Projet Pluie : JUNG Thierry

Téléchargement

On télécharge les données d'entrainement et on les parcoure. On s'apercoit qu'il y a presque une cinquantaine de variables. Celles-ci sont presque exclusivement numérique.

```
rm(list=ls())

MeteoTrain <- read.csv("meteo.train.csv", header = T, quote="")
summary(MeteoTrain)

##        X.              X..Year..        X..Month..          X..Day..
## X..Hour..
##  Min.   :   2.0   Min.   :2010   Min.   : 1.000   Min.   : 1.0   Min.   :0
##  1st Qu.: 721.5   1st Qu.:2012   1st Qu.: 3.000   1st Qu.: 8.0   1st Qu.:0
##  Median :1451.0   Median :2014   Median : 6.000   Median :16.0   Median :0
##  Mean   :1459.8   Mean   :2014   Mean   : 6.436   Mean   :15.8   Mean   :0
##  3rd Qu.:2189.0   3rd Qu.:2016   3rd Qu.: 9.000   3rd Qu.:23.0   3rd Qu.:0
##  Max.   :2940.0   Max.   :2018   Max.   :12.000   Max.   :31.0   Max.   :0
##   X..Minute.. X..Temperature.daily.mean..2.m.above.gnd...
##  Min.   :0    Min.   :-7.63
##  1st Qu.:0    1st Qu.: 6.71
##  Median :0    Median :12.08
##  Mean   :0    Mean   :12.23
##  3rd Qu.:0    3rd Qu.:17.54
##  Max.   :0    Max.   :29.45
##  X..Relative.Humidity.daily.mean..2.m.above.gnd...
##  Min.   :38.33
##  1st Qu.:64.82
##  Median :72.21
##  Mean   :71.40
##  3rd Qu.:78.63
##  Max.   :95.54
##  X..Mean.Sea.Level.Pressure.daily.mean..MSL...
##  Min.   : 978.9
##  1st Qu.:1012.4
##  Median :1017.0
##  Mean   :1017.0
##  3rd Qu.:1022.0
##  Max.   :1042.4
##  X..Total.Precipitation.daily.sum..sfc...
##  Min.   : 0.000
##  1st Qu.: 0.000
##  Median : 0.100
##  Mean   : 2.085
##  3rd Qu.: 2.300
##  Max.   :31.500
```

```
##  X..Snowfall.amount.raw.daily.sum..sfc...
##  Min.   :0.00000
##  1st Qu.:0.00000
##  Median :0.00000
##  Mean   :0.04965
##  3rd Qu.:0.00000
##  Max.   :8.61000
##  X..Total.Cloud.Cover.daily.mean..sfc...
##  Min.   :  0.00
##  1st Qu.: 23.80
##  Median : 51.67
##  Mean   : 50.76
##  3rd Qu.: 78.53
##  Max.   :100.00
##  X..High.Cloud.Cover.daily.mean..high.cld.lay...
##  Min.   :  0.000
##  1st Qu.:  1.657
##  Median : 11.880
##  Mean   : 20.284
##  3rd Qu.: 33.260
##  Max.   :100.000
##  X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...
##  Min.   :  0.00
##  1st Qu.:  1.83
##  Median : 24.98
##  Mean   : 31.50
##  3rd Qu.: 54.21
##  Max.   :100.00
##  X..Low.Cloud.Cover.daily.mean..low.cld.lay...
##  Min.   :  0.00
##  1st Qu.:  9.42
##  Median : 36.35
##  Mean   : 39.34
##  3rd Qu.: 65.76
##  Max.   :100.00
##  X..Sunshine.Duration.daily.sum..sfc...
##  Min.   :   0.0
##  1st Qu.: 114.3
##  Median : 366.8
##  Mean   : 373.1
##  3rd Qu.: 587.7
##  Max.   :1015.8
##  X..Shortwave.Radiation.daily.sum..sfc...
##  Min.   : 265.2
##  1st Qu.:2096.2
##  Median :3675.3
##  Mean   :3984.6
##  3rd Qu.:5723.6
##  Max.   :8363.3
##  X..Wind.Speed.daily.mean..10.m.above.gnd...
```

```
##   Min.   : 1.260
##   1st Qu.: 6.428
##   Median : 9.195
##   Mean   :10.707
##   3rd Qu.:12.977
##   Max.   :42.210
##   X..Wind.Direction.daily.mean..10.m.above.gnd...
##   Min.   : 11.19
##   1st Qu.:152.40
##   Median :206.36
##   Mean   :201.82
##   3rd Qu.:254.19
##   Max.   :331.67
##   X..Wind.Speed.daily.mean..80.m.above.gnd...
##   Min.   : 1.34
##   1st Qu.: 8.68
##   Median :12.41
##   Mean   :14.28
##   3rd Qu.:17.61
##   Max.   :54.03
##   X..Wind.Direction.daily.mean..80.m.above.gnd...
##   Min.   : 12.18
##   1st Qu.:157.42
##   Median :213.78
##   Mean   :206.23
##   3rd Qu.:259.06
##   Max.   :333.43
##   X..Wind.Speed.daily.mean..900.mb...
X..Wind.Direction.daily.mean..900.mb...
##   Min.   : 2.25              Min.   : 17.37
##   1st Qu.:13.02              1st Qu.:144.02
##   Median :19.57              Median :233.47
##   Mean   :24.57              Mean   :206.22
##   3rd Qu.:32.10              3rd Qu.:265.93
##   Max.   :97.06              Max.   :344.82
##   X..Wind.Gust.daily.mean..sfc...
X..Temperature.daily.max..2.m.above.gnd...
