MVA Assignment: Multiple Regression

```
#Multiple Regression
library(ggplot2)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library (stringr)
library(data.table)

##
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':
##
##     between, first, last

library(grid)
library(gridExtra)

##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##     combine

library(corrplot)

## corrplot 0.84 loaded

library(scales)
library(qqplotr)

##
## Attaching package: 'qqplotr'

## The following objects are masked from 'package:ggplot2':
##
##     stat_qq_line, StatQqLine
```

```r
library(MASS)
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

```r
library(DMwR)
```

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

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```r
library(car)
```

```
## Loading required package: carData
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```r
library(e1071)
library(caret)
library(caTools)
library(pROC)
```

```
## Type 'citation("pROC")' for a citation.
```

```
##
## Attaching package: 'pROC'
```

```
## The following objects are masked from 'package:stats':
##
##     cov, smooth, var
```

```r
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------
tidyverse 1.3.0 --
```

```
## v tibble  2.1.3     v purrr   0.3.3
## v tidyr   1.0.2     v forcats 0.5.0
## v readr   1.3.1
```

```
## -- Conflicts ------------------------------------------------
tidyverse_conflicts() --
## x data.table::between()   masks dplyr::between()
```

```
## x readr::col_factor()     masks scales::col_factor()
## x gridExtra::combine()    masks dplyr::combine()
## x purrr::discard()        masks scales::discard()
## x dplyr::filter()         masks stats::filter()
## x data.table::first()     masks dplyr::first()
## x dplyr::lag()            masks stats::lag()
## x data.table::last()      masks dplyr::last()
## x purrr::lift()           masks caret::lift()
## x car::recode()           masks dplyr::recode()
## x MASS::select()          masks dplyr::select()
## x purrr::some()           masks car::some()
## x qqplotr::stat_qq_line() masks ggplot2::stat_qq_line()
## x purrr::transpose()      masks data.table::transpose()
```

```
library(MVA)
```

```
## Loading required package: HSAUR2
```

```
## Loading required package: tools
```

```
library(GGally)
```

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

```
##
## Attaching package: 'GGally'
```

```
## The following object is masked from 'package:dplyr':
##
##     nasa
```

```
library(gvlma)
```

```
bank=read.csv("C:/Users/Shamali/Desktop/RutgersSpring/multivariat/project/New folder/bank.csv",row.names=1,fill=TRUE)
attach(bank)
```

```
# If the z-value is too big in magnitude i.e either too positive or too negative, it indicates that the corresponding true regression coefficient is not 0 and the corresponding X-variable matters.
```

```
# Performing multiple regression on bank dataset
fit <- lm(Deposit~Job+Education+Marital+Housing, data=bank)
#show the results
summary(fit)
```

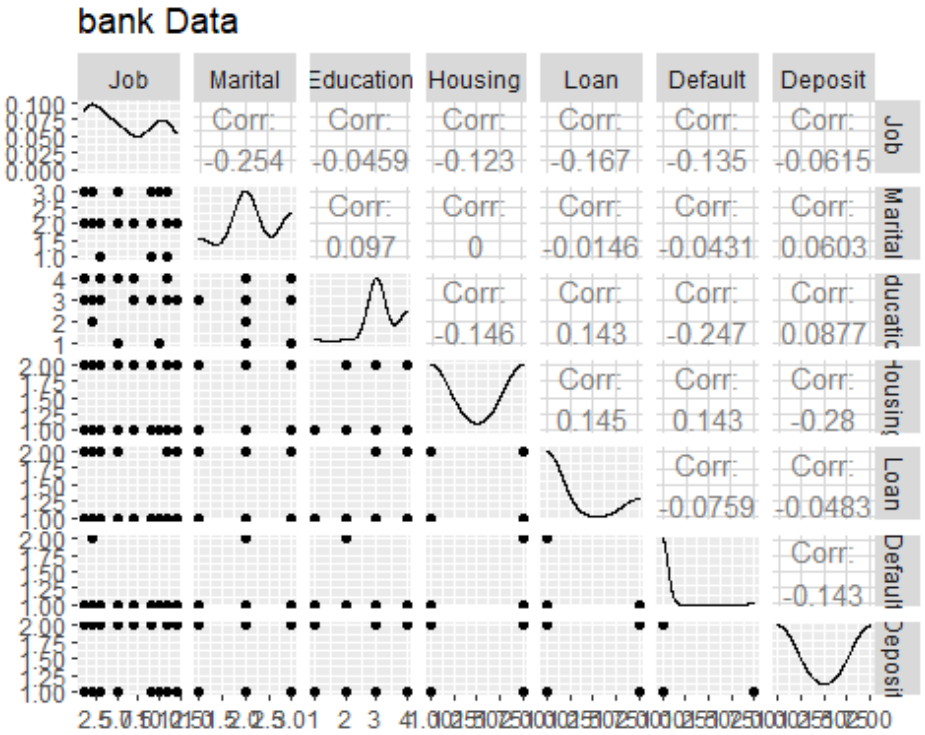
```
##
## Call:
## lm(formula = Deposit ~ Job + Education + Marital + Housing, data = bank)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69887 -0.40662 -0.00031  0.41392  0.66734
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.84223    0.51916   3.548  0.00092 ***
## Job         -0.01102    0.01901  -0.580  0.56488
## Education    0.02834    0.10568   0.268  0.78977
## Marital      0.02619    0.11123   0.235  0.81494
## Housing     -0.28491    0.14499  -1.965  0.05560 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5026 on 45 degrees of freedom
## Multiple R-squared:  0.09056,    Adjusted R-squared:  0.009723
## F-statistic:  1.12 on 4 and 45 DF,  p-value: 0.3589
```

```
#Summary has three sections. Section1: How well does the model fit the data (before Coefficients). Section2: Is the hypothesis supported? (until sifnif codes). Section3: How well does data fit the model (again).
coefficients(fit)
```

```
## (Intercept)         Job   Education     Marital     Housing
##  1.84223372 -0.01102218  0.02834262  0.02618722 -0.28491277
```

```
ggpairs(data=bank, title="bank Data")
```



```r
confint(fit,level=0.95)
```

```
##                   2.5 %      97.5 %
## (Intercept)  0.79659167 2.887875775
## Job         -0.04930482 0.027260458
## Education   -0.18450216 0.241187398
## Marital     -0.19784604 0.250220478
## Housing     -0.57693322 0.007107682
```

