Preprocessing

Admin

24 4 2019

Raw Dataset

```
data <- read.csv("./Bank.csv", header = TRUE, sep = ";")
summary(data)
```

```
##       age                 job           marital        education
##  Min.   :19.00    management :969   divorced: 528   primary  : 678
##  1st Qu.:33.00    blue-collar:946   married :2797   secondary:2306
##  Median :39.00    technician :768   single  :1196   tertiary :1350
##  Mean   :41.17    admin.     :478                   unknown  : 187
##  3rd Qu.:49.00    services   :417
##  Max.   :87.00    retired    :230
##                   (Other)    :713
##  default       balance       housing      loan           contact
##  no :4445   Min.   :-3313   no :1962   no :3830   cellular :2896
##  yes:  76   1st Qu.:   69   yes:2559   yes: 691   telephone: 301
##             Median :  444                         unknown  :1324
##             Mean   : 1423
##             3rd Qu.: 1480
##             Max.   :71188
##
##       day             month         duration         campaign
##  Min.   : 1.00    may    :1398   Min.   :   4   Min.   : 1.000
##  1st Qu.: 9.00    jul    : 706   1st Qu.: 104   1st Qu.: 1.000
##  Median :16.00    aug    : 633   Median : 185   Median : 2.000
##  Mean   :15.92    jun    : 531   Mean   : 264   Mean   : 2.794
##  3rd Qu.:21.00    nov    : 389   3rd Qu.: 329   3rd Qu.: 3.000
##  Max.   :31.00    apr    : 293   Max.   :3025   Max.   :50.000
##                   (Other): 571
##      pdays            previous          poutcome       y
##  Min.   : -1.00   Min.   : 0.0000   failure: 490   no :4000
##  1st Qu.: -1.00   1st Qu.: 0.0000   other  : 197   yes: 521
##  Median : -1.00   Median : 0.0000   success: 129
##  Mean   : 39.77   Mean   : 0.5426   unknown:3705
##  3rd Qu.: -1.00   3rd Qu.: 0.0000
##  Max.   :871.00   Max.   :25.0000
##
```

Including Plots

You can also embed plots, for example:

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.