

Bank_marketing_campaign

Daniel Sakou

Introduction

This project consist in analyzing the direct marketing campaigns to determine if the clients contacted will subscribe to a term deposit. This notebook will demonstrate the different steps: *Exploratory data analysis* Processing the data *Build the models* Select the best performing model

Load the data

```
library(MASS)

## Warning: package 'MASS' was built under R version 4.3.1

library(tidyverse)

## — Attaching core tidyverse packages — tidyverse
## 2.0.0 —
## ✓ dplyr      1.1.2      ✓ readr      2.1.4
## ✓ forcats   1.0.0      ✓ stringr    1.5.0
## ✓ ggplot2    3.4.2      ✓ tibble     3.2.1
## ✓ lubridate  1.9.2      ✓ tidyr      1.3.0
## ✓ purrr      1.0.1
## — Conflicts —
tidyverse_conflicts() —
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()     masks stats::lag()
## ✗ dplyr::select() masks MASS::select()
## ⓘ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
## conflicts to become errors

library(caret)

## Loading required package: lattice
##
## Attaching package: 'caret'
##
## The following object is masked from 'package:purrr':
##
##     lift
```

```
library(stringr)
library(forcats)
library(GGally)

## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg      ggplot2

library(dplyr)
library(pROC)

## Type 'citation("pROC")' for a citation.
##
## Attaching package: 'pROC'
##
## The following objects are masked from 'package:stats':
##
##   cov, smooth, var

library(RANN)

## Warning: package 'RANN' was built under R version 4.3.1

library(rpart)
library(rpart.plot)
library(lubridate)
library(lattice)
library(randomForest)

## randomForest 4.7-1.1
## Type rfNews() to see new features/changes/bug fixes.
##
## Attaching package: 'randomForest'
##
## The following object is masked from 'package:dplyr':
##
##   combine
##
## The following object is masked from 'package:ggplot2':
##
##   margin

library(ranger)

##
## Attaching package: 'ranger'
##
## The following object is masked from 'package:randomForest':
##
##   importance

library(glmnet)
```

```
## Warning: package 'glmnet' was built under R version 4.3.1

## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
##
## Loaded glmnet 4.1-7

library(e1071)
library(knitr)
library(MASS)
library(missRanger)

## Warning: package 'missRanger' was built under R version 4.3.1

library(missMDA)

## Warning: package 'missMDA' was built under R version 4.3.1

library(class)
set.seed(1243)
```

Upload the data

Explore the data

```
view(bank_Marketing)
head(bank_Marketing)
```

```
##   age      job marital  education default housing loan  contact month
## 1  56 housemaid married  basic.4y      no      no  no telephone  may
## 2  57 services married high.school unknown      no  no telephone  may
## 3  37 services married high.school      no  yes  no telephone  may
## 4  40 admin. married  basic.6y      no      no  no telephone  may
## 5  56 services married high.school      no      no  yes telephone  may
## 6  45 services married  basic.9y unknown      no  no telephone  may
##   day_of_week duration campaign pdays previous  poutcome emp.var.rate
## 1      mon      261         1    999         0 nonexistent         1.1
## 2      mon      149         1    999         0 nonexistent         1.1
## 3      mon      226         1    999         0 nonexistent         1.1
## 4      mon      151         1    999         0 nonexistent         1.1
## 5      mon      307         1    999         0 nonexistent         1.1
## 6      mon      198         1    999         0 nonexistent         1.1
##   cons.price.idx cons.conf.idx euribor3m nr.employed  y
## 1      93.994      -36.4      4.857      5191 no
## 2      93.994      -36.4      4.857      5191 no
## 3      93.994      -36.4      4.857      5191 no
## 4      93.994      -36.4      4.857      5191 no
```

```
## 5      93.994      -36.4      4.857      5191 no
## 6      93.994      -36.4      4.857      5191 no

tail(bank_Marketing)

##      age      job marital      education default housing loan
contact
## 41183  29  unemployed  single      basic.4y      no      yes  no
cellular
## 41184  73    retired married professional.course      no      yes  no
cellular
## 41185  46 blue-collar married professional.course      no      no  no
cellular
## 41186  56    retired married  university.degree      no      yes  no
cellular
## 41187  44  technician married professional.course      no      no  no
cellular
## 41188  74    retired married professional.course      no      yes  no
cellular
##      month day_of_week duration campaign pdays previous      poutcome
## 41183  nov      fri      112      1      9      1      success
## 41184  nov      fri      334      1     999      0 nonexistent
## 41185  nov      fri      383      1     999      0 nonexistent
## 41186  nov      fri      189      2     999      0 nonexistent
## 41187  nov      fri      442      1     999      0 nonexistent
## 41188  nov      fri      239      3     999      1      failure
##      emp.var.rate cons.price.idx cons.conf.idx euribor3m nr.employed  y
## 41183      -1.1      94.767      -50.8      1.028      4963.6  no
## 41184      -1.1      94.767      -50.8      1.028      4963.6 yes
## 41185      -1.1      94.767      -50.8      1.028      4963.6  no
## 41186      -1.1      94.767      -50.8      1.028      4963.6  no
## 41187      -1.1      94.767      -50.8      1.028      4963.6 yes
## 41188      -1.1      94.767      -50.8      1.028      4963.6  no

dim(bank_Marketing)

## [1] 41188      21

colnames(bank_Marketing)

## [1] "age"      "job"      "marital"  "education"
## [5] "default"  "housing"  "loan"     "contact"
## [9] "month"    "day_of_week" "duration" "campaign"
## [13] "pdays"   "previous"  "poutcome" "emp.var.rate"
## [17] "cons.price.idx" "cons.conf.idx" "euribor3m" "nr.employed"
## [21] "y"

sum(is.na(bank_Marketing))

## [1] 0
```

Change the outcome variable name

```
colnames(bank_Marketing)[colnames(bank_Marketing)== "y"] = "deposit"  
head(bank_Marketing)
```

```
##   age      job marital  education default housing loan  contact month  
## 1  56 housemaid married  basic.4y      no      no  no telephone  may  
## 2  57 services married high.school unknown      no  no  no telephone  may  
## 3  37 services married high.school      no  yes  no telephone  may  
## 4  40 admin. married  basic.6y      no      no  no telephone  may  
## 5  56 services married high.school      no      no  yes telephone  may  
## 6  45 services married  basic.9y unknown      no  no  no telephone  may  
##   day_of_week duration campaign pdays previous  poutcome emp.var.rate  
## 1         mon        261         1    999         0 nonexistent        1.1  
## 2         mon        149         1    999         0 nonexistent        1.1  
## 3         mon        226         1    999         0 nonexistent        1.1  
## 4         mon        151         1    999         0 nonexistent        1.1  
## 5         mon        307         1    999         0 nonexistent        1.1  
## 6         mon        198         1    999         0 nonexistent        1.1  
##   cons.price.idx cons.conf.idx euribor3m nr.employed deposit  
## 1          93.994        -36.4     4.857         5191      no  
## 2          93.994        -36.4     4.857         5191      no  
## 3          93.994        -36.4     4.857         5191      no  
## 4          93.994        -36.4     4.857         5191      no  
## 5          93.994        -36.4     4.857         5191      no  
## 6          93.994        -36.4     4.857         5191      no
```

Explore the data step 1

```
str(bank_Marketing)  
  
## 'data.frame':    41188 obs. of  21 variables:  
## $ age           : int  56 57 37 40 56 45 59 41 24 25 ...  
## $ job           : chr  "housemaid" "services" "services" "admin." ...  
## $ marital       : chr  "married" "married" "married" "married" ...  
## $ education     : chr  "basic.4y" "high.school" "high.school" "basic.6y"  
## ...  
## $ default       : chr  "no" "unknown" "no" "no" ...  
## $ housing       : chr  "no" "no" "yes" "no" ...  
## $ loan          : chr  "no" "no" "no" "no" ...  
## $ contact       : chr  "telephone" "telephone" "telephone" "telephone"  
## ...  
## $ month         : chr  "may" "may" "may" "may" ...  
## $ day_of_week   : chr  "mon" "mon" "mon" "mon" ...  
## $ duration      : int  261 149 226 151 307 198 139 217 380 50 ...  
## $ campaign      : int  1 1 1 1 1 1 1 1 1 1 ...  
## $ pdays        : int  999 999 999 999 999 999 999 999 999 999 ...  
## $ previous      : int  0 0 0 0 0 0 0 0 0 0 ...  
## $ poutcome      : chr  "nonexistent" "nonexistent" "nonexistent"  
## "nonexistent" ...  
## $ emp.var.rate  : num  1.1 1.1 1.1 1.1 1.1 1.1 1.1 1.1 1.1 1.1 ...  
## $ cons.price.idx: num  94 94 94 94 94 ...
```

```

## $ cons.conf.idx : num -36.4 -36.4 -36.4 -36.4 -36.4 -36.4 -36.4 -36.4 -
36.4 -36.4 ...
## $ euribor3m      : num  4.86 4.86 4.86 4.86 4.86 ...
## $ nr.employed    : num  5191 5191 5191 5191 5191 ...
## $ deposit        : chr   "no" "no" "no" "no" ...

bank_Marketing %>% glimpse()

## Rows: 41,188
## Columns: 21
## $ age           <int> 56, 57, 37, 40, 56, 45, 59, 41, 24, 25, 41, 25, 29,
57,...
## $ job           <chr> "housemaid", "services", "services", "admin.",
"service...
## $ marital       <chr> "married", "married", "married", "married",
"married", ...
## $ education     <chr> "basic.4y", "high.school", "high.school",
"basic.6y", "...
## $ default       <chr> "no", "unknown", "no", "no", "no", "unknown", "no",
"un...
## $ housing       <chr> "no", "no", "yes", "no", "no", "no", "no", "no",
"yes",...
## $ loan          <chr> "no", "no", "no", "no", "yes", "no", "no", "no",
"no", ...
## $ contact       <chr> "telephone", "telephone", "telephone", "telephone",
"te...
## $ month         <chr> "may", "may", "may", "may", "may", "may", "may",
"may",...
## $ day_of_week   <chr> "mon", "mon", "mon", "mon", "mon", "mon", "mon",
"mon",...
## $ duration      <int> 261, 149, 226, 151, 307, 198, 139, 217, 380, 50,
55, 22...
## $ campaign      <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1...
## $ pdays         <int> 999, 999, 999, 999, 999, 999, 999, 999, 999, 999,
999, ...
## $ previous      <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0...
## $ poutcome      <chr> "nonexistent", "nonexistent", "nonexistent",
"nonexiste...
## $ emp.var.rate   <dbl> 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1,
1.1, ...
## $ cons.price.idx <dbl> 93.994, 93.994, 93.994, 93.994, 93.994, 93.994,
93.994,...
## $ cons.conf.idx  <dbl> -36.4, -36.4, -36.4, -36.4, -36.4, -36.4, -36.4, -
36.4,...
## $ euribor3m      <dbl> 4.857, 4.857, 4.857, 4.857, 4.857, 4.857, 4.857,
4.857,...
## $ nr.employed    <dbl> 5191, 5191, 5191, 5191, 5191, 5191, 5191, 5191,
5191, 5...

