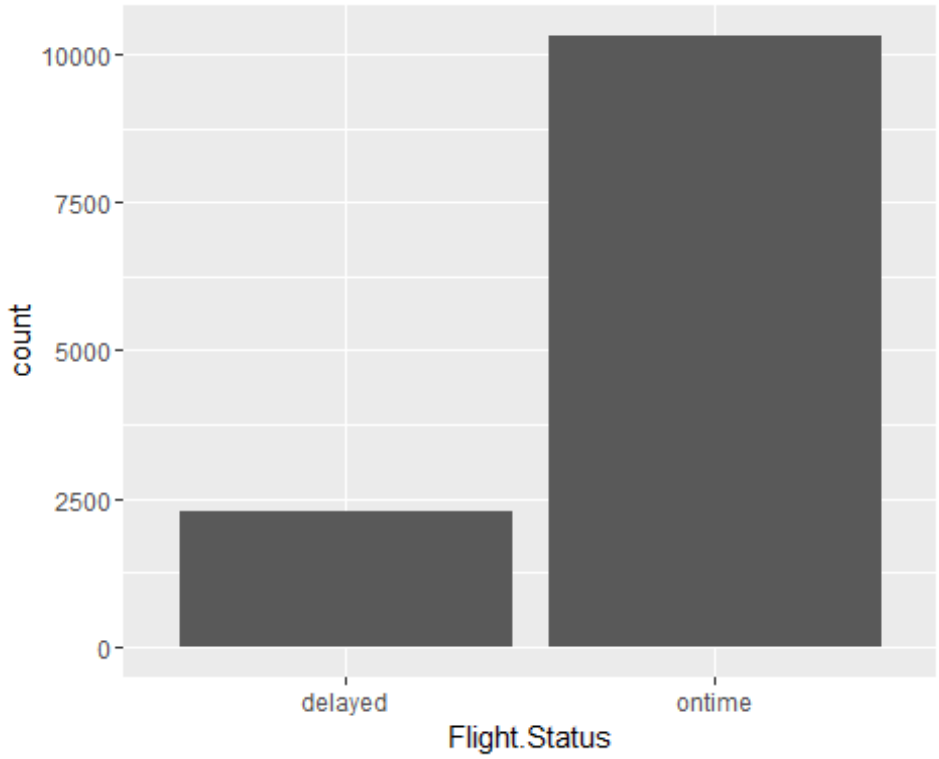
Projectcombined

Note: THIS IS NOT A PAPER in any kind acceptable format.

```
library(caret)

## Warning: package 'caret' was built under R version 3.2.5

## Loading required package: lattice

## Loading required package: ggplot2

library(ggplot2)
flights<-read.csv("projectcombined.csv")
head(flights)

##   MONTH DAY_OF_MONTH DAY_OF_WEEK CARRIER ORIGIN DEST_AIRPORT_ID DEST
## 1     1            6           2      AA    DCA           11298  DFW
## 2     1            7           3      AA    DCA           11298  DFW
## 3     1            8           4      AA    DCA           11298  DFW
## 4     1            9           5      AA    DCA           11298  DFW
## 5     1           10           6      AA    DCA           11298  DFW
## 6     1           11           7      AA    DCA           11298  DFW
##   CRS_DEP_TIME DEP_TIME DEP_DELAY_NEW WEATHER_DELAY Weather
Flight.Status
## 1         1600     1609             9            NA       0
ontime
## 2         1600     1630            30             0       0
delayed
## 3         1600     1556             0            NA       0
ontime
## 4         1600     1600             0            NA       0
ontime
## 5         1600     1555             0            NA       0
ontime
## 6         1600     1608             8            NA       0
ontime

str(flights)

## 'data.frame':    12581 obs. of  13 variables:
##  $ MONTH          : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ DAY_OF_MONTH   : int  6 7 8 9 10 11 12 13 14 15 ...
##  $ DAY_OF_WEEK    : int  2 3 4 5 6 7 1 2 3 4 ...
##  $ CARRIER        : Factor w/ 9 levels "AA","B6","DL",..: 1 1 1 1 1
1 1 1 1 1 ...
```

```
##  $ ORIGIN         : Factor w/ 3 levels "BWI","DCA","IAD": 2 2 2 2 2
2 2 2 2 2 ...
##  $ DEST_AIRPORT_ID: int  11298 11298 11298 11298 11298 11298 11298
11298 11298 11298 ...
##  $ DEST           : Factor w/ 3 levels "ATL","BOS","DFW": 3 3 3 3 3
3 3 3 3 3 ...
##  $ CRS_DEP_TIME   : int  1600 1600 1600 1600 1600 1600 1600 1600
1600 1600 ...
##  $ DEP_TIME       : int  1609 1630 1556 1600 1555 1608 1701 1559
1603 1602 ...
##  $ DEP_DELAY_NEW  : int  9 30 0 0 0 8 61 0 3 2 ...
##  $ WEATHER_DELAY  : int  NA 0 NA NA NA NA 38 NA NA NA ...
##  $ Weather        : int  0 0 0 0 0 0 1 0 0 0 ...
##  $ Flight.Status  : Factor w/ 2 levels "delayed","ontime": 2 1 2 2 2
2 1 2 2 2 ...
```

```
qplot(x=Flight.Status,data=flights)
```




```
prop.table(table(flights$CARRIER))
```

```
##
##          AA          B6          DL          EV          F9
NK
## 0.196407281 0.161672363 0.296001908 0.008504888 0.008902313
0.012797075
```

```
##          UA          US          WN
## 0.033622129 0.096653684 0.185438359
```

```
tbl <- table(flights$CARRIER)
tblFun <- function(x){
    tbl <- table(x)
    res <- cbind(tbl,round(prop.table(tbl)*100,2))
    colnames(res) <- c('Count','Percentage')
    res
}
do.call(rbind,lapply(flights[4],tblFun))
```

```
##    Count Percentage
## AA  2471      19.64
## B6  2034      16.17
## DL  3724      29.60
## EV   107       0.85
## F9   112       0.89
## NK   161       1.28
## UA   423       3.36
## US  1216       9.67
## WN  2333      18.54
```

```
###Dummy Variables
library(stats)
flights$DAY_OF_WEEK <- factor(flights$DAY_OF_WEEK)
flights$MONTH<-factor(flights$MONTH)
flights.dummy <- model.matrix(~CARRIER+DEST+ ORIGIN+
DAY_OF_WEEK+MONTH,data=flights,contrasts.arg=list(CARRIER=contrasts(fli
ghts$CARRIER, contrasts=F),DEST=contrasts(flights$DEST,
contrasts=F),ORIGIN=contrasts(flights$ORIGIN,contrasts=FALSE),DAY_OF_WE
EK=contrasts(flights$DAY_OF_WEEK,contrasts=FALSE),MONTH=contrasts(fligh
ts$MONTH,contrasts=F)))
flights.dummy<- flights.dummy[,-1]
head(flights.dummy,5)
```

```
##   CARRIERAA CARRIERB6 CARRIERDL CARRIEREV CARRIERF9 CARRIERNK
CARRIERUA
## 1         1         0         0         0         0         0
0
## 2         1         0         0         0         0         0
0
## 3         1         0         0         0         0         0
0
## 4         1         0         0         0         0         0
0
## 5         1         0         0         0         0         0
0
##   CARRIERUS CARRIERWN DESTATL DESTBOS DESTDFW ORIGINBWI ORIGINDCA
## 1         0         0       0       0       1         0         1
## 2         0         0       0       0       1         0         1
```

```
## 3         0         0       0       0       1         0         1
## 4         0         0       0       0       1         0         1
## 5         0         0       0       0       1         0         1
##   ORIGINIAD DAY_OF_WEEK1 DAY_OF_WEEK2 DAY_OF_WEEK3 DAY_OF_WEEK4
## 1         0            0            1            0            0
## 2         0            0            0            1            0
## 3         0            0            0            0            1
## 4         0            0            0            0            0
## 5         0            0            0            0            0
##   DAY_OF_WEEK5 DAY_OF_WEEK6 DAY_OF_WEEK7 MONTH1 MONTH3 MONTH6 MONTH9
## 1            0            0            0      1      0      0      0
## 2            0            0            0      1      0      0      0
## 3            0            0            0      1      0      0      0
## 4            1            0            0      1      0      0      0
## 5            0            1            0      1      0      0      0
```

```
str(flights.dummy)
```

```
##  num [1:12581, 1:26] 1 1 1 1 1 1 1 1 1 1 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : chr [1:12581] "1" "2" "3" "4" ...