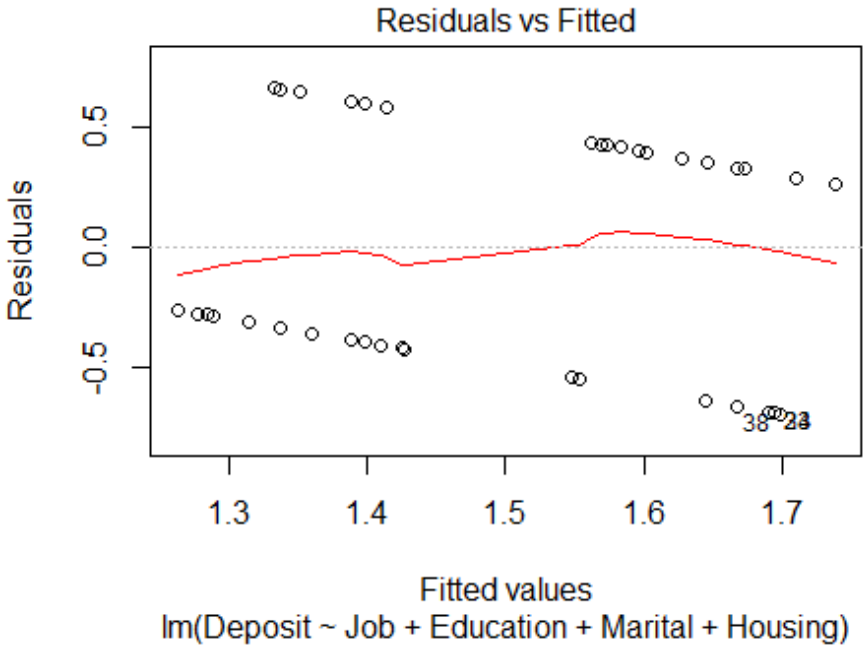
```r
# Predicted Values
fitted(fit)
```

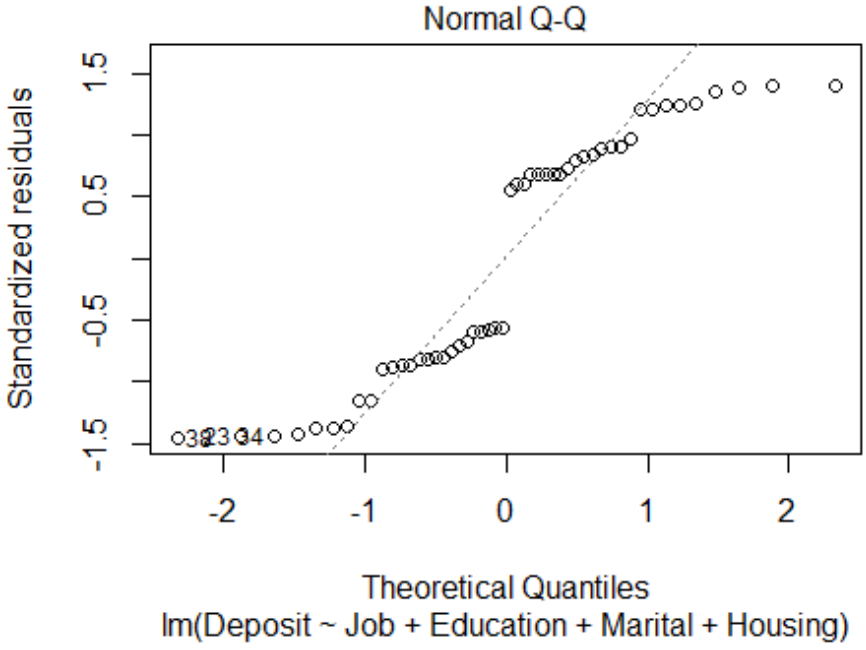
```
##        1        2        3        4        5        6        7        8
## 1.628009 1.573479 1.644336 1.582927 1.595524 1.277544 1.562457 1.409229
##        9       10       11       12       13       14       15       16
## 1.398788 1.350557 1.284424 1.262379 1.398788 1.288567 1.314754 1.398788
##       17       18       19       20       21       22       23       24
## 1.387766 1.667955 1.262379 1.554003 1.601822 1.427131 1.694142 1.359423
##       25       26       27       28       29       30       31       32
## 1.336798 1.336798 1.413953 1.569336 1.667955 1.332655 1.667955 1.413953
##       33       34       35       36       37       38       39       40
## 1.738231 1.694142 1.689999 1.667955 1.387766 1.698866 1.547292 1.672679
##       41       42       43       44       45       46       47       48
## 1.547292 1.667955 1.424976 1.709888 1.409229 1.398788 1.709888 1.645911
##       49       50
## 1.336798 1.387766
```