##   Min.   : 2.25              Min.   :-3.84
##   1st Qu.: 9.48              1st Qu.:10.58
##   Median :14.06              Median :16.54
##   Mean   :16.69              Mean   :16.54
##   3rd Qu.:21.15              3rd Qu.:22.36
##   Max.   :79.38              Max.   :35.77
##   X..Temperature.daily.min..2.m.above.gnd...
##   Min.   :-12.520
##   1st Qu.:  3.350
##   Median :  8.005
##   Mean   :  8.062
##   3rd Qu.: 13.092
##   Max.   : 23.940
```

```
##  X..Relative.Humidity.daily.max..2.m.above.gnd...
##  Min.   : 59.00
##  1st Qu.: 83.00
##  Median : 89.00
##  Mean   : 87.69
##  3rd Qu.: 94.00
##  Max.   :100.00
##  X..Relative.Humidity.daily.min..2.m.above.gnd...
##  Min.   :19.00
##  1st Qu.:45.00
##  Median :54.00
##  Mean   :54.04
##  3rd Qu.:63.00
##  Max.   :92.00
##  X..Mean.Sea.Level.Pressure.daily.max..MSL...
##  Min.   : 981.9
##  1st Qu.:1015.4
##  Median :1019.5
##  Mean   :1019.9
##  3rd Qu.:1024.7
##  Max.   :1045.4
##  X..Mean.Sea.Level.Pressure.daily.min..MSL...
##  Min.   : 977
##  1st Qu.:1009
##  Median :1015
##  Mean   :1014
##  3rd Qu.:1019
##  Max.   :1039
##  X..Total.Cloud.Cover.daily.max..sfc...
X..Total.Cloud.Cover.daily.min..sfc...
##  Min.   :  0.00                        Min.   :  0.000
##  1st Qu.:100.00                        1st Qu.:  0.000
##  Median :100.00                        Median :  0.000
##  Mean   : 88.23                        Mean   :  8.692
##  3rd Qu.:100.00                        3rd Qu.:  2.400
##  Max.   :100.00                        Max.   :100.000
##  X..High.Cloud.Cover.daily.max..high.cld.lay...
##  Min.   :  0.00
##  1st Qu.: 15.00
##  Median : 97.00
##  Mean   : 60.17
##  3rd Qu.:100.00
##  Max.   :100.00
##  X..High.Cloud.Cover.daily.min..high.cld.lay...
##  Min.   :  0.0000
##  1st Qu.:  0.0000
##  Median :  0.0000
##  Mean   :  0.9432
##  3rd Qu.:  0.0000
##  Max.   :100.0000
```

```
##   X..Medium.Cloud.Cover.daily.max..mid.cld.lay...
##   Min.    :  0.00
##   1st Qu.: 22.75
##   Median :100.00
##   Mean    : 70.94
##   3rd Qu.:100.00
##   Max.    :100.00
##   X..Medium.Cloud.Cover.daily.min..mid.cld.lay...
##   Min.    :  0.000
##   1st Qu.:  0.000
##   Median :  0.000
##   Mean    :  2.097
##   3rd Qu.:  0.000
##   Max.    :100.000
##   X..Low.Cloud.Cover.daily.max..low.cld.lay...
##   Min.    :  0
##   1st Qu.:100
##   Median :100
##   Mean    : 80
##   3rd Qu.:100
##   Max.    :100
##   X..Low.Cloud.Cover.daily.min..low.cld.lay...
##   Min.    :  0.000
##   1st Qu.:  0.000
##   Median :  0.000
##   Mean    :  3.879
##   3rd Qu.:  0.000
##   Max.    :100.000
##   X..Wind.Speed.daily.max..10.m.above.gnd...
##   Min.    : 2.52
##   1st Qu.:12.32
##   Median :17.36
##   Mean    :19.06
##   3rd Qu.:23.44
##   Max.    :79.99
##   X..Wind.Speed.daily.min..10.m.above.gnd...
##   Min.    : 0.00
##   1st Qu.: 1.14
##   Median : 2.41
##   Mean    : 3.57
##   3rd Qu.: 4.45
##   Max.    :27.73
##   X..Wind.Speed.daily.max..80.m.above.gnd...
##   Min.    : 3.98
##   1st Qu.:18.27
##   Median :23.85
##   Mean    :25.35
##   3rd Qu.:29.92
##   Max.    :93.84
##   X..Wind.Speed.daily.min..80.m.above.gnd...
```

```
X..Wind.Speed.daily.max..900.mb...
##  Min.   : 0.000                        Min.   :  4.02
##  1st Qu.: 1.140                        1st Qu.: 24.54
##  Median : 2.600                        Median : 37.12
##  Mean   : 4.727                        Mean   : 41.82
##  3rd Qu.: 5.830                        3rd Qu.: 54.37
##  Max.   :37.700                        Max.   :136.25
##  X..Wind.Speed.daily.min..900.mb... X..Wind.Gust.daily.max..sfc...
##  Min.   : 0.00                       Min.   : 4.32
##  1st Qu.: 3.05                       1st Qu.:19.08
##  Median : 6.73                       Median :26.10
##  Mean   :11.09                       Mean   :29.31
##  3rd Qu.:15.31                       3rd Qu.:37.08
##  Max.   :76.13                       Max.   :95.04
##  X..Wind.Gust.daily.min..sfc... X..pluie.demain.. X..temp.demain...
##  Min.   : 0.000                 Mode :logical    Min.   :-7.10
##  1st Qu.: 2.160                 FALSE:579        1st Qu.: 6.77
##  Median : 3.960                 TRUE :601        Median :12.37
##  Mean   : 6.502                                  Mean   :12.19
##  3rd Qu.: 8.280                                  3rd Qu.:17.34
##  Max.   :57.960                                  Max.   :29.96
```

```r
str(MeteoTrain)
```

```
## 'data.frame':    1180 obs. of  48 variables:
##  $ X.                                       : int  2 4 6 8 10 12
14 16 18 20 ...
