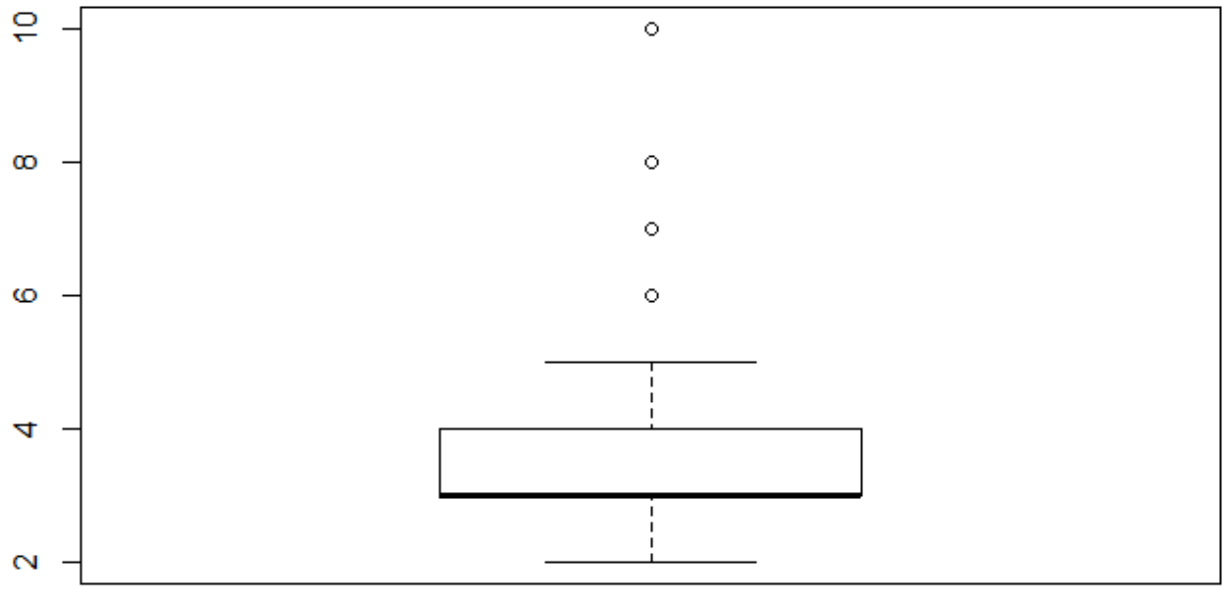
Hide

```
boxplot(HR$average_montly_hours) # no outliers
```



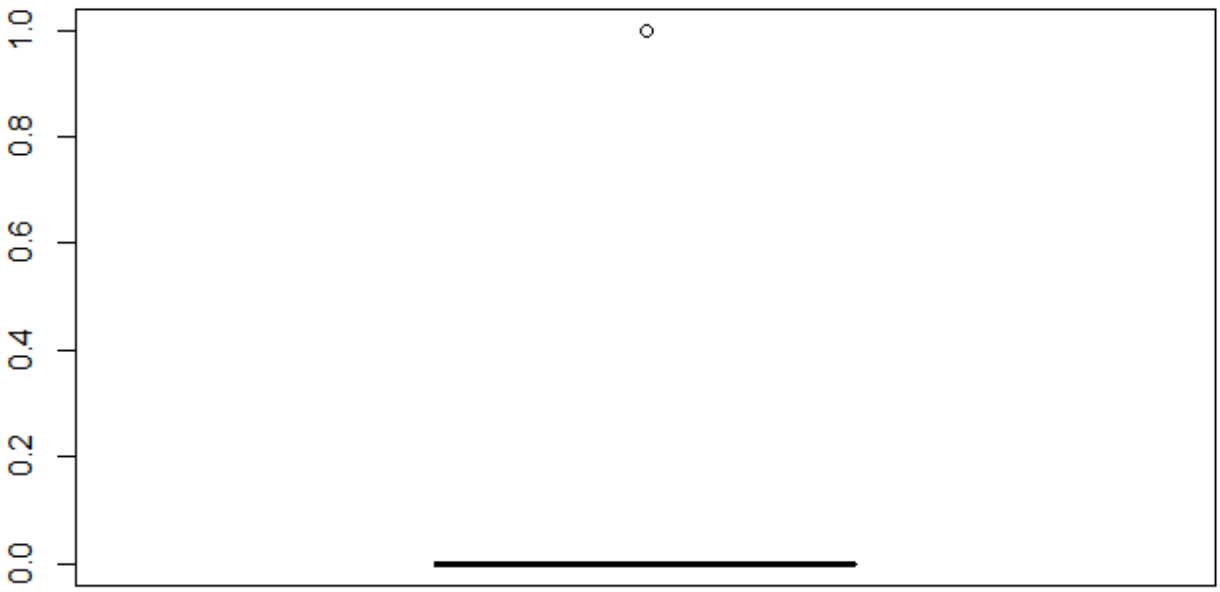
Hide

```
boxplot(HR$time_spend_company) # no outliers
```



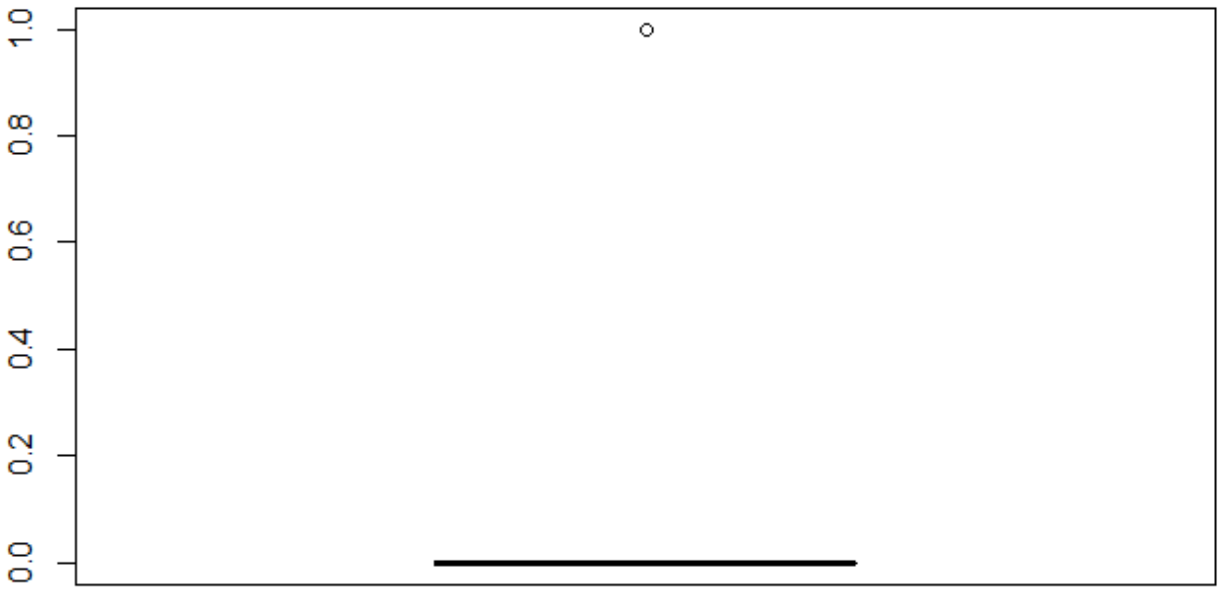
Hide

```
### here we can say that after looking at the boxplot outliers are present in time_
spend_company
boxplot(HR$Work_accident)
```



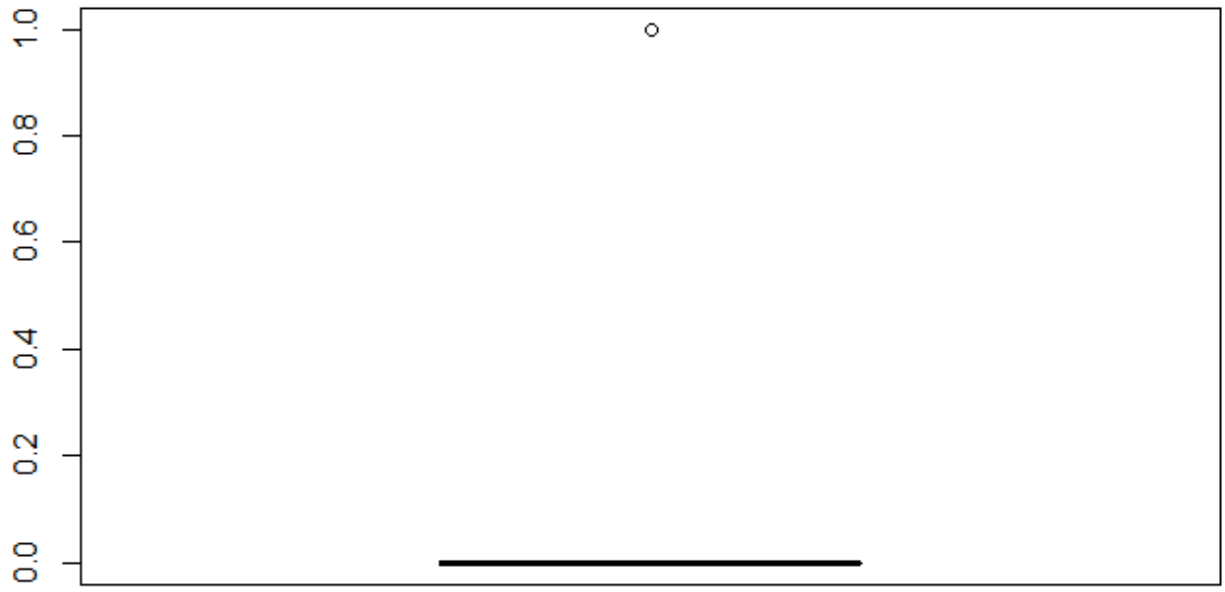
Hide

```
### here outliers are present in work_accident
boxplot(HR$left)
```



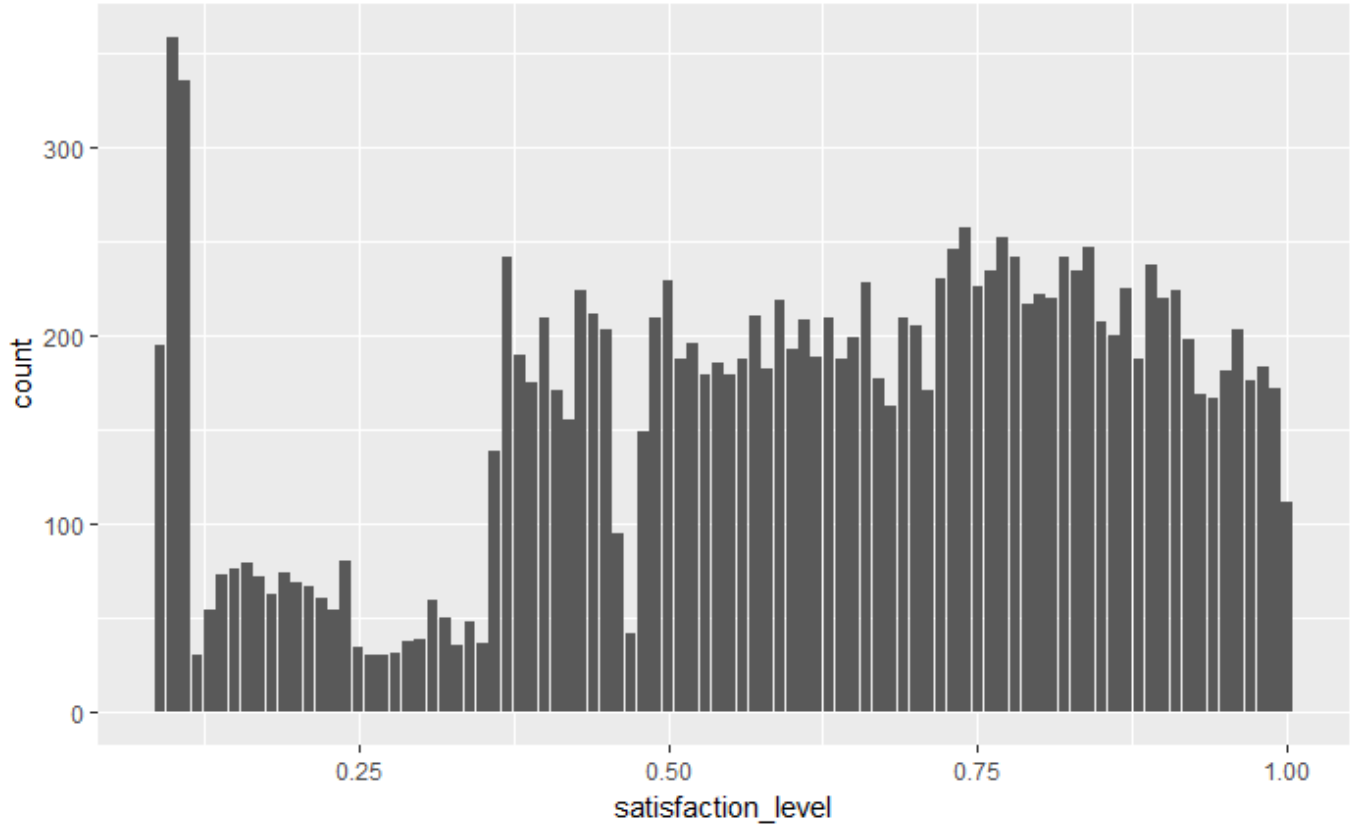
Hide

```
### outliers are present in the left column but as its important data we cannot eli
minate it
boxplot(HR$promotion_last_5years)
```



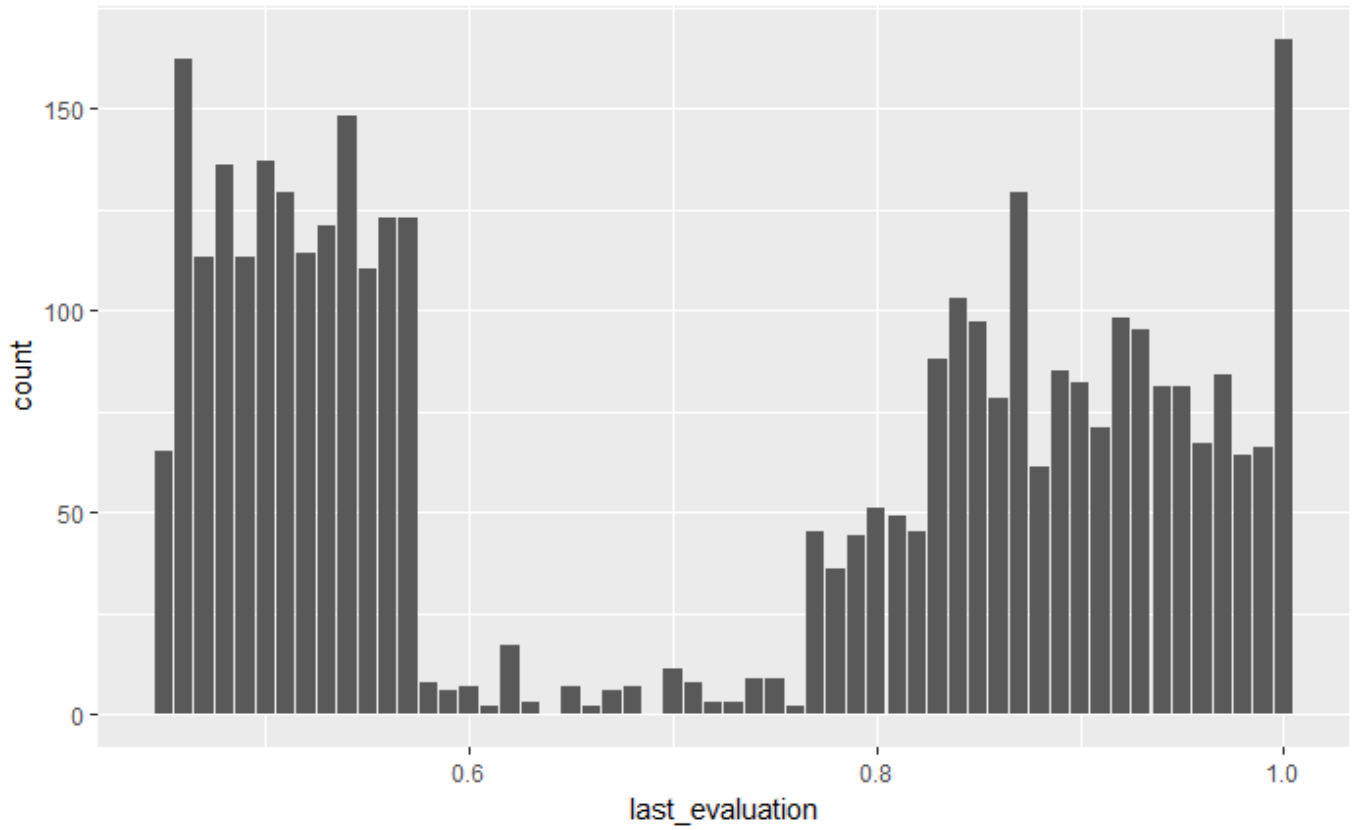
Hide

```
### above column contains outliers
#boxplot(HR$sales)
#boxplot(HR$salary)
### above two columns are character type so no outliers
### Exploratory data plots
library(ggplot2)
ggplot(HR)+ geom_bar(mapping = aes(x=satisfaction_level))
```