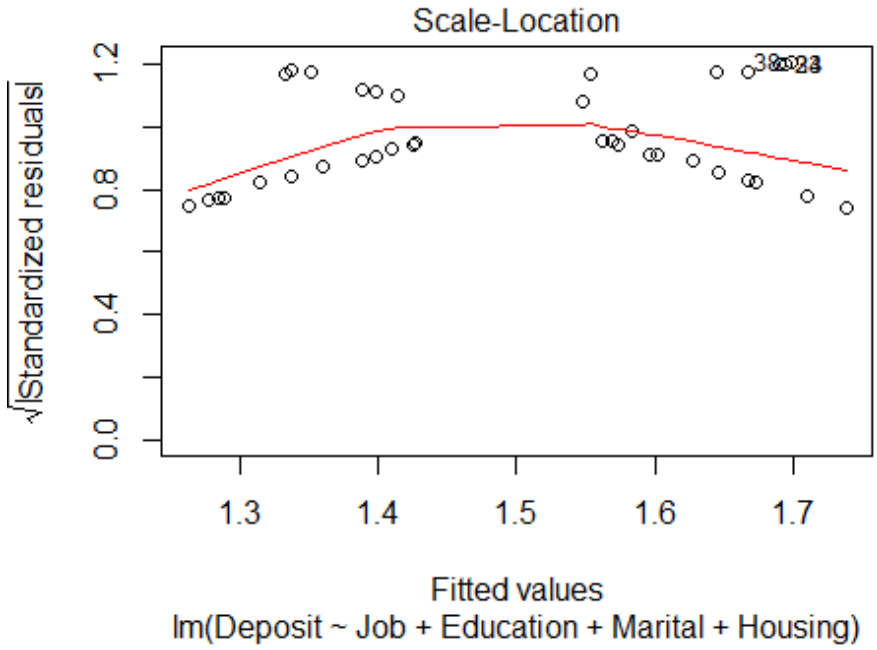
```r
residuals(fit)
```

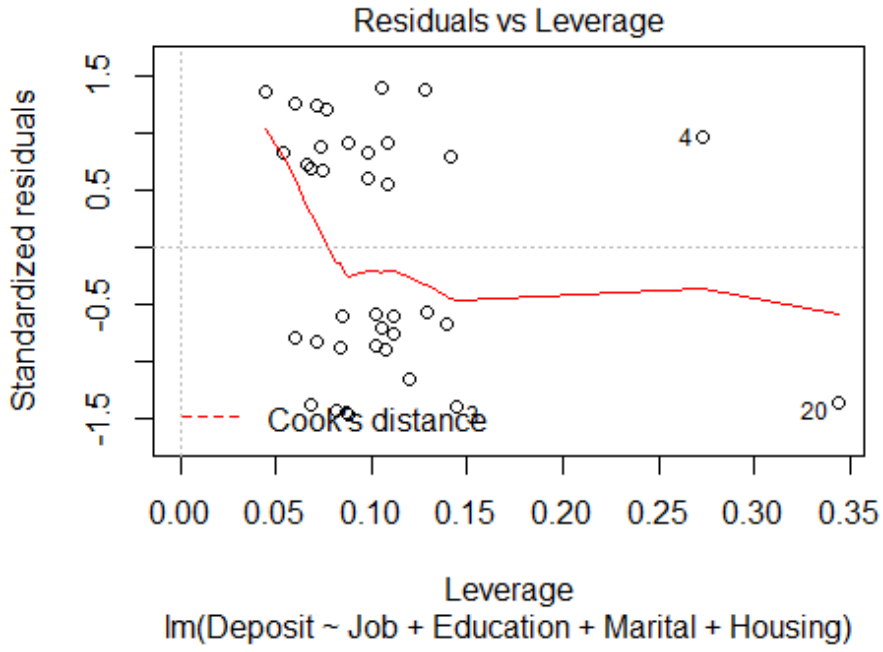
```
##          1          2          3          4          5          6
7
##  0.3719909  0.4265207 -0.6443363  0.4170729  0.4044764 -0.2775443
0.4375429
##          8          9         10         11         12         13
14
## -0.4092294 -0.3987883  0.6494433 -0.2844236 -0.2623793  0.6012117 -
0.2885665
##         15         16         17         18         19         20
21
## -0.3147537 -0.3987883 -0.3877661  0.3320450 -0.2623793 -0.5540034
0.3981781
##         22         23         24         25         26         27
28
## -0.4271309 -0.6941422 -0.3594235 -0.3367981  0.6632019  0.5860467
0.4306636
##         29         30         31         32         33         34
35
##  0.3320450  0.6673448  0.3320450  0.5860467  0.2617691 -0.6941422 -
0.6899993
##         36         37         38         39         40         41
42
```

```
##  0.3320450 -0.3877661 -0.6988661 -0.5472921  0.3273211 -0.5472921 -
0.6679550
##         43         44         45         46         47         48
49
## -0.4249755  0.2901117 -0.4092294  0.6012117  0.2901117  0.3540894
0.6632019
##         50
##  0.6122339

#Anova Table
anova(fit)

## Analysis of Variance Table
##
## Response: Deposit
##           Df  Sum Sq Mean Sq F value Pr(>F)
## Job        1  0.0473 0.04731  0.1873 0.6672
## Education  1  0.0903 0.09033  0.3576 0.5528
## Marital    1  0.0189 0.01887  0.0747 0.7859
## Housing    1  0.9755 0.97551  3.8615 0.0556 .
## Residuals 45 11.3680 0.25262
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

vcov(fit)

##              (Intercept)           Job     Education       Marital
## (Intercept)  0.269527178 -4.029631e-03 -3.727545e-02 -0.0275424263
## Job         -0.004029631  3.612769e-04  8.354329e-05  0.0005302767
## Education   -0.037275446  8.354329e-05  1.116767e-02 -0.0009922113
## Marital     -0.027542426  5.302767e-04 -9.922113e-04  0.0123726145
## Housing     -0.041638124  3.635345e-04  2.313735e-03  0.0003106233
##                   Housing
## (Intercept) -0.0416381243
## Job          0.0003635345
## Education    0.0023137351
## Marital      0.0003106233
## Housing      0.0210214665

cov2cor(vcov(fit))

##             (Intercept)         Job   Education     Marital     Housing
## (Intercept)   1.0000000 -0.40836064 -0.67942314 -0.47694706 -0.55316968
## Job          -0.4083606  1.00000000  0.04159204  0.25081395  0.13191494
## Education    -0.6794231  0.04159204  1.00000000 -0.08440967  0.15100832
## Marital      -0.4769471  0.25081395 -0.08440967  1.00000000  0.01926069
## Housing      -0.5531697  0.13191494  0.15100832  0.01926069  1.00000000

temp <- influence.measures(fit)

#diagnostic plots
plot(fit)
```


Residuals vs Fitted
Residuals
0.5
0.0
-0.5
1.3
1.4
1.5
1.6
1.7
Fitted values
lm(Deposit ~ Job + Education + Marital + Housing)


Normal Q-Q
Standardized residuals
1.5
0.5
-0.5
-1.5
-2
-1
0
1
2
Theoretical Quantiles
lm(Deposit ~ Job + Education + Marital + Housing)

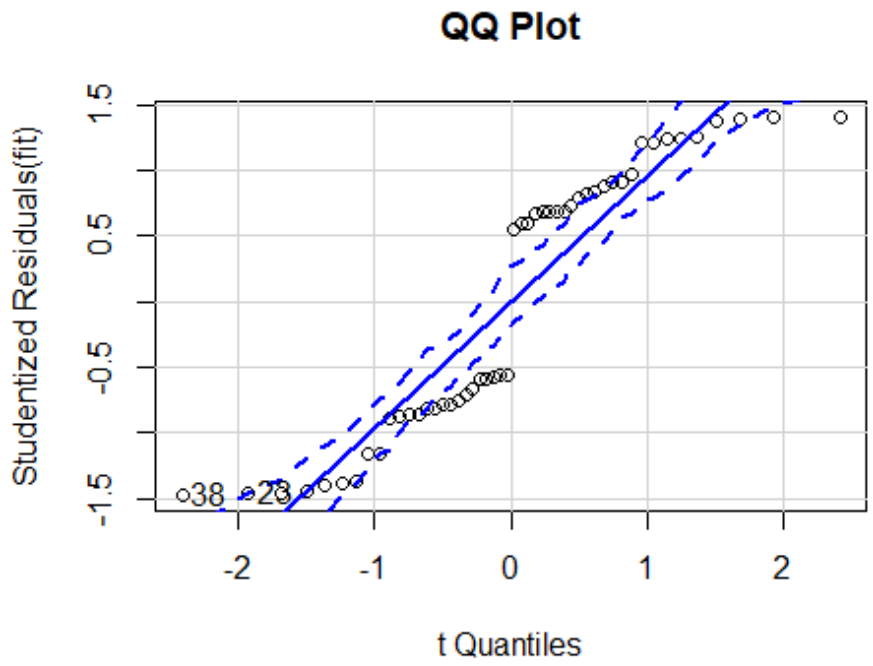





```r
# Assessing Outliers
outlierTest(fit)
```