##  $ X..Year..                                : int  2010 2010 2010
2010 2010 2010 2010 2010 2010 2010 ...
##  $ X..Month..                               : int  6 6 6 6 6 6 6 6
6 6 ...
##  $ X..Day..                                 : int  2 4 6 8 10 12
14 16 18 20 ...
##  $ X..Hour..                                : int  0 0 0 0 0 0 0 0
0 0 ...
##  $ X..Minute..                              : int  0 0 0 0 0 0 0 0
0 0 ...
##  $ X..Temperature.daily.mean..2.m.above.gnd...     : num  15 17.3 21.6
20.2 22.6 ...
##  $ X..Relative.Humidity.daily.mean..2.m.above.gnd...: num  76.5 77.6 69.5
75.1 73.5 ...
##  $ X..Mean.Sea.Level.Pressure.daily.mean..MSL...    : num  1015 1017 1015
1008 1004 ...
##  $ X..Total.Precipitation.daily.sum..sfc...         : num  1 0 3.7 0.2 0
2.2 1.8 1.8 17.5 1.2 ...
##  $ X..Snowfall.amount.raw.daily.sum..sfc...         : num  0 0 0 0 0 0 0 0
0 0 ...
##  $ X..Total.Cloud.Cover.daily.mean..sfc...          : num  79.8 4.7 42.1
67.5 56.3 ...
##  $ X..High.Cloud.Cover.daily.mean..high.cld.lay... : num  3 0.67 21.21
```

```
54.71 50.25 ...
##  $ X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... : num  31.6 0 25.9
65.8 55.3 ...
##  $ X..Low.Cloud.Cover.daily.mean..low.cld.lay...    : num  79.2 4.5 35.3
18.9 34.2 ...
##  $ X..Sunshine.Duration.daily.sum..sfc...           : num  287.2 821.4
441.3 41.9 473.2 ...
##  $ X..Shortwave.Radiation.daily.sum..sfc...          : num  6710 7974 4834
5390 7216 ...
##  $ X..Wind.Speed.daily.mean..10.m.above.gnd...      : num  11.64 6.34 8.4
5.4 9.16 ...
##  $ X..Wind.Direction.daily.mean..10.m.above.gnd...  : num  275 230 215 205
179 ...
##  $ X..Wind.Speed.daily.mean..80.m.above.gnd...      : num  14.99 8.92
10.38 6.53 11.91 ...
##  $ X..Wind.Direction.daily.mean..80.m.above.gnd...  : num  268 199 208 206
186 ...
##  $ X..Wind.Speed.daily.mean..900.mb...              : num  20.6 27.9 18.9
10.4 21.9 ...
##  $ X..Wind.Direction.daily.mean..900.mb...          : num  180.4 93.7
250.1 238.6 153 ...
##  $ X..Wind.Gust.daily.mean..sfc...                  : num  14.88 9.48 13.5
5.31 12.21 ...
##  $ X..Temperature.daily.max..2.m.above.gnd...       : num  18.5 25 26.2
24.2 30.7 ...
##  $ X..Temperature.daily.min..2.m.above.gnd...       : num  11.1 10.4 17.7
14.7 16.9 ...
##  $ X..Relative.Humidity.daily.max..2.m.above.gnd... : int  94 92 91 89 97
92 96 96 97 95 ...
##  $ X..Relative.Humidity.daily.min..2.m.above.gnd... : int  59 54 57 62 39
65 69 64 74 61 ...
##  $ X..Mean.Sea.Level.Pressure.daily.max..MSL...     : num  1017 1019 1016
1010 1006 ...
##  $ X..Mean.Sea.Level.Pressure.daily.min..MSL...     : num  1014 1016 1013
1006 1001 ...
##  $ X..Total.Cloud.Cover.daily.max..sfc...           : num  100 28 100 100
100 100 100 100 100 100 ...
##  $ X..Total.Cloud.Cover.daily.min..sfc...           : num  0 0 0 0 0 0 0
100 0 0 ...
##  $ X..High.Cloud.Cover.daily.max..high.cld.lay...   : int  16 11 100 100
100 28 100 100 100 24 ...
##  $ X..High.Cloud.Cover.daily.min..high.cld.lay...   : int  0 0 0 0 0 0 0 0
0 0 ...
##  $ X..Medium.Cloud.Cover.daily.max..mid.cld.lay...  : int  100 0 100 100
100 100 100 100 100 41 ...
##  $ X..Medium.Cloud.Cover.daily.min..mid.cld.lay...  : int  0 0 0 0 0 0 0 0
0 0 ...
##  $ X..Low.Cloud.Cover.daily.max..low.cld.lay...     : int  100 28 100 100
100 100 100 100 100 100 ...
##  $ X..Low.Cloud.Cover.daily.min..low.cld.lay...     : int  0 0 0 0 0 0 0
```

```
29 0 0 ...
##  $ X..Wind.Speed.daily.max..10.m.above.gnd...    : num  22 15.5 22.7
10.7 20.5 ...
##  $ X..Wind.Speed.daily.min..10.m.above.gnd...    : num  5.62 1.08 2.41
0 2.52 2.28 1.3 4.32 7.2 8.05 ...
##  $ X..Wind.Speed.daily.max..80.m.above.gnd...    : num  23.8 18.7 32
10.2 23.4 ...
##  $ X..Wind.Speed.daily.min..80.m.above.gnd...    : num  8.65 0 0.51
1.44 2.97 ...
