Airline Output

NIKHIL CHAVAN

Nschavan1996@gmail.com

```
#Loading the Data set
airdata.df <-read.csv(paste("SixAirlines.csv", sep=""))

##Attaching the Data set
attach(airdata.df)

#General view of the data
View(airdata.df)

#quick summary analysis of data
library(psych)
describe(airdata.df)
```

```
##                  vars   n    mean      sd  median trimmed     mad   min
## AIRLINE*            1 462    3.02    1.65    2.00    2.90    1.48  1.00
## AIRCRAFT            2 462    0.33    0.47    0.00    0.28    0.00  0.00
## FLIGHT_DURATION     3 462    7.55    3.54    7.75    7.54    4.82  1.25
## MONTH               4 462    1.67    1.05    2.00    1.71    1.48  0.00
## INTERNATIONAL       5 462    0.91    0.28    1.00    1.00    0.00  0.00
## SEATS_ECONOMY       6 462  200.71   77.96  185.00  193.76   85.99 17.00
## SEATS_PREMIUM       7 462   33.54   13.26   36.00   33.20   11.86  8.00
## PITCH_ECONOMY       8 462   31.21    0.66   31.00   31.25    0.00 30.00
## PITCH_PREMIUM       9 462   37.92    1.32   38.00   38.06    0.00 34.00
## WIDTH_ECONOMY      10 462   17.83    0.56   18.00   17.81    0.00 17.00
## WIDTH_PREMIUM      11 462   19.48    1.10   19.00   19.54    0.00 17.00
## PRICE_ECONOMY      12 462 1317.06  989.81 1224.00 1231.30 1163.84 65.00
## PRICE_PREMIUM      13 462 1832.35 1289.97 1710.00 1782.94 1852.51 86.00
## PRICE_RELATIVE     14 462    0.49    0.45    0.38    0.43    0.42  0.02
## N                  15 462  234.25   86.88  227.00  227.69   90.44 38.00
## LAMBDA             16 462    0.15    0.06    0.13    0.14    0.03  0.05
## QUALITY            17 462    6.72    1.78    7.00    6.79    0.00  2.00
##                     max   range  skew kurtosis    se
## AIRLINE*           6.00    5.00  0.59    -0.95  0.08
## AIRCRAFT           1.00    1.00  0.74    -1.46  0.02
## FLIGHT_DURATION   14.66   13.41 -0.05    -1.12  0.16
## MONTH              3.00    3.00 -0.16    -1.20  0.05
## INTERNATIONAL      1.00    1.00 -2.93     6.60  0.01
## SEATS_ECONOMY    389.00  372.00  0.61    -0.26  3.63
## SEATS_PREMIUM     66.00   58.00  0.25    -0.46  0.62
## PITCH_ECONOMY     33.00    3.00 -0.03    -0.38  0.03
## PITCH_PREMIUM     40.00    6.00 -1.48     3.43  0.06
## WIDTH_ECONOMY     19.00    2.00 -0.03    -0.12  0.03
```

```
## WIDTH_PREMIUM     21.00    4.00 -0.09    -0.34  0.05
## PRICE_ECONOMY   3593.00 3528.00  0.52    -0.88 46.05
## PRICE_PREMIUM   7414.00 7328.00  0.51     0.41 60.01
## PRICE_RELATIVE     1.89    1.87  1.14     0.61  0.02
## N                441.00  403.00  0.61    -0.44  4.04
## LAMBDA             0.55    0.50  2.70    14.02  0.00
## QUALITY           10.00    8.00 -0.51     1.67  0.08
```

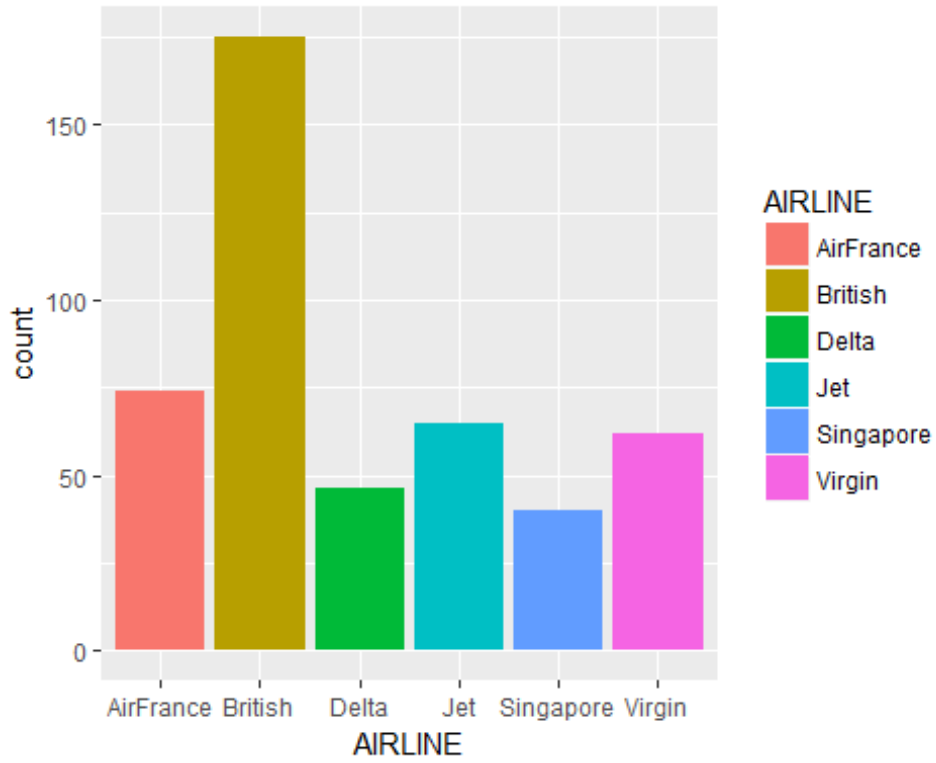
```
library(ggplot2)
```

```
##
## Attaching package: 'ggplot2'
```

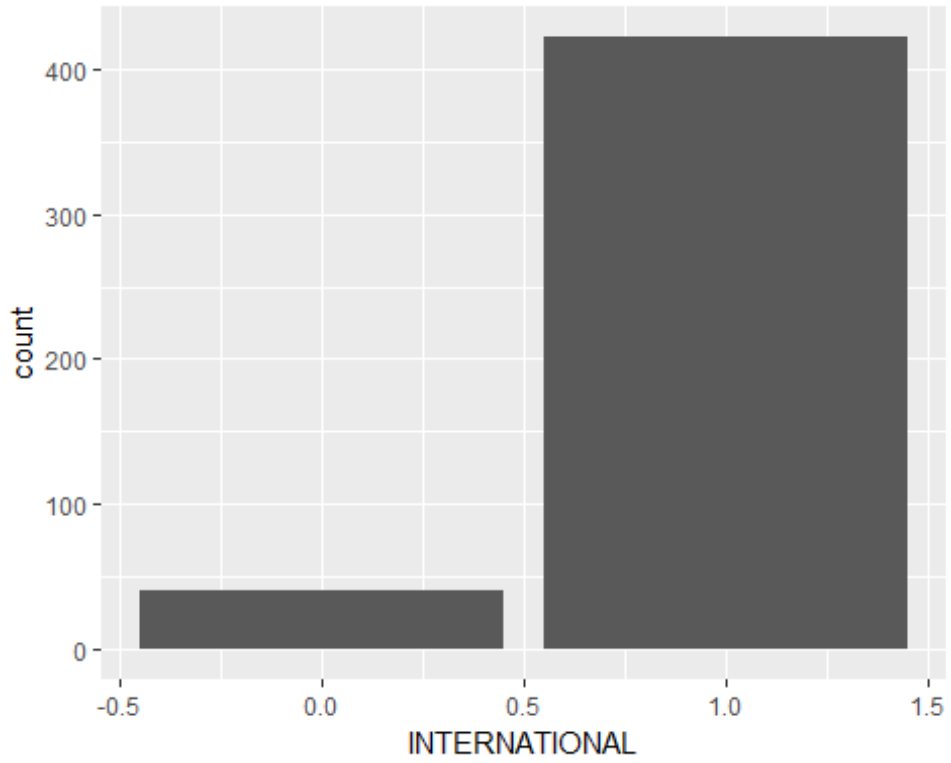
```
## The following objects are masked from 'package:psych':
##
##     %+%, alpha
```

```
## Loading required package: ggplot2
```

```
#Seggregating different flights
ggplot(airdata.df, aes(x = AIRLINE, fill = AIRLINE)) + geom_bar()
```

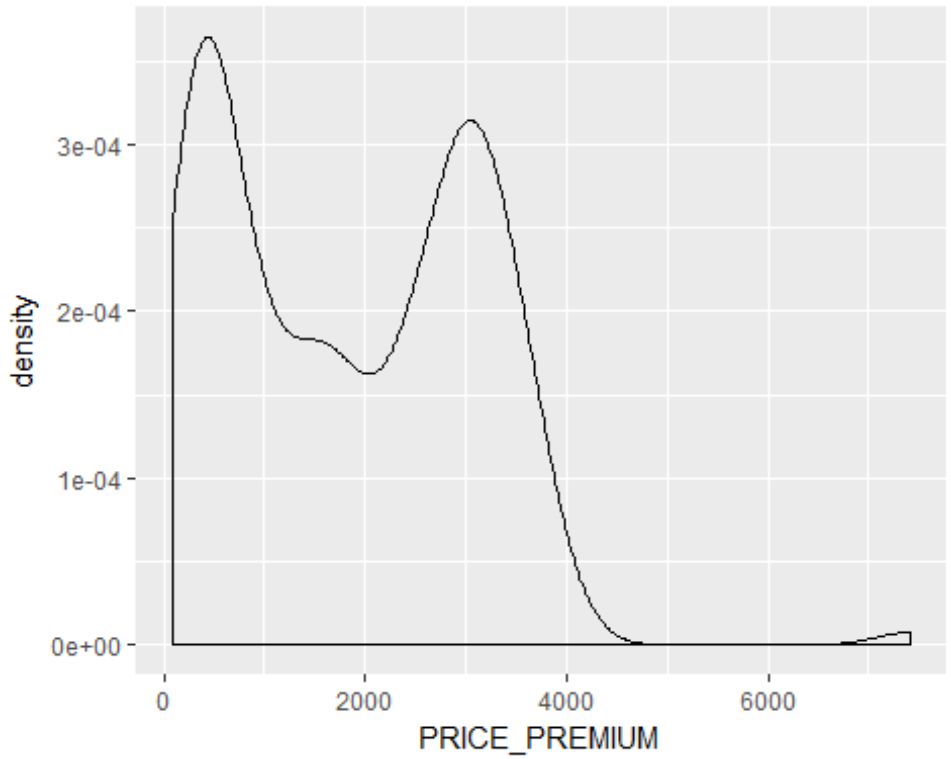



```
#Seggregating international and domestic flights
ggplot(airdata.df, aes(x =INTERNATIONAL))+ geom_bar()
```