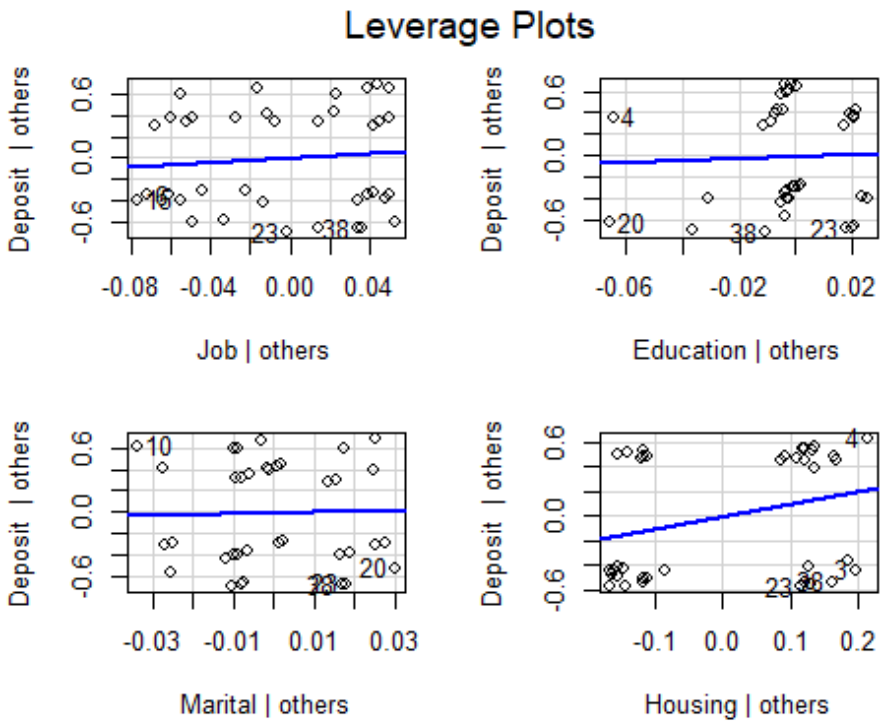
```
## No Studentized residuals with Bonferroni p < 0.05
## Largest |rstudent|:
##     rstudent unadjusted p-value Bonferroni p
## 38 -1.474992            0.14733           NA
```

```
qqPlot(fit, main="QQ Plot")
```



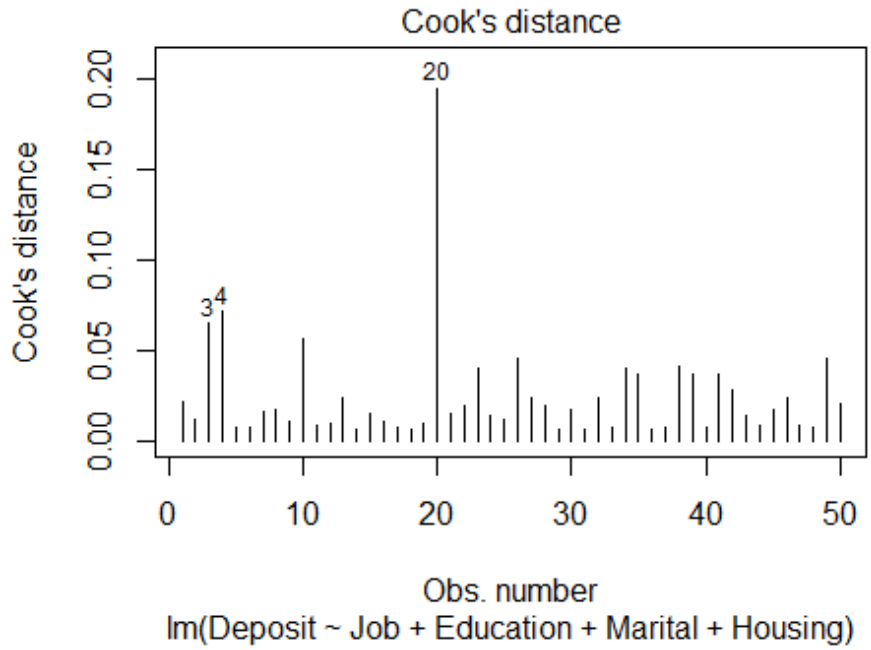

```
## [1] 23 38
```

```
leveragePlots(fit) # leverage plots
```

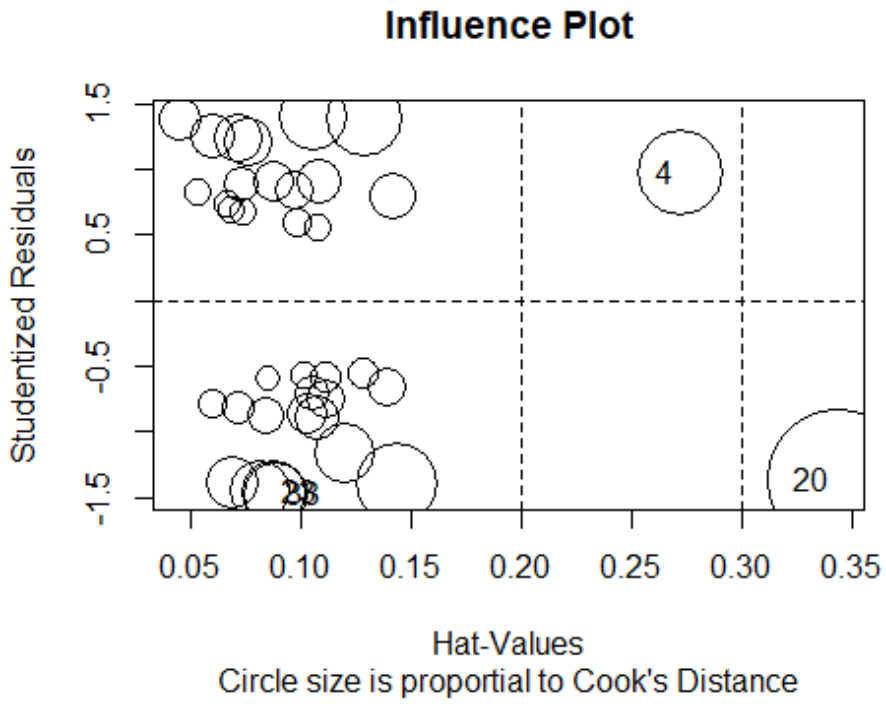



```
# Influential Observations
# added variable plots
avPlots(fit)
```

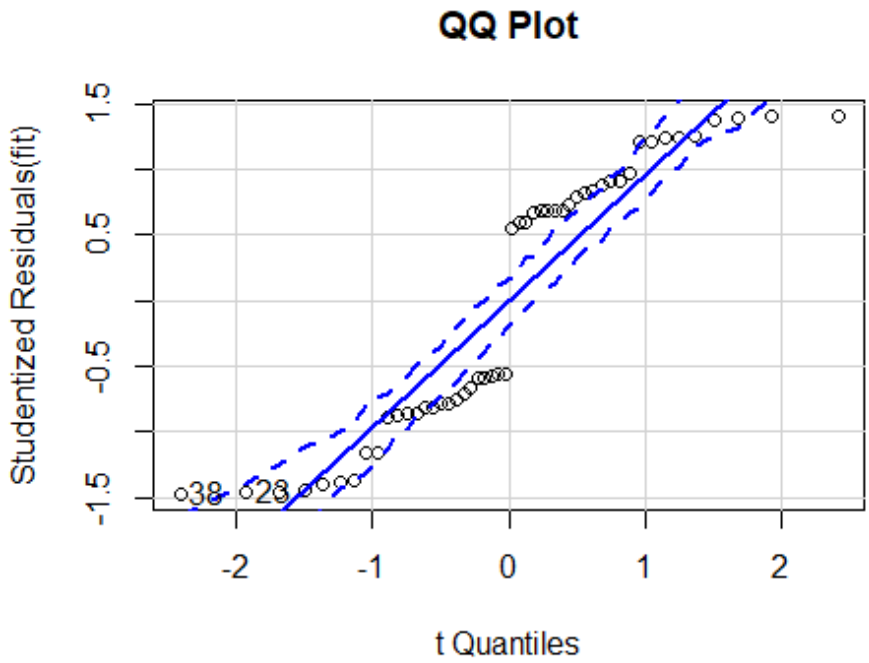
```r
# Cook's D plot
# identify D values > 4/(n-k-1)
cutoff <- 4/((nrow(bank)-length(fit$coefficients)-2))
plot(fit, which=4, cook.levels=cutoff)
```