##   ..$ : chr [1:26] "CARRIERAA" "CARRIERB6" "CARRIERDL" "CARRIEREV"
...
```

```
combinedata<-data.frame(flights.dummy)
str(combinedata)
```

```
## 'data.frame':    12581 obs. of  26 variables:
##  $ CARRIERAA   : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ CARRIERB6   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERDL   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIEREV   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERF9   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERNK   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERUA   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERUS   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERWN   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DESTATL     : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DESTBOS     : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DESTDFW     : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ ORIGINBWI   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ ORIGINDCA   : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ ORIGINIAD   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DAY_OF_WEEK1: num  0 0 0 0 0 0 1 0 0 0 ...
##  $ DAY_OF_WEEK2: num  1 0 0 0 0 0 0 1 0 0 ...
##  $ DAY_OF_WEEK3: num  0 1 0 0 0 0 0 0 1 0 ...
##  $ DAY_OF_WEEK4: num  0 0 1 0 0 0 0 0 0 1 ...
##  $ DAY_OF_WEEK5: num  0 0 0 1 0 0 0 0 0 0 ...
##  $ DAY_OF_WEEK6: num  0 0 0 0 1 0 0 0 0 0 ...
##  $ DAY_OF_WEEK7: num  0 0 0 0 0 1 0 0 0 0 ...
##  $ MONTH1      : num  1 1 1 1 1 1 1 1 1 1 ...
```

```
##  $ MONTH3      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ MONTH6      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ MONTH9      : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
deptime <- flights$DEP_TIME
flights$deptimeHourBlock<-cut(deptime, breaks =seq(100, 2400, by =
100), labels =1:23,right=FALSE)
r<-table(flights$deptimeHourBlock)
r
```

```
##
##    1    2    3    4    5    6    7    8    9   10   11   12   13
14   15
##    8    3    0    0  511  918  695 1017  625  830  719  795  792
717  730
##   16   17   18   19   20   21   22   23
##  688  874  761  755  504  297  279   63
```

```
flights.timedummy<- model.matrix(~deptimeHourBlock,data=flights)
flights.timedummy<- flights.timedummy[,-1]
head(flights.timedummy,5)
```

```
##   deptimeHourBlock2 deptimeHourBlock3 deptimeHourBlock4
deptimeHourBlock5
## 1                 0                 0                 0
0
## 2                 0                 0                 0
0
## 3                 0                 0                 0
0
## 4                 0                 0                 0
0
## 5                 0                 0                 0
0
##   deptimeHourBlock6 deptimeHourBlock7 deptimeHourBlock8
deptimeHourBlock9
## 1                 0                 0                 0
0
## 2                 0                 0                 0
0
## 3                 0                 0                 0
0
## 4                 0                 0                 0
0
## 5                 0                 0                 0
0
##   deptimeHourBlock10 deptimeHourBlock11 deptimeHourBlock12
## 1                  0                  0                  0
## 2                  0                  0                  0
## 3                  0                  0                  0
## 4                  0                  0                  0
```

```
## 5                    0                    0                    0
##   deptimeHourBlock13 deptimeHourBlock14 deptimeHourBlock15
## 1                  0                    0                    0
## 2                  0                    0                    0
## 3                  0                    0                    1
## 4                  0                    0                    0
## 5                  0                    0                    1
##   deptimeHourBlock16 deptimeHourBlock17 deptimeHourBlock18
## 1                  1                    0                    0
## 2                  1                    0                    0
## 3                  0                    0                    0
## 4                  1                    0                    0
## 5                  0                    0                    0
##   deptimeHourBlock19 deptimeHourBlock20 deptimeHourBlock21
## 1                  0                    0                    0
## 2                  0                    0                    0
## 3                  0                    0                    0
## 4                  0                    0                    0
## 5                  0                    0                    0
##   deptimeHourBlock22 deptimeHourBlock23
## 1                  0                    0
## 2                  0                    0
## 3                  0                    0
## 4                  0                    0
## 5                  0                    0
```

```
hourblock<-data.frame(flights.timedummy)
head(hourblock)
```

```
##   deptimeHourBlock2 deptimeHourBlock3 deptimeHourBlock4
deptimeHourBlock5
## 1                 0                 0                 0
0
## 2                 0                 0                 0
0
## 3                 0                 0                 0
0
## 4                 0                 0                 0
0
## 5                 0                 0                 0
0
## 6                 0                 0                 0
0
##   deptimeHourBlock6 deptimeHourBlock7 deptimeHourBlock8
deptimeHourBlock9
## 1                 0                 0                 0
0
## 2                 0                 0                 0
0
## 3                 0                 0                 0
```

```
0
## 4                  0                  0                  0
0
## 5                  0                  0                  0
0
## 6                  0                  0                  0
0
##   deptimeHourBlock10 deptimeHourBlock11 deptimeHourBlock12
## 1                  0                  0                  0
## 2                  0                  0                  0
## 3                  0                  0                  0
## 4                  0                  0                  0
## 5                  0                  0                  0
## 6                  0                  0                  0
##   deptimeHourBlock13 deptimeHourBlock14 deptimeHourBlock15
## 1                  0                  0                  0
## 2                  0                  0                  0
## 3                  0                  0                  1
## 4                  0                  0                  0
## 5                  0                  0                  1
## 6                  0                  0                  0
##   deptimeHourBlock16 deptimeHourBlock17 deptimeHourBlock18
## 1                  1                  0                  0
## 2                  1                  0                  0
## 3                  0                  0                  0
## 4                  1                  0                  0
## 5                  0                  0                  0
## 6                  1                  0                  0
##   deptimeHourBlock19 deptimeHourBlock20 deptimeHourBlock21
## 1                  0                  0                  0
## 2                  0                  0                  0
## 3                  0                  0                  0
## 4                  0                  0                  0
## 5                  0                  0                  0
## 6                  0                  0                  0
##   deptimeHourBlock22 deptimeHourBlock23
## 1                  0                  0
## 2                  0                  0
## 3                  0                  0
## 4                  0                  0
## 5                  0                  0
## 6                  0                  0

##merge data