##  $ X..Wind.Speed.daily.max..900.mb...            : num  32.1 48.1 44
22.2 40.8 ...
##  $ X..Wind.Speed.daily.min..900.mb...            : num  12.25 6.62 5.48
4.69 4.68 ...
##  $ X..Wind.Gust.daily.max..sfc...                : num  25.2 20.2 41.8
11.2 24.1 ...
##  $ X..Wind.Gust.daily.min..sfc...                : num  6.48 2.16 1.08
0.36 1.44 ...
##  $ X..pluie.demain..                             : logi   FALSE FALSE
TRUE TRUE TRUE TRUE ...
##  $ X..temp.demain...                             : num  14.3 19.9 19.3
22.1 20.9 ...
```

Analyse

En faisant une regression sur l'ensemble de celles-ci, 2 variables ressortent comme très pertinentes : "La direction du vent à l'altitude à laquelle la pression vaut 900 hPa" et "La pression".

```
gAll = glm(formula = X..pluie.demain.. ~ ., data=MeteoTrain, family=binomial)
summary(gAll)

##
## Call:
## glm(formula = X..pluie.demain.. ~ ., family = binomial, data = MeteoTrain)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.9042  -0.8201   0.2758   0.8292   2.9670
##
## Coefficients: (2 not defined because of singularities)
##                                             Estimate Std. Error z
value
## (Intercept)                                 5.120e+04  9.310e+04
0.550
## X.                                          6.984e-02  1.268e-01
0.551
## X..Year..                                  -2.544e+01  4.631e+01  -
0.549
## X..Month..                                 -2.134e+00  3.858e+00  -
0.553
```

```
## X..Day..                                         -5.962e-02  1.267e-01  -
0.471
## X..Hour..                                                NA         NA
NA
## X..Minute..                                              NA         NA
NA
## X..Temperature.daily.mean..2.m.above.gnd...       3.154e-01  1.697e-01
1.858
## X..Relative.Humidity.daily.mean..2.m.above.gnd... 2.343e-02  3.271e-02
0.716
## X..Mean.Sea.Level.Pressure.daily.mean..MSL...     5.499e-01  1.408e-01
3.905
## X..Total.Precipitation.daily.sum..sfc...          3.305e-02  2.843e-02
1.163
## X..Snowfall.amount.raw.daily.sum..sfc...          -3.253e-01 2.421e-01  -
1.343
## X..Total.Cloud.Cover.daily.mean..sfc...           9.944e-03  1.217e-02
0.817
## X..High.Cloud.Cover.daily.mean..high.cld.lay...   -1.955e-03 6.894e-03  -
0.284
## X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...  6.272e-03  6.738e-03
0.931
## X..Low.Cloud.Cover.daily.mean..low.cld.lay...     -1.870e-03 8.259e-03  -
0.226
## X..Sunshine.Duration.daily.sum..sfc...            2.371e-04  8.986e-04
0.264
## X..Shortwave.Radiation.daily.sum..sfc...          1.201e-04  1.051e-04
1.143
## X..Wind.Speed.daily.mean..10.m.above.gnd...       -5.218e-02 9.699e-02  -
0.538
## X..Wind.Direction.daily.mean..10.m.above.gnd...   5.506e-03  5.807e-03
0.948
## X..Wind.Speed.daily.mean..80.m.above.gnd...       -8.037e-02 6.972e-02  -
1.153
## X..Wind.Direction.daily.mean..80.m.above.gnd...   -9.614e-03 5.990e-03  -
1.605
## X..Wind.Speed.daily.mean..900.mb...               2.344e-02  2.612e-02
0.898
## X..Wind.Direction.daily.mean..900.mb...           5.826e-03  1.467e-03
3.972
## X..Wind.Gust.daily.mean..sfc...                   1.667e-02  3.699e-02
0.451
## X..Temperature.daily.max..2.m.above.gnd...        6.611e-03  9.698e-02
0.068
## X..Temperature.daily.min..2.m.above.gnd...        -1.674e-01 8.778e-02  -
1.907
## X..Relative.Humidity.daily.max..2.m.above.gnd...  -3.522e-03 2.083e-02  -
0.169
## X..Relative.Humidity.daily.min..2.m.above.gnd...  -4.995e-03 1.870e-02  -
0.267
```