```

```
## $ deposit      <chr> "no", "no", "no", "no", "no", "no", "no", "no",  
"no", "...
```

Explore the data step 2

```
summary(bank_Marketing)
```

```
##      age      job      marital      education  
## Min.   :17.00  Length:41188  Length:41188  Length:41188  
## 1st Qu.:32.00  Class :character  Class :character  Class :character  
## Median :38.00  Mode  :character  Mode  :character  Mode  :character  
## Mean   :40.02  
## 3rd Qu.:47.00  
## Max.   :98.00  
##      default      housing      loan      contact  
## Length:41188  Length:41188  Length:41188  Length:41188  
## Class :character  Class :character  Class :character  Class :character  
## Mode  :character  Mode  :character  Mode  :character  Mode  :character  
##  
##  
##      month      day_of_week      duration      campaign  
## Length:41188  Length:41188  Min.   : 0.0  Min.   : 1.000  
## Class :character  Class :character  1st Qu.:102.0  1st Qu.: 1.000  
## Mode  :character  Mode  :character  Median :180.0  Median : 2.000  
##                               Mean   :258.3  Mean   : 2.568  
##                               3rd Qu.:319.0  3rd Qu.: 3.000  
##                               Max.   :4918.0  Max.   :56.000  
##      pdays      previous      poutcome      emp.var.rate  
## Min.   : 0.0  Min.   :0.000  Length:41188  Min.   : -3.40000  
## 1st Qu.:999.0  1st Qu.:0.000  Class :character  1st Qu.: -1.80000  
## Median :999.0  Median :0.000  Mode  :character  Median : 1.10000  
## Mean   :962.5  Mean   :0.173                               Mean   : 0.08189  
## 3rd Qu.:999.0  3rd Qu.:0.000                               3rd Qu.: 1.40000  
## Max.   :999.0  Max.   :7.000                               Max.   : 1.40000  
## cons.price.idx  cons.conf.idx      euribor3m      nr.employed  
## Min.   :92.20  Min.   : -50.8  Min.   :0.634  Min.   :4964  
## 1st Qu.:93.08  1st Qu.: -42.7  1st Qu.:1.344  1st Qu.:5099  
## Median :93.75  Median : -41.8  Median :4.857  Median :5191  
## Mean   :93.58  Mean   : -40.5  Mean   :3.621  Mean   :5167  
## 3rd Qu.:93.99  3rd Qu.: -36.4  3rd Qu.:4.961  3rd Qu.:5228  
## Max.   :94.77  Max.   : -26.9  Max.   :5.045  Max.   :5228  
##      deposit  
## Length:41188  
## Class :character  
## Mode  :character  
##  
##  
##
```

Counts of the the outcome

```
datanew <-table(bank_Marketing$deposit)
datanew

##
##      no      yes
## 36548  4640
```

Proportion of the outcome

```
prop.table(datanew )

##
##           no           yes
## 0.8873458 0.1126542
```

Evaluation of the duration

```
bank_Marketing %>% group_by(day_of_week) %>%
  summarize(median = median(duration),
            IQR= IQR(duration),
            n= n())

## # A tibble: 5 × 4
##   day_of_week median    IQR      n
##   <chr>         <dbl> <dbl> <int>
## 1 fri           172    221  7827
## 2 mon           173    207  8514
## 3 thu           183    228  8623
## 4 tue           186.    208  8090
## 5 wed           185    221  8134
```

Evaluation of the nr.employed

```
bank_Marketing %>% group_by(day_of_week) %>%
  summarize(median = median(nr.employed),
            IQR= IQR(nr.employed),
            n= n())

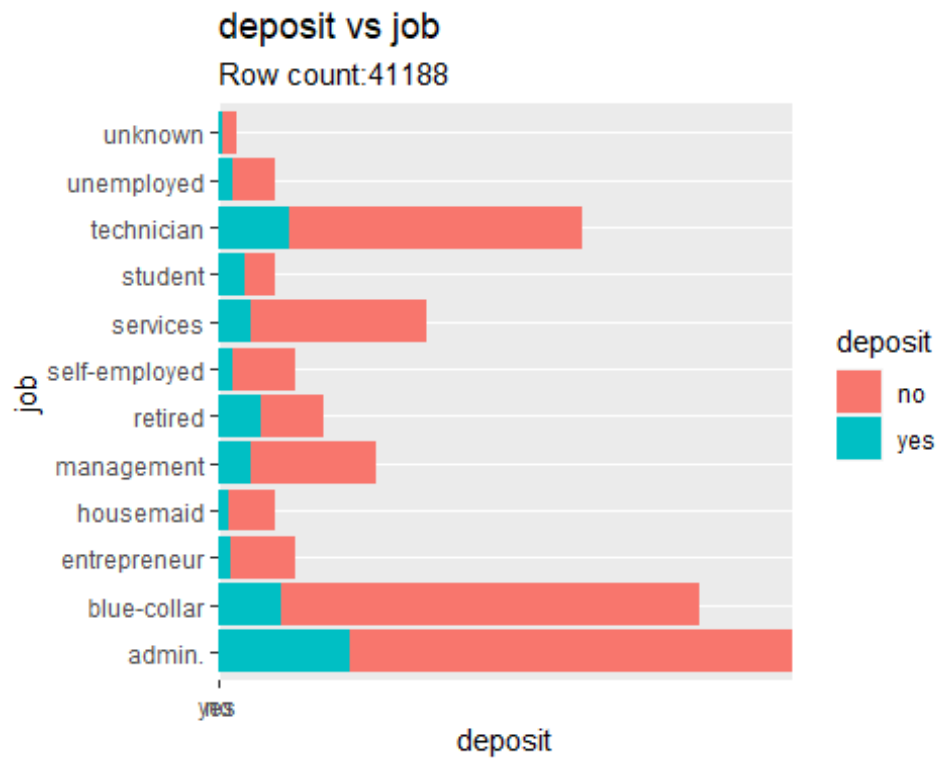
## # A tibble: 5 × 4
##   day_of_week median    IQR      n
##   <chr>         <dbl> <dbl> <int>
## 1 fri          5191    129  7827
## 2 mon          5191    129  8514
## 3 thu          5196.    129  8623
## 4 tue          5191    129  8090
## 5 wed          5191    129  8134
```

Visualize the outcome and the job variable

```
bank_Marketing %>%
  ggplot(aes(x = job, y= deposit, fill = deposit)) +
  geom_col() +
  coord_flip() +
```



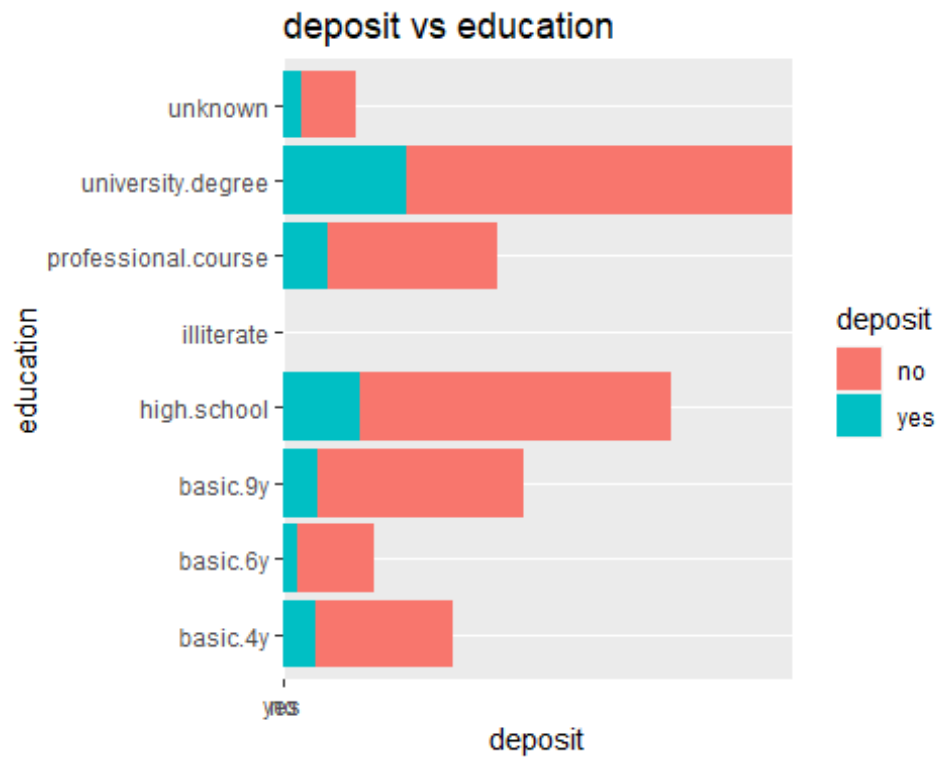
```
ggtitle("deposit vs job",
        subtitle = str_c("Row count:", bank_Marketing %>% NROW()))
```



Visualize the outcome and the education variable

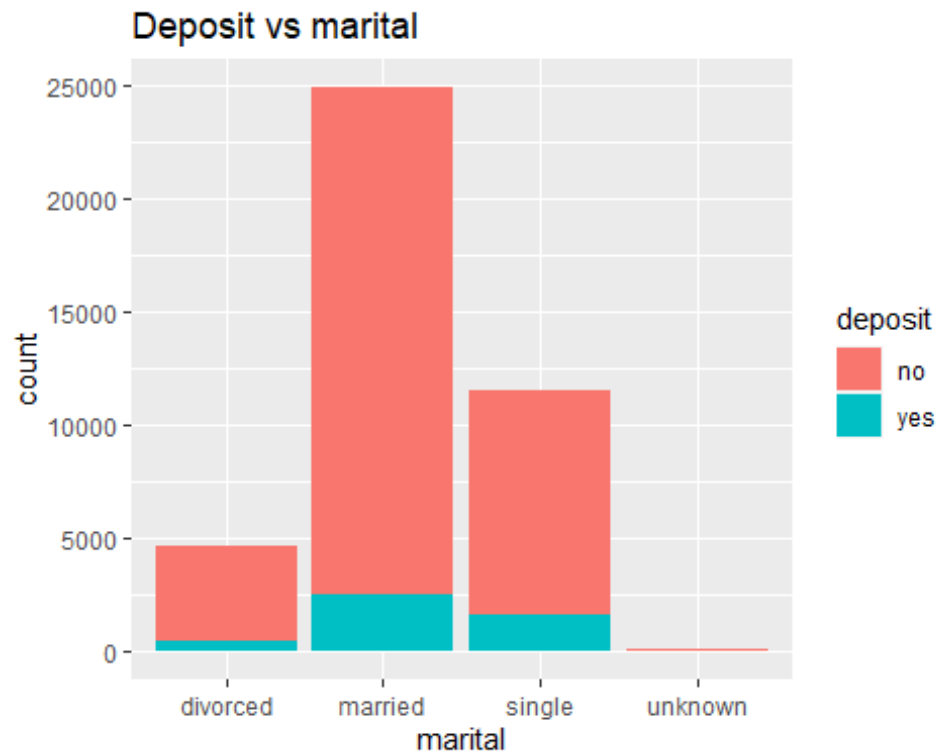
```
bank_Marketing %>%
ggplot(aes(x = education, y= deposit, fill = deposit)) +
  geom_col() +
  # facet_wrap(~ month) +
  coord_flip() +

ggtitle("deposit vs education")
```