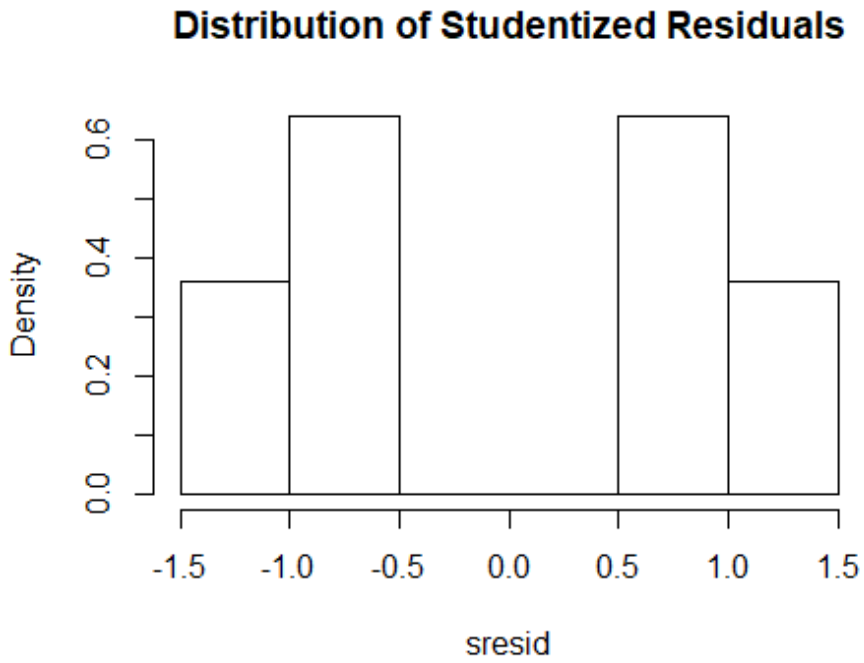

```r
# Influence Plot

influencePlot(fit,  main="Influence Plot", sub="Circle size is proportial to
Cook's Distance" )
```



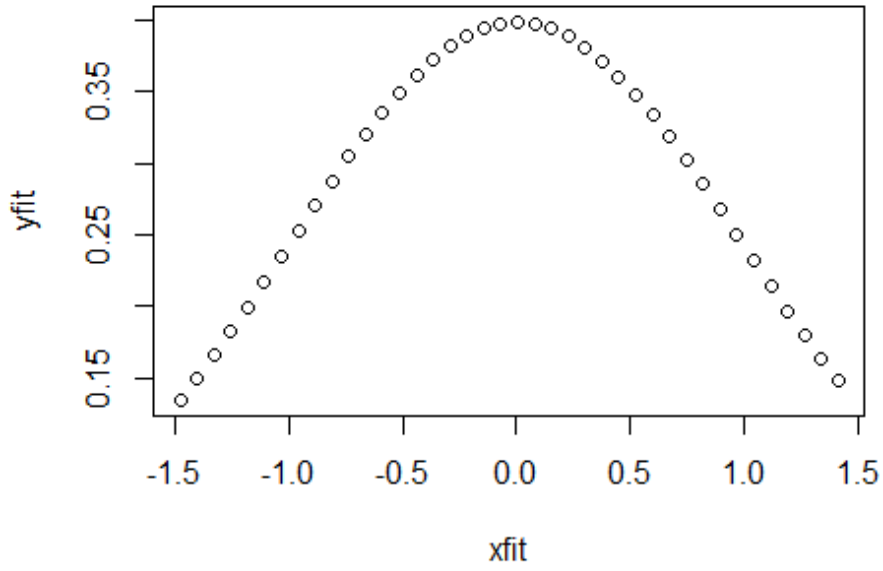

```
##       StudRes        Hat      CookD
## 4   0.9725172 0.27283088 0.07105688
## 20 -1.3739116 0.34367339 0.19386095
## 23 -1.4637723 0.08721352 0.03993023
## 38 -1.4749916 0.08811839 0.04097671

# Normality of Residuals
# qq plot for studentized resid
qqPlot(fit, main="QQ Plot")
```




```
## [1] 23 38

# distribution of studentized residuals
library(MASS)
sresid <- studres(fit)

hist(sresid, freq=FALSE,
     main="Distribution of Studentized Residuals")
```



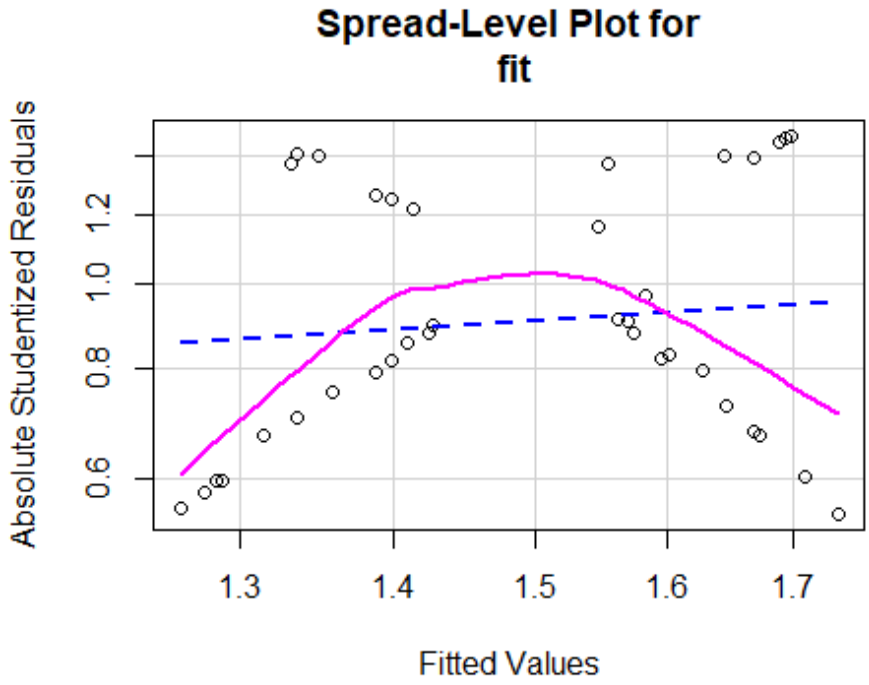

```
xfit<-seq(min(sresid),max(sresid),length=40)
yfit<-dnorm(xfit)
plot(xfit, yfit)
```



```r
#Non-constant Error Variance
# Evaluate homoscedasticity
# non-constant error variance test
ncvTest(fit)
```