```
## X..Mean.Sea.Level.Pressure.daily.max..MSL...      -2.790e-01  7.588e-02  -
3.677
## X..Mean.Sea.Level.Pressure.daily.min..MSL...      -3.330e-01  7.617e-02  -
4.372
## X..Total.Cloud.Cover.daily.max..sfc...             2.687e-03  4.901e-03
0.548
## X..Total.Cloud.Cover.daily.min..sfc...             7.628e-03  6.270e-03
1.217
## X..High.Cloud.Cover.daily.max..high.cld.lay...     3.102e-03  2.908e-03
1.067
## X..High.Cloud.Cover.daily.min..high.cld.lay...     4.344e-03  2.123e-02
0.205
## X..Medium.Cloud.Cover.daily.max..mid.cld.lay...    6.326e-03  3.179e-03
1.990
## X..Medium.Cloud.Cover.daily.min..mid.cld.lay...   -4.829e-03  9.605e-03  -
0.503
## X..Low.Cloud.Cover.daily.max..low.cld.lay...       2.935e-03  3.438e-03
0.854
## X..Low.Cloud.Cover.daily.min..low.cld.lay...       1.944e-04  7.059e-03
0.028
## X..Wind.Speed.daily.max..10.m.above.gnd...         4.765e-02  3.468e-02
1.374
## X..Wind.Speed.daily.min..10.m.above.gnd...         1.701e-01  6.433e-02
2.644
## X..Wind.Speed.daily.max..80.m.above.gnd...        -2.237e-04  2.863e-02  -
0.008
## X..Wind.Speed.daily.min..80.m.above.gnd...        -5.420e-02  4.237e-02  -
1.279
## X..Wind.Speed.daily.max..900.mb...                -1.415e-02  1.219e-02  -
1.161
## X..Wind.Speed.daily.min..900.mb...                -8.647e-03  1.923e-02  -
0.450
## X..Wind.Gust.daily.max..sfc...                     2.244e-02  1.743e-02
1.288
## X..Wind.Gust.daily.min..sfc...                     8.062e-03  2.799e-02
0.288
## X..temp.demain...                                 -1.363e-01  3.909e-02  -
3.486
##                                                    Pr(>|z|)
## (Intercept)                                        0.582372
## X.                                                 0.581688
## X..Year..                                          0.582800
## X..Month..                                         0.580184
## X..Day..                                           0.637974
## X..Hour..                                                NA
## X..Minute..                                              NA
## X..Temperature.daily.mean..2.m.above.gnd...        0.063116 .
## X..Relative.Humidity.daily.mean..2.m.above.gnd... 0.473825
## X..Mean.Sea.Level.Pressure.daily.mean..MSL...      9.42e-05 ***
## X..Total.Precipitation.daily.sum..sfc...           0.244996
```
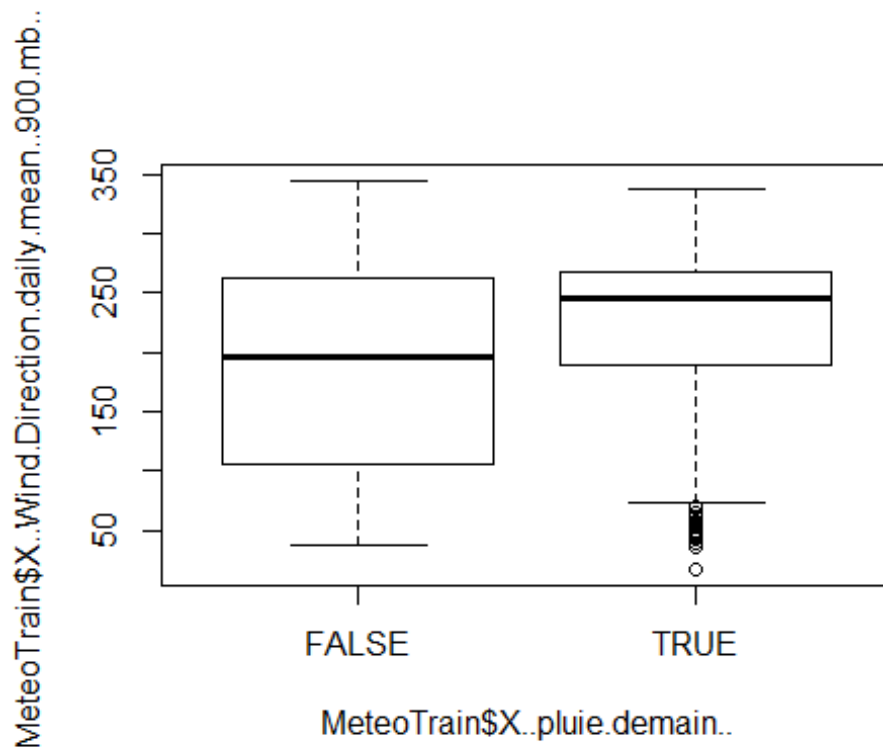
```
## X..Snowfall.amount.raw.daily.sum..sfc...              0.179203
## X..Total.Cloud.Cover.daily.mean..sfc...               0.413946
## X..High.Cloud.Cover.daily.mean..high.cld.lay...       0.776755
## X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...      0.351903
## X..Low.Cloud.Cover.daily.mean..low.cld.lay...         0.820854
## X..Sunshine.Duration.daily.sum..sfc...                0.791883
## X..Shortwave.Radiation.daily.sum..sfc...              0.253067
## X..Wind.Speed.daily.mean..10.m.above.gnd...           0.590613
## X..Wind.Direction.daily.mean..10.m.above.gnd...       0.343052
## X..Wind.Speed.daily.mean..80.m.above.gnd...           0.249019
## X..Wind.Direction.daily.mean..80.m.above.gnd...       0.108478
## X..Wind.Speed.daily.mean..900.mb...                   0.369449
## X..Wind.Direction.daily.mean..900.mb...               7.12e-05 ***
## X..Wind.Gust.daily.mean..sfc...                       0.652239
## X..Temperature.daily.max..2.m.above.gnd...            0.945655
## X..Temperature.daily.min..2.m.above.gnd...            0.056554 .