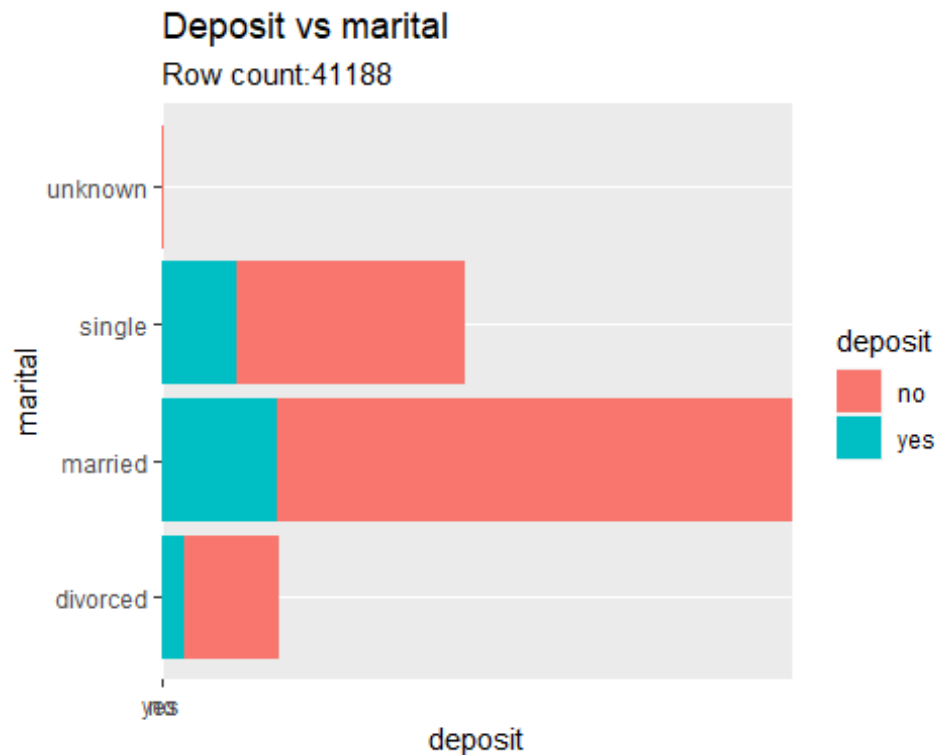
Visualize the outcome and the marital status variable

```
bank_Marketing %>%
  ggplot(aes(x = marital, fill = deposit)) +
  geom_bar() +
  ggtitle("Deposit vs marital")
```



Visualize the outcome and other predictors

```
bank_Marketing %>%  
ggplot(aes(x = marital, y = deposit, fill = deposit)) +  
geom_col() +  
coord_flip() +  
ggtitle("Deposit vs marital",  
        subtitle = str_c("Row count:", bank_Marketing %>% NROW()))
```



```
plot_scatter<-function(data, x_col, y_col,fill = y_col){
plt<-data %>% ggplot(mapping = aes_string(x = x_col, y = y_col, fill = y_col
))+
geom_col() +
coord_flip() +
ggtitle(str_c(y_col, "vs" ,x_col),
        subtitle = str_c("Row count:", data %>% NROW()))
plt %>% print()
}
```

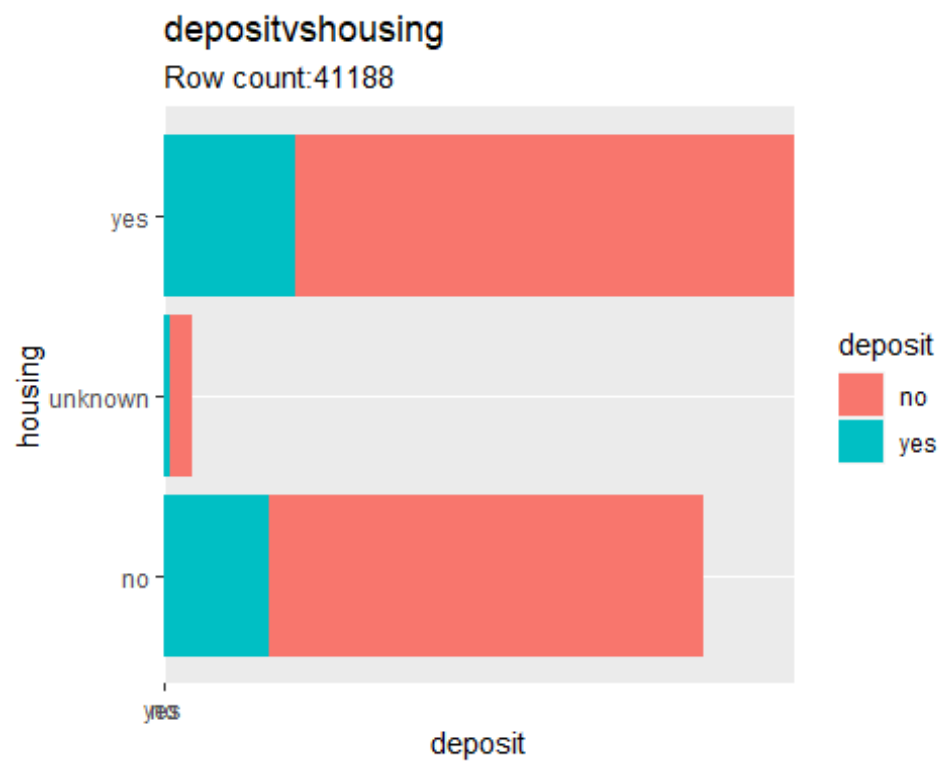
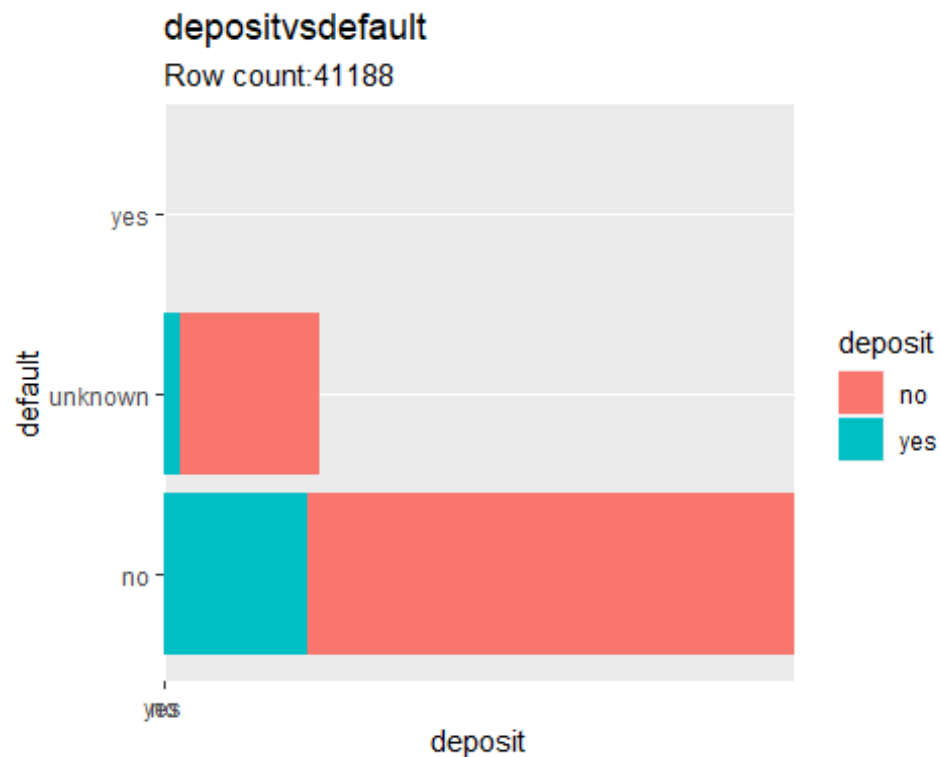
#Define an x_cols

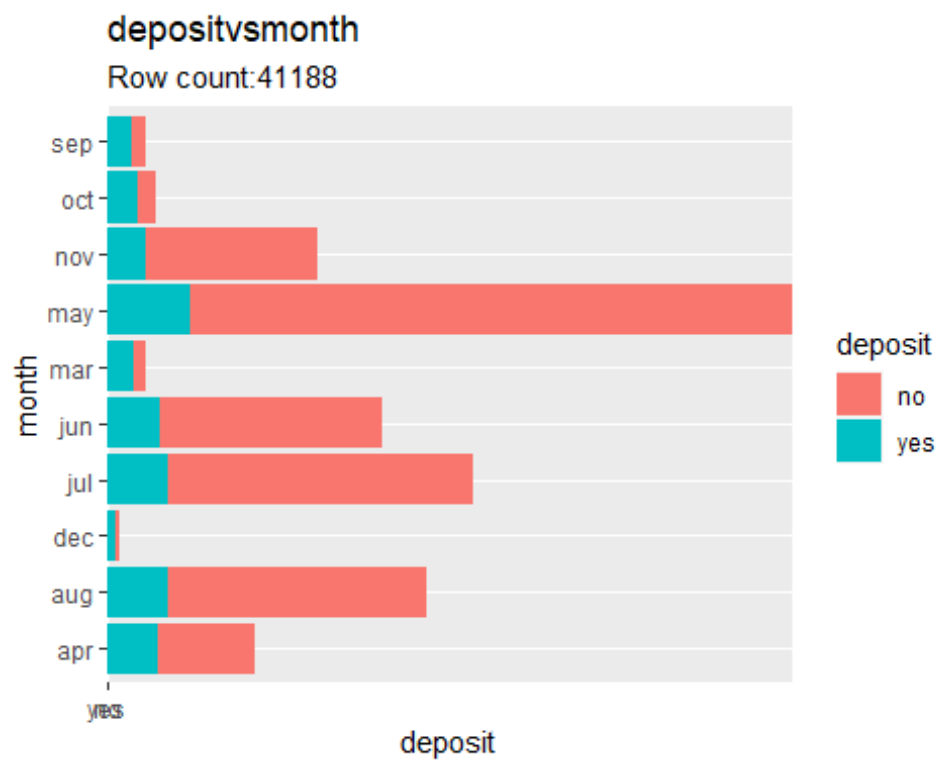
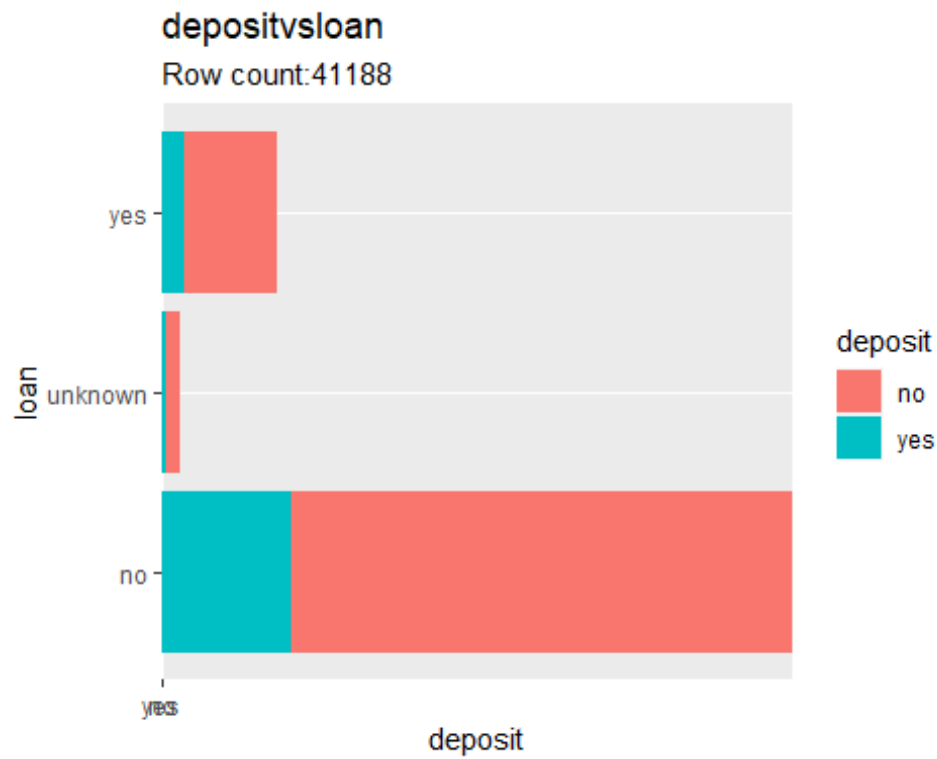
```
x_col <-c("default","housing","loan",
"month","poutcome", "contact")
```