combinedata[c("Hour5","Hour6")]<-NA
combinedata[c("Hour7","Hour8","Hour9","Hour10","Hour11","Hour12")]<-NA
combinedata[c("Hour13","Hour14","Hour15","Hour16","Hour17","Hour18")]<-
NA
combinedata[c("Hour19","Hour20","Hour21","Hour22","Hour23")]<-NA
```

```
combinedata$Hour5<-hourblock$deptimeHourBlock5
combinedata$Hour6<-hourblock$deptimeHourBlock6
combinedata$Hour7<-hourblock$deptimeHourBlock7
combinedata$Hour8<-hourblock$deptimeHourBlock8
combinedata$Hour9<-hourblock$deptimeHourBlock9
combinedata$Hour10<-hourblock$deptimeHourBlock10
combinedata$Hour11<-hourblock$deptimeHourBlock11
combinedata$Hour12<-hourblock$deptimeHourBlock12
combinedata$Hour13<-hourblock$deptimeHourBlock13
combinedata$Hour14<-hourblock$deptimeHourBlock14
combinedata$Hour15<-hourblock$deptimeHourBlock15
combinedata$Hour16<-hourblock$deptimeHourBlock16
combinedata$Hour17<-hourblock$deptimeHourBlock17
combinedata$Hour18<-hourblock$deptimeHourBlock18
combinedata$Hour19<-hourblock$deptimeHourBlock19
combinedata$Hour20<-hourblock$deptimeHourBlock20
combinedata$Hour21<-hourblock$deptimeHourBlock21
combinedata$Hour22<-hourblock$deptimeHourBlock22
combinedata$Hour23<-hourblock$deptimeHourBlock23
combinedata$Weather<-flights$Weather
combinedata$FlightStatus<-flights$Flight.Status
head(combinedata)
```

```
##   CARRIERAA CARRIERB6 CARRIERDL CARRIEREV CARRIERF9 CARRIERNK
CARRIERUA
## 1         1         0         0         0         0         0
0
## 2         1         0         0         0         0         0
0
## 3         1         0         0         0         0         0
0
## 4         1         0         0         0         0         0
0
## 5         1         0         0         0         0         0
0
## 6         1         0         0         0         0         0
0
##   CARRIERUS CARRIERWN DESTATL DESTBOS DESTDFW ORIGINBWI ORIGINDCA
## 1         0         0       0       0       1         0         1
## 2         0         0       0       0       1         0         1
## 3         0         0       0       0       1         0         1
## 4         0         0       0       0       1         0         1
## 5         0         0       0       0       1         0         1
## 6         0         0       0       0       1         0         1
##   ORIGINIAD DAY_OF_WEEK1 DAY_OF_WEEK2 DAY_OF_WEEK3 DAY_OF_WEEK4
## 1         0            0            1            0            0
## 2         0            0            0            1            0
## 3         0            0            0            0            1
## 4         0            0            0            0            0
## 5         0            0            0            0            0
```

```
## 6         0            0            0            0            0
##   DAY_OF_WEEK5 DAY_OF_WEEK6 DAY_OF_WEEK7 MONTH1 MONTH3 MONTH6 MONTH9
Hour5
## 1            0            0            0      1      0      0      0
0
## 2            0            0            0      1      0      0      0
0
## 3            0            0            0      1      0      0      0
0
## 4            1            0            0      1      0      0      0
0
## 5            0            1            0      1      0      0      0
0
## 6            0            0            1      1      0      0      0
0
##   Hour6 Hour7 Hour8 Hour9 Hour10 Hour11 Hour12 Hour13 Hour14 Hour15
Hour16
## 1     0     0     0     0      0      0      0      0      0      0
1
## 2     0     0     0     0      0      0      0      0      0      0
1
## 3     0     0     0     0      0      0      0      0      0      1
0
## 4     0     0     0     0      0      0      0      0      0      0
1
## 5     0     0     0     0      0      0      0      0      0      1
0
## 6     0     0     0     0      0      0      0      0      0      0
1
##   Hour17 Hour18 Hour19 Hour20 Hour21 Hour22 Hour23 Weather
FlightStatus
## 1      0      0      0      0      0      0      0       0
ontime
## 2      0      0      0      0      0      0      0       0
delayed
## 3      0      0      0      0      0      0      0       0
ontime
## 4      0      0      0      0      0      0      0       0
ontime
## 5      0      0      0      0      0      0      0       0
ontime
## 6      0      0      0      0      0      0      0       0
ontime
```

```
str(combinedata)
```

```
## 'data.frame':    12581 obs. of  47 variables:
##  $ CARRIERAA   : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ CARRIERB6   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERDL   : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
##  $ CARRIEREV   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERF9   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERNK   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERUA   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERUS   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ CARRIERWN   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DESTATL     : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DESTBOS     : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DESTDFW     : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ ORIGINBWI   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ ORIGINDCA   : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ ORIGINIAD   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DAY_OF_WEEK1: num  0 0 0 0 0 0 1 0 0 0 ...
##  $ DAY_OF_WEEK2: num  1 0 0 0 0 0 0 1 0 0 ...
##  $ DAY_OF_WEEK3: num  0 1 0 0 0 0 0 0 1 0 ...
##  $ DAY_OF_WEEK4: num  0 0 1 0 0 0 0 0 0 1 ...
##  $ DAY_OF_WEEK5: num  0 0 0 1 0 0 0 0 0 0 ...
##  $ DAY_OF_WEEK6: num  0 0 0 0 1 0 0 0 0 0 ...
##  $ DAY_OF_WEEK7: num  0 0 0 0 0 1 0 0 0 0 ...
##  $ MONTH1      : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ MONTH3      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ MONTH6      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ MONTH9      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour5       : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour6       : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour7       : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour8       : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour9       : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour10      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour11      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour12      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour13      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour14      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour15      : num  0 0 1 0 1 0 0 1 0 0 ...
##  $ Hour16      : num  1 1 0 1 0 1 0 0 1 1 ...
##  $ Hour17      : num  0 0 0 0 0 0 1 0 0 0 ...
##  $ Hour18      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour19      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour20      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour21      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour22      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Hour23      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Weather     : int  0 0 0 0 0 0 1 0 0 0 ...
##  $ FlightStatus: Factor w/ 2 levels "delayed","ontime": 2 1 2 2 2 2
1 2 2 2 ...