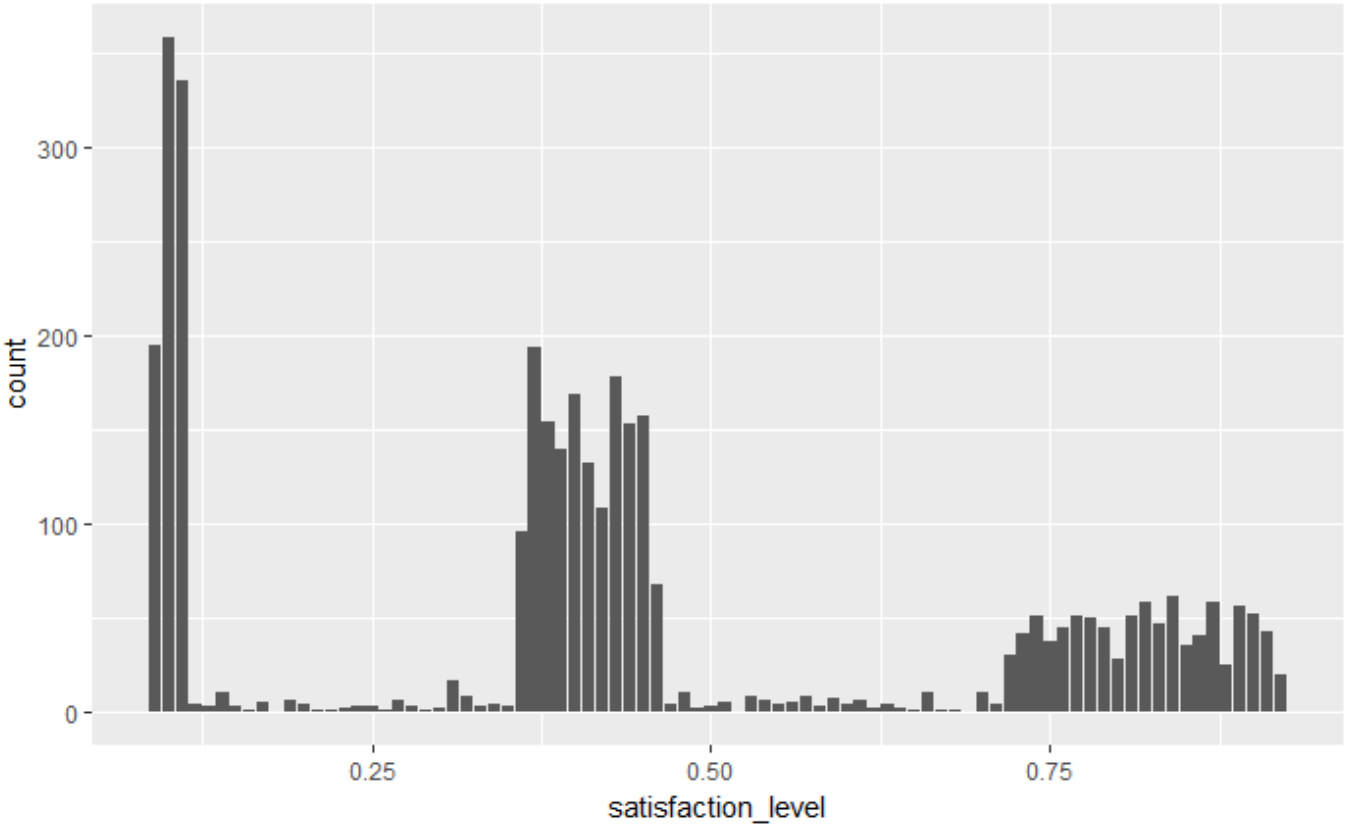
Hide

```
### creating a visualization with respect to the response variable
### for satisfaction level
HR_left <- HR%>% filter(left==1) ### selecting a dataframe containing employees tha
t have left the company to consider effect of response variable on different featur
es
ggplot(HR_left)+ geom_bar(mapping = aes(x=last_evaluation))
```



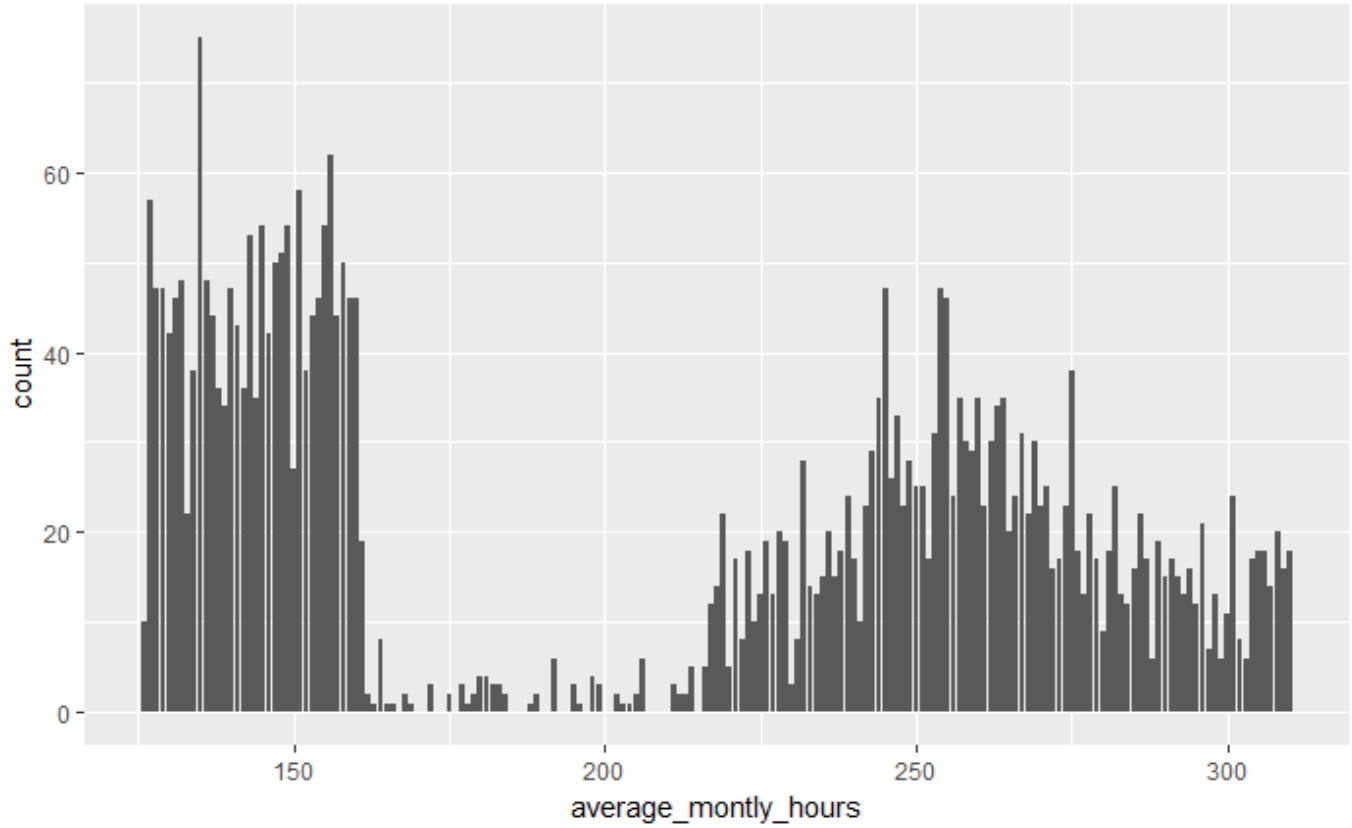
Hide

```
### Here from above plot we can see that employees with lower evaluation are leavin
g the company , but company also loosing the good employees too. So company should
retain those employees
ggplot(HR_left)+ geom_bar(mapping = aes(x=satisfaction_level))
```



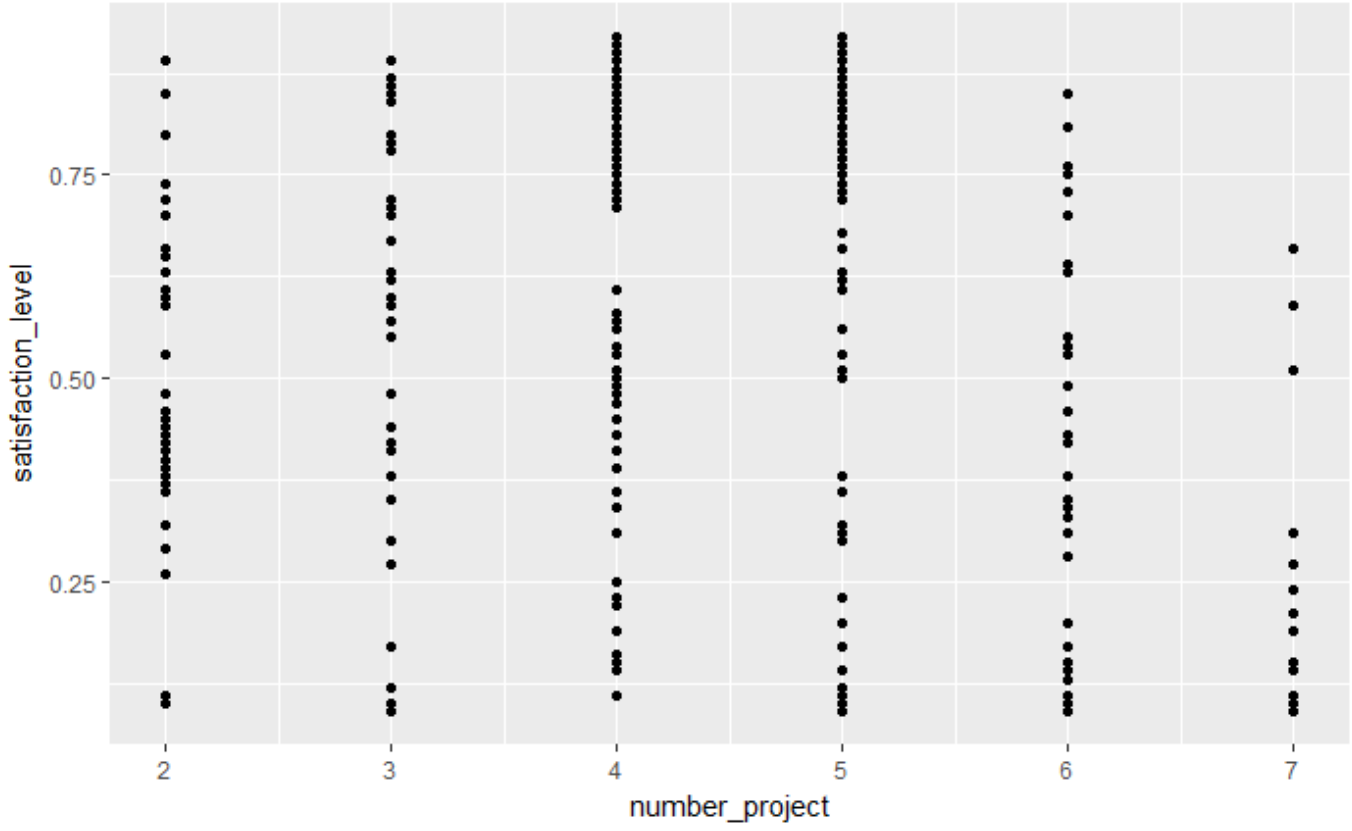
Hide

```
### Here from above plot we can say that employees with lower satisfaction level ar
e tend to leave
ggplot(HR_left)+ geom_bar(mapping = aes(x=average_montly_hours))
```



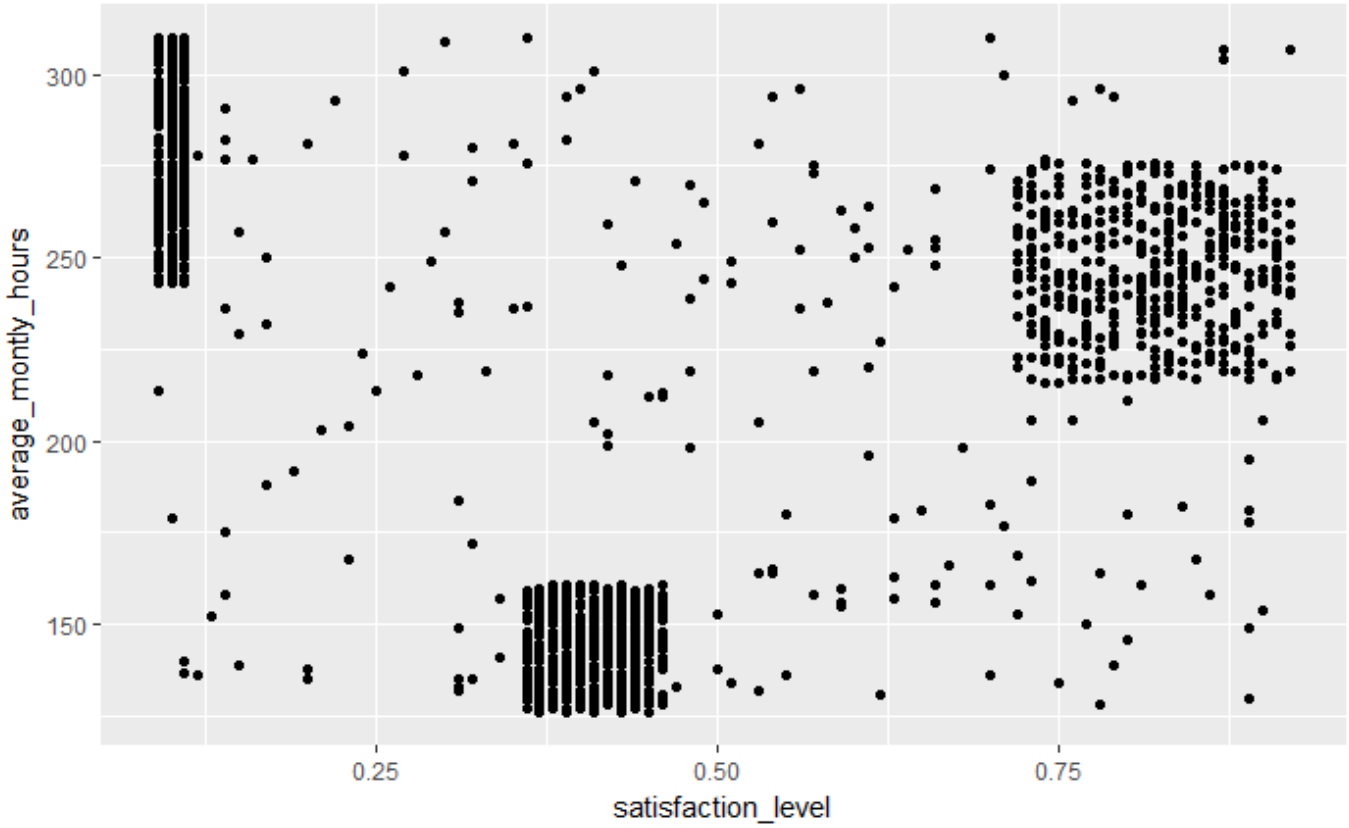
Hide

```
### employees with the less working hours and employees with more working hours are more likely to leave the company
ggplot(data = HR_left) +
  geom_point(mapping = aes(x = number_project, y = satisfaction_level))
```