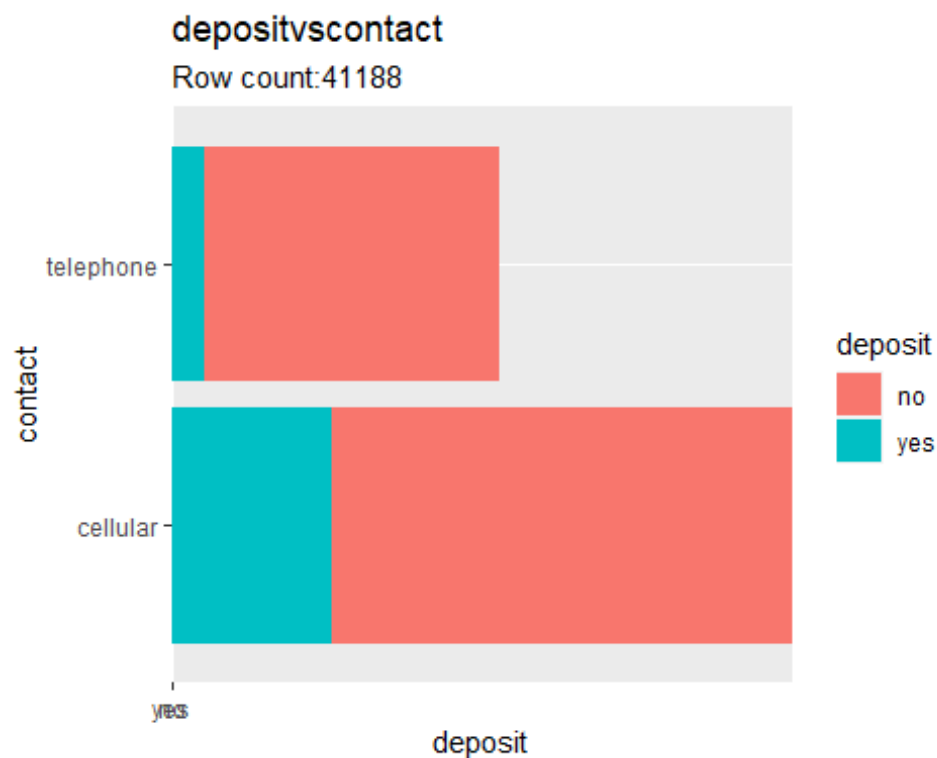
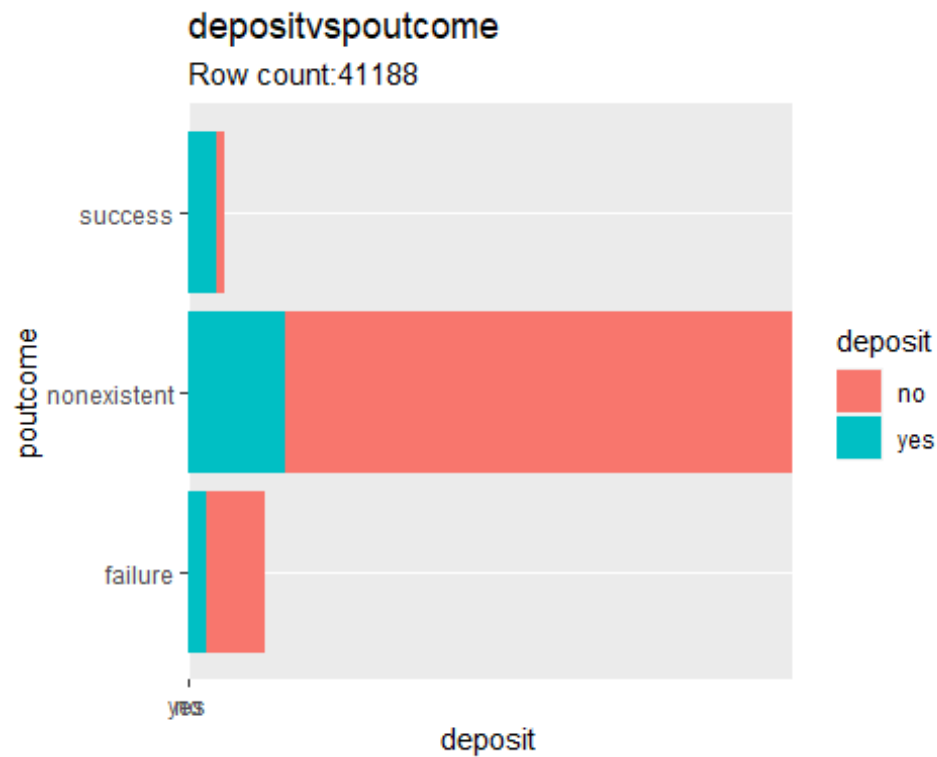
```
x_col %>% walk(plot_scatter, data = bank_Marketing, y_col = "deposit")
```

```
## Warning: `aes_string()` was deprecated in ggplot2 3.0.0.
## i Please use tidy evaluation idioms with `aes()`.
## i See also `vignette("ggplot2-in-packages")` for more information.
## This warning is displayed once every 8 hours.
```

```
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was  
## generated.
```



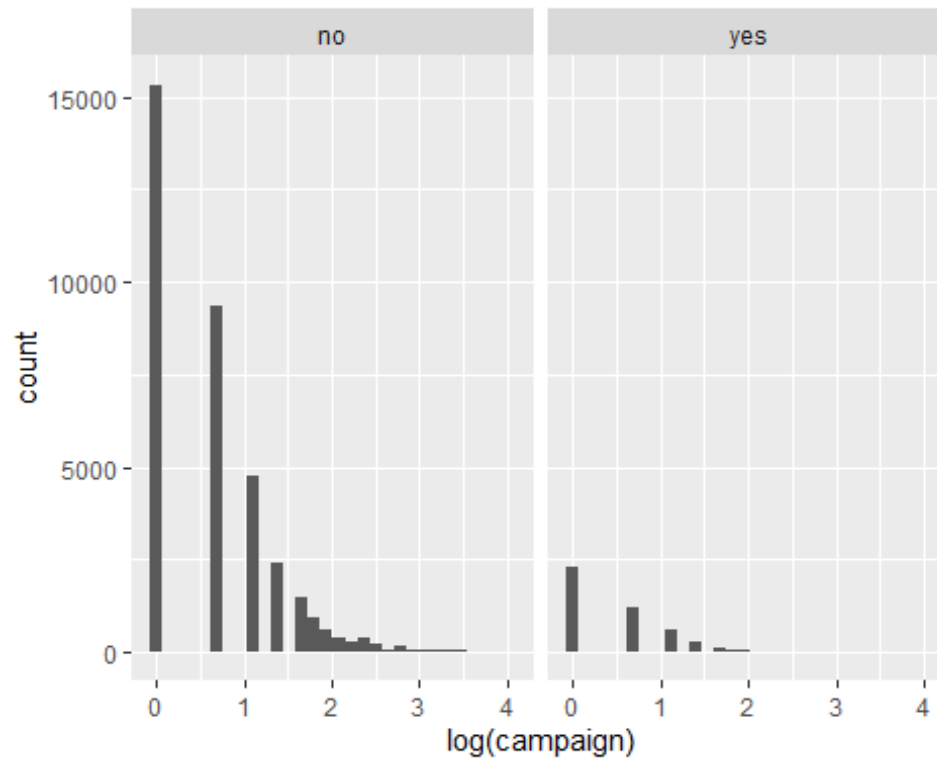




Explore the numeric variable campaign

```
ggplot(bank_Marketing, aes(x = log(campaign))) +  
  geom_histogram() +  
  facet_wrap(~deposit)
```

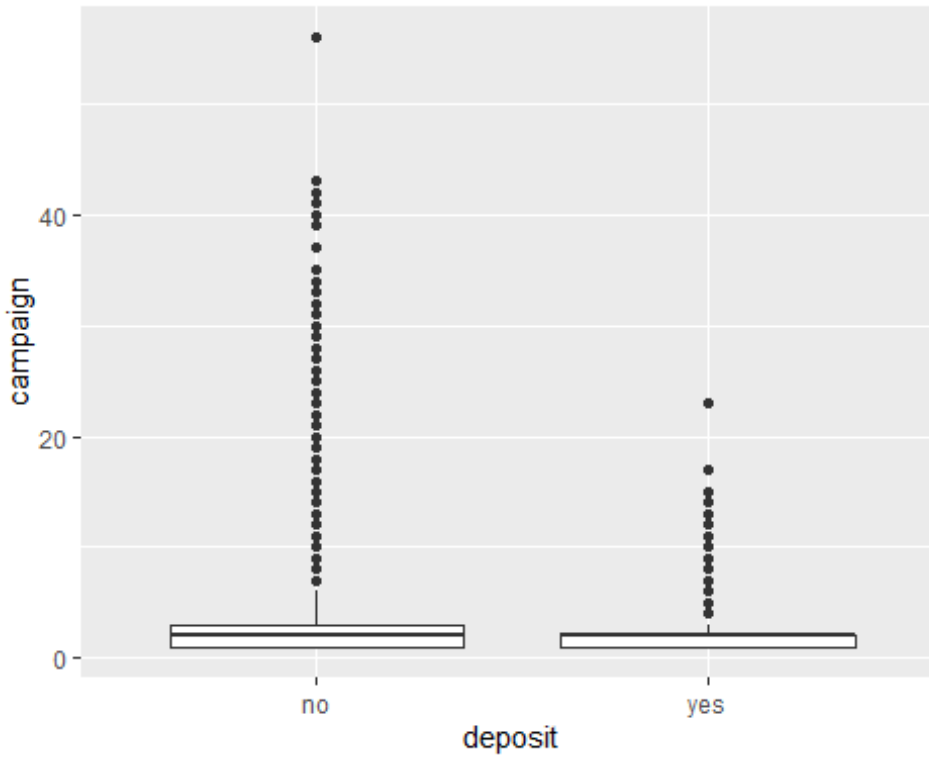
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
labs( title = "campaign")  
## $title  
## [1] "campaign"  
##  
## attr(,"class")  
## [1] "labels"
```

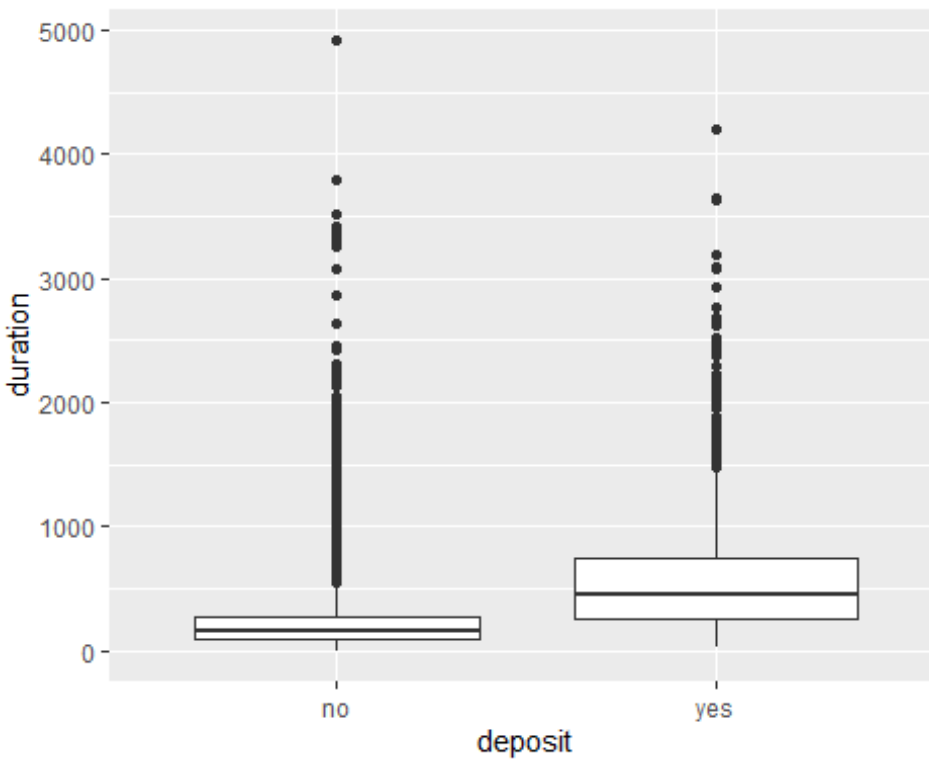
Explore the outcome variable and the campaign variable