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 0.08626249, Df = 1, p = 0.76898
```

```r
# plot studentized residuals vs. fitted values
spreadLevelPlot(fit)
```




```
##
## Suggested power transformation:  0.6605265
```

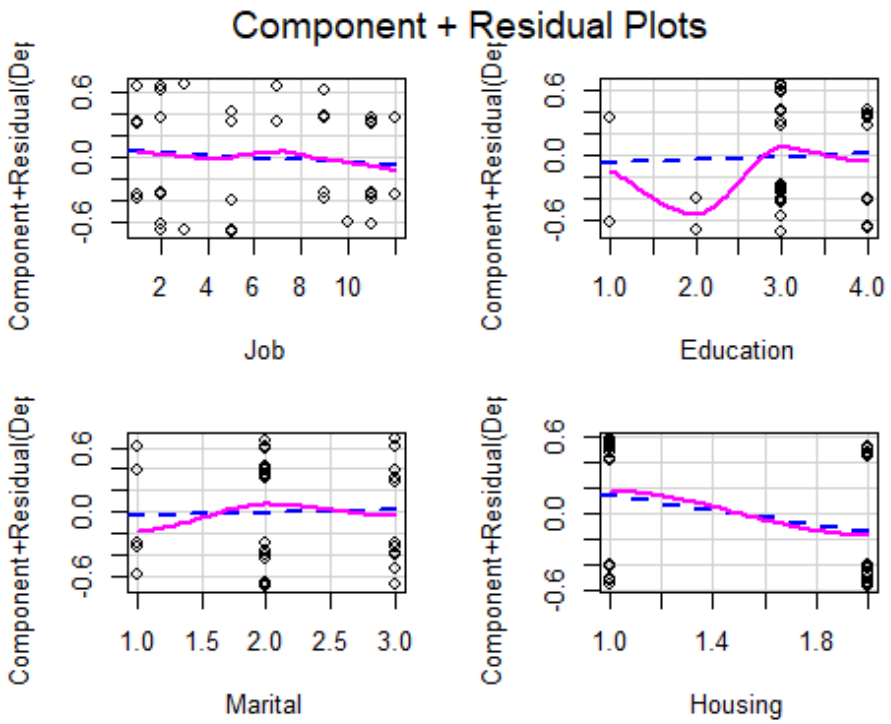
```r
#Multi-collinearity
# Evaluate Collinearity
vif(fit) # variance inflation factors
```

```
##       Job Education   Marital   Housing
##  1.088143  1.033562  1.077491  1.040166
```

```r
sqrt(vif(fit)) > 2
```

```
##       Job Education   Marital   Housing
##     FALSE     FALSE     FALSE     FALSE
```

```r
#Nonlinearity
# component + residual plot
crPlots(fit)
```




```r
#Non-independence of Errors
# Test for Autocorrelated Errors
durbinWatsonTest(fit)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1       0.1826723       1.58951   0.132
##  Alternative hypothesis: rho != 0
```

```r
library(gvlma)
gvmodel <- gvlma(fit)
summary(gvmodel)
```

```
##
## Call:
## lm(formula = Deposit ~ Job + Education + Marital + Housing, data = bank)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69887 -0.40662 -0.00031  0.41392  0.66734
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.84223    0.51916   3.548  0.00092 ***
```

```
## Job          -0.01102    0.01901  -0.580  0.56488
## Education     0.02834    0.10568   0.268  0.78977
## Marital       0.02619    0.11123   0.235  0.81494
## Housing      -0.28491    0.14499  -1.965  0.05560 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5026 on 45 degrees of freedom
## Multiple R-squared:  0.09056,    Adjusted R-squared:  0.009723
## F-statistic:  1.12 on 4 and 45 DF,  p-value: 0.3589
##
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance =  0.05
##
## Call:
##  gvlma(x = fit)
##
##                       Value p-value                   Decision
## Global Stat        6.967998 0.13759    Assumptions acceptable.
## Skewness           0.006324 0.93662    Assumptions acceptable.
## Kurtosis           5.480848 0.01923 Assumptions NOT satisfied!
## Link Function      0.977229 0.32288    Assumptions acceptable.
## Heteroscedasticity 0.503597 0.47792    Assumptions acceptable.

fit

##
## Call:
## lm(formula = Deposit ~ Job + Education + Marital + Housing, data = bank)
##
## Coefficients:
## (Intercept)          Job    Education      Marital      Housing
##     1.84223     -0.01102      0.02834      0.02619     -0.28491

summary(fit)

##
## Call:
## lm(formula = Deposit ~ Job + Education + Marital + Housing, data = bank)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69887 -0.40662 -0.00031  0.41392  0.66734
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.84223    0.51916   3.548  0.00092 ***
## Job         -0.01102    0.01901  -0.580  0.56488
## Education    0.02834    0.10568   0.268  0.78977
```

```
## Marital      0.02619    0.11123   0.235  0.81494
## Housing     -0.28491    0.14499  -1.965  0.05560 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5026 on 45 degrees of freedom
## Multiple R-squared:  0.09056,    Adjusted R-squared:  0.009723
## F-statistic:  1.12 on 4 and 45 DF,  p-value: 0.3589
```

```
fit1 <- fit
```

```
fit2 <- lm(Deposit~Job+Education+Marital, data = bank)
```

```
# compare models
anova(fit1, fit2)
```

```
## Analysis of Variance Table
##
## Model 1: Deposit ~ Job + Education + Marital + Housing
## Model 2: Deposit ~ Job + Education + Marital
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     45 11.368
## 2     46 12.344 -1  -0.97551 3.8615 0.0556 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
step <- stepAIC(fit, direction="both")
```

```
## Start:  AIC=-64.06
## Deposit ~ Job + Education + Marital + Housing
##
##             Df Sum of Sq    RSS     AIC
## - Marital    1   0.01400 11.382 -66.000
## - Education  1   0.01817 11.386 -65.981
## - Job        1   0.08495 11.453 -65.689
## <none>                   11.368 -64.061
## - Housing    1   0.97551 12.344 -61.945
##
## Step:  AIC=-66
## Deposit ~ Job + Education + Housing
##
##             Df Sum of Sq    RSS     AIC
## - Education  1   0.02111 11.403 -67.907
## - Job        1   0.11005 11.492 -67.518
## <none>                   11.382 -66.000
## + Marital    1   0.01400 11.368 -64.061
## - Housing    1   0.98038 12.362 -63.868
##
## Step:  AIC=-67.91
## Deposit ~ Job + Housing
##
```