Hide

```
### from above plot we can say that the number of employees with the good valuation and higher satisfaction level are leaving the company more as compared to others
ggplot(data = HR_left) +
  geom_point(mapping = aes(x = satisfaction_level, y = average_montly_hours))
```



Hide

```
### from above plot we can say that employees are leaving with the more working hours and more satisfaction level
### Data Cleaning and shaping
### Data Imputation
is.na(HR)
```

```
      satisfaction_level last_evaluation number_project average_montly_hours time_spend_company Work_accident  left promotion_last_5years
 [1,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
 [2,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
 [3,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
 [4,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
 [5,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
 [6,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
 [7,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
 [8,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
 [9,]              FALSE           FALSE          FALSE                 FALSE              FALSE         FALSE FALSE                FALSE
```

```
FALSE         FALSE FALSE                FALSE
   [10,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [11,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [12,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [13,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [14,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [15,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [16,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [17,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [18,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [19,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [20,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [21,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [22,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [23,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [24,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [25,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [26,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [27,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [28,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [29,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [30,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [31,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [32,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [33,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [34,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [35,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [36,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
   [37,]              FALSE          FALSE          FALSE                  FALSE
FALSE         FALSE FALSE                FALSE
```

```
   [38,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [39,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [40,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [41,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [42,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [43,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [44,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [45,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [46,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [47,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [48,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [49,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [50,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [51,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [52,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [53,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [54,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [55,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [56,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [57,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [58,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [59,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [60,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [61,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [62,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [63,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [64,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [65,]             FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [66,]             FALSE          FALSE          FALSE                FALSE
```

```
   [66,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [67,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [68,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [69,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [70,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [71,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [72,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [73,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [74,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [75,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [76,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [77,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [78,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [79,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [80,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [81,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [82,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [83,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [84,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [85,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [86,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [87,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [88,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [89,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [90,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [91,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [92,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [93,]              FALSE             FALSE          FALSE                FALSE
FALSE         FALSE FALSE                 FALSE
   [94,]              FALSE             FALSE          FALSE                FALSE
```

```
FALSE         FALSE FALSE                FALSE
   [95,]              FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [96,]              FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [97,]              FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [98,]              FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
   [99,]              FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
  [100,]              FALSE          FALSE          FALSE                FALSE
FALSE         FALSE FALSE                FALSE
        sales salary
    [1,] FALSE  FALSE
    [2,] FALSE  FALSE
    [3,] FALSE  FALSE
    [4,] FALSE  FALSE
    [5,] FALSE  FALSE
    [6,] FALSE  FALSE
    [7,] FALSE  FALSE
    [8,] FALSE  FALSE
    [9,] FALSE  FALSE
   [10,] FALSE  FALSE
   [11,] FALSE  FALSE
   [12,] FALSE  FALSE
   [13,] FALSE  FALSE
   [14,] FALSE  FALSE
   [15,] FALSE  FALSE
   [16,] FALSE  FALSE
   [17,] FALSE  FALSE
   [18,] FALSE  FALSE
   [19,] FALSE  FALSE
   [20,] FALSE  FALSE
   [21,] FALSE  FALSE
   [22,] FALSE  FALSE
   [23,] FALSE  FALSE
   [24,] FALSE  FALSE
   [25,] FALSE  FALSE
   [26,] FALSE  FALSE
   [27,] FALSE  FALSE
   [28,] FALSE  FALSE
   [29,] FALSE  FALSE
   [30,] FALSE  FALSE
   [31,] FALSE  FALSE
   [32,] FALSE  FALSE
   [33,] FALSE  FALSE
   [34,] FALSE  FALSE
   [35,] FALSE  FALSE
   [36,] FALSE  FALSE
   [37,] FALSE  FALSE
   [38,] FALSE  FALSE
   [39,] FALSE  FALSE
   [40,] FALSE  FALSE
   [41,] FALSE  FALSE
   [42,] FALSE  FALSE
   [43,] FALSE  FALSE
```

```
  [44,] FALSE  FALSE
  [45,] FALSE  FALSE
  [46,] FALSE  FALSE
  [47,] FALSE  FALSE
  [48,] FALSE  FALSE
  [49,] FALSE  FALSE
  [50,] FALSE  FALSE
  [51,] FALSE  FALSE
  [52,] FALSE  FALSE
  [53,] FALSE  FALSE
  [54,] FALSE  FALSE
  [55,] FALSE  FALSE
  [56,] FALSE  FALSE
  [57,] FALSE  FALSE
  [58,] FALSE  FALSE
  [59,] FALSE  FALSE
  [60,] FALSE  FALSE
  [61,] FALSE  FALSE
  [62,] FALSE  FALSE
  [63,] FALSE  FALSE
  [64,] FALSE  FALSE
  [65,] FALSE  FALSE
  [66,] FALSE  FALSE
  [67,] FALSE  FALSE
  [68,] FALSE  FALSE
  [69,] FALSE  FALSE
  [70,] FALSE  FALSE
  [71,] FALSE  FALSE
  [72,] FALSE  FALSE
  [73,] FALSE  FALSE
  [74,] FALSE  FALSE
  [75,] FALSE  FALSE
  [76,] FALSE  FALSE
  [77,] FALSE  FALSE
  [78,] FALSE  FALSE
  [79,] FALSE  FALSE
  [80,] FALSE  FALSE
  [81,] FALSE  FALSE
  [82,] FALSE  FALSE
  [83,] FALSE  FALSE
  [84,] FALSE  FALSE
  [85,] FALSE  FALSE
  [86,] FALSE  FALSE
  [87,] FALSE  FALSE
  [88,] FALSE  FALSE
  [89,] FALSE  FALSE
  [90,] FALSE  FALSE
  [91,] FALSE  FALSE
  [92,] FALSE  FALSE
  [93,] FALSE  FALSE
  [94,] FALSE  FALSE
  [95,] FALSE  FALSE
  [96,] FALSE  FALSE
  [97,] FALSE  FALSE
  [98,] FALSE  FALSE
  [99,] FALSE  FALSE
 [100,] FALSE  FALSE
```

```
 [ reached getOption("max.print") -- omitted 14899 rows ]
```

Hide

```
## as there are no null values in dataset so, no data imputation
### Normalization and standardization
#install.packages("clusterSim")
library(clusterSim)
HR_norm <-data.Normalization (HR,type="n0",normalization="average_montly_hours")
```

```
Data not numeric, normalization not applicableData not numeric, normalization not a
pplicable
```

Hide

```
HR_No<-scale(HR_N) # to normalize whole df without categorical variables
summary(HR_norm) #### after comparing original data set and normalized on ewe can s
ee that the data is already normalized
```

```
 satisfaction_level last_evaluation  number_project  average_montly_hours time_spen
d_company Work_accident         left
 Min.   :0.0900     Min.   :0.3600   Min.   :2.000   Min.   : 96.0        Min.   :
2.000     Min.   :0.0000   Min.   :0.0000
 1st Qu.:0.4400     1st Qu.:0.5600   1st Qu.:3.000   1st Qu.:156.0        1st Qu.:
3.000     1st Qu.:0.0000   1st Qu.:0.0000
 Median :0.6400     Median :0.7200   Median :4.000   Median :200.0        Median :
3.000     Median :0.0000   Median :0.0000
 Mean   :0.6128     Mean   :0.7161   Mean   :3.803   Mean   :201.1        Mean   :
3.498     Mean   :0.1446   Mean   :0.2381
 3rd Qu.:0.8200     3rd Qu.:0.8700   3rd Qu.:5.000   3rd Qu.:245.0        3rd Qu.:
4.000     3rd Qu.:0.0000   3rd Qu.:0.0000
 Max.   :1.0000     Max.   :1.0000   Max.   :7.000   Max.   :310.0        Max.   :
10.000     Max.   :1.0000   Max.   :1.0000
 promotion_last_5years     sales            salary
 Min.   :0.00000       Min.   : 1.000   Min.   :1.000
 1st Qu.:0.00000       1st Qu.: 5.000   1st Qu.:2.000
 Median :0.00000       Median : 8.000   Median :2.000
 Mean   :0.02127       Mean   : 6.936   Mean   :2.347
 3rd Qu.:0.00000       3rd Qu.: 9.000   3rd Qu.:3.000
 Max.   :1.00000       Max.   :10.000   Max.   :3.000
```

Hide

```
summary(HR)
```

```
 satisfaction_level last_evaluation  number_project  average_montly_hours time_spend_company Work_accident         left
 Min.   :0.0900     Min.   :0.3600   Min.   :2.000   Min.   : 96.0        Min.   : 2.000     Min.   :0.0000   Min.   :0.0000
 1st Qu.:0.4400     1st Qu.:0.5600   1st Qu.:3.000   1st Qu.:156.0        1st Qu.: 3.000     1st Qu.:0.0000   1st Qu.:0.0000
 Median :0.6400     Median :0.7200   Median :4.000   Median :200.0        Median : 3.000     Median :0.0000   Median :0.0000
 Mean   :0.6128     Mean   :0.7161   Mean   :3.803   Mean   :201.1        Mean   : 3.498     Mean   :0.1446   Mean   :0.2381
 3rd Qu.:0.8200     3rd Qu.:0.8700   3rd Qu.:5.000   3rd Qu.:245.0        3rd Qu.: 4.000     3rd Qu.:0.0000   3rd Qu.:0.0000
 Max.   :1.0000     Max.   :1.0000   Max.   :7.000   Max.   :310.0        Max.   : 10.000     Max.   :1.0000   Max.   :1.0000


 promotion_last_5years         sales         salary
 Min.   :0.00000       sales      :4140   high  :1237
 1st Qu.:0.00000       technical  :2720   low   :7316
 Median :0.00000       support    :2229   medium:6446
 Mean   :0.02127       IT         :1227
 3rd Qu.:0.00000       product_mng: 902
 Max.   :1.00000       marketing  : 858
                       (Other)    :2923
```

Hide

```
### n0 is for normalization
HR_std<-data.Normalization (HR,type="n1",normalization="average_montly_hours")
```

```
Data not numeric, normalization not applicableData not numeric, normalization not applicable
```

Hide

```
summary(HR_std)
```

```
 satisfaction_level last_evaluation    number_project    average_montly_hours time_
spend_company Work_accident          left
 Min.   :-2.1029    Min.   :-2.08041   Min.   :-1.4628   Min.   :-2.10340     Min.
:-1.0261    Min.   :-0.4112   Min.   :-0.559
 1st Qu.:-0.6951    1st Qu.:-0.91197   1st Qu.:-0.6515   1st Qu.:-0.90203     1st
Qu.:-0.3412    1st Qu.:-0.4112   1st Qu.:-0.559
 Median : 0.1093    Median : 0.02277   Median : 0.1598   Median :-0.02103     Medi
an :-0.3412    Median :-0.4112   Median :-0.559
 Mean   : 0.0000    Mean   : 0.00000   Mean   : 0.0000   Mean   : 0.00000     Mean
: 0.0000    Mean   : 0.0000   Mean   : 0.000
 3rd Qu.: 0.8332    3rd Qu.: 0.89910   3rd Qu.: 0.9711   3rd Qu.: 0.87999     3rd Q
u.: 0.3436    3rd Qu.:-0.4112   3rd Qu.:-0.559
 Max.   : 1.5572    Max.   : 1.65858   Max.   : 2.5937   Max.   : 2.18148     Max.
: 4.4528    Max.   : 2.4320   Max.   : 1.789
 promotion_last_5years     sales            salary
 Min.   :-0.1474       Min.   : 1.000   Min.   :1.000
 1st Qu.:-0.1474       1st Qu.: 5.000   1st Qu.:2.000
 Median :-0.1474       Median : 8.000   Median :2.000
 Mean   : 0.0000       Mean   : 6.936   Mean   :2.347
 3rd Qu.:-0.1474       3rd Qu.: 9.000   3rd Qu.:3.000
 Max.   : 6.7835       Max.   :10.000   Max.   :3.000
```

Hide

```
### n1=standardization
### Dummy codes
library(forecast)
set.seed(1)
HR_Data <- data.frame(sex = sample(c("male","female"), 14999, replace = TRUE))
binary_dummy<-model.matrix( ~ sex - 1, data = HR_Data)
binary_dummy
```

```
      sexfemale sexmale
1             0       1
2             0       1
3             1       0
4             1       0
5             0       1
6             1       0
7             1       0
8             1       0
9             1       0
10            0       1
11            0       1
12            0       1
13            1       0
14            0       1
15            1       0
16            0       1
17            1       0
18            1       0
19            0       1
20            1       0
21            1       0
22            0       1
23            1       0
```

23	1	0
24	0	1
25	0	1
26	0	1
27	0	1
28	0	1
29	1	0
30	0	1
31	0	1
32	1	0
33	0	1
34	0	1
35	1	0
36	1	0
37	1	0
38	0	1
39	1	0
40	0	1
41	1	0
42	1	0
43	1	0
44	1	0
45	1	0
46	1	0
47	0	1
48	0	1
49	1	0
50	1	0
51	0	1
52	1	0
53	0	1
54	0	1
55	0	1
56	0	1
57	0	1
58	1	0
59	1	0
60	0	1
61	1	0
62	0	1
63	0	1
64	0	1
65	1	0
66	0	1
67	0	1
68	1	0
69	0	1
70	1	0
71	0	1
72	1	0
73	0	1
74	0	1
75	0	1
76	1	0
77	1	0
78	0	1
79	1	0

80	1	0
81	0	1
82	1	0
83	0	1
84	0	1
85	1	0
86	0	1
87	1	0
88	0	1
89	0	1
90	0	1
91	0	1
92	0	1
93	1	0
94	1	0
95	1	0
96	1	0
97	0	1
98	0	1
99	1	0
100	1	0
101	1	0
102	0	1
103	0	1
104	1	0
105	1	0
106	0	1
107	0	1
108	0	1
109	1	0
110	1	0
111	1	0
112	1	0
113	0	1
114	0	1
115	0	1
116	0	1
117	1	0
118	0	1
119	0	1
120	1	0
121	1	0
122	0	1
123	0	1
124	0	1
125	1	0
126	0	1
127	1	0
128	0	1
129	0	1
130	1	0
131	1	0
132	0	1
133	0	1
134	1	0
135	1	0
136	1	0

```
137             1       0
138             1       0
139             1       0
140             1       0
141             1       0
142             1       0
143             0       1
144             0       1
145             1       0
146             0       1
147             0       1
148             1       0
149             0       1
150             1       0
151             1       0
152             1       0
153             0       1
154             0       1
155             1       0
156             0       1
157             1       0
158             0       1
159             0       1
160             0       1
161             0       1
162             1       0
163             0       1
164             1       0
165             1       0
166             0       1
167             0       1
168             0       1
169             1       0
170             0       1
171             1       0
172             1       0
173             1       0
174             0       1
175             0       1
176             1       0
177             1       0
178             1       0
179             1       0
180             1       0
181             0       1
182             0       1
183             1       0
184             1       0
185             1       0
186             0       1
187             1       0
188             1       0
189             1       0
190             1       0
191             1       0
192             0       1
193             0       1
```

```
194             1       0
195             0       1
196             1       0
197             0       1
198             1       0
199             0       1
200             1       0
201             0       1
202             0       1
203             1       0
204             0       1
205             0       1
206             1       0
207             1       0
208             0       1
209             0       1
210             1       0
211             1       0
212             0       1
213             1       0
214             1       0
215             1       0
216             0       1
217             1       0
218             1       0
219             1       0
220             0       1
221             0       1
222             0       1
223             0       1
224             1       0
225             1       0
226             1       0
227             0       1
228             0       1
229             0       1
230             1       0
231             0       1
232             0       1
233             0       1
234             1       0
235             0       1
236             1       0
237             1       0
238             1       0
239             0       1
240             0       1
241             0       1
242             1       0
243             1       0
244             0       1
245             0       1
246             0       1
247             0       1
248             0       1
249             0       1
250             1       0
```

250	1	0
251	1	0
252	1	0
253	0	1
254	1	0
255	0	1
256	0	1
257	0	1
258	0	1
259	0	1
260	1	0
261	1	0
262	0	1
263	0	1
264	1	0
265	1	0
266	0	1
267	0	1
268	1	0
269	1	0
270	0	1
271	0	1
272	0	1
273	0	1
274	0	1
275	0	1
276	0	1
277	0	1
278	0	1
279	0	1
280	1	0
281	0	1
282	1	0
283	1	0
284	0	1
285	0	1
286	0	1
287	0	1
288	0	1
289	0	1
290	0	1
291	0	1
292	0	1
293	1	0
294	0	1
295	1	0
296	1	0
297	0	1
298	0	1
299	0	1
300	1	0
301	1	0
302	0	1
303	0	1
304	0	1
305	0	1
306	1	0
307	0	1