## X..Relative.Humidity.daily.max..2.m.above.gnd...      0.865761
## X..Relative.Humidity.daily.min..2.m.above.gnd...      0.789351
## X..Mean.Sea.Level.Pressure.daily.max..MSL...          0.000236 ***
## X..Mean.Sea.Level.Pressure.daily.min..MSL...          1.23e-05 ***
## X..Total.Cloud.Cover.daily.max..sfc...                0.583546
## X..Total.Cloud.Cover.daily.min..sfc...                0.223713
## X..High.Cloud.Cover.daily.max..high.cld.lay...        0.286132
## X..High.Cloud.Cover.daily.min..high.cld.lay...        0.837823
## X..Medium.Cloud.Cover.daily.max..mid.cld.lay...       0.046588 *
## X..Medium.Cloud.Cover.daily.min..mid.cld.lay...       0.615105
## X..Low.Cloud.Cover.daily.max..low.cld.lay...          0.393338
## X..Low.Cloud.Cover.daily.min..low.cld.lay...          0.978027
## X..Wind.Speed.daily.max..10.m.above.gnd...            0.169528
## X..Wind.Speed.daily.min..10.m.above.gnd...            0.008187 **
## X..Wind.Speed.daily.max..80.m.above.gnd...            0.993765
## X..Wind.Speed.daily.min..80.m.above.gnd...            0.200843
## X..Wind.Speed.daily.max..900.mb...                    0.245441
## X..Wind.Speed.daily.min..900.mb...                    0.653022
## X..Wind.Gust.daily.max..sfc...                        0.197801
## X..Wind.Gust.daily.min..sfc...                        0.773367
## X..temp.demain...                                     0.000491 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1635.4  on 1179  degrees of freedom
## Residual deviance: 1220.1  on 1134  degrees of freedom
## AIC: 1312.1
##
## Number of Fisher Scoring iterations: 5
```
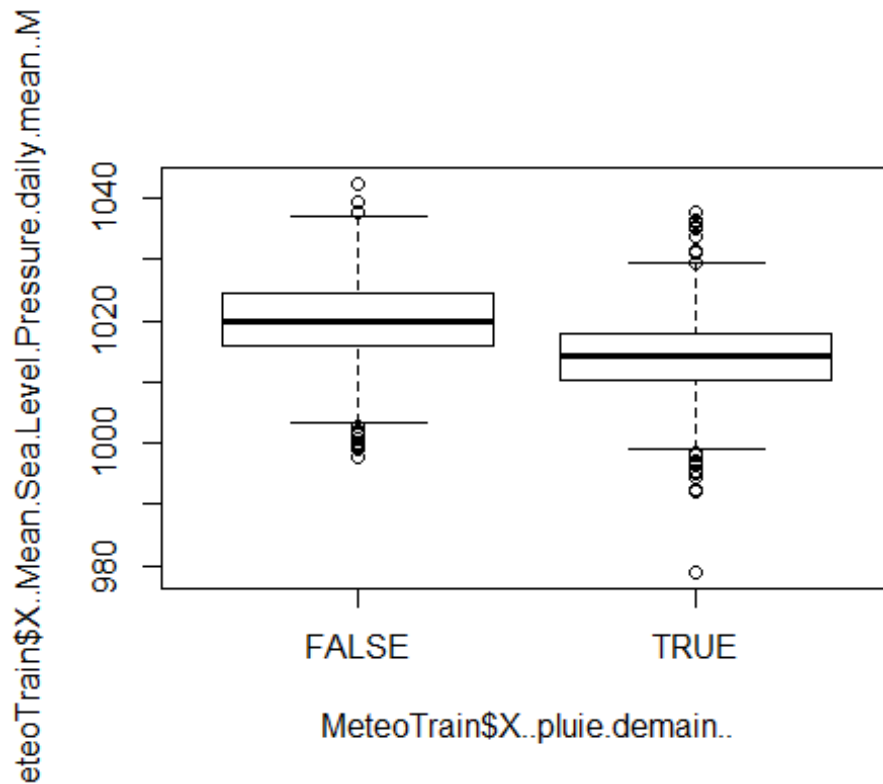
Graphique

En faisant 2 boites à moustache, on note effectivement une vraie influence de ces variables.

```
boxplot(MeteoTrain$X..Wind.Direction.daily.mean..900.mb... ~
MeteoTrain$X..pluie.demain..)
```



```
boxplot(MeteoTrain$X..Mean.Sea.Level.Pressure.daily.mean..MSL...~
MeteoTrain$X..pluie.demain..)
```

Regression

On effectue une regression sur les 2 variables.

```
g1 = glm(formula = X..pluie.demain.. ~
X..Mean.Sea.Level.Pressure.daily.mean..MSL...+
X..Wind.Direction.daily.mean..900.mb... , data=MeteoTrain, family=binomial)
summary(g1)

##
## Call:
## glm(formula = X..pluie.demain.. ~
X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
##      X..Wind.Direction.daily.mean..900.mb..., family = binomial,
##      data = MeteoTrain)
##
## Deviance Residuals:
##      Min       1Q    Median       3Q      Max
## -2.1359  -1.0090    0.4479   0.9951   2.3303
##
## Coefficients:
##                                                Estimate Std. Error z
value
## (Intercept)                                  107.531464    9.712600
11.071
## X..Mean.Sea.Level.Pressure.daily.mean..MSL...  -0.106904    0.009546 -
11.199
```

```
## X..Wind.Direction.daily.mean..900.mb...              0.005939    0.000889
6.681
##                                                    Pr(>|z|)
## (Intercept)                                         < 2e-16 ***
## X..Mean.Sea.Level.Pressure.daily.mean..MSL...       < 2e-16 ***
## X..Wind.Direction.daily.mean..900.mb...             2.37e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1635.4  on 1179  degrees of freedom
## Residual deviance: 1410.4  on 1177  degrees of freedom
## AIC: 1416.4
##
## Number of Fisher Scoring iterations: 3
```

Automatique

On peut aussi essayer de chercher de facon automatique le meilleur modèle avec la fonction
step. Celle-ci étant trop longue avec 48 variables. On effectue déja un tri en ne conservant
que les données sur les moyennes.