```
##             Df Sum of Sq    RSS     AIC
## - Job        1   0.11691 11.520 -69.397
## <none>                   11.403 -67.907
## + Education  1   0.02111 11.382 -66.000
## + Marital    1   0.01694 11.386 -65.981
## - Housing    1   1.04960 12.453 -65.504
##
## Step:  AIC=-69.4
## Deposit ~ Housing
##
##             Df Sum of Sq    RSS     AIC
## <none>                   11.520 -69.397
## + Job        1   0.11691 11.403 -67.907
## + Marital    1   0.04545 11.475 -67.595
## + Education  1   0.02797 11.492 -67.518
## - Housing    1   0.98000 12.500 -67.315
```

```
step$anova # display results
```

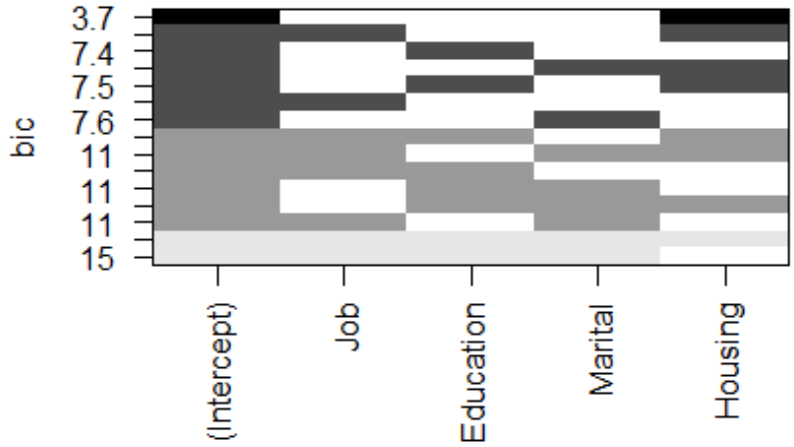
```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## Deposit ~ Job + Education + Marital + Housing
##
## Final Model:
## Deposit ~ Housing
##
##
##          Step Df   Deviance Resid. Df Resid. Dev       AIC
## 1                                  45   11.36797 -64.06116
## 2   - Marital  1 0.01400192        46   11.38197 -65.99961
## 3 - Education  1 0.02111442        47   11.40309 -67.90694
## 4       - Job  1 0.11691257        48   11.52000 -69.39692
```

```
library(leaps)
leaps<-regsubsets(Deposit~Job+Education+Marital+Housing,data=bank,nbest=10)

# view results
summary(leaps)
```

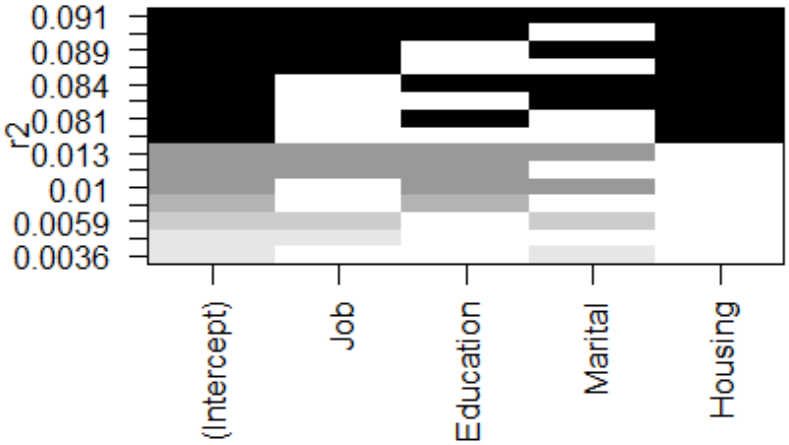
```
## Subset selection object
## Call: regsubsets.formula(Deposit ~ Job + Education + Marital + Housing,
##     data = bank, nbest = 10)
## 4 Variables  (and intercept)
##           Forced in Forced out
## Job           FALSE      FALSE
## Education     FALSE      FALSE
## Marital       FALSE      FALSE
## Housing       FALSE      FALSE
## 10 subsets of each size up to 4
```

```
## Selection Algorithm: exhaustive
##          Job Education Marital Housing
## 1  ( 1 ) " " " "       " "     "*"
## 1  ( 2 ) " " "*"       " "     " "
## 1  ( 3 ) "*" " "       " "     " "
## 1  ( 4 ) " " " "       "*"     " "
## 2  ( 1 ) "*" " "       " "     "*"
## 2  ( 2 ) " " " "       "*"     "*"
## 2  ( 3 ) " " "*"       " "     "*"
## 2  ( 4 ) "*" "*"       " "     " "
## 2  ( 5 ) " " "*"       "*"     " "
## 2  ( 6 ) "*" " "       "*"     " "
## 3  ( 1 ) "*" "*"       " "     "*"
## 3  ( 2 ) "*" " "       "*"     "*"
## 3  ( 3 ) " " "*"       "*"     "*"
## 3  ( 4 ) "*" "*"       "*"     " "
## 4  ( 1 ) "*" "*"       "*"     "*"
```

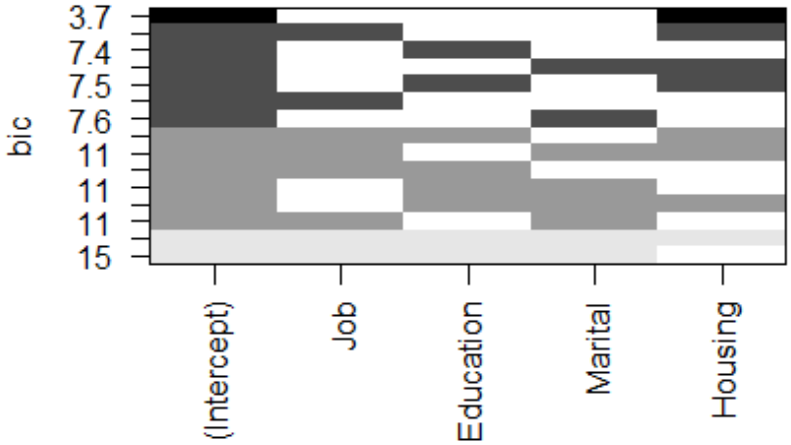
```
# plot a table of models showing variables in each model.
# models are ordered by the selection statistic.
plot(leaps)
```




```
plot(leaps,scale="r2")
```




```
#subsets(leaps, statistic="rsq")

# All Subsets Regression
plot(leaps,scale="bic")
```




```
summary(leaps)
```

```
## Subset selection object
## Call: regsubsets.formula(Deposit ~ Job + Education + Marital + Housing,
##     data = bank, nbest = 10)
## 4 Variables  (and intercept)
##           Forced in Forced out
## Job           FALSE      FALSE
## Education     FALSE      FALSE
## Marital       FALSE      FALSE
## Housing       FALSE      FALSE
## 10 subsets of each size up to 4
## Selection Algorithm: exhaustive
##          Job Education Marital Housing
## 1  ( 1 ) " " " "       " "     "*"
## 1  ( 2 ) " " "*"       " "     " "
## 1  ( 3 ) "*" " "       " "     " "
## 1  ( 4 ) " " " "       "*"     " "
## 2  ( 1 ) "*" " "       " "     "*"
## 2  ( 2 ) " " " "       "*"     "*"
## 2  ( 3 ) " " "*"       " "     "*"
## 2  ( 4 ) "*" "*"       " "     " "
## 2  ( 5 ) " " "*"       "*"     " "
## 2  ( 6 ) "*" " "       "*"     " "
## 3  ( 1 ) "*" "*"       " "     "*"
## 3  ( 2 ) "*" " "       "*"     "*"
## 3  ( 3 ) " " "*"       "*"     "*"
```