307	0	1
308	1	0
309	0	1
310	0	1
311	0	1
312	0	1
313	1	0
314	1	0
315	0	1
316	0	1
317	0	1
318	0	1
319	1	0
320	1	0
321	1	0
322	0	1
323	0	1
324	1	0
325	1	0
326	1	0
327	1	0
328	1	0
329	0	1
330	1	0
331	1	0
332	0	1
333	0	1
334	1	0
335	0	1
336	0	1
337	0	1
338	0	1
339	1	0
340	1	0
341	1	0
342	0	1
343	1	0
344	0	1
345	1	0
346	0	1
347	0	1
348	0	1
349	0	1
350	1	0
351	0	1
352	0	1
353	1	0
354	0	1
355	1	0
356	0	1
357	1	0
358	0	1
359	1	0
360	0	1
361	0	1
362	0	1
363	1	0

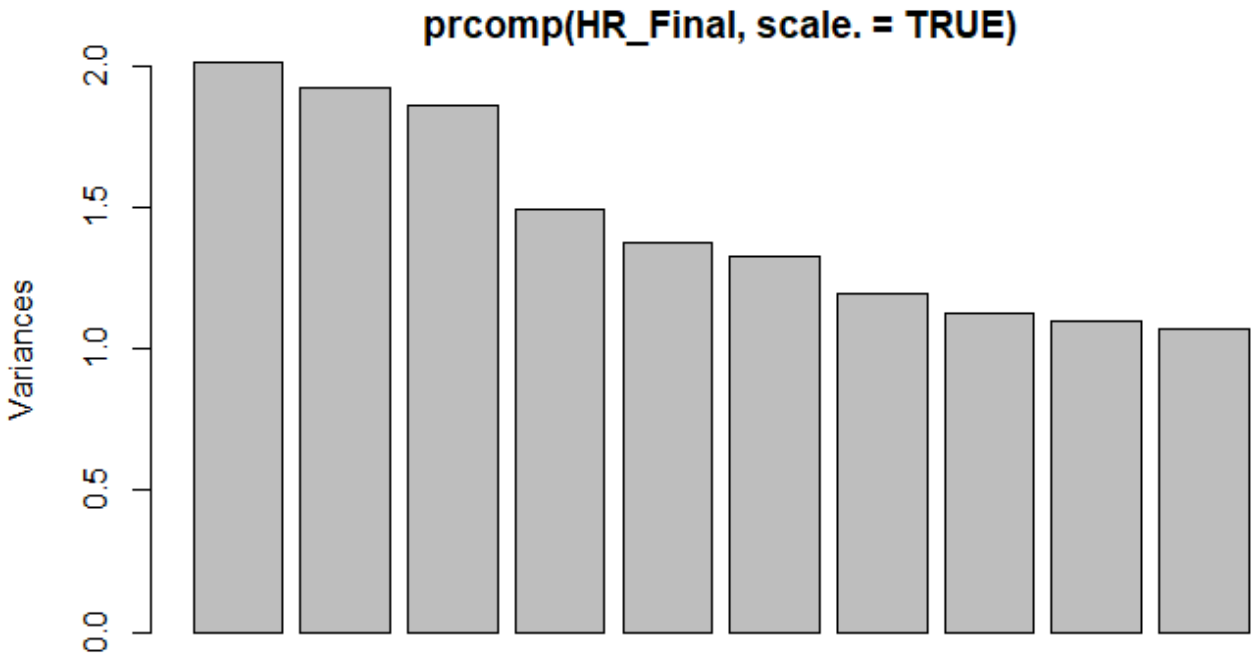
364	0	1
365	1	0
366	1	0
367	0	1
368	1	0
369	1	0
370	0	1
371	0	1
372	1	0
373	1	0
374	0	1
375	0	1
376	0	1
377	0	1
378	1	0
379	0	1
380	0	1
381	0	1
382	1	0
383	1	0
384	1	0
385	0	1
386	0	1
387	1	0
388	1	0
389	0	1
390	1	0
391	0	1
392	0	1
393	1	0
394	1	0
395	0	1
396	0	1
397	1	0
398	0	1
399	0	1
400	0	1
401	1	0
402	0	1
403	1	0
404	1	0
405	1	0
406	1	0
407	0	1
408	1	0
409	0	1
410	1	0
411	1	0
412	0	1
413	1	0
414	1	0
415	1	0
416	0	1
417	0	1
418	0	1
419	1	0
420	1	0

421	0	1
422	1	0
423	1	0
424	0	1
425	0	1
426	0	1
427	0	1
428	1	0
429	1	0
430	1	0
431	1	0
432	0	1
433	0	1
434	1	0
435	0	1
436	1	0
437	1	0
438	0	1
439	0	1
440	1	0
441	0	1
442	0	1
443	1	0
444	1	0
445	0	1
446	0	1
447	0	1
448	1	0
449	0	1
450	0	1
451	1	0
452	0	1
453	0	1
454	0	1
455	0	1
456	1	0
457	1	0
458	1	0
459	1	0
460	0	1
461	0	1
462	0	1
463	0	1
464	0	1
465	0	1
466	1	0
467	1	0
468	0	1
469	0	1
470	1	0
471	0	1
472	1	0
473	1	0
474	0	1
475	0	1
476	0	1
477	1	0

```
478           0       1
479           0       1
480           1       0
481           1       0
482           0       1
483           1       0
484           1       0
485           1       0
486           1       0
487           0       1
488           0       1
489           1       0
490           0       1
491           0       1
492           1       0
493           0       1
494           0       1
495           0       1
496           0       1
497           0       1
498           1       0
499           1       0
500           0       1
 [ reached getOption("max.print") -- omitted 14499 rows ]
attr(,"assign")
[1] 1 1
attr(,"contrasts")
attr(,"contrasts")$sex
[1] "contr.treatment"
```

Hide

```
#### I have created dummy variables male and female
### new derived features
library(forecast)
library(e1071)
library(ade4)
category <- HR[c("sales","salary")]
### as above variables as characte so , converting them to binary using following f
unction
one2n<- acm.disjonctif(category)
### creating a final data frame
HR_N<- HR[,-(9:10)]
### ommitting character variables from main df and selecting converted one
HR_Final<- data.frame(HR_N,one2n, binary_dummy)
### PCA
### to see the co relation between variables carrying out the PCA analysis
HR_pca<- prcomp(HR_Final, scale. = TRUE)
screeplot(prcomp(HR_Final, scale. = TRUE))
```



Hide

```
summary(HR_pca)
```

```
Importance of components%s:
                          PC1     PC2     PC3     PC4     PC5     PC6     PC7
PC8     PC9    PC10    PC11    PC12    PC13   PC14    PC15
Standard deviation     1.41872 1.38647 1.36281 1.22092 1.17292 1.15056 1.09430 1.06
030 1.04703 1.03608 1.03011 1.02626 1.02559 0.9934 0.96487
Proportion of Variance 0.08751 0.08358 0.08075 0.06481 0.05982 0.05756 0.05207 0.04
888 0.04766 0.04667 0.04614 0.04579 0.04573 0.0429 0.04048
Cumulative Proportion  0.08751 0.17109 0.25184 0.31665 0.37647 0.43402 0.48609 0.53
497 0.58263 0.62930 0.67544 0.72123 0.76696 0.8099 0.85034
                         PC16    PC17   PC18    PC19    PC20     PC21      PC22
PC23
Standard deviation     0.92434 0.89535 0.8293 0.79085 0.68768 6.43e-15 4.159e-15 2.
074e-16
Proportion of Variance 0.03715 0.03485 0.0299 0.02719 0.02056 0.00e+00 0.000e+00 0.
000e+00
Cumulative Proportion  0.88749 0.92235 0.9523 0.97944 1.00000 1.00e+00 1.000e+00 1.
000e+00
```

Hide

```
#### creation of training and test data
library(caret)
HR_sample <- sample(2,
                    nrow(HR),
                    replace = T,
                    prob = c(0.7,0.3)) ### splitting the dataset in training and
testing
HR_train <- HR[HR_sample==1,]
HR_test <- HR[HR_sample==2,]
summary(HR_test)
```

```
 satisfaction_level last_evaluation  number_project  average_montly_hours time_spen
d_company Work_accident         left
 Min.   :0.0900     Min.   :0.3600   Min.   :2.000   Min.   : 97.0        Min.   :
2.000     Min.   :0.0000   Min.   :0.0000
 1st Qu.:0.4400     1st Qu.:0.5600   1st Qu.:3.000   1st Qu.:156.0        1st Qu.:
3.000     1st Qu.:0.0000   1st Qu.:0.0000
 Median :0.6400     Median :0.7200   Median :4.000   Median :198.0        Median :
3.000     Median :0.0000   Median :0.0000
 Mean   :0.6118     Mean   :0.7141   Mean   :3.796   Mean   :200.0        Mean   :
3.487     Mean   :0.1478   Mean   :0.2365
 3rd Qu.:0.8200     3rd Qu.:0.8700   3rd Qu.:5.000   3rd Qu.:244.2        3rd Qu.:
4.000     3rd Qu.:0.0000   3rd Qu.:0.0000
 Max.   :1.0000     Max.   :1.0000   Max.   :7.000   Max.   :310.0        Max.   :
10.000     Max.   :1.0000   Max.   :1.0000


 promotion_last_5years         sales         salary
 Min.   :0.00000       sales      :1235   high  : 362
 1st Qu.:0.00000       technical  : 810   low   :2156
 Median :0.00000       support    : 671   medium:1914
 Mean   :0.02301       IT         : 370
 3rd Qu.:0.00000       marketing  : 263
 Max.   :1.00000       product_mng: 259
                       (Other)    : 824
```

Hide

```
summary(HR_train)
```

```
 satisfaction_level last_evaluation  number_project  average_montly_hours time_spen
d_company Work_accident         left
 Min.   :0.0900     Min.   :0.3600   Min.   :2.000   Min.   : 96.0        Min.   :
2.000     Min.   :0.0000   Min.   :0.0000
 1st Qu.:0.4400     1st Qu.:0.5600   1st Qu.:3.000   1st Qu.:156.0        1st Qu.:
3.000     1st Qu.:0.0000   1st Qu.:0.0000
 Median :0.6400     Median :0.7200   Median :4.000   Median :201.0        Median :
3.000     Median :0.0000   Median :0.0000
 Mean   :0.6133     Mean   :0.7169   Mean   :3.806   Mean   :201.5        Mean   :
3.503     Mean   :0.1433   Mean   :0.2388
 3rd Qu.:0.8200     3rd Qu.:0.8700   3rd Qu.:5.000   3rd Qu.:245.0        3rd Qu.:
4.000     3rd Qu.:0.0000   3rd Qu.:0.0000
 Max.   :1.0000     Max.   :1.0000   Max.   :7.000   Max.   :310.0        Max.   :
10.000     Max.   :1.0000   Max.   :1.0000


 promotion_last_5years         sales         salary
 Min.   :0.00000       sales      :2905   high  : 875
 1st Qu.:0.00000       technical  :1910   low   :5160
 Median :0.00000       support    :1558   medium:4532
 Mean   :0.02054       IT         : 857
 3rd Qu.:0.00000       product_mng: 643
 Max.   :1.00000       marketing  : 595
                       (Other)    :2099
```

Hide

```
# cross-validation
HR_train$left <- as.factor(HR_train$left) ### setting the variable as factor
train_control<- trainControl(method="cv", )
head(train_control)
```

```
$method
[1] "cv"

$number
[1] 10

$repeats
[1] NA

$search
[1] "grid"

$p
[1] 0.75

$initialWindow
NULL
```

Hide

```
### Building a tree model
library("rpart") ## loading the necessary libraries
```

```
package <U+393C><U+3E31>rpart<U+393C><U+3E32> was built under R version 3.4.2
```

Hide

```
library("rpart.plot")
```

```
package <U+393C><U+3E31>rpart.plot<U+393C><U+3E32> was built under R version 3.4.2
```

Hide

```
# training the rpart model
library(C50)
```

```
package <U+393C><U+3E31>C50<U+393C><U+3E32> was built under R version 3.4.4
```

Hide

```
HR_rpartmodel<- train(left ~ ., data=HR_train, trControl=train_control, method="rpa
rt", metric="Accuracy")
# making the predictions
predictions<- predict(HR_rpartmodel,HR_train)
HR_model_tree<- cbind(HR_train,predictions)
# summarize results
confusionMatrix<- confusionMatrix(HR_model_tree$predictions,HR_model_tree$left)
confusionMatrix
```

```
Confusion Matrix and Statistics

          Reference
Prediction    0    1
         0 7897  831
         1  147 1692

               Accuracy : 0.9074
                 95% CI : (0.9018, 0.9129)
    No Information Rate : 0.7612
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.7193
 Mcnemar's Test P-Value : < 2.2e-16

            Sensitivity : 0.9817
            Specificity : 0.6706
         Pos Pred Value : 0.9048
         Neg Pred Value : 0.9201
             Prevalence : 0.7612
         Detection Rate : 0.7473
   Detection Prevalence : 0.8260
      Balanced Accuracy : 0.8262

       'Positive' Class : 0
```

Hide

```
### Naive Bayes
library(e1071)
###install.packages("rminer")
library(rminer)
```

```
package <U+393C><U+3E31>rminer<U+393C><U+3E32> was built under R version 3.4.3
```

Hide

```
### training the Naive bayes model on the train data set
HR_nb <- train(left~., data=HR_train, trControl=train_control, method="nb")
```

```
Numerical 0 probability for all classes with observation 34Numerical 0 probability
for all classes with observation 105Numerical 0 probability for all classes with ob
servation 112Numerical 0 probability for all classes with observation 166Numerical
0 probability for all classes with observation 169Numerical 0 probability for all c
lasses with observation 241Numerical 0 probability for all classes with observation
261Numerical 0 probability for all classes with observation 303Numerical 0 probabil
ity for all classes with observation 305Numerical 0 probability for all classes wit
h observation 306Numerical 0 probability for all classes with observation 307Numeri
cal 0 probability for all classes with observation 335Numerical 0 probability for a
ll classes with observation 359Numerical 0 probability for all classes with observa
tion 367Numerical 0 probability for all classes with observation 457Numerical 0 pro
bability for all classes with observation 469Numerical 0 probability for all classe
s with observation 595Numerical 0 probability for all classes with observation 754N
umerical 0 probability for all classes with observation 764Numerical 0 probability
for all classes with observation 779Numerical 0 probability for all classes with ob
servation 789Numerical 0 probability for all classes with observation 791Numerical
0 probability for all classes with observation 792Numerical 0 probability for all c
lasses with observation 800Numerical 0 probability for all classes with observation
801Numerical 0 probability for all classes with observation 804Numerical 0 probabil
ity for all classes with observation 805Numerical 0 probability for all classes wit
h observation 811Numerical 0 probability for all classes with observation 814Numeri
cal 0 probability for all classes with observation 827Numerical 0 probability for a
ll classes with observation 875Numerical 0 probability for all classes with observa
tion 920Numerical 0 probability for all classes with observation 921Numerical 0 pro
bability for all classes with observation 925Numerical 0 probability for all classe
s with observation 926Numerical 0 probability for all classes with observation 927N
umerical 0 probability for all classes with observation 937Numerical 0 probability
for all classes with observation 946Numerical 0 probability for all classes with ob
servation 950Numerical 0 probability for all classes with observation 951Numerical
0 probability for all classes with observation 952Numerical 0 probability for all c
lasses with observation 958Numerical 0 probability for all classes with observation
961Numerical 0 probability for all classes with observation 962Numerical 0 probabil
ity for all classes with observation 969Numerical 0 probability for all classes wit
h observation 972Numerical 0 probability for all classes with observation 978Numeri
cal 0 probability for all classes with observation 979Numerical 0 probability for a
ll classes with observation 980Numerical 0 probability for all classes with observa
tion 985Numerical 0 probability for all classes with observation 987Numerical 0 pro
bability for all classes with observation 991Numerical 0 probability for all classe
s with observation 999Numerical 0 probability for all classes with observation 1000
Numerical 0 probability for all classes with observation 1004Numerical 0 probabilit
y for all classes with observation 1053Numerical 0 probability for all classes with
observation 10Numerical 0 probability for all classes with observation 57Numerical
```