##testandtrain
shuff<-runif(nrow(combinedata))
flightdata<-combinedata[order(shuff),]
##model
```

```
train<- flightdata[1:7549,]
head(train)
```

```
##       CARRIERAA CARRIERB6 CARRIERDL CARRIEREV CARRIERF9 CARRIERNK
## 7462          0         0         1         0         0         0
## 8749          0         0         0         0         0         0
## 7348          0         0         1         0         0         0
## 10872         0         0         1         0         0         0
## 19            1         0         0         0         0         0
## 4875          0         0         0         0         1         0
##       CARRIERUA CARRIERUS CARRIERWN DESTATL DESTBOS DESTDFW
ORIGINBWI
## 7462          0         0         0       1       0       0
0
## 8749          0         0         1       1       0       0
1
## 7348          0         0         0       1       0       0
0
## 10872         0         0         0       1       0       0
1
## 19            0         0         0       0       0       1
0
## 4875          0         0         0       1       0       0
0
##       ORIGINDCA ORIGINIAD DAY_OF_WEEK1 DAY_OF_WEEK2 DAY_OF_WEEK3
## 7462          1         0            0            0            0
## 8749          0         0            0            0            0
## 7348          1         0            0            1            0
## 10872         0         0            1            0            0
## 19            1         0            0            0            0
## 4875          0         1            0            1            0
##       DAY_OF_WEEK4 DAY_OF_WEEK5 DAY_OF_WEEK6 DAY_OF_WEEK7 MONTH1
MONTH3
## 7462             0            0            1            0      0
0
## 8749             0            1            0            0      0
0
## 7348             0            0            0            0      0
0
## 10872            0            0            0            0      0
0
## 19               0            0            1            0      1
0
## 4875             0            0            0            0      0
1
##       MONTH6 MONTH9 Hour5 Hour6 Hour7 Hour8 Hour9 Hour10 Hour11
Hour12
## 7462       1      0     0     0     0     0     0      0      0
0
## 8749       1      0     0     0     0     0     0      0      0
```

```
0
## 7348       1      0     0     0     0     0     0      0      0
0
## 10872      0      1     0     0     0     0     0      0      0
0
## 19         0      0     0     0     0     0     0      0      0
0
## 4875       0      0     0     0     0     0     0      0      0
0
##       Hour13 Hour14 Hour15 Hour16 Hour17 Hour18 Hour19 Hour20 Hour21
## 7462       0      0      1      0      0      0      0      0      0
## 8749       0      0      1      0      0      0      0      0      0
## 7348       0      0      0      0      0      0      0      0      1
## 10872      1      0      0      0      0      0      0      0      0
## 19         0      0      1      0      0      0      0      0      0
## 4875       1      0      0      0      0      0      0      0      0
##       Hour22 Hour23 Weather FlightStatus
## 7462       0      0       0       ontime
## 8749       0      0       0       ontime
## 7348       0      0       0      delayed
## 10872      0      0       0       ontime
## 19         0      0       0       ontime
## 4875       0      0       0      delayed

test<-flightdata[7550:12581,1:46]
head(test)

##      CARRIERAA CARRIERB6 CARRIERDL CARRIEREV CARRIERF9 CARRIERNK
CARRIERUA
## 8680         0         0         0         0         0         0
0
## 9905         1         0         0         0         0         0
0
## 5978         0         0         0         0         0         0
0
## 3968         0         0         1         0         0         0
0
## 4029         0         0         1         0         0         0
0
## 6453         1         0         0         0         0         0
0
##      CARRIERUS CARRIERWN DESTATL DESTBOS DESTDFW ORIGINBWI ORIGINDCA
## 8680         0         1       1       0       0         0         1
## 9905         0         0       0       1       0         0         1
## 5978         0         1       0       1       0         1         0
## 3968         0         0       1       0       0         1         0
## 4029         0         0       1       0       0         0         1
## 6453         0         0       0       0       1         0         0
##      ORIGINIAD DAY_OF_WEEK1 DAY_OF_WEEK2 DAY_OF_WEEK3 DAY_OF_WEEK4
## 8680         0            1            0            0            0
```

```
## 9905          0             0             1             0             0
## 5978          0             0             1             0             0
## 3968          0             0             1             0             0
## 4029          0             0             0             0             0
## 6453          1             0             0             0             0
##      DAY_OF_WEEK5 DAY_OF_WEEK6 DAY_OF_WEEK7 MONTH1 MONTH3 MONTH6
MONTH9
## 8680            0            0            0      0      0      1
0
## 9905            0            0            0      0      0      0
1
## 5978            0            0            0      0      1      0
0
## 3968            0            0            0      0      1      0
0
## 4029            0            0            1      0      1      0
0
## 6453            0            0            1      0      0      1
0
##      Hour5 Hour6 Hour7 Hour8 Hour9 Hour10 Hour11 Hour12 Hour13
Hour14
## 8680     1     0     0     0     0      0      0      0      0
0
## 9905     0     0     0     0     0      0      0      0      0
0
## 5978     0     0     0     0     0      0      0      0      0
0
## 3968     0     0     0     0     0      0      0      0      0
0
## 4029     0     0     0     0     0      0      0      0      0
0
## 6453     0     0     0     0     0      0      0      0      0
0
##      Hour15 Hour16 Hour17 Hour18 Hour19 Hour20 Hour21 Hour22 Hour23
## 8680      0      0      0      0      0      0      0      0      0
## 9905      0      0      0      0      1      0      0      0      0
## 5978      0      1      0      0      0      0      0      0      0
## 3968      0      0      1      0      0      0      0      0      0
## 4029      1      0      0      0      0      0      0      0      0
## 6453      1      0      0      0      0      0      0      0      0
##      Weather
## 8680       0
## 9905       0
## 5978       0
## 3968       0
## 4029       0
## 6453       0

test.target<-flightdata[7550:12581,c(47)]
str(test.target)
```

```
##  Factor w/ 2 levels "delayed","ontime": 2 2 2 2 2 2 1 2 2 2 ...

##linear model

model<-glm(formula = FlightStatus ~ ., family = binomial, data = train)
summary(model)

##
## Call:
## glm(formula = FlightStatus ~ ., family = binomial, data = train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.0192   0.2127   0.4425   0.6431   2.6702
##
## Coefficients: (5 not defined because of singularities)
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -15.43051  509.70856  -0.030 0.975849