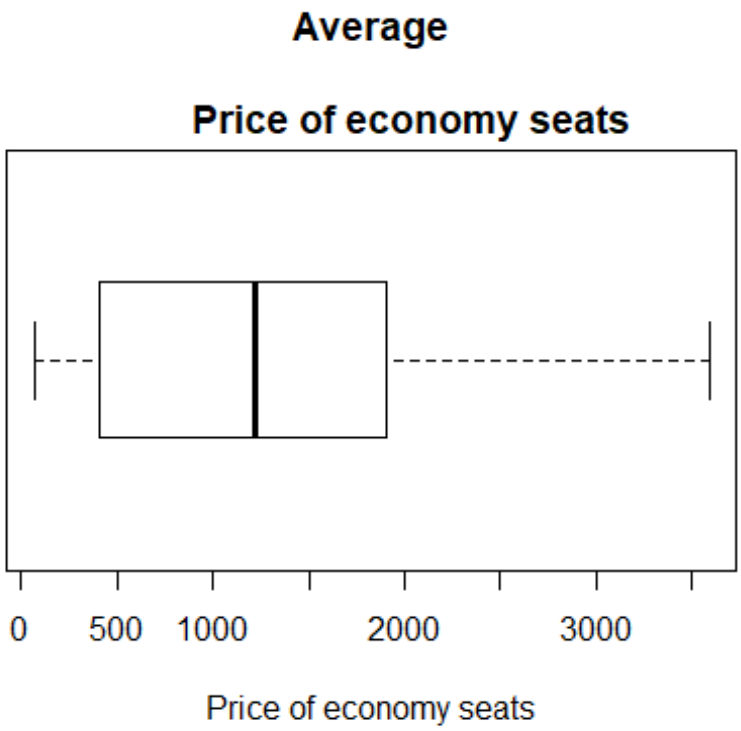



```r
#Prices of Economy and Premium tickets
ggplot(airdata.df, aes(x = PRICE_ECONOMY)) + geom_density()
```

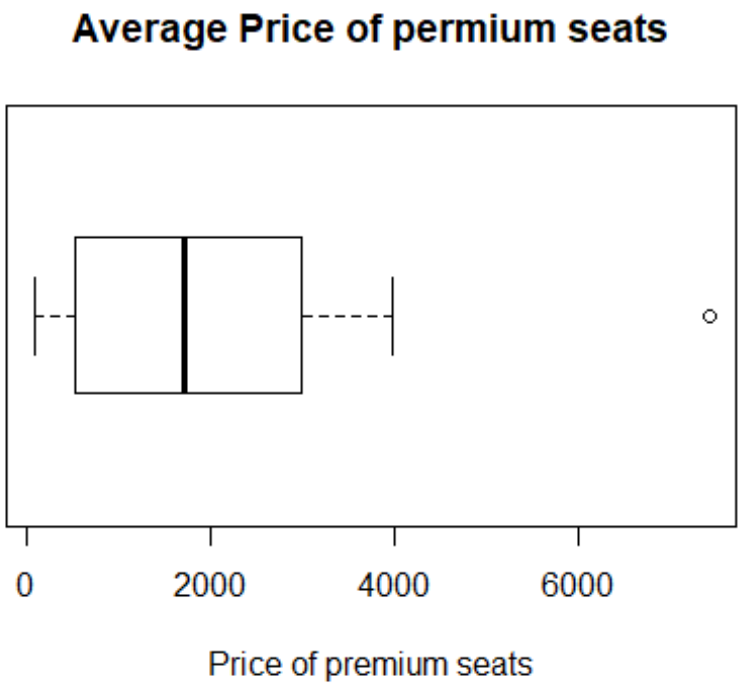
```
ggplot(airdata.df, aes(x = PRICE_PREMIUM)) + geom_density()
```



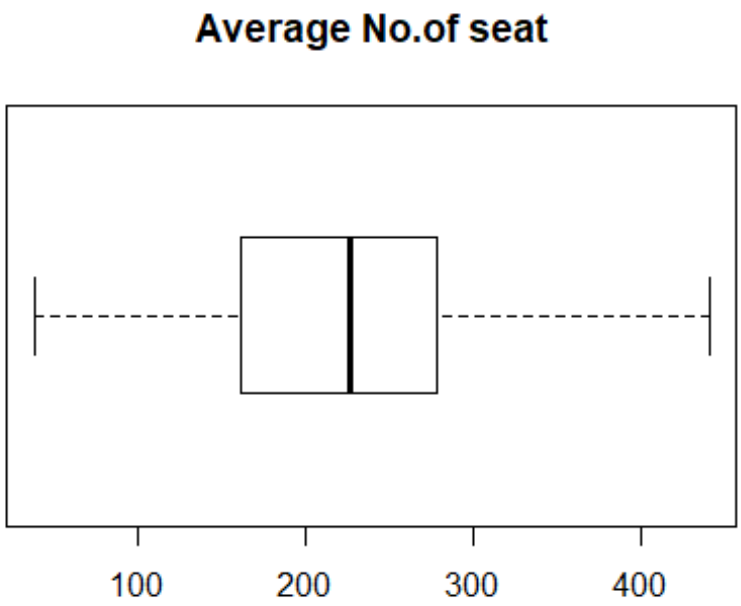

```
##BoxPlots to visualize data better

boxplot(PRICE_ECONOMY,horizontal=TRUE,xlab="Price of economy seats",main="Ave
rage

        Price of economy seats")
```



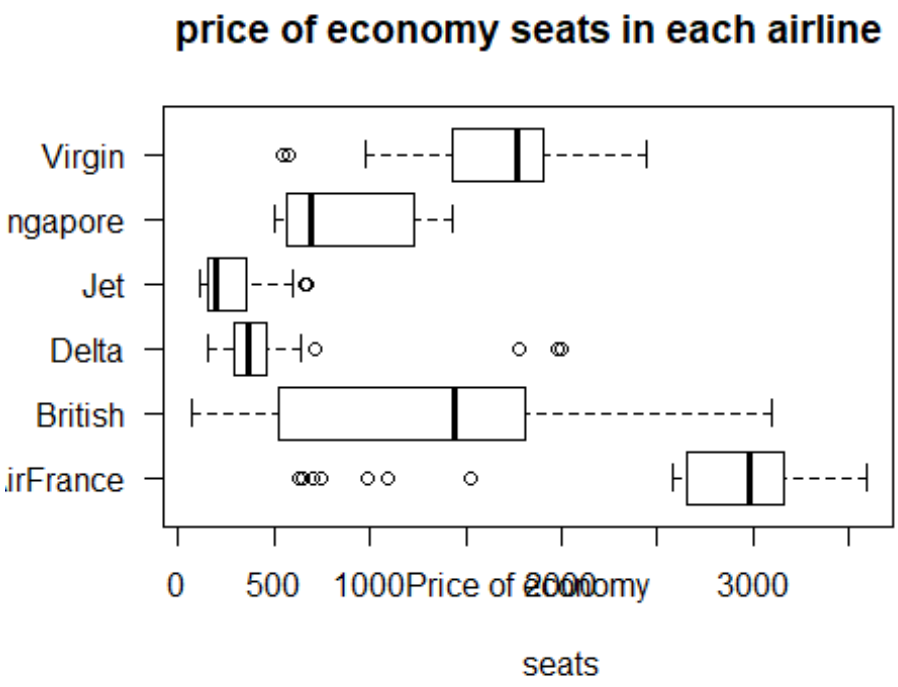

```
boxplot(PRICE_PREMIUM,horizontal=TRUE,xlab="Price of premium seats",main="Ave
rage Price of permium seats")
```



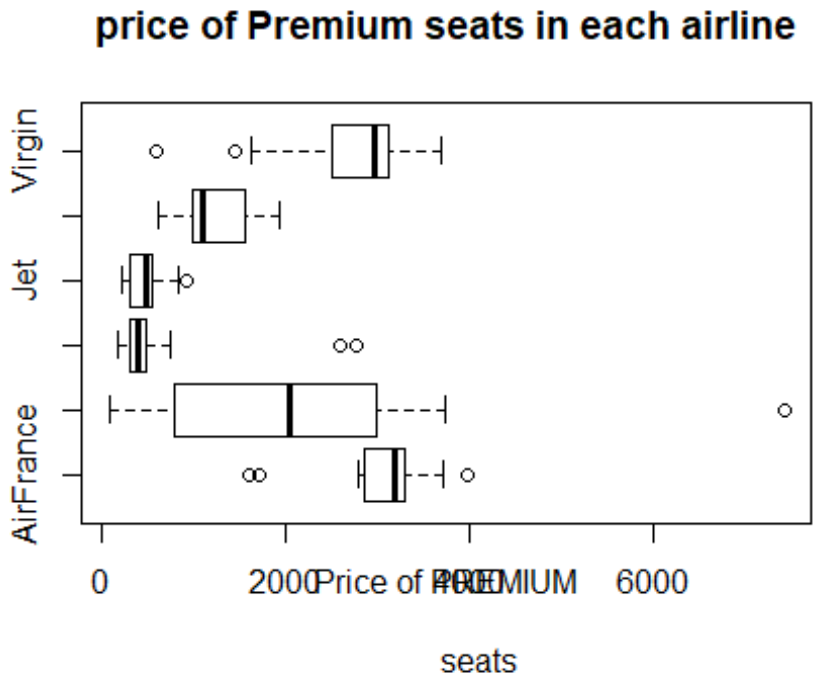
```
boxplot(N,horizontal=TRUE,main="Average No.of seat")
```