```
## 3  ( 4 ) "*" "*"       "*"     " "
## 4  ( 1 ) "*" "*"       "*"     "*"
```

```
summary(leaps)
```

```
## Subset selection object
## Call: regsubsets.formula(Deposit ~ Job + Education + Marital + Housing,
##     data = bank, nbest = 10)
## 4 Variables  (and intercept)
##           Forced in Forced out
## Job           FALSE      FALSE
## Education     FALSE      FALSE
## Marital       FALSE      FALSE
## Housing       FALSE      FALSE
## 10 subsets of each size up to 4
## Selection Algorithm: exhaustive
##          Job Education Marital Housing
## 1  ( 1 ) " " " "       " "     "*"
## 1  ( 2 ) " " "*"       " "     " "
## 1  ( 3 ) "*" " "       " "     " "
## 1  ( 4 ) " " " "       "*"     " "
## 2  ( 1 ) "*" " "       " "     "*"
## 2  ( 2 ) " " " "       "*"     "*"
## 2  ( 3 ) " " "*"       " "     "*"
## 2  ( 4 ) "*" "*"       " "     " "
## 2  ( 5 ) " " "*"       "*"     " "
## 2  ( 6 ) "*" " "       "*"     " "
## 3  ( 1 ) "*" "*"       " "     "*"
## 3  ( 2 ) "*" " "       "*"     "*"
## 3  ( 3 ) " " "*"       "*"     "*"
## 3  ( 4 ) "*" "*"       "*"     " "
## 4  ( 1 ) "*" "*"       "*"     "*"
```

```
leaps
```

```
## Subset selection object
## Call: regsubsets.formula(Deposit ~ Job + Education + Marital + Housing,
##     data = bank, nbest = 10)
## 4 Variables  (and intercept)
##           Forced in Forced out
## Job           FALSE      FALSE
## Education     FALSE      FALSE
## Marital       FALSE      FALSE
## Housing       FALSE      FALSE
## 10 subsets of each size up to 4
## Selection Algorithm: exhaustive
```

```
coef(leaps,1:4)
```

```
## [[1]]
## (Intercept)     Housing
```

```
##        1.92       -0.28
##
## [[2]]
## (Intercept)   Education
##   1.2959795   0.0641574
##
## [[3]]
##  (Intercept)          Job
##  1.544790243 -0.007885606
##
## [[4]]
## (Intercept)     Marital
##  1.40000000  0.04545455
```

```
# Calculate Relative Importance for Each Predictor
library(relaimpo)
```

```
## Loading required package: boot
```

```
##
## Attaching package: 'boot'
```

```
## The following object is masked from 'package:car':
##
##     logit
```

```
## The following object is masked from 'package:lattice':
##
##     melanoma
```

```
## Loading required package: survey
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
## Loading required package: survival
```

```
##
## Attaching package: 'survival'
```

```
## The following object is masked from 'package:boot':
##
##     aml
```

```
## The following object is masked from 'package:caret':
##
##     cluster
```

```
##
## Attaching package: 'survey'

## The following object is masked from 'package:graphics':
##
##     dotchart

## Loading required package: mitools

## This is the global version of package relaimpo.

## If you are a non-US user, a version with the interesting additional metric
pmvd is available

## from Ulrike Groempings web site at prof.beuth-hochschule.de/groemping.

calc.relimp(fit,type=c("lmg","last","first","pratt"),
            rela=TRUE)

## Warning in rev(variances[[p]]) - variances[[p + 1]]: Recycling array of
length 1 in vector-array arithmetic is deprecated.
##   Use c() or as.vector() instead.

## Response variable: Deposit
## Total response variance: 0.255102
## Analysis based on 50 observations
##
## 4 Regressors:
## Job Education Marital Housing
## Proportion of variance explained by model: 9.06%
## Metrics are normalized to sum to 100% (rela=TRUE).
##
## Relative importance metrics:
##
##                   lmg       last      first      pratt
## Job       0.05952197 0.07774833 0.04047345 0.05841995
## Education 0.04942479 0.01663083 0.08232314 0.03755552
## Marital   0.02644906 0.01281485 0.03888313 0.02313299
## Housing   0.86460417 0.89280600 0.83832028 0.88089154
##
## Average coefficients for different model sizes:
##
##                      1X          2Xs          3Xs         4Xs
## Job       -0.007885606 -0.008736321 -0.009824594 -0.01102218
## Education  0.064157399  0.052546113  0.040345255  0.02834262
## Marital    0.045454545  0.040292175  0.033822664  0.02618722
## Housing   -0.280000000 -0.281666008 -0.283392255 -0.28491277

summary(fit)

##
## Call:
```

```
## lm(formula = Deposit ~ Job + Education + Marital + Housing, data = bank)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69887 -0.40662 -0.00031  0.41392  0.66734
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.84223    0.51916   3.548  0.00092 ***
## Job         -0.01102    0.01901  -0.580  0.56488
## Education    0.02834    0.10568   0.268  0.78977
## Marital      0.02619    0.11123   0.235  0.81494
## Housing     -0.28491    0.14499  -1.965  0.05560 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5026 on 45 degrees of freedom
## Multiple R-squared:  0.09056,    Adjusted R-squared:  0.009723
## F-statistic:  1.12 on 4 and 45 DF,  p-value: 0.3589
```

```
predict.lm(fit, data.frame(wt =3.2 ,drat=3.9,hp=130,disp=150) )
```

```
## Warning: 'newdata' had 1 row but variables found have 50 rows
```

```
##        1        2        3        4        5        6        7        8
## 1.628009 1.573479 1.644336 1.582927 1.595524 1.277544 1.562457 1.409229
##        9       10       11       12       13       14       15       16
## 1.398788 1.350557 1.284424 1.262379 1.398788 1.288567 1.314754 1.398788
##       17       18       19       20       21       22       23       24
## 1.387766 1.667955 1.262379 1.554003 1.601822 1.427131 1.694142 1.359423
##       25       26       27       28       29       30       31       32
## 1.336798 1.336798 1.413953 1.569336 1.667955 1.332655 1.667955 1.413953
##       33       34       35       36       37       38       39       40
## 1.738231 1.694142 1.689999 1.667955 1.387766 1.698866 1.547292 1.672679
##       41       42       43       44       45       46       47       48
## 1.547292 1.667955 1.424976 1.709888 1.409229 1.398788 1.709888 1.645911
##       49       50
## 1.336798 1.387766
```