```
ggplot(bank_Marketing, aes(x = deposit, y = campaign ))+  
  geom_boxplot()
```

Check the duration variable with the outcome

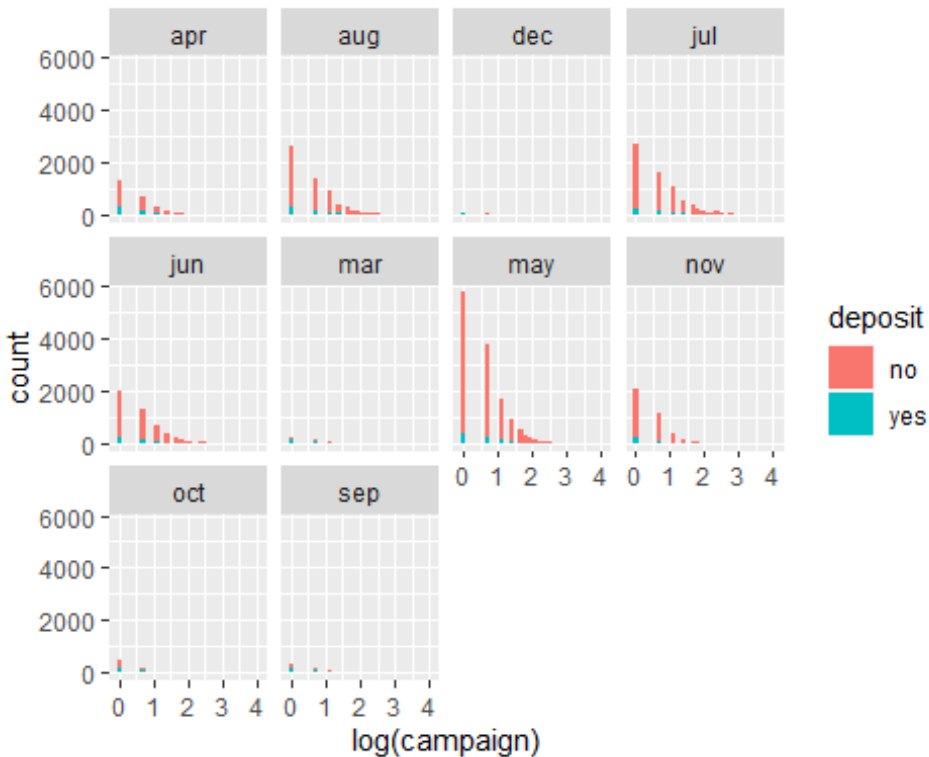
```
ggplot(bank_Marketing, aes(x = deposit, y = duration ))+  
  geom_boxplot()
```



Explore the campaign data and the outcome

```
ggplot(bank_Marketing, aes(x = log(campaign), fill = deposit ))+  
  geom_histogram() +  
  facet_wrap(~ month)
```

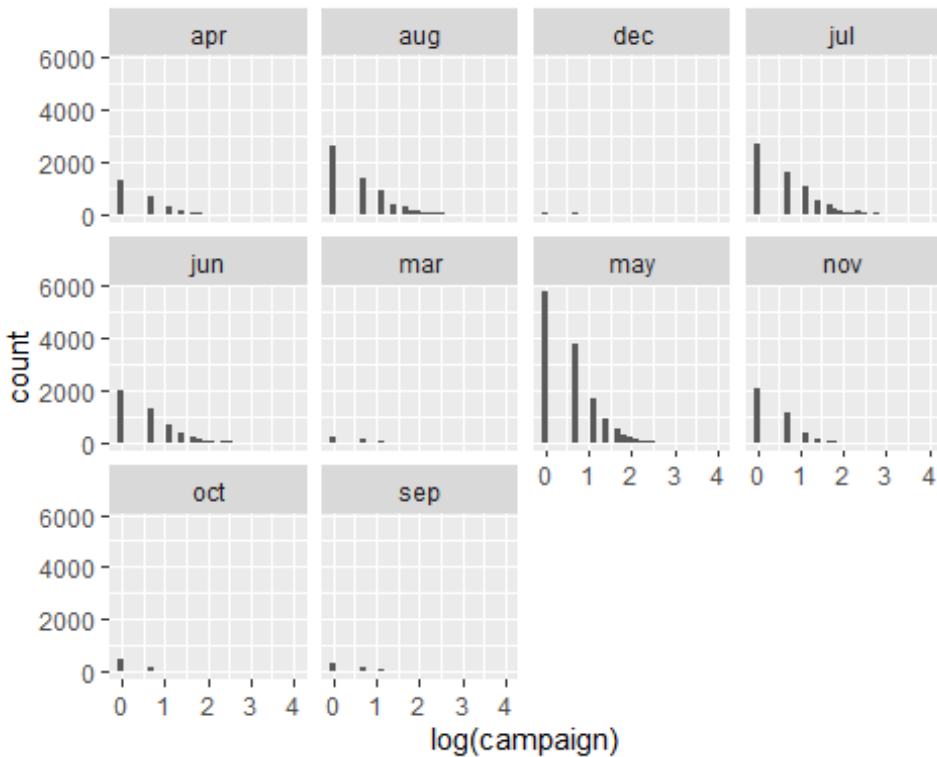
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.



Explore the campaign variable by month

```
ggplot(bank_Marketing, aes(x = log(campaign)))+  
  geom_histogram() +  
  facet_wrap(~ month)
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.



Create a new featute age category variable

```
bank_Marketing$age_category = ifelse(bank_Marketing$age %in% c(0,14)
,"children",
                                     ifelse(bank_Marketing$age
%in%c(15,24),"youth",
                                     ifelse(bank_Marketing$age %in%c(25,64),
"Adults",
                                     "senior")))
```

```
bank_Marketing$age_category = factor(bank_Marketing$age_category)
head(bank_Marketing$age_category)
```

```
## [1] senior senior senior senior senior senior
## Levels: Adults senior youth
```

```
table(bank_Marketing$age_category)
```

```
##
## Adults senior youth
##    655 40070  463
```

```
head(bank_Marketing)
```

```
##   age      job marital  education default housing loan  contact month
## 1  56 housemaid married  basic.4y      no      no  no telephone  may
## 2  57 services married high.school unknown    no  no telephone  may
## 3  37 services married high.school      no    yes  no telephone  may
```

```
## 4 40 admin. married basic.6y no no no telephone may
## 5 56 services married high.school no no yes telephone may
## 6 45 services married basic.9y unknown no no telephone may
## day_of_week duration campaign pdays previous poutcome emp.var.rate
## 1 mon 261 1 999 0 nonexistent 1.1
## 2 mon 149 1 999 0 nonexistent 1.1
## 3 mon 226 1 999 0 nonexistent 1.1
## 4 mon 151 1 999 0 nonexistent 1.1
## 5 mon 307 1 999 0 nonexistent 1.1
## 6 mon 198 1 999 0 nonexistent 1.1
## cons.price.idx cons.conf.idx euribor3m nr.employed deposit age_category
## 1 93.994 -36.4 4.857 5191 no senior
## 2 93.994 -36.4 4.857 5191 no senior
## 3 93.994 -36.4 4.857 5191 no senior
## 4 93.994 -36.4 4.857 5191 no senior
## 5 93.994 -36.4 4.857 5191 no senior
## 6 93.994 -36.4 4.857 5191 no senior
```

```
summary(bank_Marketing$age_category)
```

```
## Adults senior youth
## 655 40070 463
```

```
glimpse(bank_Marketing)
```

```
## Rows: 41,188
## Columns: 22
## $ age <int> 56, 57, 37, 40, 56, 45, 59, 41, 24, 25, 41, 25, 29,
57,...
## $ job <chr> "housemaid", "services", "services", "admin.",
"service...
## $ marital <chr> "married", "married", "married", "married",
"married", ...
## $ education <chr> "basic.4y", "high.school", "high.school",
"basic.6y", "...
## $ default <chr> "no", "unknown", "no", "no", "no", "unknown", "no",
"un...
## $ housing <chr> "no", "no", "yes", "no", "no", "no", "no", "no",
"yes",...
## $ loan <chr> "no", "no", "no", "no", "yes", "no", "no", "no",
"no", ...
## $ contact <chr> "telephone", "telephone", "telephone", "telephone",
"te...
## $ month <chr> "may", "may", "may", "may", "may", "may", "may",
"may",...
## $ day_of_week <chr> "mon", "mon", "mon", "mon", "mon", "mon", "mon",
"mon",...
## $ duration <int> 261, 149, 226, 151, 307, 198, 139, 217, 380, 50,
55, 22...
## $ campaign <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1...
```

```
## $ pdays          <int> 999, 999, 999, 999, 999, 999, 999, 999, 999, 999,
999, ...
## $ previous        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0...
## $ poutcome        <chr> "nonexistent", "nonexistent", "nonexistent",
"nonexiste...
## $ emp.var.rate     <dbl> 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1,
1.1, ...
## $ cons.price.idx   <dbl> 93.994, 93.994, 93.994, 93.994, 93.994, 93.994,
93.994,...
## $ cons.conf.idx    <dbl> -36.4, -36.4, -36.4, -36.4, -36.4, -36.4, -36.4, -
36.4,...
## $ euribor3m        <dbl> 4.857, 4.857, 4.857, 4.857, 4.857, 4.857, 4.857,
4.857,...
## $ nr.employed      <dbl> 5191, 5191, 5191, 5191, 5191, 5191, 5191, 5191,
5191, 5...
## $ deposit          <chr> "no", "no", "no", "no", "no", "no", "no", "no",
"no", "...
## $ age_category     <fct> senior, senior, senior, senior, senior, senior,
senior,...
```

#senior are the most interviewed for this campaign

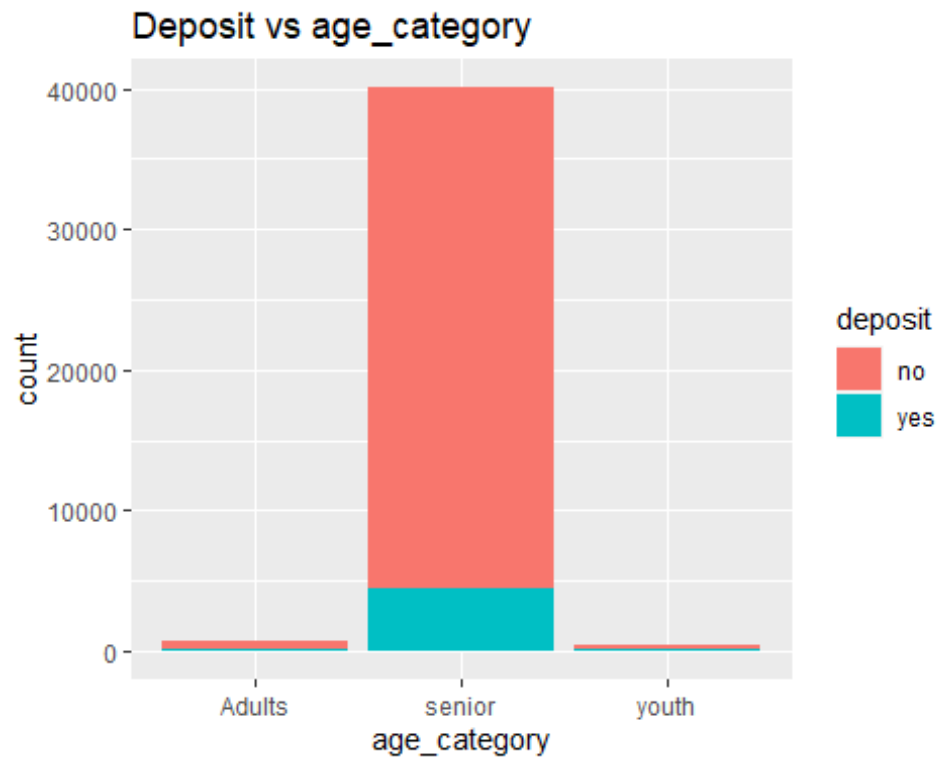
Age_category counts

```
table(bank_Marketing$age_category)

##
## Adults senior  youth
##    655  40070    463
```