```
g2 = step(glm(formula = X..pluie.demain.. ~
X..Temperature.daily.mean..2.m.above.gnd... +
 X..Relative.Humidity.daily.mean..2.m.above.gnd... +
X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
X..Total.Precipitation.daily.sum..sfc... +
X..Snowfall.amount.raw.daily.sum..sfc... +
X..Total.Cloud.Cover.daily.mean..sfc... +
X..High.Cloud.Cover.daily.mean..high.cld.lay... +
X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
X..Low.Cloud.Cover.daily.mean..low.cld.lay... +
X..Sunshine.Duration.daily.sum..sfc... +
X..Shortwave.Radiation.daily.sum..sfc... +
X..Wind.Speed.daily.mean..10.m.above.gnd... +
X..Wind.Direction.daily.mean..10.m.above.gnd... +
X..Wind.Speed.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..80.m.above.gnd... +
X..Wind.Speed.daily.mean..900.mb... +
X..Wind.Direction.daily.mean..900.mb... +
X..Wind.Gust.daily.mean..sfc... , data=MeteoTrain, family=binomial))

## Start:  AIC=1350.25
## X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##     X..Relative.Humidity.daily.mean..2.m.above.gnd... +
X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
##     X..Total.Precipitation.daily.sum..sfc... +
X..Snowfall.amount.raw.daily.sum..sfc... +
##     X..Total.Cloud.Cover.daily.mean..sfc... +
```

```
X..High.Cloud.Cover.daily.mean..high.cld.lay... +
##      X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
X..Low.Cloud.Cover.daily.mean..low.cld.lay... +
##      X..Sunshine.Duration.daily.sum..sfc... +
X..Shortwave.Radiation.daily.sum..sfc... +
##      X..Wind.Speed.daily.mean..10.m.above.gnd... +
X..Wind.Direction.daily.mean..10.m.above.gnd... +
##      X..Wind.Speed.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..80.m.above.gnd... +
##      X..Wind.Speed.daily.mean..900.mb... +
X..Wind.Direction.daily.mean..900.mb... +
##      X..Wind.Gust.daily.mean..sfc...
##
##                                                  Df Deviance    AIC
## - X..Wind.Speed.daily.mean..900.mb...             1   1312.3 1348.3
## - X..Relative.Humidity.daily.mean..2.m.above.gnd...  1   1312.5 1348.5
## - X..Sunshine.Duration.daily.sum..sfc...          1   1312.6 1348.6
## - X..Total.Precipitation.daily.sum..sfc...        1   1312.8 1348.8
## - X..Wind.Direction.daily.mean..10.m.above.gnd... 1   1313.0 1349.0
## - X..High.Cloud.Cover.daily.mean..high.cld.lay... 1   1313.1 1349.1
## - X..Low.Cloud.Cover.daily.mean..low.cld.lay...   1   1313.6 1349.6
## - X..Shortwave.Radiation.daily.sum..sfc...        1   1313.8 1349.8
## <none>                                                1312.2 1350.2
## - X..Wind.Gust.daily.mean..sfc...                 1   1314.6 1350.6
## - X..Wind.Direction.daily.mean..80.m.above.gnd... 1   1314.8 1350.8
## - X..Snowfall.amount.raw.daily.sum..sfc...        1   1314.8 1350.8
## - X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... 1   1314.9 1350.9
## - X..Temperature.daily.mean..2.m.above.gnd...     1   1315.2 1351.2
## - X..Wind.Speed.daily.mean..10.m.above.gnd...     1   1315.7 1351.7
## - X..Total.Cloud.Cover.daily.mean..sfc...         1   1316.5 1352.5
## - X..Wind.Speed.daily.mean..80.m.above.gnd...     1   1317.2 1353.2
## - X..Wind.Direction.daily.mean..900.mb...         1   1329.7 1365.7
## - X..Mean.Sea.Level.Pressure.daily.mean..MSL...   1   1346.3 1382.3
##
## Step:  AIC=1348.28
## X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##      X..Relative.Humidity.daily.mean..2.m.above.gnd... +
X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
##      X..Total.Precipitation.daily.sum..sfc... +
X..Snowfall.amount.raw.daily.sum..sfc... +
##      X..Total.Cloud.Cover.daily.mean..sfc... +
X..High.Cloud.Cover.daily.mean..high.cld.lay... +
##      X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
X..Low.Cloud.Cover.daily.mean..low.cld.lay... +
##      X..Sunshine.Duration.daily.sum..sfc... +
X..Shortwave.Radiation.daily.sum..sfc... +
##      X..Wind.Speed.daily.mean..10.m.above.gnd... +
X..Wind.Direction.daily.mean..10.m.above.gnd... +
##      X..Wind.Speed.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..80.m.above.gnd... +
```

```
##      X..Wind.Direction.daily.mean..900.mb... +
X..Wind.Gust.daily.mean..sfc...