## CARRIERAA      0.72227    0.22545   3.204 0.001357 **
## CARRIERB6      0.35635    0.14914   2.389 0.016878 *
## CARRIERDL      0.24065    0.11818   2.036 0.041723 *
## CARRIEREV     -0.31658    0.34300  -0.923 0.356018
## CARRIERF9     -0.96140    0.33663  -2.856 0.004291 **
## CARRIERNK     -0.46056    0.29200  -1.577 0.114739
## CARRIERUA     -0.10073    0.22775  -0.442 0.658288
## CARRIERUS      0.68445    0.17524   3.906 9.39e-05 ***
## CARRIERWN           NA         NA      NA       NA
## DESTATL        0.83649    0.23677   3.533 0.000411 ***
## DESTBOS        0.64415    0.20268   3.178 0.001482 **
## DESTDFW             NA         NA      NA       NA
## ORIGINBWI     -0.19738    0.12426  -1.588 0.112177
## ORIGINDCA     -0.04885    0.11643  -0.420 0.674813
## ORIGINIAD           NA         NA      NA       NA
## DAY_OF_WEEK1  -0.21406    0.12055  -1.776 0.075785 .
## DAY_OF_WEEK2   0.05493    0.12429   0.442 0.658519
## DAY_OF_WEEK3  -0.16797    0.12354  -1.360 0.173926
## DAY_OF_WEEK4  -0.15120    0.12558  -1.204 0.228572
## DAY_OF_WEEK5  -0.13851    0.12276  -1.128 0.259212
## DAY_OF_WEEK6  -0.03860    0.14292  -0.270 0.787081
## DAY_OF_WEEK7        NA         NA      NA       NA
## MONTH1        -0.31752    0.10443  -3.040 0.002363 **
## MONTH3        -0.39092    0.10138  -3.856 0.000115 ***
## MONTH6        -0.65888    0.09897  -6.658 2.79e-11 ***
## MONTH9              NA         NA      NA       NA
## Hour5         19.39803  509.70874   0.038 0.969642
## Hour6         18.59220  509.70855   0.036 0.970903
## Hour7         18.46595  509.70856   0.036 0.971100
## Hour8         17.98407  509.70851   0.035 0.971854
## Hour9         16.81735  509.70851   0.033 0.973679
## Hour10        17.09179  509.70851   0.034 0.973250
```

```
## Hour11        16.90589  509.70850   0.033 0.973541
## Hour12        17.35700  509.70851   0.034 0.972835
## Hour13        16.88581  509.70850   0.033 0.973572
## Hour14        16.31913  509.70850   0.032 0.974459
## Hour15        16.79231  509.70850   0.033 0.973719
## Hour16        16.43221  509.70850   0.032 0.974282
## Hour17        16.14358  509.70849   0.032 0.974733
## Hour18        16.13246  509.70850   0.032 0.974751
## Hour19        15.79463  509.70850   0.031 0.975279
## Hour20        15.74527  509.70850   0.031 0.975357
## Hour21        15.00902  509.70851   0.029 0.976509
## Hour22        15.60233  509.70850   0.031 0.975580
## Hour23        -0.21846  549.96127   0.000 0.999683
## Weather       -3.88304    0.38720 -10.029  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 7132.8  on 7548  degrees of freedom
## Residual deviance: 5844.1  on 7507  degrees of freedom
## AIC: 5928.1
##
## Number of Fisher Scoring iterations: 14
```

```
##evaluation
anova(model, test="Chisq")
```

```
## Analysis of Deviance Table
##
## Model: binomial, link: logit
##
## Response: FlightStatus
##
## Terms added sequentially (first to last)
##
##
##              Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
## NULL                          7548     7132.8
## CARRIERAA     1    1.145      7547     7131.6 0.2846827
## CARRIERB6     1    5.188      7546     7126.5 0.0227478 *
## CARRIERDL     1   55.043      7545     7071.4 1.179e-13 ***
## CARRIEREV     1    6.063      7544     7065.3 0.0138033 *
## CARRIERF9     1    3.019      7543     7062.3 0.0823029 .
## CARRIERNK     1    2.899      7542     7059.4 0.0886130 .
## CARRIERUA     1    7.305      7541     7052.1 0.0068754 **
## CARRIERUS     1   20.393      7540     7031.7 6.307e-06 ***
## CARRIERWN     0    0.000      7540     7031.7
## DESTATL       1   12.338      7539     7019.4 0.0004439 ***
## DESTBOS       1   13.257      7538     7006.1 0.0002715 ***
```

```
## DESTDFW       0    0.000      7538     7006.1
## ORIGINBWI     1    0.983      7537     7005.2 0.3213585
## ORIGINDCA     1    0.059      7536     7005.1 0.8086281
## ORIGINIAD     0    0.000      7536     7005.1
## DAY_OF_WEEK1  1    1.693      7535     7003.4 0.1931870
## DAY_OF_WEEK2  1    1.655      7534     7001.7 0.1982844
## DAY_OF_WEEK3  1    0.063      7533     7001.7 0.8012207
## DAY_OF_WEEK4  1    5.347      7532     6996.3 0.0207521 *
## DAY_OF_WEEK5  1    1.269      7531     6995.1 0.2598772
## DAY_OF_WEEK6  1    4.347      7530     6990.7 0.0370698 *
## DAY_OF_WEEK7  0    0.000      7530     6990.7
## MONTH1        1    0.151      7529     6990.6 0.6977028
## MONTH3        1    1.919      7528     6988.6 0.1659535
## MONTH6        1   60.674      7527     6928.0 6.737e-15 ***
## MONTH9        0    0.000      7527     6928.0
## Hour5         1   86.340      7526     6841.6 < 2.2e-16 ***
## Hour6         1  132.128      7525     6709.5 < 2.2e-16 ***
## Hour7         1   99.257      7524     6610.2 < 2.2e-16 ***
## Hour8         1  131.997      7523     6478.3 < 2.2e-16 ***
## Hour9         1   13.354      7522     6464.9 0.0002579 ***
## Hour10        1   38.671      7521     6426.2 5.017e-10 ***
## Hour11        1   32.451      7520     6393.8 1.222e-08 ***
## Hour12        1   72.851      7519     6320.9 < 2.2e-16 ***
## Hour13        1   43.761      7518     6277.2 3.710e-11 ***
## Hour14        1    9.188      7517     6268.0 0.0024365 **
## Hour15        1   48.127      7516     6219.9 3.995e-12 ***
## Hour16        1   27.422      7515     6192.4 1.636e-07 ***
## Hour17        1   18.390      7514     6174.0 1.800e-05 ***
## Hour18        1   21.277      7513     6152.8 3.974e-06 ***
## Hour19        1   13.532      7512     6139.2 0.0002345 ***
## Hour20        1   20.546      7511     6118.7 5.821e-06 ***
## Hour21        1    0.063      7510     6118.6 0.8024748
## Hour22        1   68.038      7509     6050.6 < 2.2e-16 ***
## Hour23        1    0.000      7508     6050.6 0.9998052
## Weather       1  206.475      7507     5844.1 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
library(caret)

y<-varImp(model)
y
```

```
##                  Overall
## CARRIERAA    3.203710e+00
## CARRIERB6    2.389353e+00
## CARRIERDL    2.036271e+00
## CARRIEREV    9.229792e-01
## CARRIERF9    2.855964e+00
## CARRIERNK    1.577247e+00
```

```