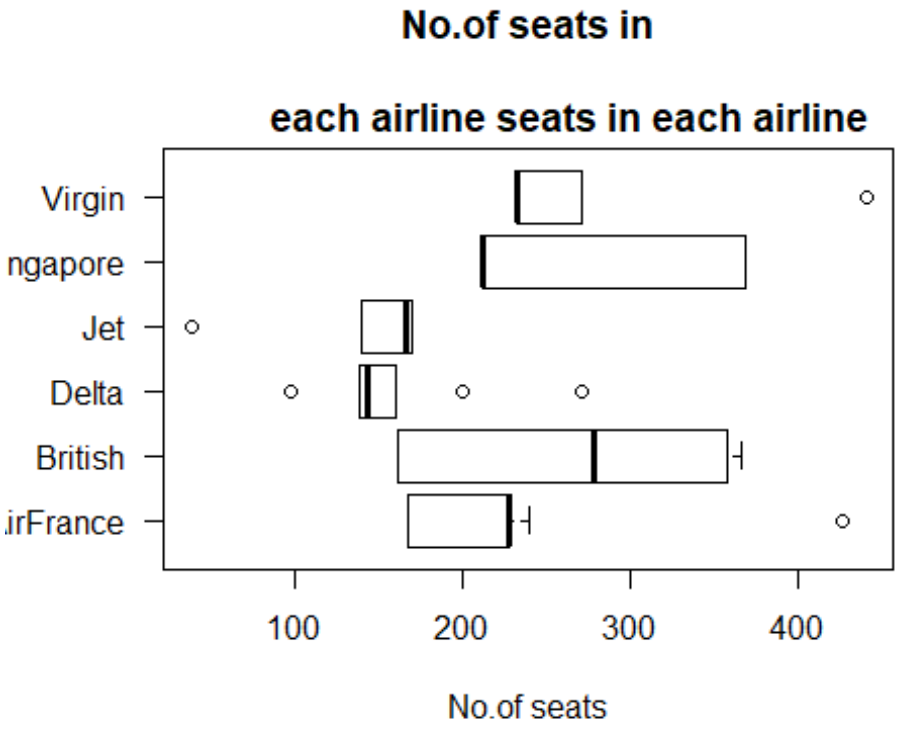

```
boxplot(PRICE_ECONOMY~AIRLINE,horizontal=TRUE,xlab="Price of economy

        seats",ytab="Airline",main="price of economy seats in each airline",l
as=1)
```



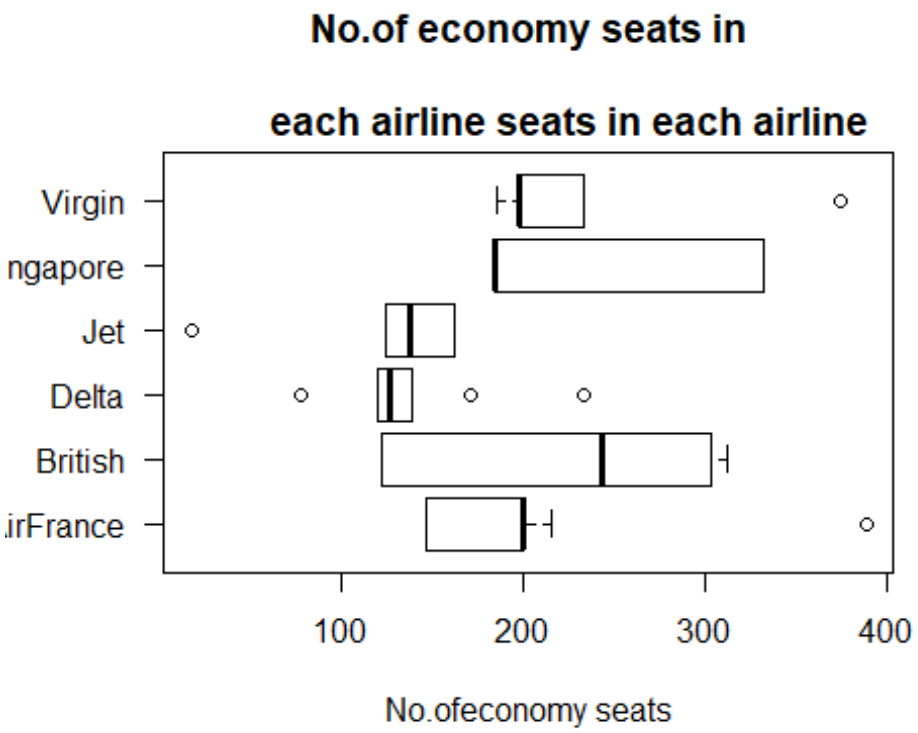

```
boxplot(PRICE_PREMIUM~AIRLINE,horizontal=TRUE,xlab="Price of PREMIUM
        seats",ytab="Airline",main="price of Premium seats in each airline")
```



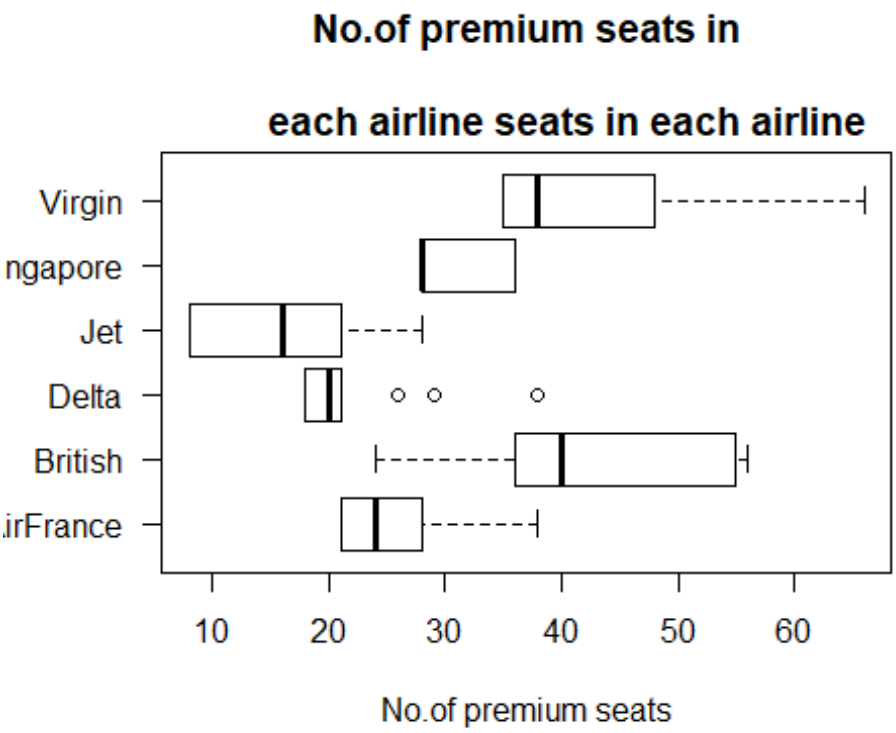
```
boxplot(N~AIRLINE,horizontal=TRUE,xlab="No.of seats",ytab="Airline",main="No.
of seats in

        each airline seats in each airline",las=1)
```



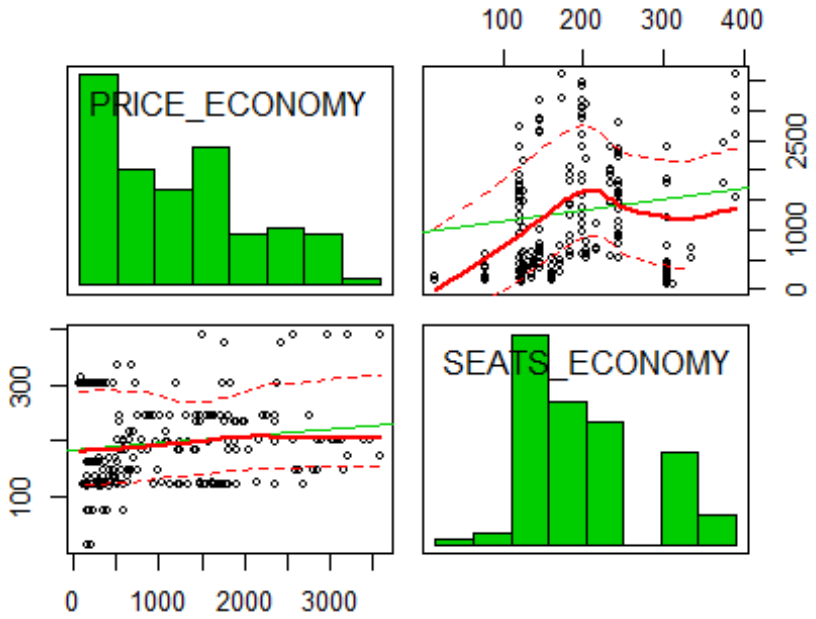

```
boxplot(SEATS_ECONOMY~AIRLINE,horizontal=TRUE,xlab="No.ofeconomy seats",ytab=
"Airline",main="No.of economy seats in

        each airline seats in each airline",las=1)
```