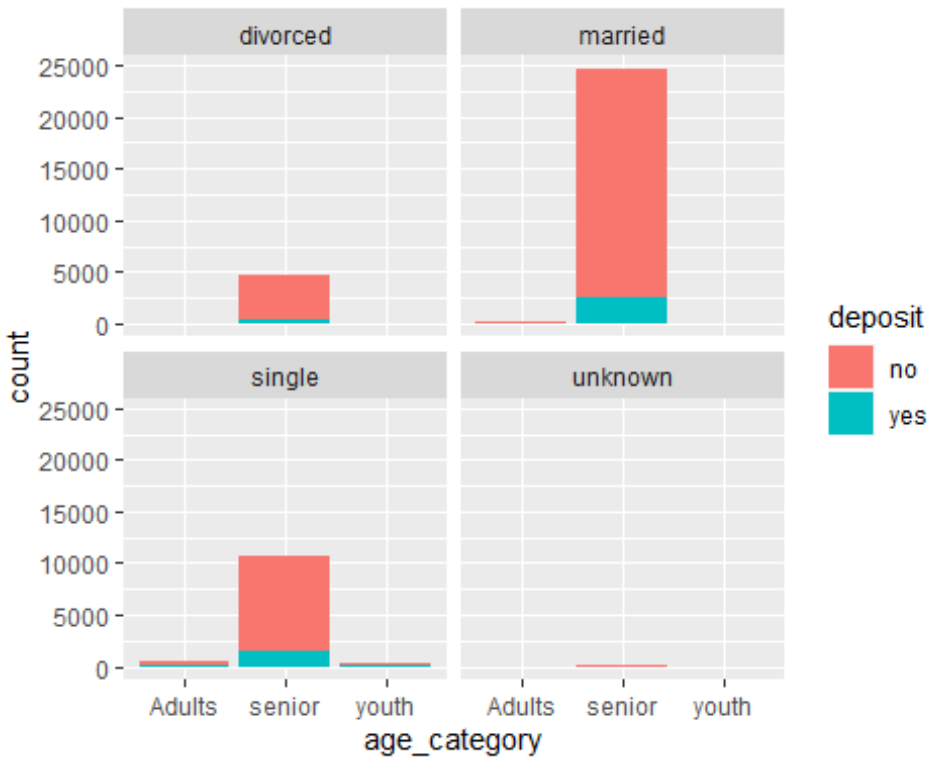
Visualize the age category variable with the outcome

```
bank_Marketing %>%
ggplot(aes(x = age_category, fill = deposit)) +
geom_bar() +
ggtitle("Deposit vs age_category")
```



Visualize the new feature with the outcome variable and marital status

```
bank_Marketing %>%  
ggplot(aes(x = age_category, fill = deposit)) +  
geom_bar() +  
facet_wrap(~marital )
```



```
ggtitle("Deposit vs age_category")
```

```
## $title
## [1] "Deposit vs age_category"
##
## attr(,"class")
## [1] "labels"
```

Dependant variable data frame

```
deposit_y <- bank_Marketing$deposit
head(deposit_y)
```

```
## [1] "no" "no" "no" "no" "no" "no"
```

```
head(bank_Marketing)
```

```
##   age      job marital  education default housing loan  contact month
## 1  56 housemaid married  basic.4y    no      no   no telephone  may
## 2  57 services married high.school unknown    no   no   no telephone  may
## 3  37 services married high.school    no    yes   no telephone  may
## 4  40 admin. married  basic.6y    no      no   no telephone  may
## 5  56 services married high.school    no      no  yes telephone  may
## 6  45 services married  basic.9y unknown    no   no   no telephone  may
##   day_of_week duration campaign pdays previous  poutcome emp.var.rate
## 1         mon      261         1    999         0 nonexistent         1.1
## 2         mon      149         1    999         0 nonexistent         1.1
## 3         mon      226         1    999         0 nonexistent         1.1
```

```
## 4      mon      151      1  999      0 nonexistent      1.1
## 5      mon      307      1  999      0 nonexistent      1.1
## 6      mon      198      1  999      0 nonexistent      1.1
##   cons.price.idx cons.conf.idx euribor3m nr.employed deposit age_category
## 1      93.994      -36.4      4.857      5191      no      senior
## 2      93.994      -36.4      4.857      5191      no      senior
## 3      93.994      -36.4      4.857      5191      no      senior
## 4      93.994      -36.4      4.857      5191      no      senior
## 5      93.994      -36.4      4.857      5191      no      senior
## 6      93.994      -36.4      4.857      5191      no      senior
```

```
names(bank_Marketing)
```

```
## [1] "age"      "job"      "marital"  "education"
## [5] "default"  "housing"  "loan"     "contact"
## [9] "month"    "day_of_week" "duration" "campaign"
## [13] "pdays"   "previous"  "poutcome" "emp.var.rate"
## [17] "cons.price.idx" "cons.conf.idx" "euribor3m" "nr.employed"
## [21] "deposit"   "age_category"
```

Extract the predictors as data frame

```
deposit_x <- bank_Marketing[,c(1:20,22)]
head(deposit_x )
```

```
##   age      job marital  education default housing loan  contact month
## 1  56 housemaid married  basic.4y      no      no  no telephone  may
## 2  57 services married high.school unknown      no  no telephone  may
## 3  37 services married high.school      no  yes  no telephone  may
## 4  40 admin. married  basic.6y      no      no  no telephone  may
## 5  56 services married high.school      no      no  yes telephone  may
## 6  45 services married  basic.9y unknown      no  no telephone  may
##   day_of_week duration campaign pdays previous  poutcome emp.var.rate
## 1      mon      261      1  999      0 nonexistent      1.1
## 2      mon      149      1  999      0 nonexistent      1.1
## 3      mon      226      1  999      0 nonexistent      1.1
## 4      mon      151      1  999      0 nonexistent      1.1
## 5      mon      307      1  999      0 nonexistent      1.1
## 6      mon      198      1  999      0 nonexistent      1.1
##   cons.price.idx cons.conf.idx euribor3m nr.employed age_category
## 1      93.994      -36.4      4.857      5191      senior
## 2      93.994      -36.4      4.857      5191      senior
## 3      93.994      -36.4      4.857      5191      senior
## 4      93.994      -36.4      4.857      5191      senior
## 5      93.994      -36.4      4.857      5191      senior
## 6      93.994      -36.4      4.857      5191      senior
```

Transform the character variables in fator

```
bank_Marketing = bank_Marketing %>% mutate(job = as.factor(job) , deposit =
as.factor(deposit),marital= as.factor(marital) , education =
as.factor(education),default =as.factor(default),housing = as.factor(housing
```



```

),loan = as.factor(loan) ,contact = as.factor(contact),month =
as.factor(month),day_of_week = as.factor(day_of_week), poutcome =
as.factor(poutcome))

glimpse(bank_Marketing)

## Rows: 41,188
## Columns: 22
## $ age          <int> 56, 57, 37, 40, 56, 45, 59, 41, 24, 25, 41, 25, 29,
57,...
## $ job          <fct> housemaid, services, services, admin., services,
servic...
## $ marital      <fct> married, married, married, married, married,
married, m...
## $ education    <fct> basic.4y, high.school, high.school, basic.6y,
high.scho...
## $ default      <fct> no, unknown, no, no, no, no, unknown, no, unknown, no,
no, ...
## $ housing      <fct> no, no, yes, no, no, no, no, no, yes, yes, no, yes,
no,...
## $ loan         <fct> no, no, no, no, yes, no, no, no, no, no, no, no,
yes, n...
## $ contact      <fct> telephone, telephone, telephone, telephone,
telephone, ...
## $ month        <fct> may, may, may, may, may, may, may, may, may, may,
may, ...
## $ day_of_week  <fct> mon, mon, mon, mon, mon, mon, mon, mon, mon, mon,
mon, ...
## $ duration     <int> 261, 149, 226, 151, 307, 198, 139, 217, 380, 50,
55, 22...
## $ campaign     <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1...
## $ pdays        <int> 999, 999, 999, 999, 999, 999, 999, 999, 999, 999,
999, ...
## $ previous     <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0...
## $ poutcome     <fct> nonexistent, nonexistent, nonexistent, nonexistent,
non...
## $ emp.var.rate <dbl> 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1, 1.1,
1.1, ...
## $ cons.price.idx <dbl> 93.994, 93.994, 93.994, 93.994, 93.994, 93.994,
93.994,...
## $ cons.conf.idx <dbl> -36.4, -36.4, -36.4, -36.4, -36.4, -36.4, -36.4, -
36.4,...
## $ euribor3m    <dbl> 4.857, 4.857, 4.857, 4.857, 4.857, 4.857, 4.857,
4.857,...
## $ nr.employed  <dbl> 5191, 5191, 5191, 5191, 5191, 5191, 5191, 5191,
5191, 5...
## $ deposit      <fct> no, no, no, no, no, no, no, no, no, no, no, no,
no, no,
no,...

```

```
## $ age_category    <fct> senior, senior, senior, senior, senior, senior,
senior,...
```

Create a fold data

```
gfold<-createFolds(deposit_y, k =5)
```

Build a control parameter

```
gcontrol<-trainControl(
  summaryFunction = twoClassSummary,
  classProbs = TRUE,
  verboseIter = TRUE,
  savePredictions = TRUE,
  index= gfold
)
```

Build a randomForest model

```
forestmodel<-train(
  x= deposit_x ,
  y= deposit_y,
  metric="ROC",
  method = "ranger",
  trControl = gcontrol
)
```

```
## + Fold1: mtry= 2, min.node.size=1, splitrule=gini
## - Fold1: mtry= 2, min.node.size=1, splitrule=gini
## + Fold1: mtry=11, min.node.size=1, splitrule=gini
## - Fold1: mtry=11, min.node.size=1, splitrule=gini
## + Fold1: mtry=21, min.node.size=1, splitrule=gini
## - Fold1: mtry=21, min.node.size=1, splitrule=gini
## + Fold1: mtry= 2, min.node.size=1, splitrule=extratrees
## - Fold1: mtry= 2, min.node.size=1, splitrule=extratrees
## + Fold1: mtry=11, min.node.size=1, splitrule=extratrees
## - Fold1: mtry=11, min.node.size=1, splitrule=extratrees
## + Fold1: mtry=21, min.node.size=1, splitrule=extratrees
## - Fold1: mtry=21, min.node.size=1, splitrule=extratrees
## + Fold2: mtry= 2, min.node.size=1, splitrule=gini
## - Fold2: mtry= 2, min.node.size=1, splitrule=gini
## + Fold2: mtry=11, min.node.size=1, splitrule=gini
## - Fold2: mtry=11, min.node.size=1, splitrule=gini
## + Fold2: mtry=21, min.node.size=1, splitrule=gini
## - Fold2: mtry=21, min.node.size=1, splitrule=gini
## + Fold2: mtry= 2, min.node.size=1, splitrule=extratrees
## - Fold2: mtry= 2, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=11, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=11, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=21, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=21, min.node.size=1, splitrule=extratrees
## + Fold3: mtry= 2, min.node.size=1, splitrule=gini
```

```

## - Fold3: mtry= 2, min.node.size=1, splitrule=gini
## + Fold3: mtry=11, min.node.size=1, splitrule=gini
## - Fold3: mtry=11, min.node.size=1, splitrule=gini
## + Fold3: mtry=21, min.node.size=1, splitrule=gini
## - Fold3: mtry=21, min.node.size=1, splitrule=gini
## + Fold3: mtry= 2, min.node.size=1, splitrule=extratrees
## - Fold3: mtry= 2, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=11, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=11, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=21, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=21, min.node.size=1, splitrule=extratrees
## + Fold4: mtry= 2, min.node.size=1, splitrule=gini
## - Fold4: mtry= 2, min.node.size=1, splitrule=gini
## + Fold4: mtry=11, min.node.size=1, splitrule=gini
## - Fold4: mtry=11, min.node.size=1, splitrule=gini
## + Fold4: mtry=21, min.node.size=1, splitrule=gini
## - Fold4: mtry=21, min.node.size=1, splitrule=gini
## + Fold4: mtry= 2, min.node.size=1, splitrule=extratrees
## - Fold4: mtry= 2, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=11, min.node.size=1, splitrule=extratrees
## - Fold4: mtry=11, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=21, min.node.size=1, splitrule=extratrees
## - Fold4: mtry=21, min.node.size=1, splitrule=extratrees
## + Fold5: mtry= 2, min.node.size=1, splitrule=gini
## - Fold5: mtry= 2, min.node.size=1, splitrule=gini
## + Fold5: mtry=11, min.node.size=1, splitrule=gini
## - Fold5: mtry=11, min.node.size=1, splitrule=gini
## + Fold5: mtry=21, min.node.size=1, splitrule=gini
## - Fold5: mtry=21, min.node.size=1, splitrule=gini
## + Fold5: mtry= 2, min.node.size=1, splitrule=extratrees
## - Fold5: mtry= 2, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=11, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=11, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=21, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=21, min.node.size=1, splitrule=extratrees
## Aggregating results
## Selecting tuning parameters
## Fitting mtry = 11, splitrule = gini, min.node.size = 1 on full training
set