0 probability for all classes with observation 150Numerical 0 probability for all c
lasses with observation 167Numerical 0 probability for all classes with observation
204Numerical 0 probability for all classes with observation 301Numerical 0 probabil
ity for all classes with observation 332Numerical 0 probability for all classes wit
h observation 333Numerical 0 probability for all classes with observation 334Numeri
cal 0 probability for all classes with observation 364Numerical 0 probability for a
ll classes with observation 477Numerical 0 probability for all classes with observa
tion 489Numerical 0 probability for all classes with observation 543Numerical 0 pro
bability for all classes with observation 617Numerical 0 probability for all classe
s with observation 752Numerical 0 probability for all classes with observation 764N
umerical 0 probability for all classes with observation 765Numerical 0 probability
for all classes with observation 791Numerical 0 probability for all classes with ob
servation 792Numerical 0 probability for all classes with observation 793Numerical
0 probability for all classes with observation 797Numerical 0 probability for all c
lasses with observation 798Numerical 0 probability for all classes with observation
800Numerical 0 probability for all classes with observation 805Numerical 0 probabil
ity for all classes with observation 807Numerical 0 probability for all classes wit
h observation 808Numerical 0 probability for all classes with observation 812Numeri
cal 0 probability for all classes with observation 816Numerical 0 probability for a
ll classes with observation 818Numerical 0 probability for all classes with observa
tion 831Numerical 0 probability for all classes with observation 837Numerical 0 pro
bability for all classes with observation 844Numerical 0 probability for all classe
s with observation 847Numerical 0 probability for all classes with observation 853N
umerical 0 probability for all classes with observation 859Numerical 0 probability
for all classes with observation 863Numerical 0 probability for all classes with ob
servation 895Numerical 0 probability for all classes with observation 914Numerical
0 probability for all classes with observation 931Numerical 0 probability for all c
lasses with observation 932Numerical 0 probability for all classes with observation
962Numerical 0 probability for all classes with observation 973Numerical 0 probabil
ity for all classes with observation 976Numerical 0 probability for all classes wit
h observation 980Numerical 0 probability for all classes with observation 983Numeri
cal 0 probability for all classes with observation 984Numerical 0 probability for a
ll classes with observation 987Numerical 0 probability for all classes with observa
tion 998Numerical 0 probability for all classes with observation 1007Numerical 0 pr
obability for all classes with observation 30Numerical 0 probability for all classe
s with observation 33Numerical 0 probability for all classes with observation 207Nu
merical 0 probability for all classes with observation 256Numerical 0 probability f
or all classes with observation 311Numerical 0 probability for all classes with obs
ervation 344Numerical 0 probability for all classes with observation 345Numerical 0
probability for all classes with observation 346Numerical 0 probability for all cla
sses with observation 364Numerical 0 probability for all classes with observation 4
82Numerical 0 probability for all classes with observation 491Numerical 0 probabili
ty for all classes with observation 527Numerical 0 probability for all classes with
observation 530Numerical 0 probability for all classes with observation 753Numerica
l 0 probability for all classes with observation 754Numerical 0 probability for all
classes with observation 764Numerical 0 probability for all classes with observatio
n 774Numerical 0 probability for all classes with observation 791Numerical 0 probab
ility for all classes with observation 797Numerical 0 probability for all classes w
ith observation 801Numerical 0 probability for all classes with observation 804Nume
rical 0 probability for all classes with observation 809Numerical 0 probability for
all classes with observation 811Numerical 0 probability for all classes with observ
ation 812Numerical 0 probability for all classes with observation 817Numerical 0 pr
obability for all classes with observation 820Numerical 0 probability for all class
es with observation 821Numerical 0 probability for all classes with observation 829
Numerical 0 probability for all classes with observation 830Numerical 0 probability
for all classes with observation 903Numerical 0 probability for all classes with ob
servation 905Numerical 0 probability for all classes with observation 911Numerical

0 probability for all classes with observation 925Numerical 0 probability for all c
lasses with observation 926Numerical 0 probability for all classes with observation
932Numerical 0 probability for all classes with observation 934Numerical 0 probabil
ity for all classes with observation 935Numerical 0 probability for all classes wit
h observation 942Numerical 0 probability for all classes with observation 945Numeri
cal 0 probability for all classes with observation 959Numerical 0 probability for a
ll classes with observation 960Numerical 0 probability for all classes with observa
tion 970Numerical 0 probability for all classes with observation 977Numerical 0 pro
bability for all classes with observation 73Numerical 0 probability for all classes
with observation 138Numerical 0 probability for all classes with observation 150Num
erical 0 probability for all classes with observation 324Numerical 0 probability fo
r all classes with observation 362Numerical 0 probability for all classes with obse
rvation 363Numerical 0 probability for all classes with observation 364Numerical 0
probability for all classes with observation 417Numerical 0 probability for all cla
sses with observation 490Numerical 0 probability for all classes with observation 5
12Numerical 0 probability for all classes with observation 543Numerical 0 probabili
ty for all classes with observation 646Numerical 0 probability for all classes with
observation 679Numerical 0 probability for all classes with observation 697Numerica
l 0 probability for all classes with observation 782Numerical 0 probability for all
classes with observation 792Numerical 0 probability for all classes with observatio
n 799Numerical 0 probability for all classes with observation 800Numerical 0 probab
ility for all classes with observation 801Numerical 0 probability for all classes w
ith observation 805Numerical 0 probability for all classes with observation 808Nume
rical 0 probability for all classes with observation 818Numerical 0 probability for
all classes with observation 819Numerical 0 probability for all classes with observ
ation 822Numerical 0 probability for all classes with observation 823Numerical 0 pr
obability for all classes with observation 831Numerical 0 probability for all class
es with observation 840Numerical 0 probability for all classes with observation 845
Numerical 0 probability for all classes with observation 869Numerical 0 probability
for all classes with observation 887Numerical 0 probability for all classes with ob
servation 888Numerical 0 probability for all classes with observation 947Numerical
0 probability for all classes with observation 952Numerical 0 probability for all c
lasses with observation 979Numerical 0 probability for all classes with observation
981Numerical 0 probability for all classes with observation 982Numerical 0 probabil
ity for all classes with observation 984Numerical 0 probability for all classes wit
h observation 986Numerical 0 probability for all classes with observation 987Numeri
cal 0 probability for all classes with observation 992Numerical 0 probability for a
ll classes with observation 993Numerical 0 probability for all classes with observa
tion 994Numerical 0 probability for all classes with observation 1033Numerical 0 pr
obability for all classes with observation 27Numerical 0 probability for all classe
s with observation 58Numerical 0 probability for all classes with observation 101Nu
merical 0 probability for all classes with observation 275Numerical 0 probability f
or all classes with observation 299Numerical 0 probability for all classes with obs
ervation 301Numerical 0 probability for all classes with observation 331Numerical 0
probability for all classes with observation 332Numerical 0 probability for all cla
sses with observation 335Numerical 0 probability for all classes with observation 3
58Numerical 0 probability for all classes with observation 359Numerical 0 probabili
ty for all classes with observation 360Numerical 0 probability for all classes with
observation 374Numerical 0 probability for all classes with observation 709Numerica
l 0 probability for all classes with observation 747Numerical 0 probability for all
classes with observation 756Numerical 0 probability for all classes with observatio
n 771Numerical 0 probability for all classes with observation 788Numerical 0 probab
ility for all classes with observation 794Numerical 0 probability for all classes w
ith observation 797Numerical 0 probability for all classes with observation 800Nume
rical 0 probability for all classes with observation 803Numerical 0 probability for
all classes with observation 804Numerical 0 probability for all classes with observ
ation 824Numerical 0 probability for all classes with observation 854Numerical 0 pr

obability for all classes with observation 877Numerical 0 probability for all class
es with observation 938Numerical 0 probability for all classes with observation 941
Numerical 0 probability for all classes with observation 943Numerical 0 probability
for all classes with observation 944Numerical 0 probability for all classes with ob
servation 947Numerical 0 probability for all classes with observation 950Numerical
0 probability for all classes with observation 960Numerical 0 probability for all c
lasses with observation 961Numerical 0 probability for all classes with observation
963Numerical 0 probability for all classes with observation 964Numerical 0 probabil
ity for all classes with observation 968Numerical 0 probability for all classes wit
h observation 969Numerical 0 probability for all classes with observation 971Numeri
cal 0 probability for all classes with observation 974Numerical 0 probability for a
ll classes with observation 1023Numerical 0 probability for all classes with observ
ation 42Numerical 0 probability for all classes with observation 65Numerical 0 prob
ability for all classes with observation 67Numerical 0 probability for all classes
with observation 128Numerical 0 probability for all classes with observation 227Num
erical 0 probability for all classes with observation 331Numerical 0 probability fo
r all classes with observation 342Numerical 0 probability for all classes with obse
rvation 343Numerical 0 probability for all classes with observation 346Numerical 0
probability for all classes with observation 347Numerical 0 probability for all cla
sses with observation 383Numerical 0 probability for all classes with observation 4
14Numerical 0 probability for all classes with observation 417Numerical 0 probabili
ty for all classes with observation 456Numerical 0 probability for all classes with
observation 483Numerical 0 probability for all classes with observation 763Numerica
l 0 probability for all classes with observation 767Numerical 0 probability for all
classes with observation 769Numerical 0 probability for all classes with observatio
n 788Numerical 0 probability for all classes with observation 793Numerical 0 probab
ility for all classes with observation 794Numerical 0 probability for all classes w
ith observation 795Numerical 0 probability for all classes with observation 802Nume
rical 0 probability for all classes with observation 803Numerical 0 probability for
all classes with observation 804Numerical 0 probability for all classes with observ
ation 807Numerical 0 probability for all classes with observation 816Numerical 0 pr
obability for all classes with observation 817Numerical 0 probability for all class
es with observation 831Numerical 0 probability for all classes with observation 833
Numerical 0 probability for all classes with observation 838Numerical 0 probability
for all classes with observation 839Numerical 0 probability for all classes with ob
servation 842Numerical 0 probability for all classes with observation 844Numerical
0 probability for all classes with observation 924Numerical 0 probability for all c
lasses with observation 932Numerical 0 probability for all classes with observation
934Numerical 0 probability for all classes with observation 940Numerical 0 probabil
ity for all classes with observation 943Numerical 0 probability for all classes wit
h observation 946Numerical 0 probability for all classes with observation 947Numeri
cal 0 probability for all classes with observation 953Numerical 0 probability for a
ll classes with observation 957Numerical 0 probability for all classes with observa
tion 967Numerical 0 probability for all classes with observation 968Numerical 0 pro
bability for all classes with observation 969Numerical 0 probability for all classe
s with observation 971Numerical 0 probability for all classes with observation 974N
umerical 0 probability for all classes with observation 975Numerical 0 probability
for all classes with observation 976Numerical 0 probability for all classes with ob
servation 984Numerical 0 probability for all classes with observation 986Numerical
0 probability for all classes with observation 989Numerical 0 probability for all c
lasses with observation 991Numerical 0 probability for all classes with observation
22Numerical 0 probability for all classes with observation 23Numerical 0 probabilit
y for all classes with observation 38Numerical 0 probability for all classes with o
bservation 40Numerical 0 probability for all classes with observation 101Numerical
0 probability for all classes with observation 104Numerical 0 probability for all c
lasses with observation 114Numerical 0 probability for all classes with observation
115Numerical 0 probability for all classes with observation 244Numerical 0 probabil

115Numerical 0 probability for all classes with observation 244Numerical 0 probabil
ity for all classes with observation 272Numerical 0 probability for all classes wit
h observation 297Numerical 0 probability for all classes with observation 310Numeri
cal 0 probability for all classes with observation 311Numerical 0 probability for a
ll classes with observation 334Numerical 0 probability for all classes with observa
tion 379Numerical 0 probability for all classes with observation 420Numerical 0 pro
bability for all classes with observation 545Numerical 0 probability for all classe
s with observation 571Numerical 0 probability for all classes with observation 639N
umerical 0 probability for all classes with observation 643Numerical 0 probability
for all classes with observation 760Numerical 0 probability for all classes with ob
servation 769Numerical 0 probability for all classes with observation 784Numerical
0 probability for all classes with observation 786Numerical 0 probability for all c
lasses with observation 793Numerical 0 probability for all classes with observation
802Numerical 0 probability for all classes with observation 808Numerical 0 probabil
ity for all classes with observation 810Numerical 0 probability for all classes wit
h observation 811Numerical 0 probability for all classes with observation 812Numeri
cal 0 probability for all classes with observation 815Numerical 0 probability for a
ll classes with observation 819Numerical 0 probability for all classes with observa
tion 822Numerical 0 probability for all classes with observation 829Numerical 0 pro
bability for all classes with observation 830Numerical 0 probability for all classe
s with observation 833Numerical 0 probability for all classes with observation 916N
umerical 0 probability for all classes with observation 925Numerical 0 probability
for all classes with observation 929Numerical 0 probability for all classes with ob
servation 937Numerical 0 probability for all classes with observation 941Numerical
0 probability for all classes with observation 953Numerical 0 probability for all c
lasses with observation 957Numerical 0 probability for all classes with observation
958Numerical 0 probability for all classes with observation 966Numerical 0 probabil
ity for all classes with observation 970Numerical 0 probability for all classes wit
h observation 976Numerical 0 probability for all classes with observation 977Numeri
cal 0 probability for all classes with observation 981Numerical 0 probability for a
ll classes with observation 983Numerical 0 probability for all classes with observa
tion 996Numerical 0 probability for all classes with observation 1041Numerical 0 pr
obability for all classes with observation 1050Numerical 0 probability for all clas
ses with observation 40Numerical 0 probability for all classes with observation 83N
umerical 0 probability for all classes with observation 84Numerical 0 probability f
or all classes with observation 126Numerical 0 probability for all classes with obs
ervation 286Numerical 0 probability for all classes with observation 317Numerical 0
probability for all classes with observation 319Numerical 0 probability for all cla
sses with observation 320Numerical 0 probability for all classes with observation 3
21Numerical 0 probability for all classes with observation 337Numerical 0 probabili
ty for all classes with observation 338Numerical 0 probability for all classes with
observation 363Numerical 0 probability for all classes with observation 364Numerica
l 0 probability for all classes with observation 367Numerical 0 probability for all
classes with observation 368Numerical 0 probability for all classes with observatio
n 404Numerical 0 probability for all classes with observation 405Numerical 0 probab
ility for all classes with observation 477Numerical 0 probability for all classes w
ith observation 592Numerical 0 probability for all classes with observation 602Nume
rical 0 probability for all classes with observation 639Numerical 0 probability for
all classes with observation 725Numerical 0 probability for all classes with observ
ation 765Numerical 0 probability for all classes with observation 797Numerical 0 pr
obability for all classes with observation 798Numerical 0 probability for all class
es with observation 799Numerical 0 probability for all classes with observation 806
Numerical 0 probability for all classes with observation 807Numerical 0 probability
for all classes with observation 818Numerical 0 probability for all classes with ob
servation 819Numerical 0 probability for all classes with observation 825Numerical
0 probability for all classes with observation 827Numerical 0 probability for all c
lasses with observation 828Numerical 0 probability for all classes with observation