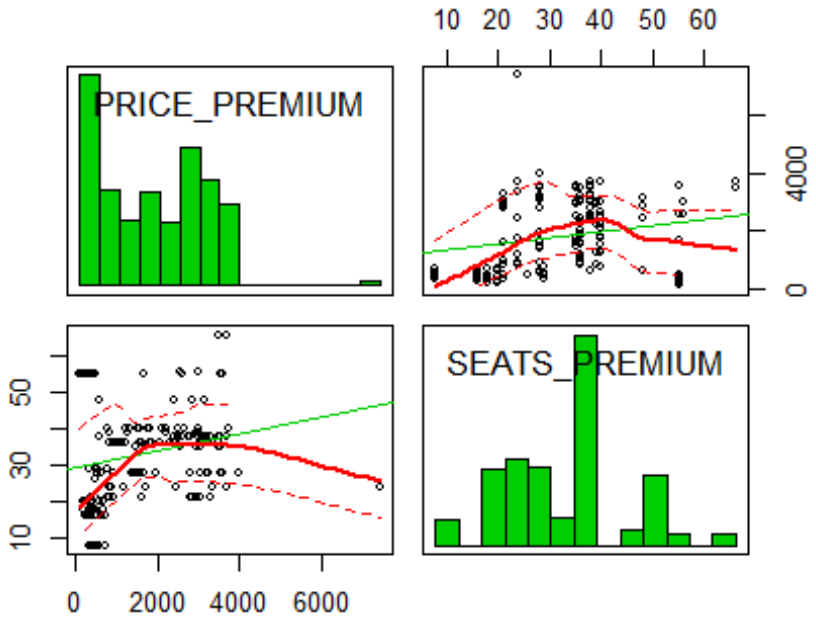

```
boxplot(SEATS_PREMIUM~AIRLINE,horizontal=TRUE,xlab="No.of premium seats",ytab
="Airline",main="No.of premium seats in

        each airline seats in each airline",las=1)
```



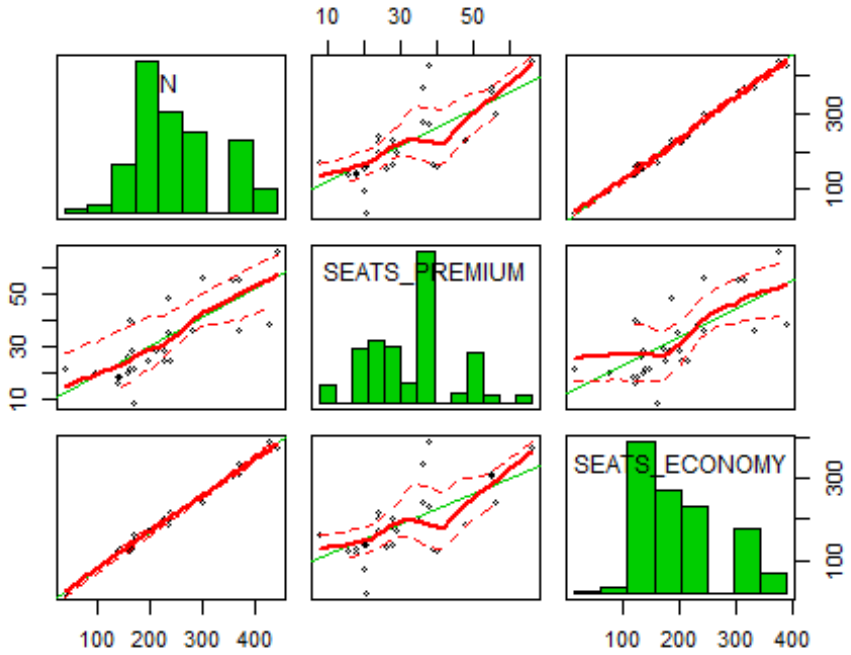

```
###scatterplots to understand correlations between variables

library(car)

##
## Attaching package: 'car'

## The following object is masked from 'package:psych':
##
##     logit

## Loading required package: car

scatterplotMatrix(formula=~PRICE_ECONOMY+SEATS_ECONOMY,cex=0.6,diagonal="hist
ogram")
```




```
scatterplotMatrix(formula=~PRICE_PREMIUM+SEATS_PREMIUM,cex=0.6,diagonal="hist
ogram")
```



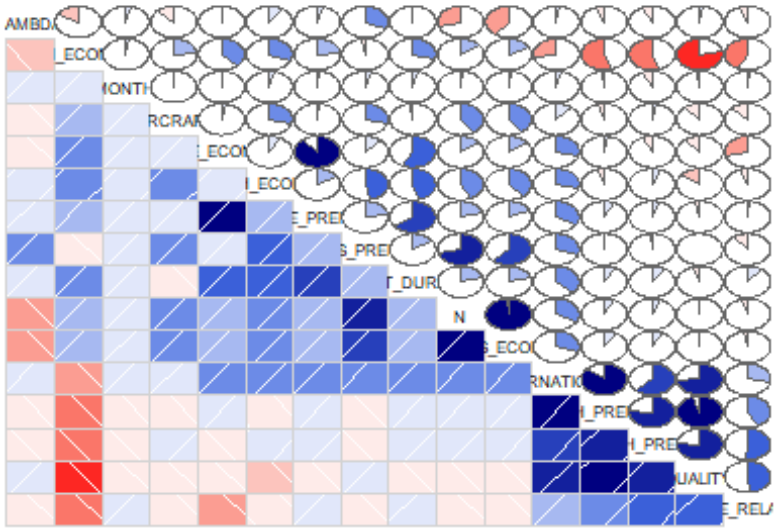
```
scatterplotMatrix(formula=~N+SEATS_PREMIUM+SEATS_ECONOMY,cex=0.6,diagonal="histogram")
```




```
#Calculating correlations between Prices of Economy and Premium in correlation to other factors

cor.test(PRICE_ECONOMY, PITCH_ECONOMY)

##
##  Pearson's product-moment correlation
##
## data:  PRICE_ECONOMY and PITCH_ECONOMY
## t = 8.8003, df = 460, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.2987210 0.4550742
## sample estimates:
##      cor
## 0.379605

cor.test(PRICE_ECONOMY, WIDTH_ECONOMY)

##
##  Pearson's product-moment correlation
##
## data:  PRICE_ECONOMY and WIDTH_ECONOMY
## t = 1.764, df = 460, p-value = 0.0784
```

```
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.009330795  0.171911298
## sample estimates:
##       cor
## 0.0819679

cor.test(PRICE_PREMIUM, PITCH_PREMIUM)

##
##  Pearson's product-moment correlation
##
## data:  PRICE_PREMIUM and PITCH_PREMIUM
## t = 1.5338, df = 460, p-value = 0.1258
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.02002801  0.16150915
## sample estimates:
##       cor
## 0.07133125

cor.test(PRICE_PREMIUM, WIDTH_PREMIUM)

##
##  Pearson's product-moment correlation
##
## data:  PRICE_PREMIUM and WIDTH_PREMIUM
## t = 1.0592, df = 460, p-value = 0.2901
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.04209336  0.13992426
## sample estimates:
##       cor
## 0.04932498

##Drawing corrgram

library(corrgram)
corrgram(airdata.df, main = "corrgram of Sixairplane variables", lower.panel
= panel.shade, upper.panel =

           panel.pie, text.panel = panel.txt,order=TRUE)
```