## CARRIERUA    4.422785e-01
## CARRIERUS    3.905773e+00
## DESTATL      3.532903e+00
## DESTBOS      3.178098e+00
## ORIGINBWI    1.588482e+00
## ORIGINDCA    4.195522e-01
## DAY_OF_WEEK1 1.775687e+00
## DAY_OF_WEEK2 4.419593e-01
## DAY_OF_WEEK3 1.359696e+00
## DAY_OF_WEEK4 1.204045e+00
## DAY_OF_WEEK5 1.128255e+00
## DAY_OF_WEEK6 2.701026e-01
## MONTH1       3.040351e+00
## MONTH3       3.856088e+00
## MONTH6       6.657512e+00
## Hour5        3.805709e-02
## Hour6        3.647615e-02
## Hour7        3.622846e-02
## Hour8        3.528306e-02
## Hour9        3.299406e-02
## Hour10       3.353248e-02
## Hour11       3.316776e-02
## Hour12       3.405279e-02
## Hour13       3.312836e-02
## Hour14       3.201659e-02
## Hour15       3.294493e-02
## Hour16       3.223844e-02
## Hour17       3.167218e-02
## Hour18       3.165036e-02
## Hour19       3.098757e-02
## Hour20       3.089073e-02
## Hour21       2.944629e-02
## Hour22       3.061030e-02
## Hour23       3.972346e-04
## Weather      1.002858e+01

imp<-data.frame(y)
print(imp)

##                  Overall
## CARRIERAA    3.203710e+00
## CARRIERB6    2.389353e+00
## CARRIERDL    2.036271e+00
## CARRIEREV    9.229792e-01
## CARRIERF9    2.855964e+00
## CARRIERNK    1.577247e+00
## CARRIERUA    4.422785e-01
## CARRIERUS    3.905773e+00
## DESTATL      3.532903e+00
## DESTBOS      3.178098e+00
```

```
## ORIGINBWI    1.588482e+00
## ORIGINDCA    4.195522e-01
## DAY_OF_WEEK1 1.775687e+00
## DAY_OF_WEEK2 4.419593e-01
## DAY_OF_WEEK3 1.359696e+00
## DAY_OF_WEEK4 1.204045e+00
## DAY_OF_WEEK5 1.128255e+00
## DAY_OF_WEEK6 2.701026e-01
## MONTH1       3.040351e+00
## MONTH3       3.856088e+00
## MONTH6       6.657512e+00
## Hour5        3.805709e-02
## Hour6        3.647615e-02
## Hour7        3.622846e-02
## Hour8        3.528306e-02
## Hour9        3.299406e-02
## Hour10       3.353248e-02
## Hour11       3.316776e-02
## Hour12       3.405279e-02
## Hour13       3.312836e-02
## Hour14       3.201659e-02
## Hour15       3.294493e-02
## Hour16       3.223844e-02
## Hour17       3.167218e-02
## Hour18       3.165036e-02
## Hour19       3.098757e-02
## Hour20       3.089073e-02
## Hour21       2.944629e-02
## Hour22       3.061030e-02
## Hour23       3.972346e-04
## Weather      1.002858e+01
```

```
library(base)
fitted.results <- predict(model,newdata=test,type='response')
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be
misleading
```

```
fittedresults <- ifelse(fitted.results >= 0.5,"ontime","delayed")
confusionMatrix(fittedresults,test.target)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction delayed ontime
##    delayed     151     41
##    ontime      757   4083
##
##                Accuracy : 0.8414
##                  95% CI : (0.831, 0.8514)
```

```
##     No Information Rate : 0.8196
##     P-Value [Acc > NIR] : 2.284e-05
##
##                   Kappa : 0.2258
##  Mcnemar's Test P-Value : < 2.2e-16
##
##             Sensitivity : 0.16630
##             Specificity : 0.99006
##          Pos Pred Value : 0.78646
##          Neg Pred Value : 0.84360
##              Prevalence : 0.18045
##          Detection Rate : 0.03001
##    Detection Prevalence : 0.03816
##       Balanced Accuracy : 0.57818
##
##        'Positive' Class : delayed
##
```

```
misClasificError <- mean(fitted.results != test.target)
print(paste('Accuracy',1-misClasificError))
```

```
## [1] "Accuracy 0"
```

```
library(ROCR)
```

```
## Warning: package 'ROCR' was built under R version 3.2.4
```

```
## Loading required package: gplots
```

```
##
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':
##
##     lowess
```

```
p <- predict(model, newdata=test,type="response")
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be
misleading
```

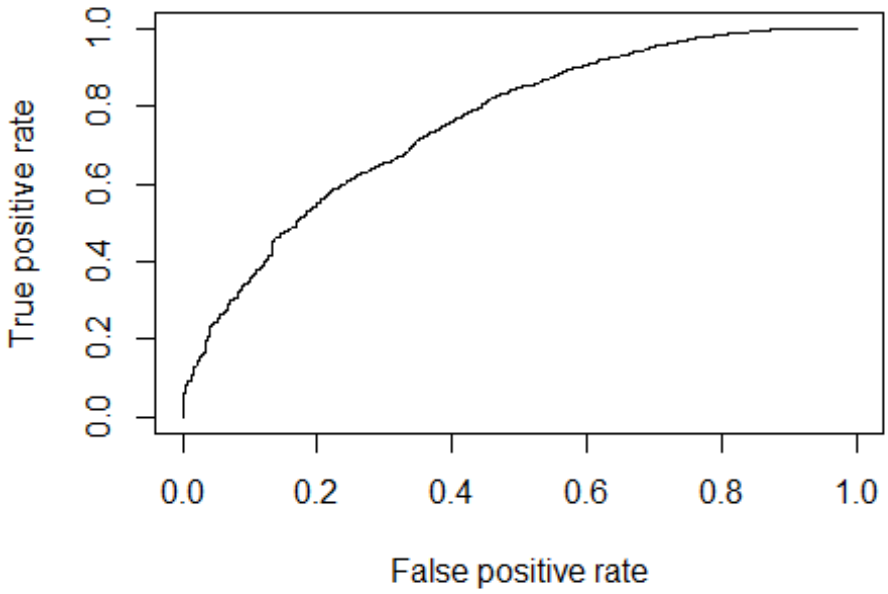
```
summary(p)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.7551  0.8543  0.8169  0.9329  0.9939
```

```
pr <- prediction(p, test.target)
summary(pr)
```

```
##     Length      Class       Mode
##          1 prediction         S4
```

```
prf <- performance(pr, measure = "tpr", x.measure = "fpr")
plot(prf)
```




```
auc <- performance(pr, measure = "auc")

auc <- auc@y.values[[1]]
auc

## [1] 0.758228

###RANDOM FORESTS
library(randomForest)

## Warning: package 'randomForest' was built under R version 3.2.4

## 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

samp <- sample(nrow(combinedata), 0.6 * nrow(combinedata),replace=TRUE)
trainrf <- combinedata[samp, ]
```

```
testrf <- combinedata[-samp, ]
head(testrf)
```

```
##    CARRIERAA CARRIERB6 CARRIERDL CARRIEREV CARRIERF9 CARRIERNK
CARRIERUA
## 1          1         0         0         0         0         0
0
## 2          1         0         0         0         0         0
0
## 5          1         0         0         0         0         0
0
## 7          1         0         0         0         0         0
0
## 8          1         0         0         0         0         0
0
## 11         1         0         0         0         0         0
0
##    CARRIERUS CARRIERWN DESTATL DESTBOS DESTDFW ORIGINBWI ORIGINDCA
## 1          0         0       0       0       1         0         1
## 2          0         0       0       0       1         0         1
## 5          0         0       0       0       1         0         1
## 7          0         0       0       0       1         0         1
## 8          0         0       0       0       1         0         1
## 11         0         0       0       0       1         0         1
##    ORIGINIAD DAY_OF_WEEK1 DAY_OF_WEEK2 DAY_OF_WEEK3 DAY_OF_WEEK4
## 1          0            0            1            0            0
## 2          0            0            0            1            0
## 5          0            0            0            0            0
## 7          0            1            0            0            0
## 8          0            0            1            0            0
## 11         0            0            0            0            0
##    DAY_OF_WEEK5 DAY_OF_WEEK6 DAY_OF_WEEK7 MONTH1 MONTH3 MONTH6
MONTH9
## 1             0            0            0      1      0      0
0
## 2             0            0            0      1      0      0
0