```

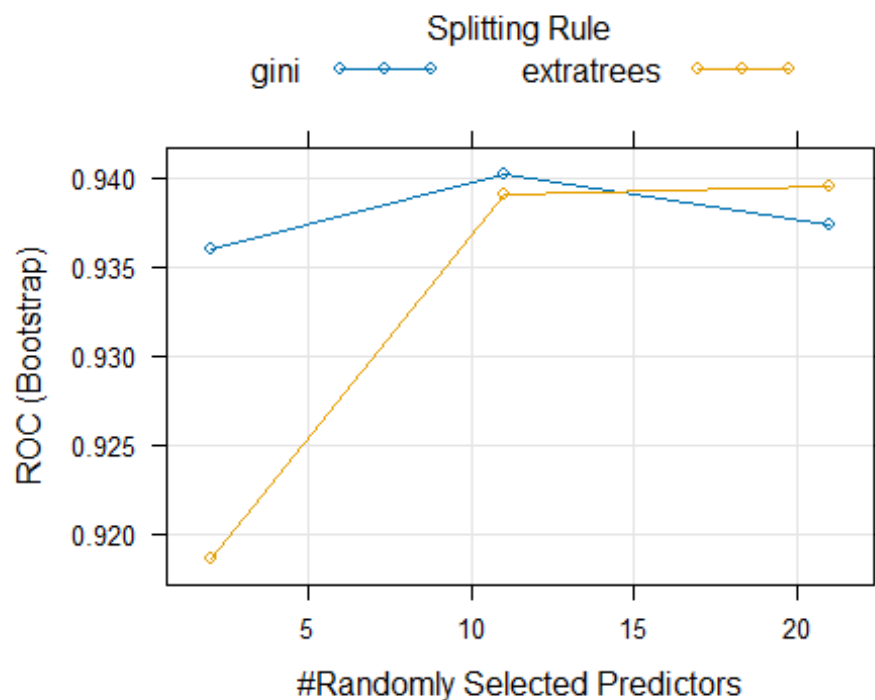
Visualize the randomforest model

```
summary(forestmodel)
```

	Length	Class	Mode
## predictions	82376	-none-	numeric
## num.trees	1	-none-	numeric
## num.independent.variables	1	-none-	numeric
## mtry	1	-none-	numeric
## min.node.size	1	-none-	numeric
## prediction.error	1	-none-	numeric

```
## forest          11 ranger.forest list
## splitrule       1 -none-         character
## treetype        1 -none-         character
## call            9 -none-         call
## importance.mode  1 -none-         character
## num.samples     1 -none-         numeric
## replace         1 -none-         logical
## xNames          21 -none-         character
## problemType     1 -none-         character
## tuneValue       3 data.frame     list
## obsLevels       2 -none-         character
## param           0 -none-         list
```

```
plot(forestmodel)
```



Confusion matrix for logistic randomForest model

```
confusionMatrix(forestmodel , positive = "1")
```

```
## Bootstrapped (5 reps) Confusion Matrix
##
## (entries are percentual average cell counts across resamples)
##
##           Reference
## Prediction  no  yes
##           no 85.2 5.3
##           yes 3.5 5.9
```

```
##
## Accuracy (average) : 0.9112

confusionMatrix(forestmodel)

## Bootstrapped (5 reps) Confusion Matrix
##
## (entries are percentual average cell counts across resamples)
##
##           Reference
## Prediction   no  yes
##           no 85.2 5.3
##           yes 3.5 5.9
##
## Accuracy (average) : 0.9112
```

Check the default variable

```
levels(bank_Marketing$default)

## [1] "no"      "unknown" "yes"

levels(deposit_x$default)

## NULL
```

Check for NA's variable

```
cmp<-is.na(deposit_x)
for( j in 1:ncol(cmp)) cmp[,j]<- as.numeric(cmp[,j])
colnames(cmp) <-paste0("MVP_", colnames(cmp))
```

Pass value to NA

```
deposit_x$MVP<- as.numeric(rowSums(cmp) > 0)
deposit_x$default<-as.numeric(deposit_x$default == "yes")
```

Input value into the missing variable

```
bankimpute <- caret::preProcess(
  x = deposit_x,
  method = "center","scale","medianImpute")
```

Fit a logistic regression model

```
model_glm<-train(
  x= deposit_x,
  y= deposit_y,
  metric="ROC",
  method = "glm",
  trControl = gcontrol)

## + Fold1: parameter=none

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
```

```
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## - Fold1: parameter=none
## + Fold2: parameter=none

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## - Fold2: parameter=none
## + Fold3: parameter=none

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## - Fold3: parameter=none
## + Fold4: parameter=none

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## - Fold4: parameter=none
## + Fold5: parameter=none
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

## - Fold5: parameter=none
## Aggregating results
## Fitting final model on full training set
```

Evaluate the logistic regressioun model

```
summary(model_glm)
```

```
##
## Call:
## NULL
##
## Coefficients: (2 not defined because of singularities)
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.374e+02  3.837e+01 -6.187 6.11e-10 ***
## age         -4.836e-04  2.444e-03 -0.198 0.843167
## jobblue-collar -2.628e-01  7.966e-02 -3.299 0.000971 ***
## jobentrepreneur -1.865e-01  1.260e-01 -1.480 0.138920
## jobhousemaid   -3.376e-02  1.477e-01 -0.229 0.819212
## jobmanagement -5.699e-02  8.534e-02 -0.668 0.504267
## jobretired     2.900e-01  1.072e-01  2.704 0.006861 **
## jobself-employed -1.675e-01  1.178e-01 -1.422 0.155054
## jobservices    -1.565e-01  8.613e-02 -1.817 0.069157 .
## jobstudent     1.699e-01  1.121e-01  1.516 0.129394
## jobtechnician  -1.558e-02  7.114e-02 -0.219 0.826611
## jobunemployed   5.166e-03  1.279e-01  0.040 0.967776
## jobunknown     -1.122e-01  2.383e-01 -0.471 0.637877
## maritalmarried -1.453e-02  6.846e-02 -0.212 0.831977
## maritalsingle  4.320e-02  7.829e-02  0.552 0.581141
## maritalunknown 4.521e-02  4.160e-01  0.109 0.913468
## educationbasic.6y 1.233e-01  1.203e-01  1.025 0.305308
## educationbasic.9y 2.369e-02  9.501e-02  0.249 0.803052
## educationhigh.school 7.934e-02  9.165e-02  0.866 0.386700
## educationilliterate 1.036e+00  7.645e-01  1.355 0.175541
## educationprofessional.course 1.485e-01  1.011e-01  1.469 0.141738
## educationuniversity.degree 2.305e-01  9.170e-02  2.514 0.011953 *
## educationunknown 1.590e-01  1.193e-01  1.333 0.182597
## default       -7.244e+00  1.135e+02 -0.064 0.949092
## housingunknown -9.477e-02  1.395e-01 -0.679 0.496898
## housingyes     -4.612e-03  4.134e-02 -0.112 0.911177
## loanunknown    NA          NA      NA      NA
```

```

## loanyes -4.889e-02 5.745e-02 -0.851 0.394761
## contacttelephone -6.501e-01 7.701e-02 -8.441 < 2e-16 ***
## monthaug 8.782e-01 1.206e-01 7.279 3.36e-13 ***
## monthdec 3.247e-01 2.095e-01 1.550 0.121102
## monthjul 1.425e-01 9.622e-02 1.481 0.138717
## monthjun -5.299e-01 1.264e-01 -4.192 2.77e-05 ***
## monthmar 2.032e+00 1.445e-01 14.067 < 2e-16 ***
## monthmay -4.501e-01 8.256e-02 -5.452 4.97e-08 ***
## monthnov -3.989e-01 1.210e-01 -3.297 0.000976 ***
## monthoct 2.044e-01 1.539e-01 1.328 0.184070
## monthsep 3.874e-01 1.797e-01 2.156 0.031080 *
## day_of_weekmon -1.152e-01 6.612e-02 -1.742 0.081497 .
## day_of_weekthu 5.959e-02 6.408e-02 0.930 0.352435
## day_of_wektue 1.015e-01 6.589e-02 1.540 0.123568
## day_of_weekwed 1.839e-01 6.565e-02 2.801 0.005102 **
## duration 4.698e-03 7.439e-05 63.158 < 2e-16 ***
## campaign -4.003e-02 1.156e-02 -3.464 0.000533 ***
## pdays -9.364e-04 2.169e-04 -4.318 1.57e-05 ***
## previous -6.346e-02 5.913e-02 -1.073 0.283151
## poutcomenonexistent 4.250e-01 9.420e-02 4.512 6.43e-06 ***
## poutcomesuccess 9.653e-01 2.113e-01 4.568 4.92e-06 ***
## emp.var.rate -1.773e+00 1.423e-01 -12.458 < 2e-16 ***
## cons.price.idx 2.203e+00 2.529e-01 8.712 < 2e-16 ***
## cons.conf.idx 2.051e-02 7.772e-03 2.640 0.008298 **
## euribor3m 3.330e-01 1.300e-01 2.562 0.010410 *
## nr.employed 5.374e-03 3.119e-03 1.723 0.084861 .
## age_categorysenior -2.019e-01 1.359e-01 -1.486 0.137355
## age_categoryyouth -1.243e-01 2.056e-01 -0.604 0.545546
## MVP NA NA NA NA
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 28999 on 41187 degrees of freedom
## Residual deviance: 17096 on 41134 degrees of freedom
## AIC: 17204
##
## Number of Fisher Scoring iterations: 10

```

Confusion matrix for logistic regression

```
confusionMatrix(model_glm)
```

```
## Bootstrapped (5 reps) Confusion Matrix
```

```
##
```

```
## (entries are percentual average cell counts across resamples)
```

```
##
```

```
##           Reference
```

```
## Prediction  no  yes
```

```
##           no 86.2 6.5
```



```
##           yes  2.5  4.8
##
## Accuracy (average) : 0.9096
```