corrgram of Sixairplane variables




```
### Performing tTests
air1.df <--airdata.df[- c(1)]

cov(air1.df)

##                     AIRCRAFT FLIGHT_DURATION        MONTH INTERNATIONAL
## AIRCRAFT         0.220492812   -3.701285e-02  0.003643500  1.534402e-02
## FLIGHT_DURATION -0.037012846    1.253262e+01  0.159093163  3.690261e-01
## MONTH            0.003643500    1.590932e-01  1.106271892  1.266774e-02
## INTERNATIONAL    0.015344020    3.690261e-01  0.012667737  7.925552e-02
## SEATS_ECONOMY   14.696579993    5.776481e+01  1.120573570  6.755557e+00
## SEATS_PREMIUM    1.838648336    7.870071e+00  0.756880863  1.153055e+00
## PITCH_ECONOMY    0.075104938    7.107017e-01  0.014282897 -5.355382e-02
## PITCH_PREMIUM   -0.025077237    3.839608e-01 -0.011963452  3.209755e-01
## WIDTH_ECONOMY    0.076992422    9.176009e-01  0.020255233  4.608840e-02
## WIDTH_PREMIUM    0.019062644    3.547466e-01 -0.076579241  1.895747e-01
## PRICE_ECONOMY   17.323632983    1.999429e+03  3.101670564  8.333535e+01
## PRICE_PREMIUM   12.170728043    2.976982e+03 25.684686969  1.255051e+02
## PRICE_RELATIVE  -0.025507038    1.735896e-01  0.009525782  3.539182e-02
## N               16.535228329    6.563488e+01  1.877454433  7.908612e+00
## LAMBDA          -0.003692613    7.567372e-04  0.002478613  3.120451e-04
## QUALITY         -0.100182175   -3.267410e-01 -0.026246349  3.745293e-01
##                 SEATS_ECONOMY SEATS_PREMIUM PITCH_ECONOMY PITCH_PREMIUM
## AIRCRAFT            14.6965800    1.83864834   0.075104938  -0.025077237
## FLIGHT_DURATION     57.7648124    7.87007057   0.710701749   0.383960757
## MONTH                1.1205736    0.75688086   0.014282897  -0.011963452
## INTERNATIONAL        6.7555568    1.15305519  -0.053553821   0.320975481
```

```
## SEATS_ECONOMY    6077.6910021  647.73917514   9.091218976   8.522283573
## SEATS_PREMIUM     647.7391751  175.71539848  -0.162126377  -0.143462828
## PITCH_ECONOMY       9.0912190   -0.16212638   0.438290560  -0.491816210
## PITCH_PREMIUM       8.5222836   -0.14346283  -0.491816210   1.749129034
## WIDTH_ECONOMY      17.1089106    3.43175480   0.115418204  -0.032340761
## WIDTH_PREMIUM       6.0725695   -0.20548215  -0.400315520   1.099717347
## PRICE_ECONOMY   11433.0342001 1602.26644505 248.750035214  44.030960363
## PRICE_PREMIUM   19639.2185537 3847.35524129 204.823759754 121.694406100
## PRICE_RELATIVE     -0.8657126   -0.64405105  -0.130376370   0.256361007
## N                6725.4301772  823.45457363   8.929092599   8.378820745
## LAMBDA             -1.9029572    0.25651576  -0.006991858  -0.004355485
## QUALITY            -0.5689354    0.01866355  -0.930106770   2.240945244
##                  WIDTH_ECONOMY WIDTH_PREMIUM PRICE_ECONOMY PRICE_PREMIUM
## AIRCRAFT          0.076992422   0.019062644  1.732363e+01  1.217073e+01
## FLIGHT_DURATION   0.917600924   0.354746598  1.999429e+03  2.976982e+03
## MONTH             0.020255233  -0.076579241  3.101671e+00  2.568469e+01
## INTERNATIONAL     0.046088402   0.189574706  8.333535e+01  1.255051e+02
## SEATS_ECONOMY    17.108910612   6.072569513  1.143303e+04  1.963922e+04
## SEATS_PREMIUM     3.431754796  -0.205482153  1.602266e+03  3.847355e+03
## PITCH_ECONOMY     0.115418204  -0.400315520  2.487500e+02  2.048238e+02
## PITCH_PREMIUM    -0.032340761   1.099717347  4.403096e+01  1.216944e+02
## WIDTH_ECONOMY     0.314167394   0.038651154  4.547520e+01  1.179328e+02
## WIDTH_PREMIUM     0.038651154   1.213435877 -7.651962e+01  7.008986e+01
## PRICE_ECONOMY    45.475195087 -76.519621376  9.797165e+05  1.152371e+06
## PRICE_PREMIUM   117.932783052  70.089857359  1.152371e+06  1.664025e+06
## PRICE_RELATIVE   -0.015475111   0.255376323 -1.336887e+02  1.027099e+01
## N                20.540665408   5.867087359  1.303530e+04  2.348657e+04
## LAMBDA            0.003334085  -0.004853744 -3.434499e-01  2.801268e+00
## QUALITY          -0.147758966   1.500032867 -2.047191e+02 -8.312935e+01
##                  PRICE_RELATIVE             N        LAMBDA       QUALITY
## AIRCRAFT          -2.550704e-02    16.5352283 -0.0036926125 -1.001822e-01
## FLIGHT_DURATION    1.735896e-01    65.6348829  0.0007567372 -3.267410e-01
## MONTH              9.525782e-03     1.8774544  0.0024786132 -2.624635e-02
## INTERNATIONAL      3.539182e-02     7.9086120  0.0003120451  3.745293e-01
## SEATS_ECONOMY     -8.657126e-01  6725.4301772 -1.9029571513 -5.689354e-01
## SEATS_PREMIUM     -6.440510e-01   823.4545736  0.2565157619  1.866355e-02
## PITCH_ECONOMY     -1.303764e-01     8.9290926 -0.0069918585 -9.301068e-01
## PITCH_PREMIUM      2.563610e-01     8.3788207 -0.0043554854  2.240945e+00
## WIDTH_ECONOMY     -1.547511e-02    20.5406654  0.0033340846 -1.477590e-01
## WIDTH_PREMIUM      2.553763e-01     5.8670874 -0.0048537435  1.500033e+00
## PRICE_ECONOMY     -1.336887e+02 13035.3006451 -0.3434498690 -2.047191e+02
## PRICE_PREMIUM      1.027099e+01 23486.5737950  2.8012680884 -8.312935e+01
## PRICE_RELATIVE     2.052828e-01    -1.5097636 -0.0013872506  3.867374e-01
## N                 -1.509764e+00  7548.8847508 -1.6464413894 -5.502719e-01
## LAMBDA            -1.387251e-03    -1.6464414  0.0037413115  2.636373e-03
## QUALITY            3.867374e-01    -0.5502719  0.0026363730  3.171052e+00
```

```
t.test(N~AIRCRAFT)
```

```
##
##  Welch Two Sample t-test
##
## data:  N by AIRCRAFT
## t = -8.9886, df = 258.16, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -91.42114 -58.56314
## sample estimates:
## mean in group 0 mean in group 1
##        209.7363        284.7285
```

```
t.test(SEATS_PREMIUM~AIRCRAFT)
```

```
##
##  Welch Two Sample t-test
##
## data:  SEATS_PREMIUM by AIRCRAFT
## t = -6.4301, df = 274.73, p-value = 5.627e-10
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -10.891813  -5.785813
## sample estimates:
## mean in group 0 mean in group 1
##        30.81350        39.15232
```

```
t.test(SEATS_ECONOMY~AIRCRAFT)
```

```
##
##  Welch Two Sample t-test
##
## data:  SEATS_ECONOMY by AIRCRAFT
## t = -8.9447, df = 262.2, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -81.32609 -51.98057
## sample estimates:
## mean in group 0 mean in group 1
##        178.9228        245.5762
```

```
t.test(PRICE_PREMIUM~AIRCRAFT)
```

```
##
##  Welch Two Sample t-test
##
## data:  PRICE_PREMIUM by AIRCRAFT
## t = -0.43785, df = 309.53, p-value = 0.6618
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -303.2518  192.8561
## sample estimates:
```

```
## mean in group 0 mean in group 1
##        1814.305        1869.503

t.test(PRICE_ECONOMY~AIRCRAFT)

##
##  Welch Two Sample t-test
##
## data:  PRICE_ECONOMY by AIRCRAFT
## t = -0.79106, df = 288.69, p-value = 0.4296
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -274.0507  116.9151
## sample estimates:
## mean in group 0 mean in group 1
##        1291.386        1369.954

#Regression models

reg1 <-lm(PRICE_ECONOMY~FLIGHT_DURATION+SEATS_ECONOMY+PITCH_ECONOMY+WIDTH_ECO
NOMY+QUALITY+MONTH+AIRCRAFT+AIRLINE,data=airdata.df)
summary(reg1)

##
## Call:
## lm(formula = PRICE_ECONOMY ~ FLIGHT_DURATION + SEATS_ECONOMY +
##     PITCH_ECONOMY + WIDTH_ECONOMY + QUALITY + MONTH + AIRCRAFT +
##     AIRLINE, data = airdata.df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2130.19  -258.90    92.13   334.17  1194.38
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -1.079e+04  5.034e+03  -2.143 0.032632 *
## FLIGHT_DURATION   1.055e+02  9.317e+00  11.320  < 2e-16 ***
## SEATS_ECONOMY    -7.401e-01  3.912e-01  -1.892 0.059128 .
## PITCH_ECONOMY     3.324e+02  1.391e+02   2.390 0.017261 *
## WIDTH_ECONOMY     3.904e+01  1.001e+02   0.390 0.696744
## QUALITY           2.640e+02  6.126e+01   4.310 2.00e-05 ***
## MONTH            -3.562e+01  2.346e+01  -1.518 0.129607
## AIRCRAFT         -1.620e+02  6.773e+01  -2.392 0.017182 *
## AIRLINEBritish   -1.334e+03  1.285e+02 -10.376  < 2e-16 ***
## AIRLINEDelta     -8.897e+02  2.312e+02  -3.847 0.000137 ***
## AIRLINEJet       -2.449e+03  1.769e+02 -13.843  < 2e-16 ***
## AIRLINESingapore -2.119e+03  1.772e+02 -11.959  < 2e-16 ***
## AIRLINEVirgin    -1.128e+03  1.392e+02  -8.104 5.08e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 525 on 449 degrees of freedom
## Multiple R-squared:  0.726,  Adjusted R-squared:  0.7187
## F-statistic: 99.13 on 12 and 449 DF,  p-value: < 2.2e-16
```

```
fitted(reg1)
```

```
##          1          2          3          4          5          6
## 2594.51708 2558.90135 2760.18814 2630.13280 2795.80386 2985.84813
##          7          8          9         10         11         12
## 2950.23240 2831.41959 2665.74852 2791.74543 2756.12971 2720.51398
##         13         14         15         16         17         18
## 2743.63267 2708.01695 2672.40123 2636.78550 2743.63267 2708.01695
##         19         20         21         22         23         24
## 2672.40123 2636.78550 2602.54461 2566.92889 2531.31316 2441.36599
##         25         26         27         28         29         30
## 2405.75026 2632.97194 2514.19272 2478.57699 2633.53692 2871.65954
##         31         32         33         34         35         36
## 2836.04381 2800.42809 2936.65646 2901.04074 2865.42502 2769.19091
##         37         38         39         40         41         42
## 2733.57518 2697.95946 2721.07897 2597.92120 2562.30548 2572.96154
##         43         44         45         46         47         48
## 2546.83833 2511.22260 2637.60477 2646.03314 2950.76379 2915.14807
##         49         50         51         52         53         54
## 2879.53234 2843.91662 2785.43031 2749.81458 2714.19886 2687.02092
##         55         56         57         58         59         60
## 2833.28416 2797.66844 2769.43578 2733.82005 2831.17472 2795.55899
##         61         62         63         64         65         66
## 2769.43578 2733.82005 2905.65556 2870.03983 2834.42411 3207.31966
##         67         68         69         70         71         72
## 3171.70394 2989.63750 2954.02178 3204.95213 3169.33640 3133.72068
##         73         74         75         76         77         78
## 3098.10496 3057.07957 2006.55206 1970.93633 1935.32061 1935.47117
##         79         80         81         82         83         84
## 1899.85544 1864.23972 1828.62400 1524.54680 1488.93108 1453.31535
##         85         86         87         88         89         90
## 1417.69963 1454.93506 1419.31933 1383.70361 1348.08789 1616.23188
##         91         92         93         94         95         96
## 1545.00043 1580.61616 1678.46056 1642.84484 1607.22911 1599.11144
##         97         98         99        100        101        102
## 1563.49571 1527.87999 1492.26426 1423.21750 1387.60178 1351.98605
##        103        104        105        106        107        108
## 1757.31994 1721.70422 1686.08850 1650.47277 1468.47358 1432.85786
##        109        110        111        112        113        114
## 1397.24213 1293.38950 1257.77377 1218.45745 1785.13547 1749.51974
##        115        116        117        118        119        120
## 1820.75119 1856.36691 1274.25663 1238.64091 1203.02519 1167.40946
##        121        122        123        124        125        126
## 1772.75221 1737.13649 1701.52076 1829.18898 1793.57325 1757.95753
##        127        128        129        130        131        132
## 1691.43797 1655.82224 1620.20652 1317.74516 1282.12944 1246.51372
```

```
##        133        134        135        136        137        138
## 1353.36089 1573.94592 1538.33019 1502.71447 1502.32176 1466.70603
##        139        140        141        142        143        144
## 1431.09031 1953.81589 1918.20017 1882.58444 1422.40765 1431.65529
##        145        146        147        148        149        150
## 1396.03957 1360.42384 1282.93929 1247.32357 1211.70784 1245.39052
##        151        152        153        154        155        156
## 1209.77480 1174.15907 1319.75758 1284.14186 1248.52614 1310.26507
##        157        158        159        160        161        162
## 1274.64935 1239.03362 1530.70226 1495.08654 1459.47081 1458.02337
##        163        164        165        166        167        168
## 1265.00899 1229.39327 1193.77755 1125.68848 1090.07276 1054.45703
##        169        170        171        172        173        174
## 1721.70422 1650.47277 1757.31994 1686.08850 1047.39408 1011.77836
##        175        176        177        178        179        180
##  976.16263  940.54691 1453.31535 1417.69963 1488.93108 1386.79193
##        181        182        183        184        185        186
## 1722.51407 2031.69587 1996.08014 1960.46442 1047.39408 1011.77836
##        187        188        189        190        191        192
##  976.16263  940.54691 1178.42465 1142.80893 1107.19320 1922.52293
##        193        194        195        196        197        198
## 1839.44465 1477.96609 1442.35037 1406.73464 1686.89835 1651.28262
##        199        200        201        202        203        204
## 1487.62815 1229.39327 1452.01243 1416.39671 1615.66690 1558.29462
##        205        206        207        208        209        210
## 1851.29148 1487.06317 1522.67890 1952.59161 1916.97589 1881.36016
##        211        212        213        214        215        216
##  706.57120  689.45075  653.83503  601.66384  654.64488  619.02916
##        217        218        219        220        221        222
##  689.45075  521.18475  566.04812  530.43239  566.04812  583.41343
##        223        224        225        226        227        228
##  540.48989  618.21931  786.48530  494.81667  715.25386  504.87416
##        229        230        231        232        233        234
##  601.66384  653.83503  742.18692  574.73078  750.86958  479.31593
##        235        236        237        238        239        240
##  670.95547  605.69809  530.43239  469.25844  433.64271  443.70021
##        241        242        243        244        245        246
##  408.08448  570.08237  618.21931  539.11505  614.13588  578.52015
##        247        248        249        250        251        252
##  534.46664  488.15559  542.90443 1686.33452 1615.10307 1809.73715
##        253        254        255        256        257        258
## 1774.12143 1738.50570 1650.71879  437.88979  399.10990  748.56418
##        259        260        261        262        263        264
##  510.42083  476.91456  442.35355  475.42227  404.19082  368.57509
##        265        266        267        268        269        270
##  446.32758  399.10990  418.09492  439.80654  767.54920  471.39607
##        271        272        273        274        275        276
##  450.54647  784.17991  502.57107  477.32659   34.58207  245.54993
##        277        278        279        280        281        282
##   67.27850  382.83100  616.06034  238.41174  546.03656  648.75676
```

```