830Numerical 0 probability for all classes with observation 838Numerical 0 probabil
ity for all classes with observation 839Numerical 0 probability for all classes wit
h observation 841Numerical 0 probability for all classes with observation 846Numeri
cal 0 probability for all classes with observation 872Numerical 0 probability for a
ll classes with observation 885Numerical 0 probability for all classes with observa
tion 947Numerical 0 probability for all classes with observation 948Numerical 0 pro
bability for all classes with observation 953Numerical 0 probability for all classe
s with observation 960Numerical 0 probability for all classes with observation 961N
umerical 0 probability for all classes with observation 967Numerical 0 probability
for all classes with observation 968Numerical 0 probability for all classes with ob
servation 972Numerical 0 probability for all classes with observation 973Numerical
0 probability for all classes with observation 974Numerical 0 probability for all c
lasses with observation 982Numerical 0 probability for all classes with observation
984Numerical 0 probability for all classes with observation 993Numerical 0 probabil
ity for all classes with observation 995Numerical 0 probability for all classes wit
h observation 1000Numerical 0 probability for all classes with observation 62Numeri
cal 0 probability for all classes with observation 76Numerical 0 probability for al
l classes with observation 87Numerical 0 probability for all classes with observati
on 89Numerical 0 probability for all classes with observation 94Numerical 0 probabi
lity for all classes with observation 99Numerical 0 probability for all classes wit
h observation 101Numerical 0 probability for all classes with observation 125Numeri
cal 0 probability for all classes with observation 132Numerical 0 probability for a
ll classes with observation 150Numerical 0 probability for all classes with observa
tion 174Numerical 0 probability for all classes with observation 180Numerical 0 pro
bability for all classes with observation 188Numerical 0 probability for all classe
s with observation 204Numerical 0 probability for all classes with observation 209N
umerical 0 probability for all classes with observation 234Numerical 0 probability
for all classes with observation 267Numerical 0 probability for all classes with ob
servation 338Numerical 0 probability for all classes with observation 339Numerical
0 probability for all classes with observation 356Numerical 0 probability for all c
lasses with observation 373Numerical 0 probability for all classes with observation
375Numerical 0 probability for all classes with observation 452Numerical 0 probabil
ity for all classes with observation 464Numerical 0 probability for all classes wit
h observation 488Numerical 0 probability for all classes with observation 570Numeri
cal 0 probability for all classes with observation 621Numerical 0 probability for a
ll classes with observation 701Numerical 0 probability for all classes with observa
tion 710Numerical 0 probability for all classes with observation 779Numerical 0 pro
bability for all classes with observation 780Numerical 0 probability for all classe
s with observation 792Numerical 0 probability for all classes with observation 796N
umerical 0 probability for all classes with observation 798Numerical 0 probability
for all classes with observation 805Numerical 0 probability for all classes with ob
servation 813Numerical 0 probability for all classes with observation 816Numerical
0 probability for all classes with observation 817Numerical 0 probability for all c
lasses with observation 824Numerical 0 probability for all classes with observation
835Numerical 0 probability for all classes with observation 836Numerical 0 probabil
ity for all classes with observation 838Numerical 0 probability for all classes wit
h observation 839Numerical 0 probability for all classes with observation 845Numeri
cal 0 probability for all classes with observation 856Numerical 0 probability for a
ll classes with observation 859Numerical 0 probability for all classes with observa
tion 921Numerical 0 probability for all classes with observation 925Numerical 0 pro
bability for all classes with observation 954Numerical 0 probability for all classe
s with observation 955Numerical 0 probability for all classes with observation 965N
umerical 0 probability for all classes with observation 966Numerical 0 probability
for all classes with observation 978Numerical 0 probability for all classes with ob
servation 991Numerical 0 probability for all classes with observation 997Numerical
0 probability for all classes with observation 1003Numerical 0 probability for all
classes with observation 1005Numerical 0 probability for all classes with observati

on 1006Numerical 0 probability for all classes with observation 1009Numerical 0 pro
bability for all classes with observation 1010Numerical 0 probability for all class
es with observation 1017Numerical 0 probability for all classes with observation 10
25Numerical 0 probability for all classes with observation 1028Numerical 0 probabil
ity for all classes with observation 1030Numerical 0 probability for all classes wi
th observation 1041Numerical 0 probability for all classes with observation 1054Num
erical 0 probability for all classes with observation 6Numerical 0 probability for
all classes with observation 18Numerical 0 probability for all classes with observa
tion 60Numerical 0 probability for all classes with observation 111Numerical 0 prob
ability for all classes with observation 117Numerical 0 probability for all classes
with observation 164Numerical 0 probability for all classes with observation 228Num
erical 0 probability for all classes with observation 298Numerical 0 probability fo
r all classes with observation 302Numerical 0 probability for all classes with obse
rvation 334Numerical 0 probability for all classes with observation 335Numerical 0
probability for all classes with observation 339Numerical 0 probability for all cla
sses with observation 365Numerical 0 probability for all classes with observation 3
66Numerical 0 probability for all classes with observation 460Numerical 0 probabili
ty for all classes with observation 500Numerical 0 probability for all classes with
observation 558Numerical 0 probability for all classes with observation 618Numerica
l 0 probability for all classes with observation 636Numerical 0 probability for all
classes with observation 650Numerical 0 probability for all classes with observatio
n 726Numerical 0 probability for all classes with observation 746Numerical 0 probab
ility for all classes with observation 763Numerical 0 probability for all classes w
ith observation 777Numerical 0 probability for all classes with observation 781Nume
rical 0 probability for all classes with observation 784Numerical 0 probability for
all classes with observation 804Numerical 0 probability for all classes with observ
ation 812Numerical 0 probability for all classes with observation 815Numerical 0 pr
obability for all classes with observation 817Numerical 0 probability for all class
es with observation 820Numerical 0 probability for all classes with observation 821
Numerical 0 probability for all classes with observation 823Numerical 0 probability
for all classes with observation 824Numerical 0 probability for all classes with ob
servation 832Numerical 0 probability for all classes with observation 840Numerical
0 probability for all classes with observation 844Numerical 0 probability for all c
lasses with observation 881Numerical 0 probability for all classes with observation
885Numerical 0 probability for all classes with observation 917Numerical 0 probabil
ity for all classes with observation 922Numerical 0 probability for all classes wit
h observation 929Numerical 0 probability for all classes with observation 938Numeri
cal 0 probability for all classes with observation 944Numerical 0 probability for a
ll classes with observation 945Numerical 0 probability for all classes with observa
tion 952Numerical 0 probability for all classes with observation 966Numerical 0 pro
bability for all classes with observation 975Numerical 0 probability for all classe
s with observation 976Numerical 0 probability for all classes with observation 980N
umerical 0 probability for all classes with observation 981Numerical 0 probability
for all classes with observation 984Numerical 0 probability for all classes with ob
servation 986Numerical 0 probability for all classes with observation 1001Numerical
0 probability for all classes with observation 1004Numerical 0 probability for all
classes with observation 1005Numerical 0 probability for all classes with observati
on 1043Numerical 0 probability for all classes with observation 1053

Hide

```
### predictions
predictions_nb <- predict(HR_nb, HR_train)
HR_nb_model <- cbind(HR_train, predictions_nb)
### summarize
confusionMatrix_nb <- confusionMatrix(HR_nb_model$predictions_nb, HR_nb_model$left)
confusionMatrix_nb
```

```
Confusion Matrix and Statistics

          Reference
Prediction    0    1
         0 7942  833
         1  102 1690

               Accuracy : 0.9115
                 95% CI : (0.9059, 0.9169)
    No Information Rate : 0.7612
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.7297
 Mcnemar's Test P-Value : < 2.2e-16

            Sensitivity : 0.9873
            Specificity : 0.6698
         Pos Pred Value : 0.9051
         Neg Pred Value : 0.9431
             Prevalence : 0.7612
         Detection Rate : 0.7516
   Detection Prevalence : 0.8304
      Balanced Accuracy : 0.8286

       'Positive' Class : 0
```

Hide

```
### logistic regression
### training the model on the train dataset
HR_lr <- train(left~., data=HR_train, trControl=train_control, method="LogitBoost")
# make predictions
predictions_lr<- predict(HR_lr,HR_train)
HR_lr_model <- cbind(HR_train,predictions_lr)
# summarize results
confusionMatrix_lr<- confusionMatrix(HR_lr_model$predictions, HR_lr_model$left)
confusionMatrix_lr
```

```
Confusion Matrix and Statistics

          Reference
Prediction    0    1
         0 7854  349
         1  190 2174

               Accuracy : 0.949
                 95% CI : (0.9446, 0.9531)
    No Information Rate : 0.7612
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.8566
 Mcnemar's Test P-Value : 1.007e-11

            Sensitivity : 0.9764
            Specificity : 0.8617
         Pos Pred Value : 0.9575
         Neg Pred Value : 0.9196
             Prevalence : 0.7612
         Detection Rate : 0.7433
   Detection Prevalence : 0.7763
      Balanced Accuracy : 0.9190

       'Positive' Class : 0
```

Hide

```
install.packages("Metrics")
```

```
Installing package into <U+393C><U+3E31>C:/Users/CK/Documents/R/win-library/3.4<U+3
93C><U+3E32>
(as <U+393C><U+3E31>lib<U+393C><U+3E32> is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.4/Metrics_0.1.3.zip'
Content type 'application/zip' length 65116 bytes (63 KB)
downloaded 63 KB
```

```
package  Metrics  successfully unpacked and MD5 sums checked

The downloaded binary packages are in
    C:\Users\CK\AppData\Local\Temp\RtmpUpqpWf\downloaded_packages
```

Hide

```
### buidlingn the KNN classification model
library(caret)
knn_fit <- train(left ~., data = HR_train, method = "knn",
 trControl= train_control,
 preProcess = c("center", "scale"),
 tuneLength = 10)
knn_fit
```

```
k-Nearest Neighbors

10567 samples
    9 predictor
    2 classes: '0', '1'

Pre-processing: centered (18), scaled (18)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 9511, 9510, 9511, 9510, 9511, 9510, ...
Resampling results across tuning parameters:

  k   Accuracy   Kappa
   5  0.9394355  0.8370847
   7  0.9381087  0.8322196
   9  0.9360269  0.8266642
  11  0.9325258  0.8169862
  13  0.9287407  0.8066047
  15  0.9268478  0.8011237
  17  0.9245766  0.7944939
  19  0.9225896  0.7890570
  21  0.9222105  0.7874503
  23  0.9202235  0.7814156

Accuracy was used to select the optimal model using the largest value.
The final value used for the model was k = 5.
```

Hide

```
### tuning the model by tune length of the from 10 to 20
library(caret)
knn_fit_tune <- train(left ~., data = HR_train, method = "knn",
 trControl= train_control,
 preProcess = c("center", "scale"),
 tuneLength = 20)
knn_fit_tune
```

```
k-Nearest Neighbors

10567 samples
    9 predictor
    2 classes: '0', '1'

Pre-processing: centered (18), scaled (18)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 9511, 9510, 9510, 9511, 9510, 9510, ...
Resampling results across tuning parameters:

  k   Accuracy   Kappa
   5  0.9394331  0.8366644
   7  0.9387715  0.8339517
   9  0.9343245  0.8218798
  11  0.9317690  0.8150993
  13  0.9303501  0.8106512
  15  0.9262809  0.7995495
  17  0.9258076  0.7979445
  19  0.9232523  0.7906805
  21  0.9221165  0.7871919
  23  0.9201291  0.7809981
  25  0.9173854  0.7727626
  27  0.9157763  0.7681618
  29  0.9131268  0.7602976
  31  0.9114227  0.7556736
  33  0.9110445  0.7542659
  35  0.9104771  0.7525085
  37  0.9083948  0.7462728
  39  0.9065967  0.7408794
  41  0.9036638  0.7313860
  43  0.9027182  0.7278233

Accuracy was used to select the optimal model using the largest value.
The final value used for the model was k = 5.
```

Hide

```
### after increasing the tune length from 10 to 20 accuracy increased from
###using random forest to see the most important variables with respect to the resp
onse variable
library(randomForest)
```

```
package <U+393C><U+3E31>randomForest<U+393C><U+3E32> was built under R version 3.4.
4randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.

Attaching package: <U+393C><U+3E31>randomForest<U+393C><U+3E32>

The following object is masked from <U+393C><U+3E31>package:ggplot2<U+393C><U+3E32>
:

    margin

The following object is masked from <U+393C><U+3E31>package:dplyr<U+393C><U+3E32>:

    combine
```

Hide

```
randomforest<-randomForest(left ~ satisfaction_level+last_evaluation+number_project
+average_montly_hours+time_spend_company+Work_accident+promotion_last_5years+sales+
salary, HR_train, ntree=500, trcontrol=train_control)
randomforest
```

```
Call:
 randomForest(formula = left ~ satisfaction_level + last_evaluation +      number_p
roject + average_montly_hours + time_spend_company +      Work_accident + promotion
_last_5years + sales + salary, data = HR_train,      ntree = 500, trcontrol = trai
n_control)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 3

        OOB estimate of  error rate: 0.92%
Confusion matrix:
     0    1 class.error
0 8030   14 0.001740428
1   83 2440 0.032897344
```

Hide

```
importance(randomforest) #### through this command we can lay out the mmost importa
nt variables with respect to response variable and they are; satisfaction level, la
st_evaluation, number_project, avergae_montly_hours and time_spend_company
```

```
                     MeanDecreaseGini
satisfaction_level         1302.973003
last_evaluation             445.435157
number_project              682.623900
average_montly_hours        543.174761
time_spend_company          727.081072
Work_accident                22.443770
promotion_last_5years         3.067834
sales                        61.541132
salary                       31.880204
```

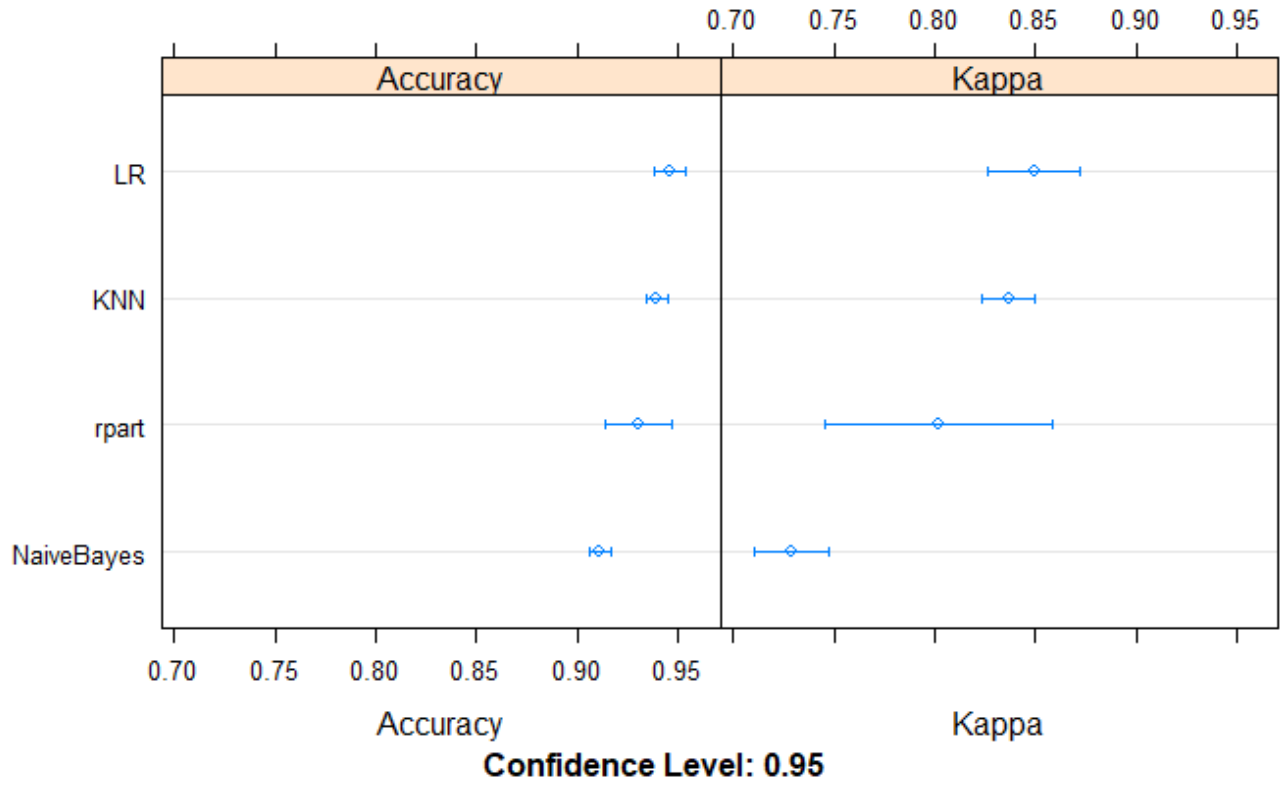
Hide

```
### comparision of model accuracy with dotplot
accuracy <- resamples(list(rpart=HR_rpartmodel, NaiveBayes= HR_nb, LR= HR_lr, KNN=k
nn_fit_tune))
### checking the accuracy for 4 different models
summary(accuracy)
```

```
Call:
summary.resamples(object = accuracy)

Models: rpart, NaiveBayes, LR, KNN
Number of resamples: 10

Accuracy
                Min.   1st Qu.    Median      Mean   3rd Qu.      Max. NA's
rpart      0.8939394 0.9122517 0.9398952 0.9303530 0.9436690 0.9526963    0
NaiveBayes 0.8968780 0.9086174 0.9128788 0.9113254 0.9162926 0.9205298    0
LR         0.9243141 0.9441288 0.9479668 0.9457758 0.9526963 0.9583333    0
KNN        0.9280303 0.9363163 0.9404251 0.9394331 0.9451277 0.9479659    0

Kappa
                Min.   1st Qu.    Median      Mean   3rd Qu.      Max. NA's
rpart      0.6724275 0.7391350 0.8398790 0.8016785 0.8492733 0.8721688    0
NaiveBayes 0.6788174 0.7230532 0.7325226 0.7288770 0.7461864 0.7566105    0
LR         0.7898619 0.8417775 0.8593908 0.8495374 0.8686941 0.8847070    0
KNN        0.8071956 0.8294605 0.8393826 0.8366644 0.8520999 0.8569066    0
```