Build a decision model

```
tree_model<-train(y = deposit_y,
                  x = deposit_x,
                  metric= "ROC",
                  method = 'rpart',
                  trControl = gcontrol
                  )

## + Fold1: cp=0.01595
## - Fold1: cp=0.01595
## + Fold2: cp=0.01595
## - Fold2: cp=0.01595
## + Fold3: cp=0.01595
## - Fold3: cp=0.01595
## + Fold4: cp=0.01595
## - Fold4: cp=0.01595
## + Fold5: cp=0.01595
## - Fold5: cp=0.01595
## Aggregating results
## Selecting tuning parameters
## Fitting cp = 0.0159 on full training set
```

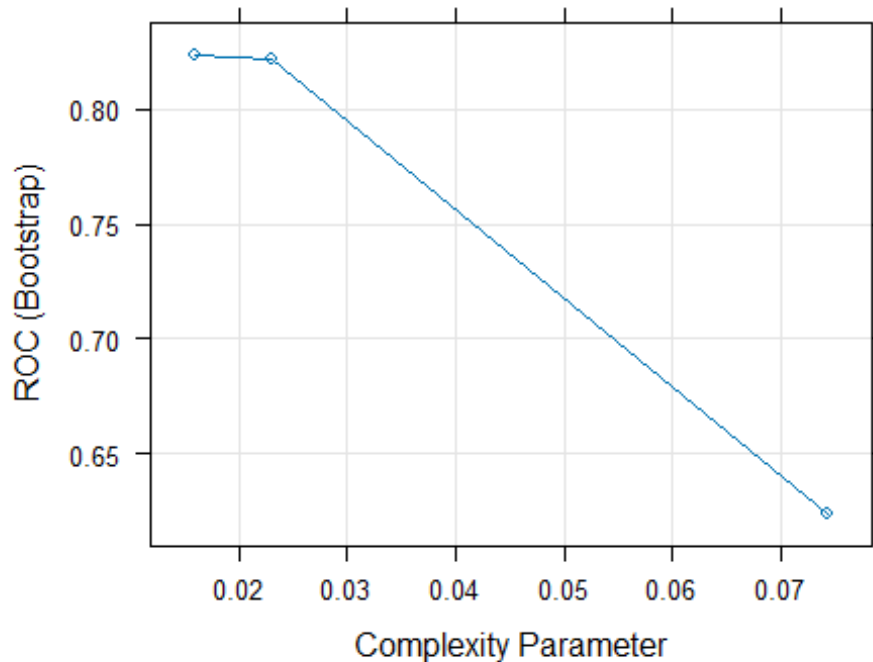
Decision tree model

```
tree_model

## CART
##
## 41188 samples
##    22 predictor
##    2 classes: 'no', 'yes'
##
## No pre-processing
## Resampling: Bootstrapped (5 reps)
## Summary of sample sizes: 8238, 8238, 8237, 8237, 8238
## Resampling results across tuning parameters:
##
##    cp          ROC          Sens          Spec
## 0.01594828 0.8234410 0.9673579 0.4443966
## 0.02295259 0.8216496 0.9587461 0.4987069
## 0.07424569 0.6237691 0.9792122 0.2462823
##
## ROC was used to select the optimal model using the largest value.
## The final value used for the model was cp = 0.01594828.
```

Plot the tree model

```
plot(tree_model)
```



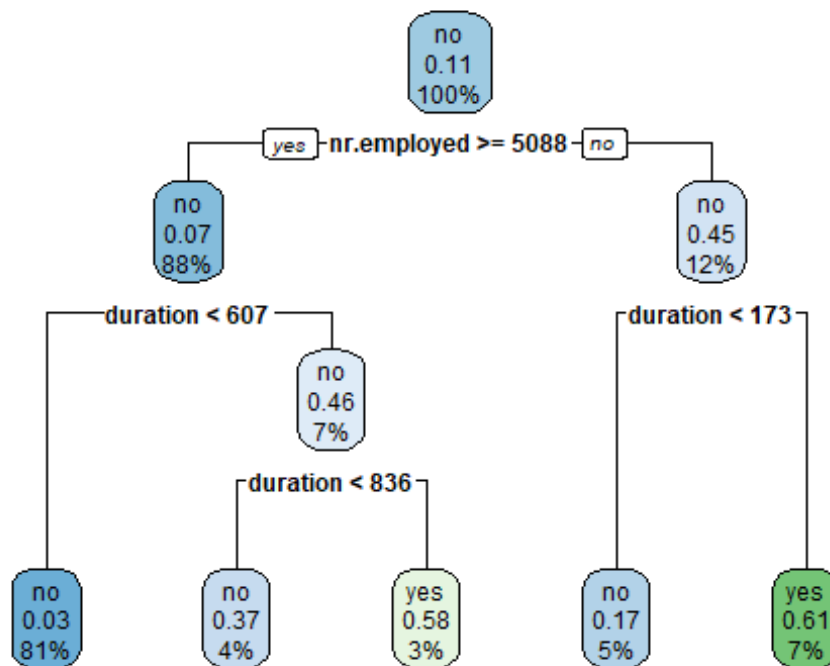
See the final model

```
tree_model$finalModel
```

```
## n= 41188
##
## node), split, n, loss, yval, (yprob)
##      * denotes terminal node
##
## 1) root 41188 4640 no (0.8873458 0.1126542)
##    2) nr.employed>=5087.65 36224 2431 no (0.9328898 0.0671102)
##      4) duration< 606.5 33232 1044 no (0.9685845 0.0314155) *
##      5) duration>=606.5 2992 1387 no (0.5364305 0.4635695)
##        10) duration< 835.5 1601 585 no (0.6346034 0.3653966) *
##        11) duration>=835.5 1391 589 yes (0.4234364 0.5765636) *
##    3) nr.employed< 5087.65 4964 2209 no (0.5549960 0.4450040)
##      6) duration< 172.5 1891 328 no (0.8265468 0.1734532) *
##      7) duration>=172.5 3073 1192 yes (0.3878946 0.6121054) *
```

Plot the tree model

```
rpart.plot(tree_model$finalModel)
```



Confusion matrix tree model

```
confusionMatrix(tree_model)
```

```
## Bootstrapped (5 reps) Confusion Matrix
##
## (entries are percentual average cell counts across resamples)
##
##           Reference
## Prediction  no  yes
##           no 85.8  6.3
##           yes  2.9  5.0
##
## Accuracy (average) : 0.9084
```

Comparing the model

```
model_list<-list(rf = forestmodel,
                 glm = model_glm,
                 tree = tree_model)
```

Resamples

```
resamp<-resamples(model_list)
```

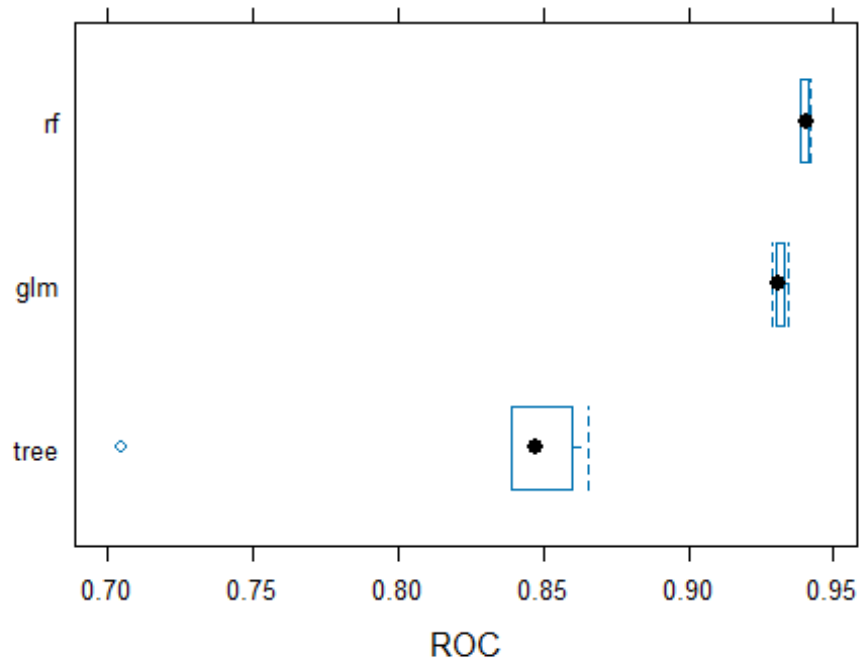
See the model

```
summary(resamp)
```

```
##
## Call:
## summary.resamples(object = resamp)
##
## Models: rf, glm, tree
## Number of resamples: 5
##
## ROC
##           Min.   1st Qu.   Median     Mean   3rd Qu.     Max. NA's
## rf    0.9386569 0.9388596 0.9403848 0.9401855 0.9412530 0.9417731    0
## glm   0.9285250 0.9304293 0.9311636 0.9315327 0.9330099 0.9345357    0
## tree  0.7052195 0.8389073 0.8473067 0.8234410 0.8600441 0.8657274    0
##
## Sens
##           Min.   1st Qu.   Median     Mean   3rd Qu.     Max. NA's
## rf    0.9595732 0.9596074 0.9599850 0.9600183 0.9602914 0.9606348    0
## glm   0.9685694 0.9714755 0.9714765 0.9712980 0.9715439 0.9734250    0
## tree  0.9616266 0.9621725 0.9632670 0.9673579 0.9691839 0.9805397    0
##
## Spec
##           Min.   1st Qu.   Median     Mean   3rd Qu.     Max. NA's
## rf    0.5102371 0.5126616 0.5304418 0.5262392 0.5355603 0.5422953    0
## glm   0.4059806 0.4194504 0.4218750 0.4233297 0.4329203 0.4364224    0
## tree  0.3205819 0.4148707 0.4797953 0.4443966 0.4892241 0.5175108    0
```

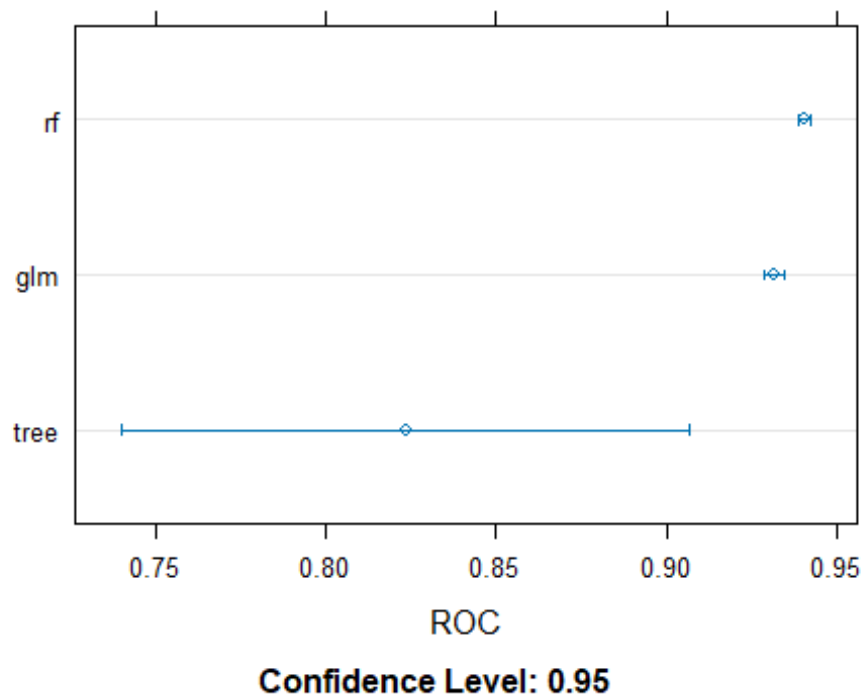
Model comparison with boxplot

```
bwplot(resamp, metric = "ROC")
```



Model comparison with dotplot

```
dotplot(resamp, metric = "ROC")
```



Model comparison with densityplot

```
densityplot(resamp, metric = "ROC")
```