##        283        284        285        286        287        288
##  281.16565  169.04486  202.79601  803.16493   19.81595  -13.69033
##        289        290        291        292        293        294
##  187.67548  646.40244  -16.85450  248.46923  223.29120  270.86329
##        295        296        297        298        299        300
##  306.47901  254.07908  218.46336  182.84764  289.69481  480.14288
##        301        302        303        304        305        306
##  252.07776  145.23058  182.84764  385.67464  278.82746  310.72907
##        307        308        309        310        311        312
##  515.75860  480.14288  444.52715  408.91143  341.62113  438.74421
##        313        314        315        316        317        318
##  403.12849  367.51276  210.59055  139.35911  103.74338  515.75860
##        319        320        321        322        323        324
##  227.46613  156.23468  120.61896  350.05891  216.46203  254.07908
##        325        326        327        328        329        330
##  306.00540  278.82746  381.96052  314.44319  417.57624  174.97483
##        331        332        333        334        335        336
##  191.85040  155.28808  119.67235   84.05663  403.60493  367.98921
##        337        338        339        340        341        342
##  332.37349  296.75776  180.84631  474.35993  190.90380   84.05663
##        343        344        345        346        347        348
##  218.46336  261.40284  225.78711  190.17139  154.55567  155.28808
##        349        350        351        352        353        354
##  119.67235  284.11612  248.50039  346.34480  319.73184  261.57027
##        355        356        357        358        359        360
##  225.78711  190.17139  154.55567  261.40284  289.69481  190.90380
##        361        362        363        364        365        366
## 1351.30397 1053.47004 1017.85432  982.23859  946.62287  297.23338
##        367        368        369        370        371        372
##  261.61765  226.00193  190.38620 1315.68824 1280.07252 1386.91969
##        373        374        375        376        377        378
##  982.80357  947.18785  911.57213  875.95640  333.09397  297.47825
##        379        380        381        382        383        384
##  261.86252  226.24680 1193.09546 1157.47974 1121.86401 1062.06489
##        385        386        387        388        389        390
## 1026.44916  990.83344 1430.40822 1394.79250 1359.17677  903.04653
##        391        392        393        394        395        396
##  867.43080  831.81508  323.75852  288.14280  252.52707  216.91135
##        397        398        399        400        401        402
## 1228.95605 1193.34033 1157.72461 1264.57178 1979.61264 1943.99692
##        403        404        405        406        407        408
## 1908.38120 1872.76547 2031.53896 1995.92324 1960.30751 2067.15469
##        409        410        411        412        413        414
## 1333.96063 1298.34490 1262.72918 1227.11345 1299.15476 1263.53903
##        415        416        417        418        419        420
## 1227.92331 1192.30758 1707.27856 1671.66283 1636.04711 1600.43139
##        421        422        423        424        425        426
## 1935.31426 1899.69854 1864.08281 1828.46709 1703.11381 1667.49809
##        427        428        429        430        431        432
## 1631.88236 1784.16871 1484.45121 1448.83549 1413.21976 1377.60404
```

```
##        433        434        435        436        437        438
## 2084.27513 1795.87532 1760.25960 1724.64387 1689.02815 2013.04368
##        439        440        441        442        443        444
## 1764.53264 1728.91691 1881.20326 1386.69680 1351.08107 1315.46535
##        445        446        447        448        449        450
## 1279.84962 1600.96695 1565.35123 1529.73550 1494.11978 1571.99325
##        451        452        453        454        455        456
## 1536.37753 1500.76180 1465.14608 1351.89092 1351.80094 1316.18521
##        457        458        459        460        461        462
## 1280.56949 2048.65941 1422.55739 1422.46740 1351.32594 1315.71022
```

```
reg1$coefficients
```

```
##      (Intercept)  FLIGHT_DURATION    SEATS_ECONOMY    PITCH_ECONOMY
##    -1.078876e+04     1.054723e+02    -7.401195e-01     3.324232e+02
##    WIDTH_ECONOMY          QUALITY            MONTH         AIRCRAFT
##     3.904307e+01     2.640474e+02    -3.561572e+01    -1.619979e+02
##   AIRLINEBritish     AIRLINEDelta       AIRLINEJet AIRLINESingapore
##    -1.333738e+03    -8.896564e+02    -2.449459e+03    -2.118905e+03
##    AIRLINEVirgin
##    -1.127875e+03
```

```
reg1$residuals
```

```
##            1            2            3            4            5
## -1946.517076 -1910.901351 -2130.188138 -1930.132800 -2052.803862
##            6            7            8            9           10
## -1463.848125 -1428.232401 -1841.419586 -1571.748524   126.254568
##           11           12           13           14           15
##   161.870292   197.486016   -84.632674   -49.016950   -13.401226
##           16           17           18           19           20
##    22.214498   -84.632674   -49.016950   -13.401226    22.214498
##           21           22           23           24           25
##     4.455388    40.071112    75.686836   139.634011   175.249736
##           26           27           28           29           30
##   -51.971943   345.807281   381.423006   226.463075   154.340462
##           31           32           33           34           35
##   189.956186   225.571910    59.343536    94.959260   130.574984
##           36           37           38           39           40
##  -160.190906  -124.575182   -88.959458   335.921034   459.078799
##           41           42           43           44           45
##   494.694524   406.038459   432.161673   467.777397   341.395234
##           46           47           48           49           50
##   166.966859   463.236208   498.851932   534.467656   570.083380
##           51           52           53           54           55
##   429.569693   465.185417   500.801141   527.979078   331.715838
##           56           57           58           59           60
##   367.331562   395.564223   431.179947   333.825285   369.441009
##           61           62           63           64           65
##   395.564223   431.179947   574.344442   609.960167   645.575891
##           66           67           68           69           70
```

```
##   385.680339   421.296064   603.362495   638.978220   -45.952127
##           71           72           73           74           75
##   -10.336403    25.279321    60.895045   162.920427  1095.447941
##           76           77           78           79           80
##  1131.063666  1166.679390   771.528832   807.144556   842.760280
##           81           82           83           84           85
##   878.376004   859.453198   895.068922   930.684646   966.300371
##           86           87           88           89           90
##   393.064941   428.680666   464.296390   499.912114    34.768119
##           91           92           93           94           95
##   105.999567  1194.383843   551.539439   587.155163   622.770887
##           96           97           98           99          100
##   756.888564   792.504288   828.120012   863.735736   138.782498
##          101          102          103          104          105
##   174.398222   210.013946   523.680056   559.295780   594.911505
##          106          107          108          109          110
##   630.527229   324.526420   360.142144   395.757868   182.610501
##          111          112          113          114          115
##   218.226225   257.542547   -80.135466   -44.519742    90.248810
##          116          117          118          119          120
##   521.633086   169.743366   205.359090   240.974814   276.590538
##          121          122          123          124          125
##   383.247789   418.863513   454.479238   -79.188977   -43.573253
##          126          127          128          129          130
##    -7.957529    66.562033   102.177757   137.793481   506.254836
##          131          132          133          134          135
##   541.870560   577.486284   469.639112   239.054081   274.669805
##          136          137          138          139          140
##   310.285530   310.678244   346.293968   381.909692   212.184111
##          141          142          143          144          145
##   247.799835   283.415559  -282.407650   177.344711   212.960435
##          146          147          148          149          150
##   248.576159   349.060708   384.676432   420.292156   388.609481
##          151          152          153          154          155
##   424.225205   459.840929   314.242417   349.858141   385.473865
##          156          157          158          159          160
##   340.734927   376.350651   411.966375    11.297740    70.913464
##          161          162          163          164          165
##   106.529188  -318.023374   470.991005   506.606729   542.222454
##          166          167          168          169          170
##   230.311519   265.927243   301.542967  -875.704220  -804.472771
##          171          172          173          174          175
##  -820.319944  -201.088495   194.605920   230.221644   265.837369
##          176          177          178          179          180
##   301.453093  -734.315354  -698.699629  -290.931078  -495.791926
##          181          182          183          184          185
##  -399.514072 -1091.695867 -1056.080143 -1020.464419   176.605920
##          186          187          188          189          190
##   212.221644   247.837369   283.453093   -51.424650   -15.808926
##          191          192          193          194          195
```

```
##    19.806798  -884.522931  -801.444655  -354.966091  -319.350367
##          196          197          198          199          200
##  -283.734643  -663.898348  -628.282624  -838.628154  -472.393271
##          201          202          203          204          205
##  -877.012430  -841.396706 -1082.666900  -761.294621 -1342.291483
##          206          207          208          209          210
##  -963.063172  -940.678897 -1443.591613 -1407.975889 -1372.360165
##          211          212          213          214          215
##  -249.571198  -287.450753  -251.835029  -209.663842  -298.644882
##          216          217          218          219          220
##  -263.029158  -367.450753  -224.184754  -263.048118  -227.432393
##          221          222          223          224          225
##  -290.048118  -334.413433  -302.489886  -380.219305  -558.485305
##          226          227          228          229          230
##  -263.816669  -512.253856  -303.874161  -394.663842  -446.835029
##          231          232          233          234          235
##  -560.186923  -403.730776  -582.869580  -339.315929  -523.955474
##          236          237          238          239          240
##  -468.698092  -392.432393  -343.258437  -307.642713  -334.700205
##          241          242          243          244          245
##  -299.084481  -461.082368  -514.219305  -442.115051  -537.135879
##          246          247          248          249          250
##  -501.520155  -465.466644  -414.155594  -477.904431    91.665485
##          251          252          253          254          255
##   162.896933   189.262849   224.878573   260.494297   334.281209
##          256          257          258          259          260
##  -149.889791  -111.109896  -450.564184  -147.420833  -113.914556
##          261          262          263          264          265
##   -79.353555   -62.422265     8.809183    44.424907   -33.327578
##          266          267          268          269          270
##    13.890104   -78.094917   -16.806541  -344.549204  -143.396068
##          271          272          273          274          275
##  -122.546471  -301.179908   -44.571075  -311.326591   131.417928
##          276          277          278          279          280
##   -87.549929   261.721503  -139.831000  -373.060335   -49.411736
##          281          282          283          284          285
##    79.963443  -294.756760   -53.165653    52.955136    13.203988
##          286          287          288          289          290
##   -90.164929   273.184055   306.690332   105.324522   -10.402443
##          291          292          293          294          295
##   432.854502   142.530772   125.708798    78.136710   274.520986
##          296          297          298          299          300
##   -87.079084   -51.463360   -15.847636  -150.694808  -279.142878
##          301          302          303          304          305
##  -104.077757     2.769416   -33.847636  -188.674636   -67.827464
##          306          307          308          309          310
##    25.270928  -328.758602  -293.142878  -257.527154  -221.911430
##          311          312          313          314          315
##  -202.621125  -282.744210  -247.128486  -211.512762   -92.590555
##          316          317          318          319          320
```

```
##   -21.359106    14.256618  -270.758602  -119.466129   -48.234681
##          321           322           323           324           325
##   -12.618956   -53.058912    17.537967   -20.079084  -150.005401
##          326           327           328           329           330
##  -122.827464    47.039480     9.556812    44.423756   -27.974831
##          331           332           333           334           335
##   -64.850405    16.711924    52.327648    87.943372  -249.604934
##          336           337           338           339           340
##  -213.989210  -178.373486  -142.757761   112.153692  -259.359934
##          341           342           343           344           345
##    90.096199   210.943372   103.536640    92.597162   128.212886
##          346           347           348           349           350
##   163.828610   199.444335   224.711924   260.327648   272.883883
##          351           352           353           354           355
##   308.499608   314.655204   356.268159   243.429733   238.212886
##          356           357           358           359           360
##   273.828610   309.444335   227.597162   304.305192   319.096199
##          361           362           363           364           365
##  -557.303967  -548.470041  -512.854317  -477.238593  -441.622869
##          366           367           368           369           370
##   207.766624   243.382348   278.998073   314.613797  -521.688243
##          371           372           373           374           375
##  -486.072518  -592.919691  -292.803575  -257.187850  -221.572126
##          376           377           378           379           380
##  -185.956402   356.906029   392.521753   428.137478   463.753202
##          381           382           383           384           385
##    21.904541    57.520265    93.135989  -186.064889  -417.449165
##          386           387           388           389           390
##  -381.833441   -24.408220    11.207504    46.823228   343.953471
##          391           392           393           394           395
##   379.569195   415.184919   239.241479   274.857203   310.472927
##          396           397           398           399           400
##   346.088652   202.043946   237.659670   273.275394   166.428222
##          401           402           403           404           405
## -1405.612645 -1369.996921 -1334.381197 -1298.765472  -945.538961
##          406           407           408           409           410
##  -909.923237  -874.307513  -820.154686   100.039373   135.655097
##          411           412           413           414           415
##   171.270821   206.886545   176.845245   212.460969   248.076693
##          416           417           418           419           420
##   283.692417    73.721441   109.337165   144.952890   180.568614
##          421           422           423           424           425
##  -355.314263  -319.698539  -284.082814  -248.467090   199.886189
##          426           427           428           429           430
##   235.501913   271.117637   118.831292   328.548788   364.164512
##          431           432           433           434           435
##   399.780236   435.395960  -988.275131   649.124677   684.740401
##          436           437           438           439           440
##   720.356125   755.971849 -1038.043682   604.467362   640.083086
##          441           442           443           444           445
```

```
##   487.796741   380.303204   415.918928   451.534652   487.150376
##          446          447          448          449          450
##   210.033050   245.648774   281.264498   316.880222   480.006747
##          451          452          453          454          455
##   515.622471   551.238195   586.853920   567.109075   567.199063
##          456          457          458          459          460
##   602.814787   638.430512  -692.659406  -882.557391  -882.467403
##          461          462
##  -811.325943  -775.710219
```

```
#Dividing the Data set into Test and Training Data ste
ratio = sample(1:nrow(airdata.df), size = 0.25*nrow(airdata.df))
Test = airdata.df[ratio,] #Test dataset 25% of total
Training = airdata.df[-ratio,] #Train dataset 75% of total
dim(Training)
```

```
## [1] 347  17
```