## 5             0            1            0      1      0      0
0
## 7             0            0            0      1      0      0
0
## 8             0            0            0      1      0      0
0
## 11            1            0            0      1      0      0
0
##    Hour5 Hour6 Hour7 Hour8 Hour9 Hour10 Hour11 Hour12 Hour13 Hour14
Hour15
## 1      0     0     0     0     0      0      0      0      0      0
0
## 2      0     0     0     0     0      0      0      0      0      0
```

```
0
## 5      0     0     0     0     0      0      0      0      0      0
1
## 7      0     0     0     0     0      0      0      0      0      0
0
## 8      0     0     0     0     0      0      0      0      0      0
1
## 11     0     0     0     0     0      0      0      0      0      0
1
##    Hour16 Hour17 Hour18 Hour19 Hour20 Hour21 Hour22 Hour23 Weather
## 1       1      0      0      0      0      0      0      0       0
## 2       1      0      0      0      0      0      0      0       0
## 5       0      0      0      0      0      0      0      0       0
## 7       0      1      0      0      0      0      0      0       1
## 8       0      0      0      0      0      0      0      0       0
## 11      0      0      0      0      0      0      0      0       0
##    FlightStatus
## 1        ontime
## 2       delayed
## 5        ontime
## 7       delayed
## 8        ontime
## 11       ontime
```

```
testrf.target<-combinedata[-samp,c(47)]
head(testrf.target)
```

```
## [1] ontime  delayed ontime  delayed ontime  ontime
## Levels: delayed ontime
```

```
modelrf<-randomForest(FlightStatus ~ .,data=trainrf,mtry=3, ntree=500)
```

```
fitted.results2 <- predict(modelrf,newdata=testrf,type='response')
confusionMatrix(fitted.results2,testrf.target)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction delayed ontime
##    delayed     151     24
##    ontime     1097   5614
##
##                Accuracy : 0.8372
##                  95% CI : (0.8283, 0.8459)
##     No Information Rate : 0.8188
##     P-Value [Acc > NIR] : 3.054e-05
##
##                   Kappa : 0.1755
##  Mcnemar's Test P-Value : < 2.2e-16
##
```

```
##             Sensitivity : 0.12099
##             Specificity : 0.99574
##          Pos Pred Value : 0.86286
##          Neg Pred Value : 0.83654
##              Prevalence : 0.18124
##          Detection Rate : 0.02193
##    Detection Prevalence : 0.02541
##       Balanced Accuracy : 0.55837
##
##        'Positive' Class : delayed
##
```

```
table(fitted.results2, testrf.target)
```

```
##                 testrf.target
## fitted.results2 delayed ontime
##         delayed     151     24
##         ontime     1097   5614
```

```
nrow(testrf)
```

```
## [1] 6886
```

```
precision2<-posPredValue(fitted.results2,testrf.target)
precision2
```

```
## [1] 0.8628571
```

```
recall2<-sensitivity(fitted.results2,testrf.target)
recall2
```

```
## [1] 0.1209936
```

```
f2<-(2* precision2 *recall2)/(precision2 + recall2)
f2
```

```
## [1] 0.2122277
```

```
##NAIVE BAYES
```

```
library(e1071)
```

```
## Warning: package 'e1071' was built under R version 3.2.4
```

```
library(klaR)
```

```
## Warning: package 'klaR' was built under R version 3.2.5
```

```
## Loading required package: MASS
```

```
modelnb<-naiveBayes(FlightStatus ~ .,data=train)
modelnb
```

```
##
## Naive Bayes Classifier for Discrete Predictors
```

```
##
## Call:
## naiveBayes.default(x = X, y = Y, laplace = laplace)
##
## A-priori probabilities:
## Y
##   delayed    ontime
## 0.1806862 0.8193138
##
## Conditional probabilities:
##          CARRIERAA
## Y             [,1]      [,2]
##   delayed 0.1862170 0.3894245
##   ontime  0.1988682 0.3991809
##
##          CARRIERB6
## Y             [,1]      [,2]
##   delayed 0.1444282 0.3516523
##   ontime  0.1662086 0.3722979
##
##          CARRIERDL
## Y             [,1]      [,2]
##   delayed 0.2368035 0.4252766
##   ontime  0.3102668 0.4626402
##
##          CARRIEREV
## Y               [,1]       [,2]
##   delayed 0.017595308 0.13152336
##   ontime  0.006952304 0.08309685
##
##          CARRIERF9
## Y               [,1]       [,2]
##   delayed 0.014662757 0.12024292
##   ontime  0.006952304 0.08309685
##
##          CARRIERNK
## Y              [,1]      [,2]
##   delayed 0.02052786 0.1418493
##   ontime  0.01083266 0.1035232
##
##          CARRIERUA
## Y              [,1]      [,2]
##   delayed 0.05351906 0.2251487
##   ontime  0.02942603 0.1690111
##
##          CARRIERUS
## Y              [,1]      [,2]
##   delayed 0.08577713 0.2801374
##   ontime  0.10169766 0.3022747
##
```

```
##          CARRIERWN
## Y             [,1]      [,2]
##   delayed 0.2404692 0.4275252
##   ontime  0.1687955 0.3746015
##
##          DESTATL
## Y             [,1]      [,2]
##   delayed 0.4002933 0.4901374
##   ontime  0.4342765 0.4957017
##
##          DESTBOS
## Y             [,1]      [,2]
##   delayed 0.4068915 0.4914345
##   ontime  0.3933711 0.4885375
##
##          DESTDFW
## Y             [,1]      [,2]
##   delayed 0.1928152 0.3946539
##   ontime  0.1723525 0.3777170
##
##          ORIGINBWI
## Y             [,1]      [,2]
##   delayed 0.3731672 0.4838234
##   ontime  0.3214228 0.4670604
##
##          ORIGINDCA
## Y             [,1]      [,2]
##   delayed 0.4369501 0.4961907
##   ontime  0.5084883 0.4999684
##
##          ORIGINIAD
## Y             [,1]      [,2]
##   delayed 0.1898827 0.3923520
##   ontime  0.1700889 0.3757413
##
##          DAY_OF_WEEK1
## Y             [,1]      [,2]
##   delayed 0.1686217 0.3745548
##   ontime  0.1565077 0.3633653
##
##          DAY_OF_WEEK2
## Y             [,1]      [,2]
##   delayed 0.1480938 0.3553233
##   ontime  0.1660469 0.3721528
##
##          DAY_OF_WEEK3
## Y             [,1]      [,2]
##   delayed 0.1473607 0.3545951
##   ontime  0.1495554 0.3566639
##
```

```