Hide

```
dotplot(accuracy) ### plotting the accuracy for 4 different models
```



Hide

```
### Ensmebling the model
library(randomForest)
library(caret)
library(caretEnsemble)
```

```
package <U+393C><U+3E31>caretEnsemble<U+393C><U+3E32> was built under R version 3.4
.3
Attaching package: <U+393C><U+3E31>caretEnsemble<U+393C><U+3E32>

The following object is masked from <U+393C><U+3E31>package:ggplot2<U+393C><U+3E32>
:

    autoplot
```

Hide

```
Model_list <- c('rpart', 'nb', 'knn', 'LogitBoost')
set.seed(10)
models <- caretList(left~., data=HR_train, trControl=train_control, methodList=Mode
l_list)
```

```
trControl$savePredictions not 'all' or 'final'.  Setting to 'final' so we can ensem
ble the models.indexes not defined in trControl.  Attempting to set them ourselves,
so each model in the ensemble will have the same resampling indexes.Numerical 0 pro
bability for all classes with observation 35Numerical 0 probability for all classes
with observation 56Numerical 0 probability for all classes with observation 87Numer
ical 0 probability for all classes with observation 129Numerical 0 probability for
all classes with observation 168Numerical 0 probability for all classes with observ
ation 272Numerical 0 probability for all classes with observation 323Numerical 0 pr
obability for all classes with observation 324Numerical 0 probability for all class
```

es with observation 342Numerical 0 probability for all classes with observation 343 Numerical 0 probability for all classes with observation 361Numerical 0 probability for all classes with observation 365Numerical 0 probability for all classes with ob servation 366Numerical 0 probability for all classes with observation 392Numerical 0 probability for all classes with observation 393Numerical 0 probability for all c lasses with observation 402Numerical 0 probability for all classes with observation 433Numerical 0 probability for all classes with observation 499Numerical 0 probabil ity for all classes with observation 520Numerical 0 probability for all classes wit h observation 768Numerical 0 probability for all classes with observation 776Numeri cal 0 probability for all classes with observation 782Numerical 0 probability for a ll classes with observation 787Numerical 0 probability for all classes with observa tion 789Numerical 0 probability for all classes with observation 794Numerical 0 pro bability for all classes with observation 800Numerical 0 probability for all classe s with observation 807Numerical 0 probability for all classes with observation 809N umerical 0 probability for all classes with observation 822Numerical 0 probability for all classes with observation 828Numerical 0 probability for all classes with ob servation 833Numerical 0 probability for all classes with observation 835Numerical 0 probability for all classes with observation 911Numerical 0 probability for all c lasses with observation 925Numerical 0 probability for all classes with observation 926Numerical 0 probability for all classes with observation 949Numerical 0 probabil ity for all classes with observation 950Numerical 0 probability for all classes wit h observation 951Numerical 0 probability for all classes with observation 956Numeri cal 0 probability for all classes with observation 957Numerical 0 probability for a ll classes with observation 961Numerical 0 probability for all classes with observa tion 968Numerical 0 probability for all classes with observation 976Numerical 0 pro bability for all classes with observation 977Numerical 0 probability for all classe s with observation 982Numerical 0 probability for all classes with observation 985N umerical 0 probability for all classes with observation 988Numerical 0 probability for all classes with observation 990Numerical 0 probability for all classes with ob servation 995Numerical 0 probability for all classes with observation 996Numerical 0 probability for all classes with observation 997Numerical 0 probability for all c lasses with observation 1003Numerical 0 probability for all classes with observatio n 20Numerical 0 probability for all classes with observation 36Numerical 0 probabil ity for all classes with observation 52Numerical 0 probability for all classes with observation 77Numerical 0 probability for all classes with observation 149Numerical 0 probability for all classes with observation 215Numerical 0 probability for all c lasses with observation 285Numerical 0 probability for all classes with observation 305Numerical 0 probability for all classes with observation 318Numerical 0 probabil ity for all classes with observation 319Numerical 0 probability for all classes wit h observation 320Numerical 0 probability for all classes with observation 347Numeri cal 0 probability for all classes with observation 428Numerical 0 probability for a ll classes with observation 460Numerical 0 probability for all classes with observa tion 522Numerical 0 probability for all classes with observation 567Numerical 0 pro bability for all classes with observation 759Numerical 0 probability for all classe s with observation 764Numerical 0 probability for all classes with observation 773N umerical 0 probability for all classes with observation 774Numerical 0 probability for all classes with observation 776Numerical 0 probability for all classes with ob servation 781Numerical 0 probability for all classes with observation 782Numerical 0 probability for all classes with observation 783Numerical 0 probability for all c lasses with observation 784Numerical 0 probability for all classes with observation 797Numerical 0 probability for all classes with observation 800Numerical 0 probabil ity for all classes with observation 808Numerical 0 probability for all classes wit h observation 809Numerical 0 probability for all classes with observation 814Numeri cal 0 probability for all classes with observation 822Numerical 0 probability for a ll classes with observation 826Numerical 0 probability for all classes with observa tion 831Numerical 0 probability for all classes with observation 833Numerical 0 pro bability for all classes with observation 911Numerical 0 probability for all classe

bability for all classes with observation 911Numerical 0 probability for all classe
s with observation 912Numerical 0 probability for all classes with observation 930N
umerical 0 probability for all classes with observation 940Numerical 0 probability
for all classes with observation 950Numerical 0 probability for all classes with ob
servation 956Numerical 0 probability for all classes with observation 960Numerical
0 probability for all classes with observation 961Numerical 0 probability for all c
lasses with observation 968Numerical 0 probability for all classes with observation
969Numerical 0 probability for all classes with observation 971Numerical 0 probabil
ity for all classes with observation 975Numerical 0 probability for all classes wit
h observation 979Numerical 0 probability for all classes with observation 986Numeri
cal 0 probability for all classes with observation 992Numerical 0 probability for a
ll classes with observation 1042Numerical 0 probability for all classes with observ
ation 1044Numerical 0 probability for all classes with observation 1048Numerical 0
probability for all classes with observation 5Numerical 0 probability for all class
es with observation 83Numerical 0 probability for all classes with observation 129N
umerical 0 probability for all classes with observation 184Numerical 0 probability
for all classes with observation 260Numerical 0 probability for all classes with ob
servation 286Numerical 0 probability for all classes with observation 307Numerical
0 probability for all classes with observation 310Numerical 0 probability for all c
lasses with observation 311Numerical 0 probability for all classes with observation
352Numerical 0 probability for all classes with observation 356Numerical 0 probabil
ity for all classes with observation 360Numerical 0 probability for all classes wit
h observation 361Numerical 0 probability for all classes with observation 380Numeri
cal 0 probability for all classes with observation 441Numerical 0 probability for a
ll classes with observation 477Numerical 0 probability for all classes with observa
tion 665Numerical 0 probability for all classes with observation 761Numerical 0 pro
bability for all classes with observation 769Numerical 0 probability for all classe
s with observation 788Numerical 0 probability for all classes with observation 790N
umerical 0 probability for all classes with observation 791Numerical 0 probability
for all classes with observation 792Numerical 0 probability for all classes with ob
servation 794Numerical 0 probability for all classes with observation 799Numerical
0 probability for all classes with observation 803Numerical 0 probability for all c
lasses with observation 809Numerical 0 probability for all classes with observation
814Numerical 0 probability for all classes with observation 818Numerical 0 probabil
ity for all classes with observation 830Numerical 0 probability for all classes wit
h observation 831Numerical 0 probability for all classes with observation 839Numeri
cal 0 probability for all classes with observation 847Numerical 0 probability for a
ll classes with observation 848Numerical 0 probability for all classes with observa
tion 853Numerical 0 probability for all classes with observation 856Numerical 0 pro
bability for all classes with observation 950Numerical 0 probability for all classe
s with observation 953Numerical 0 probability for all classes with observation 955N
umerical 0 probability for all classes with observation 960Numerical 0 probability
for all classes with observation 961Numerical 0 probability for all classes with ob
servation 966Numerical 0 probability for all classes with observation 967Numerical
0 probability for all classes with observation 968Numerical 0 probability for all c
lasses with observation 969Numerical 0 probability for all classes with observation
975Numerical 0 probability for all classes with observation 981Numerical 0 probabil
ity for all classes with observation 982Numerical 0 probability for all classes wit
h observation 983Numerical 0 probability for all classes with observation 986Numeri
cal 0 probability for all classes with observation 994Numerical 0 probability for a
ll classes with observation 1004Numerical 0 probability for all classes with observ
ation 1041Numerical 0 probability for all classes with observation 1050Numerical 0
probability for all classes with observation 21Numerical 0 probability for all clas
ses with observation 28Numerical 0 probability for all classes with observation 44N
umerical 0 probability for all classes with observation 73Numerical 0 probability f
or all classes with observation 120Numerical 0 probability for all classes with obs
ervation 126Numerical 0 probability for all classes with observation 212Numerical 0
probability for all classes with observation 267Numerical 0 probability for all cla

probability for all classes with observation 267Numerical 0 probability for all cla
sses with observation 342Numerical 0 probability for all classes with observation 3
55Numerical 0 probability for all classes with observation 378Numerical 0 probabili
ty for all classes with observation 461Numerical 0 probability for all classes with
observation 536Numerical 0 probability for all classes with observation 584Numerica
l 0 probability for all classes with observation 681Numerical 0 probability for all
classes with observation 702Numerical 0 probability for all classes with observatio
n 782Numerical 0 probability for all classes with observation 783Numerical 0 probab
ility for all classes with observation 791Numerical 0 probability for all classes w
ith observation 806Numerical 0 probability for all classes with observation 828Nume
rical 0 probability for all classes with observation 838Numerical 0 probability for
all classes with observation 841Numerical 0 probability for all classes with observ
ation 842Numerical 0 probability for all classes with observation 843Numerical 0 pr
obability for all classes with observation 846Numerical 0 probability for all class
es with observation 849Numerical 0 probability for all classes with observation 855
Numerical 0 probability for all classes with observation 860Numerical 0 probability
for all classes with observation 861Numerical 0 probability for all classes with ob
servation 903Numerical 0 probability for all classes with observation 905Numerical
0 probability for all classes with observation 938Numerical 0 probability for all c
lasses with observation 940Numerical 0 probability for all classes with observation
954Numerical 0 probability for all classes with observation 955Numerical 0 probabil
ity for all classes with observation 959Numerical 0 probability for all classes wit
h observation 960Numerical 0 probability for all classes with observation 962Numeri
cal 0 probability for all classes with observation 963Numerical 0 probability for a
ll classes with observation 979Numerical 0 probability for all classes with observa
tion 981Numerical 0 probability for all classes with observation 984Numerical 0 pro
bability for all classes with observation 985Numerical 0 probability for all classe
s with observation 988Numerical 0 probability for all classes with observation 997N
umerical 0 probability for all classes with observation 1007Numerical 0 probability
for all classes with observation 1008Numerical 0 probability for all classes with o
bservation 1021Numerical 0 probability for all classes with observation 1035Numeric
al 0 probability for all classes with observation 31Numerical 0 probability for all
classes with observation 65Numerical 0 probability for all classes with observation
126Numerical 0 probability for all classes with observation 147Numerical 0 probabil
ity for all classes with observation 176Numerical 0 probability for all classes wit
h observation 256Numerical 0 probability for all classes with observation 323Numeri
cal 0 probability for all classes with observation 396Numerical 0 probability for a
ll classes with observation 468Numerical 0 probability for all classes with observa
tion 549Numerical 0 probability for all classes with observation 555Numerical 0 pro
bability for all classes with observation 602Numerical 0 probability for all classe
s with observation 682Numerical 0 probability for all classes with observation 778N
umerical 0 probability for all classes with observation 782Numerical 0 probability
for all classes with observation 808Numerical 0 probability for all classes with ob
servation 814Numerical 0 probability for all classes with observation 821Numerical
0 probability for all classes with observation 822Numerical 0 probability for all c
lasses with observation 827Numerical 0 probability for all classes with observation
829Numerical 0 probability for all classes with observation 834Numerical 0 probabil
ity for all classes with observation 841Numerical 0 probability for all classes wit
h observation 876Numerical 0 probability for all classes with observation 900Numeri
cal 0 probability for all classes with observation 933Numerical 0 probability for a
ll classes with observation 934Numerical 0 probability for all classes with observa
tion 945Numerical 0 probability for all classes with observation 961Numerical 0 pro
bability for all classes with observation 962Numerical 0 probability for all classe
s with observation 966Numerical 0 probability for all classes with observation 977N
umerical 0 probability for all classes with observation 978Numerical 0 probability
for all classes with observation 981Numerical 0 probability for all classes with ob
servation 985Numerical 0 probability for all classes with observation 986Numerical

0 probability for all classes with observation 988Numerical 0 probability for all c
lasses with observation 989Numerical 0 probability for all classes with observation
993Numerical 0 probability for all classes with observation 996Numerical 0 probabil
ity for all classes with observation 1000Numerical 0 probability for all classes wi
th observation 1001Numerical 0 probability for all classes with observation 1010Num
erical 0 probability for all classes with observation 1013Numerical 0 probability f
or all classes with observation 119Numerical 0 probability for all classes with obs
ervation 125Numerical 0 probability for all classes with observation 127Numerical 0
probability for all classes with observation 208Numerical 0 probability for all cla
sses with observation 216Numerical 0 probability for all classes with observation 2
87Numerical 0 probability for all classes with observation 300Numerical 0 probabili
ty for all classes with observation 313Numerical 0 probability for all classes with
observation 315Numerical 0 probability for all classes with observation 327Numerica
l 0 probability for all classes with observation 350Numerical 0 probability for all
classes with observation 395Numerical 0 probability for all classes with observatio
n 412Numerical 0 probability for all classes with observation 475Numerical 0 probab
ility for all classes with observation 483Numerical 0 probability for all classes w
ith observation 488Numerical 0 probability for all classes with observation 620Nume
rical 0 probability for all classes with observation 752Numerical 0 probability for
all classes with observation 758Numerical 0 probability for all classes with observ
ation 766Numerical 0 probability for all classes with observation 775Numerical 0 pr
obability for all classes with observation 784Numerical 0 probability for all class
es with observation 791Numerical 0 probability for all classes with observation 799
Numerical 0 probability for all classes with observation 800Numerical 0 probability
for all classes with observation 801Numerical 0 probability for all classes with ob
servation 802Numerical 0 probability for all classes with observation 807Numerical
0 probability for all classes with observation 809Numerical 0 probability for all c
lasses with observation 810Numerical 0 probability for all classes with observation
818Numerical 0 probability for all classes with observation 820Numerical 0 probabil
ity for all classes with observation 834Numerical 0 probability for all classes wit
h observation 869Numerical 0 probability for all classes with observation 916Numeri
cal 0 probability for all classes with observation 917Numerical 0 probability for a
ll classes with observation 942Numerical 0 probability for all classes with observa
tion 943Numerical 0 probability for all classes with observation 946Numerical 0 pro
bability for all classes with observation 954Numerical 0 probability for all classe
s with observation 955Numerical 0 probability for all classes with observation 959N
umerical 0 probability for all classes with observation 966Numerical 0 probability
for all classes with observation 967Numerical 0 probability for all classes with ob
servation 968Numerical 0 probability for all classes with observation 971Numerical
0 probability for all classes with observation 988Numerical 0 probability for all c
lasses with observation 991Numerical 0 probability for all classes with observation
993Numerical 0 probability for all classes with observation 999Numerical 0 probabil
ity for all classes with observation 1003Numerical 0 probability for all classes wi
th observation 33Numerical 0 probability for all classes with observation 69Numeric
al 0 probability for all classes with observation 85Numerical 0 probability for all
classes with observation 109Numerical 0 probability for all classes with observatio
n 120Numerical 0 probability for all classes with observation 331Numerical 0 probab
ility for all classes with observation 334Numerical 0 probability for all classes w
ith observation 339Numerical 0 probability for all classes with observation 340Nume
rical 0 probability for all classes with observation 341Numerical 0 probability for
all classes with observation 354Numerical 0 probability for all classes with observ
ation 362Numerical 0 probability for all classes with observation 370Numerical 0 pr
obability for all classes with observation 371Numerical 0 probability for all class
es with observation 530Numerical 0 probability for all classes with observation 575
Numerical 0 probability for all classes with observation 739Numerical 0 probability
for all classes with observation 770Numerical 0 probability for all classes with ob
servation 779Numerical 0 probability for all classes with observation 780Numerical