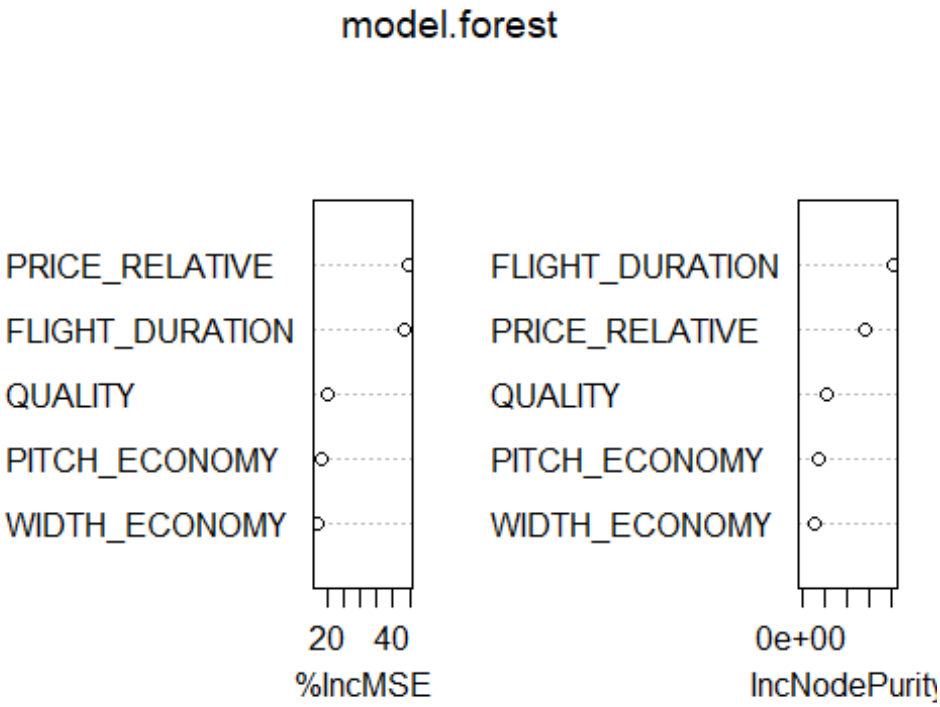
```
dim(Test)
```

```
## [1] 115  17
```

```
#Generating A Multi Variable Linear Regressional Model for Price of Economy F
lights
linear.mod<- lm(PRICE_ECONOMY~ PITCH_ECONOMY + WIDTH_ECONOMY + FLIGHT_DURATIO
N + QUALITY + PRICE_RELATIVE, data = Training)
summary(linear.mod)
```

```
##
## Call:
## lm(formula = PRICE_ECONOMY ~ PITCH_ECONOMY + WIDTH_ECONOMY +
##     FLIGHT_DURATION + QUALITY + PRICE_RELATIVE, data = Training)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1604.82  -476.33    37.91   551.17  1590.60
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -8667.04    3153.20  -2.749   0.0063 **
## PITCH_ECONOMY     556.23      94.93   5.859 1.10e-08 ***
## WIDTH_ECONOMY    -544.57      72.65  -7.496 5.72e-13 ***
## FLIGHT_DURATION   187.21      12.16  15.391  < 2e-16 ***
## QUALITY           204.86      33.33   6.147 2.21e-09 ***
## PRICE_RELATIVE   -924.84      89.98 -10.279  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 649.3 on 341 degrees of freedom
## Multiple R-squared:  0.5694, Adjusted R-squared:  0.5631
## F-statistic: 90.18 on 5 and 341 DF,  p-value: < 2.2e-16
```

```r
#the t value of Pitch_economy and quality is positive indicating that these p
redictors are associated with
#Price_economy. A larger t-value indicates that that it is less likely that t
he coefficient is not equal to zero purely by chance.
#Again, as the p-value for Flight_Duration and Price_Relative is less than 0.
05 they are both statistically significant in the multiple linear regression
model for Price_Economy response variable.
#The model's, p-value: < 2.2e-16 is also lower than the statistical significa
nce level of 0.05, this indicates that we can safely reject the null hypothes
is that the value for the coefficient is zero
#(or in other words, the predictor variable has no explanatory relationship w
ith the response variable).
#The model has a F Statistic of 90, which is considerably high
library(rpart)
library(randomForest)

## Warning: package 'randomForest' was built under R version 3.4.1

## randomForest 4.6-12

## Type rfNews() to see new features/changes/bug fixes.

##
## Attaching package: 'randomForest'

## The following object is masked from 'package:ggplot2':
##
##     margin

## The following object is masked from 'package:psych':
##
##     outlier

model.forest <- randomForest(PRICE_ECONOMY~ PITCH_ECONOMY + WIDTH_ECONOMY + F
LIGHT_DURATION + QUALITY + PRICE_RELATIVE, data = Training, method = "anova",
                             ntree = 300,
                             mtry = 2, #mtry is sqrt(6)
                             replace = F,
                             nodesize = 1,
                             importance = T)

varImpPlot(model.forest)
```




```
#From the VIF plot we see that Flight Duration and Price Relative are most im
portant factors in predicitng Price Economy.

#We test the model using Random Forest
prediction <- predict(model.forest,Test)
rmse <- sqrt(mean((log(prediction)-log(Test$PRICE_ECONOMY))^2))
round(rmse, digits = 3)
```