##          DAY_OF_WEEK4
## Y             [,1]      [,2]
##   delayed 0.1590909 0.3658950
##   ontime  0.1354891 0.3422729
##
##          DAY_OF_WEEK5
## Y             [,1]      [,2]
##   delayed 0.1539589 0.3610418
##   ontime  0.1506871 0.3577726
##
##          DAY_OF_WEEK6
## Y              [,1]      [,2]
##   delayed 0.08504399 0.2790495
##   ontime  0.10573969 0.3075290
##
##          DAY_OF_WEEK7
## Y             [,1]      [,2]
##   delayed 0.1378299 0.3448478
##   ontime  0.1359741 0.3427888
##
##          MONTH1
## Y             [,1]      [,2]
##   delayed 0.2316716 0.4220550
##   ontime  0.2307195 0.4213273
##
##          MONTH3
## Y             [,1]      [,2]
##   delayed 0.2646628 0.4413153
##   ontime  0.2462409 0.4308554
##
##          MONTH6
## Y             [,1]      [,2]
##   delayed 0.3189150 0.4662269
##   ontime  0.2452708 0.4302824
##
##          MONTH9
## Y             [,1]      [,2]
##   delayed 0.1847507 0.3882376
##   ontime  0.2777688 0.4479350
##
##          Hour5
## Y               [,1]       [,2]
##   delayed 0.002932551 0.05409341
##   ontime  0.048989491 0.21586351
##
##          Hour6
## Y              [,1]      [,2]
##   delayed 0.01173021 0.1077085
##   ontime  0.08666128 0.2813608
##
```

```
##          Hour7
## Y               [,1]      [,2]
##   delayed 0.01026393 0.1008268
##   ontime  0.06273242 0.2425007
##
##          Hour8
## Y               [,1]      [,2]
##   delayed 0.02639296 0.1603597
##   ontime  0.09312854 0.2906360
##
##          Hour9
## Y               [,1]      [,2]
##   delayed 0.04032258 0.1967868
##   ontime  0.05109135 0.2202019
##
##          Hour10
## Y               [,1]      [,2]
##   delayed 0.04838710 0.2146615
##   ontime  0.07033145 0.2557255
##
##          Hour11
## Y               [,1]      [,2]
##   delayed 0.04472141 0.2067674
##   ontime  0.06451091 0.2456807
##
##          Hour12
## Y               [,1]      [,2]
##   delayed 0.03592375 0.1861683
##   ontime  0.06628941 0.2488074
##
##          Hour13
## Y               [,1]      [,2]
##   delayed 0.05131965 0.2207298
##   ontime  0.06451091 0.2456807
##
##          Hour14
## Y               [,1]      [,2]
##   delayed 0.07404692 0.2619433
##   ontime  0.05529507 0.2285738
##
##          Hour15
## Y               [,1]      [,2]
##   delayed 0.05278592 0.2236878
##   ontime  0.05933711 0.2362737
##
##          Hour16
## Y               [,1]      [,2]
##   delayed 0.06085044 0.2391435
##   ontime  0.05157639 0.2211881
##
```

```
##          Hour17
## Y              [,1]      [,2]
##   delayed 0.1055718 0.3074015
##   ontime  0.0654810 0.2473927
##
##          Hour18
## Y               [,1]      [,2]
##   delayed 0.09017595 0.2865387
##   ontime  0.05286985 0.2237917
##
##          Hour19
## Y               [,1]      [,2]
##   delayed 0.10557185 0.3074015
##   ontime  0.04721099 0.2121070
##
##          Hour20
## Y               [,1]      [,2]
##   delayed 0.07844575 0.2689704
##   ontime  0.03088116 0.1730097
##
##          Hour21
## Y               [,1]      [,2]
##   delayed 0.07111437 0.2571101
##   ontime  0.01438965 0.1191003
##
##          Hour22
## Y               [,1]      [,2]
##   delayed 0.05205279 0.2222150
##   ontime  0.01471302 0.1204113
##
##          Hour23
## Y               [,1]      [,2]
##   delayed 0.03152493 0.1747956
##   ontime  0.00000000 0.0000000
##
##          Weather
## Y                [,1]       [,2]
##   delayed 0.070381232 0.25588223
##   ontime  0.001293452 0.03594423
```

```
fitted.results3<-predict(modelnb,newdata=test)
confusionMatrix(fitted.results3,test.target)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction delayed ontime
##    delayed     305    394
##    ontime      603   3730
##
```

```
##                Accuracy : 0.8019
##                  95% CI : (0.7906, 0.8128)
##     No Information Rate : 0.8196
##     P-Value [Acc > NIR] : 0.9994
##
##                   Kappa : 0.2641
##  Mcnemar's Test P-Value : 4.475e-11
##
##             Sensitivity : 0.33590
##             Specificity : 0.90446
##          Pos Pred Value : 0.43634
##          Neg Pred Value : 0.86084
##              Prevalence : 0.18045
##          Detection Rate : 0.06061
##    Detection Prevalence : 0.13891
##       Balanced Accuracy : 0.62018
##
##        'Positive' Class : delayed
##
```

```
table(fitted.results3,test.target)
```

```
##                test.target
## fitted.results3 delayed ontime
##         delayed     305    394
##         ontime      603   3730
```

```
nrow(test)
```

```
## [1] 5032
```

```
precision3<-posPredValue(fitted.results3,test.target)
precision3
```

```
## [1] 0.4363376
```

```
recall3<-sensitivity(fitted.results3,test.target)
recall3
```

```
## [1] 0.3359031
```

```
f3<-(2*precision3*recall3)/(precision3+recall3)
f3
```

```
## [1] 0.3795893
```

```
###SVM

modelsvm<-svm(FlightStatus ~ .,data=train)
modelsvm
```

```
##
## Call:
```

```
## svm(formula = FlightStatus ~ ., data = train)
##
##
## Parameters:
##    SVM-Type:  C-classification
##  SVM-Kernel:  radial
##        cost:  1
##       gamma:  0.02173913
##
## Number of Support Vectors:  3373
```

```
fitted.results4<-predict(modelsvm,newdata=test)
confusionMatrix(fitted.results4,test.target)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction delayed ontime
##    delayed     191     31
##    ontime      717   4093
##
##                Accuracy : 0.8514
##                  95% CI : (0.8412, 0.8611)
##     No Information Rate : 0.8196
##     P-Value [Acc > NIR] : 1.059e-09
##
##                   Kappa : 0.2875
##  Mcnemar's Test P-Value : < 2.2e-16
##
##             Sensitivity : 0.21035
##             Specificity : 0.99248
##          Pos Pred Value : 0.86036
##          Neg Pred Value : 0.85094
##              Prevalence : 0.18045
##          Detection Rate : 0.03796
##    Detection Prevalence : 0.04412
##       Balanced Accuracy : 0.60142
##
##        'Positive' Class : delayed
##
```

```
table(fitted.results4,test.target)
```

```
##                test.target
## fitted.results4 delayed ontime
##         delayed     191     31
##         ontime      717   4093
```

```
nrow(test)
```

```
## [1] 5032
```

```
precision4<-posPredValue(fitted.results4,test.target)
precision4
```

```
## [1] 0.8603604
```

```
recall4<-sensitivity(fitted.results4,test.target)
recall4
```

```
## [1] 0.2103524
```

```
f4<-(2*precision4*recall4)/(precision4+recall4)
f4
```

```
## [1] 0.3380531
```