0 probability for all classes with observation 781Numerical 0 probability for all c
lasses with observation 784Numerical 0 probability for all classes with observation
785Numerical 0 probability for all classes with observation 790Numerical 0 probabil
ity for all classes with observation 793Numerical 0 probability for all classes wit
h observation 797Numerical 0 probability for all classes with observation 802Numeri
cal 0 probability for all classes with observation 803Numerical 0 probability for a
ll classes with observation 812Numerical 0 probability for all classes with observa
tion 817Numerical 0 probability for all classes with observation 823Numerical 0 pro
bability for all classes with observation 824Numerical 0 probability for all classe
s with observation 825Numerical 0 probability for all classes with observation 828N
umerical 0 probability for all classes with observation 835Numerical 0 probability
for all classes with observation 879Numerical 0 probability for all classes with ob
servation 907Numerical 0 probability for all classes with observation 914Numerical
0 probability for all classes with observation 928Numerical 0 probability for all c
lasses with observation 930Numerical 0 probability for all classes with observation
935Numerical 0 probability for all classes with observation 938Numerical 0 probabil
ity for all classes with observation 941Numerical 0 probability for all classes wit
h observation 944Numerical 0 probability for all classes with observation 964Numeri
cal 0 probability for all classes with observation 967Numerical 0 probability for a
ll classes with observation 969Numerical 0 probability for all classes with observa
tion 974Numerical 0 probability for all classes with observation 980Numerical 0 pro
bability for all classes with observation 984Numerical 0 probability for all classe
s with observation 992Numerical 0 probability for all classes with observation 994N
umerical 0 probability for all classes with observation 1016Numerical 0 probability
for all classes with observation 39Numerical 0 probability for all classes with obs
ervation 65Numerical 0 probability for all classes with observation 113Numerical 0
probability for all classes with observation 120Numerical 0 probability for all cla
sses with observation 132Numerical 0 probability for all classes with observation 1
56Numerical 0 probability for all classes with observation 174Numerical 0 probabili
ty for all classes with observation 181Numerical 0 probability for all classes with
observation 224Numerical 0 probability for all classes with observation 279Numerica
l 0 probability for all classes with observation 307Numerical 0 probability for all
classes with observation 308Numerical 0 probability for all classes with observatio
n 343Numerical 0 probability for all classes with observation 378Numerical 0 probab
ility for all classes with observation 379Numerical 0 probability for all classes w
ith observation 380Numerical 0 probability for all classes with observation 458Nume
rical 0 probability for all classes with observation 527Numerical 0 probability for
all classes with observation 593Numerical 0 probability for all classes with observ
ation 621Numerical 0 probability for all classes with observation 643Numerical 0 pr
obability for all classes with observation 704Numerical 0 probability for all class
es with observation 746Numerical 0 probability for all classes with observation 766
Numerical 0 probability for all classes with observation 772Numerical 0 probability
for all classes with observation 784Numerical 0 probability for all classes with ob
servation 789Numerical 0 probability for all classes with observation 798Numerical
0 probability for all classes with observation 801Numerical 0 probability for all c
lasses with observation 808Numerical 0 probability for all classes with observation
817Numerical 0 probability for all classes with observation 818Numerical 0 probabil
ity for all classes with observation 819Numerical 0 probability for all classes wit
h observation 820Numerical 0 probability for all classes with observation 830Numeri
cal 0 probability for all classes with observation 834Numerical 0 probability for a
ll classes with observation 835Numerical 0 probability for all classes with observa
tion 843Numerical 0 probability for all classes with observation 846Numerical 0 pro
bability for all classes with observation 847Numerical 0 probability for all classe
s with observation 855Numerical 0 probability for all classes with observation 881N
umerical 0 probability for all classes with observation 930Numerical 0 probability
for all classes with observation 938Numerical 0 probability for all classes with ob
servation 946Numerical 0 probability for all classes with observation 953Numerical

0 probability for all classes with observation 955Numerical 0 probability for all c
lasses with observation 970Numerical 0 probability for all classes with observation
974Numerical 0 probability for all classes with observation 976Numerical 0 probabil
ity for all classes with observation 982Numerical 0 probability for all classes wit
h observation 987Numerical 0 probability for all classes with observation 992Numeri
cal 0 probability for all classes with observation 996Numerical 0 probability for a
ll classes with observation 999Numerical 0 probability for all classes with observa
tion 1054Numerical 0 probability for all classes with observation 12Numerical 0 pro
bability for all classes with observation 22Numerical 0 probability for all classes
with observation 29Numerical 0 probability for all classes with observation 40Numer
ical 0 probability for all classes with observation 90Numerical 0 probability for a
ll classes with observation 156Numerical 0 probability for all classes with observa
tion 172Numerical 0 probability for all classes with observation 298Numerical 0 pro
bability for all classes with observation 301Numerical 0 probability for all classe
s with observation 310Numerical 0 probability for all classes with observation 311N
umerical 0 probability for all classes with observation 327Numerical 0 probability
for all classes with observation 361Numerical 0 probability for all classes with ob
servation 362Numerical 0 probability for all classes with observation 451Numerical
0 probability for all classes with observation 457Numerical 0 probability for all c
lasses with observation 514Numerical 0 probability for all classes with observation
535Numerical 0 probability for all classes with observation 614Numerical 0 probabil
ity for all classes with observation 625Numerical 0 probability for all classes wit
h observation 627Numerical 0 probability for all classes with observation 702Numeri
cal 0 probability for all classes with observation 721Numerical 0 probability for a
ll classes with observation 737Numerical 0 probability for all classes with observa
tion 741Numerical 0 probability for all classes with observation 751Numerical 0 pro
bability for all classes with observation 754Numerical 0 probability for all classe
s with observation 767Numerical 0 probability for all classes with observation 768N
umerical 0 probability for all classes with observation 784Numerical 0 probability
for all classes with observation 789Numerical 0 probability for all classes with ob
servation 797Numerical 0 probability for all classes with observation 802Numerical
0 probability for all classes with observation 803Numerical 0 probability for all c
lasses with observation 805Numerical 0 probability for all classes with observation
814Numerical 0 probability for all classes with observation 828Numerical 0 probabil
ity for all classes with observation 829Numerical 0 probability for all classes wit
h observation 842Numerical 0 probability for all classes with observation 864Numeri
cal 0 probability for all classes with observation 877Numerical 0 probability for a
ll classes with observation 904Numerical 0 probability for all classes with observa
tion 919Numerical 0 probability for all classes with observation 938Numerical 0 pro
bability for all classes with observation 939Numerical 0 probability for all classe
s with observation 944Numerical 0 probability for all classes with observation 961N
umerical 0 probability for all classes with observation 963Numerical 0 probability
for all classes with observation 972Numerical 0 probability for all classes with ob
servation 973Numerical 0 probability for all classes with observation 974Numerical
0 probability for all classes with observation 990Numerical 0 probability for all c
lasses with observation 60Numerical 0 probability for all classes with observation
62Numerical 0 probability for all classes with observation 78Numerical 0 probabilit
y for all classes with observation 103Numerical 0 probability for all classes with
observation 105Numerical 0 probability for all classes with observation 160Numerica
l 0 probability for all classes with observation 287Numerical 0 probability for all
classes with observation 312Numerical 0 probability for all classes with observatio
n 314Numerical 0 probability for all classes with observation 327Numerical 0 probab
ility for all classes with observation 337Numerical 0 probability for all classes w
ith observation 349Numerical 0 probability for all classes with observation 350Nume
rical 0 probability for all classes with observation 351Numerical 0 probability for
all classes with observation 358Numerical 0 probability for all classes with observ
ation 359Numerical 0 probability for all classes with observation 388Numerical 0 pr

```
ation 339Numerical 0 probability for all classes with observation 366Numerical 0 pr
obability for all classes with observation 392Numerical 0 probability for all class
es with observation 453Numerical 0 probability for all classes with observation 582
Numerical 0 probability for all classes with observation 636Numerical 0 probability
for all classes with observation 769Numerical 0 probability for all classes with ob
servation 782Numerical 0 probability for all classes with observation 789Numerical
0 probability for all classes with observation 799Numerical 0 probability for all c
lasses with observation 800Numerical 0 probability for all classes with observation
801Numerical 0 probability for all classes with observation 802Numerical 0 probabil
ity for all classes with observation 808Numerical 0 probability for all classes wit
h observation 809Numerical 0 probability for all classes with observation 814Numeri
cal 0 probability for all classes with observation 818Numerical 0 probability for a
ll classes with observation 822Numerical 0 probability for all classes with observa
tion 825Numerical 0 probability for all classes with observation 830Numerical 0 pro
bability for all classes with observation 837Numerical 0 probability for all classe
s with observation 861Numerical 0 probability for all classes with observation 876N
umerical 0 probability for all classes with observation 911Numerical 0 probability
for all classes with observation 924Numerical 0 probability for all classes with ob
servation 936Numerical 0 probability for all classes with observation 953Numerical
0 probability for all classes with observation 955Numerical 0 probability for all c
lasses with observation 956Numerical 0 probability for all classes with observation
958Numerical 0 probability for all classes with observation 962Numerical 0 probabil
ity for all classes with observation 975Numerical 0 probability for all classes wit
h observation 981Numerical 0 probability for all classes with observation 1012Numer
ical 0 probability for all classes with observation 1021Numerical 0 probability for
all classes with observation 1046Numerical 0 probability for all classes with obser
vation 1054
```

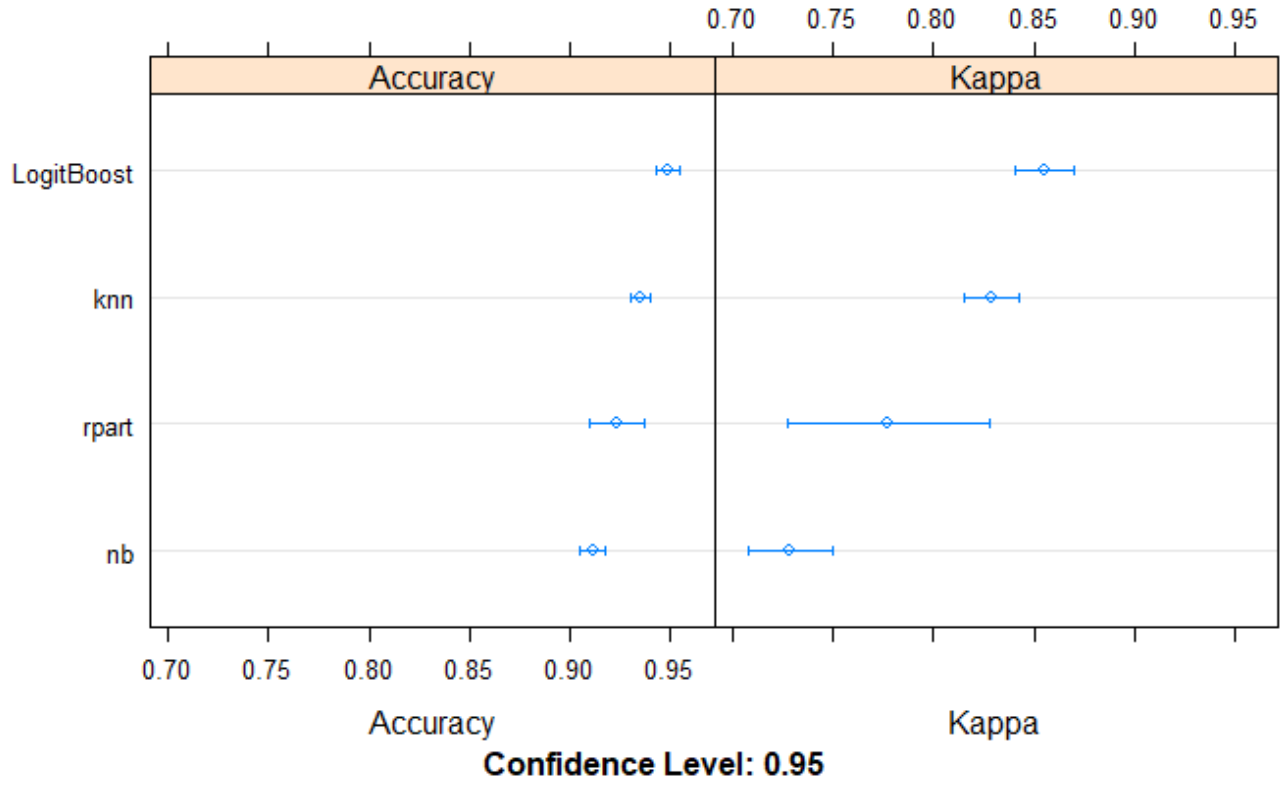
Hide

```
res_mod <- resamples(models)
summary(res_mod)
```

```
Call:
summary.resamples(object = res_mod)

Models: rpart, nb, knn, LogitBoost
Number of resamples: 10

Accuracy
                Min.   1st Qu.    Median      Mean   3rd Qu.      Max. NA's
rpart      0.8996212 0.9059334 0.9232955 0.9232479 0.9403409 0.9480151    0
nb         0.8996212 0.9041193 0.9110252 0.9112274 0.9155232 0.9291115    0
knn        0.9270833 0.9308712 0.9337757 0.9353635 0.9384470 0.9470699    0
LogitBoost 0.9346591 0.9460482 0.9512311 0.9485169 0.9539225 0.9583333    0

Kappa
                Min.   1st Qu.    Median      Mean   3rd Qu.      Max. NA's
rpart      0.6890252 0.7157091 0.7791268 0.7773943 0.8399521 0.8629024    0
nb         0.6899761 0.7075683 0.7279966 0.7287509 0.7424670 0.7876400    0
knn        0.8038784 0.8168684 0.8247881 0.8288628 0.8375766 0.8609522    0
LogitBoost 0.8204294 0.8455835 0.8605602 0.8552726 0.8699733 0.8821237    0
```

Hide

```
dotplot(res_mod)
```



Hide

```
### by looking at the accuaracy values we can see that logistic regression model is best among the all 4 models
### Using the test set for best fit model, after seeing the accuracy and checking the kappa of different models, I came to conclusion that logistic regression model is best with accuaracy of 95% and kappa of 84%
HR_lr_test = glm(left ~ ., family=binomial(logit), data=HR_train) ### glm is used to fit generilized linear moldels
EmployeemayLeave=predict(logreg,newdata=HR_test,type="response") ### to predict which employee may leave in order to retain them
Employeeattrition = data.frame(EmployeemayLeave) ### adding data of possible leaving employees to new data frame
Employeeattrition$performance=HR_test$last_evaluation ###  adding performance as a parameter in the Employeeattrition data frame
plot(Employeeattrition$EmployeemayLeave,Employeeattrition$performance)
```

Hide

```
#### Installing the package DT
library(DT)
```

```
package <U+393C><U+3E31>DT<U+393C><U+3E32> was built under R version 3.4.4
```

Hide

```
Employeeattrition$retain = Employeeattrition$performance*Employeeattrition$Employee
mayLeave
### assigning data obtained form above to new variable so, that the list of employe
es that we should retain should be in order
Employeeretainlist = Employeeattrition[order(Employeeattrition$retain ,decreasing =
TRUE),]
### ordering the list according to performance so, that top performed employee will
be on top of the list
Employeeretainlist <- head(Employeeretainlist, n=500) ### selecting a top 500 emplo
yees according to the performance
datatable(Employeeretainlist)
```

Hide

```
### Datatable function gives us the list of top 500 employees that company should r
etain
```