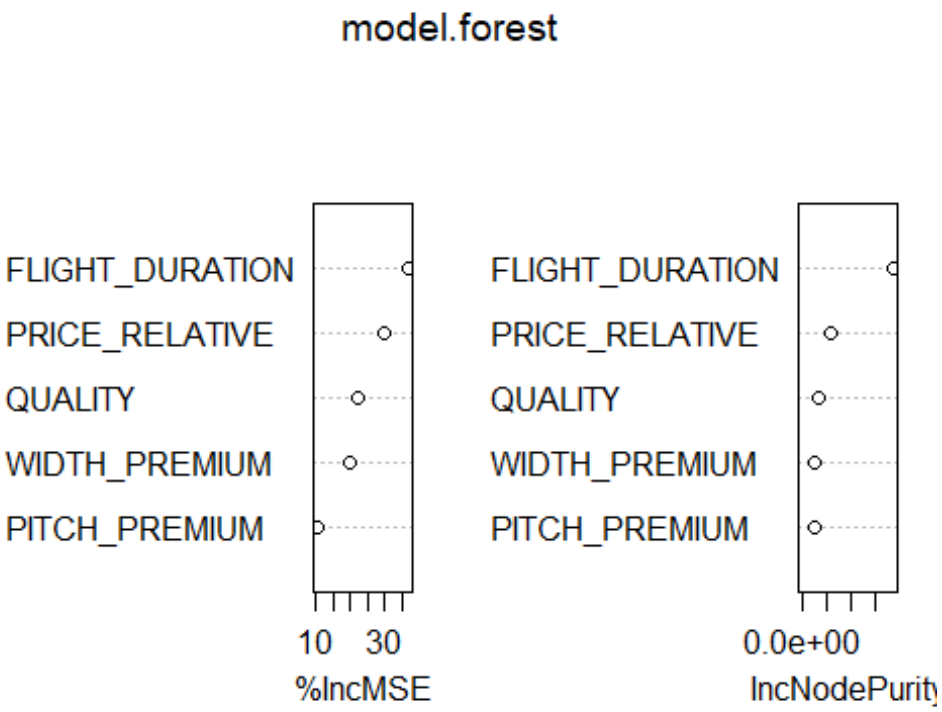
```
## [1] 0.411
```

```
# Evaluation metric function
#A custom root mean Square Function to evaluate the performance of our model
RMSE <- function(x,y)
{
  a <- sqrt(sum((log(x)-log(y))^2)/length(y))
  return(a)
}

#Implementing the Regression Tree Model
model <- rpart(PRICE_ECONOMY~ PITCH_ECONOMY + WIDTH_ECONOMY + FLIGHT_DURATION
+ QUALITY + PRICE_RELATIVE, data = Training, method = "anova")
predict <- predict(model, Test)
RMSE1 <- RMSE(predict, Test$PRICE_ECONOMY)
RMSE1 <- round(RMSE1, digits = 3)
RMSE1
```

```
## [1] 0.532
```

```r
#For Premium Class Tickets

#Generating A Multi Variable Linear Regressional Model for Price of Premium F
lights
linear.mod<- lm(PRICE_PREMIUM~ PITCH_PREMIUM + WIDTH_PREMIUM + FLIGHT_DURATIO
N + QUALITY + PRICE_RELATIVE, data = Training)
summary(linear.mod)

##
## Call:
## lm(formula = PRICE_PREMIUM ~ PITCH_PREMIUM + WIDTH_PREMIUM +
##     FLIGHT_DURATION + QUALITY + PRICE_RELATIVE, data = Training)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2269.6  -553.2   -51.1   698.5  4322.6
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -4867.873   4718.473  -1.032   0.3030
## PITCH_PREMIUM     143.241    143.274   1.000   0.3181
## WIDTH_PREMIUM       4.175     78.720   0.053   0.9577
## FLIGHT_DURATION   238.273     17.106  13.929   <2e-16 ***
## QUALITY           -75.043    112.545  -0.667   0.5054
## PRICE_RELATIVE   -239.715    140.603  -1.705   0.0891 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 987.7 on 341 degrees of freedom
## Multiple R-squared:  0.436,  Adjusted R-squared:  0.4277
## F-statistic: 52.72 on 5 and 341 DF,  p-value: < 2.2e-16

#The model has an F Statistic of 48.4 which is mediumly high
#the t value of Pitch_premium, width_premium, Price_relative and quality is p
ositive indicating that these predictors are associated with
#Price_Premium. A larger t-value indicates that that it is less likely that t
he coefficient is not equal to zero purely by chance.
#Again, as the p-value for Flight_Duration  is less than 0.05 they are both s
tatistically significant in the multiple linear regression model for Price_Ec
onomy response variable.
#The model's, p-value: < 2.2e-16 is also lower than the statistical significa
nce level of 0.05, this indicates that we can safely reject the null hypothes
is that the value for the coefficient is zero
#(or in other words, the predictor variable has no explanatory relationship w
ith the response variable).

library(rpart)
library(randomForest)
model.forest <- randomForest(PRICE_PREMIUM~ PITCH_PREMIUM + WIDTH_PREMIUM + F
LIGHT_DURATION + QUALITY + PRICE_RELATIVE, data = Training, method = "anova",
```

```
                              ntree = 300,
                              mtry = 2, #mtry is sqrt(6)
                              replace = F,
                              nodesize = 1,
                              importance = T)

varImpPlot(model.forest)
```




```
#From the VIF plot we see that Flight Duration and Price Relative are most im
portant factors in predicitng Price Economy.

# Evaluation metric function
#A custom root mean Square Function to evaluate the performance of our model
RMSE <- function(x,y)
{
  a <- sqrt(sum((log(x)-log(y))^2)/length(y))
  return(a)
}

#Implementing the Regression Tree Model
model <- rpart(PRICE_ECONOMY~ PITCH_ECONOMY + WIDTH_ECONOMY + FLIGHT_DURATION
+ QUALITY + PRICE_RELATIVE, data = Training, method = "anova")
predict <- predict(model, Test)
RMSE1 <- RMSE(predict, Test$PRICE_ECONOMY)
RMSE1 <- round(RMSE1, digits = 3)
RMSE1
```

```
## [1] 0.532
```