##
##                                                  Df Deviance    AIC
## - X..Relative.Humidity.daily.mean..2.m.above.gnd...  1   1312.5 1346.5
## - X..Sunshine.Duration.daily.sum..sfc...             1   1312.6 1346.6
## - X..Total.Precipitation.daily.sum..sfc...           1   1312.8 1346.8
## - X..Wind.Direction.daily.mean..10.m.above.gnd...    1   1313.0 1347.0
## - X..High.Cloud.Cover.daily.mean..high.cld.lay...    1   1313.2 1347.2
## - X..Low.Cloud.Cover.daily.mean..low.cld.lay...      1   1313.7 1347.7
## - X..Shortwave.Radiation.daily.sum..sfc...           1   1313.8 1347.8
## <none>                                                   1312.3 1348.3
## - X..Wind.Direction.daily.mean..80.m.above.gnd...    1   1314.8 1348.8
## - X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...   1   1314.9 1348.9
## - X..Snowfall.amount.raw.daily.sum..sfc...           1   1315.0 1349.0
## - X..Temperature.daily.mean..2.m.above.gnd...        1   1315.2 1349.2
## - X..Wind.Speed.daily.mean..10.m.above.gnd...        1   1315.7 1349.7
## - X..Total.Cloud.Cover.daily.mean..sfc...            1   1316.5 1350.5
## - X..Wind.Gust.daily.mean..sfc...                    1   1317.0 1351.0
## - X..Wind.Speed.daily.mean..80.m.above.gnd...        1   1317.3 1351.3
## - X..Wind.Direction.daily.mean..900.mb...            1   1329.8 1363.8
## - X..Mean.Sea.Level.Pressure.daily.mean..MSL...      1   1346.4 1380.4
##
## Step:  AIC=1346.53
## X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##      X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
X..Total.Precipitation.daily.sum..sfc... +
##      X..Snowfall.amount.raw.daily.sum..sfc... +
X..Total.Cloud.Cover.daily.mean..sfc... +
##      X..High.Cloud.Cover.daily.mean..high.cld.lay... +
X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
##      X..Low.Cloud.Cover.daily.mean..low.cld.lay... +
X..Sunshine.Duration.daily.sum..sfc... +
##      X..Shortwave.Radiation.daily.sum..sfc... +
X..Wind.Speed.daily.mean..10.m.above.gnd... +
##      X..Wind.Direction.daily.mean..10.m.above.gnd... +
X..Wind.Speed.daily.mean..80.m.above.gnd... +
##      X..Wind.Direction.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..900.mb... +
##      X..Wind.Gust.daily.mean..sfc...
##
##                                                  Df Deviance    AIC
## - X..Sunshine.Duration.daily.sum..sfc...             1   1312.9 1344.9
## - X..Total.Precipitation.daily.sum..sfc...           1   1313.2 1345.2
## - X..Wind.Direction.daily.mean..10.m.above.gnd...    1   1313.4 1345.4
## - X..High.Cloud.Cover.daily.mean..high.cld.lay...    1   1313.4 1345.4
## - X..Low.Cloud.Cover.daily.mean..low.cld.lay...      1   1313.8 1345.8
## - X..Shortwave.Radiation.daily.sum..sfc...           1   1313.8 1345.8
## <none>                                                   1312.5 1346.5
## - X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...   1   1315.1 1347.1
```

```
## - X..Wind.Direction.daily.mean..80.m.above.gnd...    1   1315.1 1347.1
## - X..Snowfall.amount.raw.daily.sum..sfc...           1   1315.2 1347.2
## - X..Temperature.daily.mean..2.m.above.gnd...         1   1315.5 1347.5
## - X..Wind.Speed.daily.mean..10.m.above.gnd...         1   1315.9 1347.9
## - X..Total.Cloud.Cover.daily.mean..sfc...             1   1316.8 1348.8
## - X..Wind.Gust.daily.mean..sfc...                     1   1317.0 1349.0
## - X..Wind.Speed.daily.mean..80.m.above.gnd...         1   1317.4 1349.4
## - X..Wind.Direction.daily.mean..900.mb...             1   1329.8 1361.8
## - X..Mean.Sea.Level.Pressure.daily.mean..MSL...       1   1346.6 1378.6
##
## Step:  AIC=1344.89
## X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##      X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
X..Total.Precipitation.daily.sum..sfc... +
##      X..Snowfall.amount.raw.daily.sum..sfc... +
X..Total.Cloud.Cover.daily.mean..sfc... +
##      X..High.Cloud.Cover.daily.mean..high.cld.lay... +
X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
##      X..Low.Cloud.Cover.daily.mean..low.cld.lay... +
X..Shortwave.Radiation.daily.sum..sfc... +
##      X..Wind.Speed.daily.mean..10.m.above.gnd... +
X..Wind.Direction.daily.mean..10.m.above.gnd... +
##      X..Wind.Speed.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..80.m.above.gnd... +
##      X..Wind.Direction.daily.mean..900.mb... +
X..Wind.Gust.daily.mean..sfc...
##
##                                                       Df Deviance    AIC
## - X..Total.Precipitation.daily.sum..sfc...             1   1313.6 1343.6
## - X..High.Cloud.Cover.daily.mean..high.cld.lay...      1   1313.7 1343.7
## - X..Wind.Direction.daily.mean..10.m.above.gnd...      1   1313.7 1343.7
## - X..Low.Cloud.Cover.daily.mean..low.cld.lay...        1   1314.0 1344.0
## <none>                                                     1312.9 1344.9
## - X..Wind.Direction.daily.mean..80.m.above.gnd...      1   1315.4 1345.4
## - X..Snowfall.amount.raw.daily.sum..sfc...             1   1315.5 1345.5
## - X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...     1   1315.5 1345.5
## - X..Shortwave.Radiation.daily.sum..sfc...             1   1315.7 1345.7
## - X..Temperature.daily.mean..2.m.above.gnd...          1   1315.9 1345.9
## - X..Wind.Speed.daily.mean..10.m.above.gnd...          1   1316.3 1346.3
## - X..Wind.Gust.daily.mean..sfc...                      1   1317.3 1347.3
## - X..Total.Cloud.Cover.daily.mean..sfc...              1   1317.3 1347.3
## - X..Wind.Speed.daily.mean..80.m.above.gnd...          1   1317.7 1347.7
## - X..Wind.Direction.daily.mean..900.mb...              1   1330.2 1360.2
## - X..Mean.Sea.Level.Pressure.daily.mean..MSL...        1   1347.2 1377.2
##
## Step:  AIC=1343.59
## X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##      X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
X..Snowfall.amount.raw.daily.sum..sfc... +
##      X..Total.Cloud.Cover.daily.mean..sfc... +
```

```
X..High.Cloud.Cover.daily.mean..high.cld.lay... +
##     X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
X..Low.Cloud.Cover.daily.mean..low.cld.lay... +
##     X..Shortwave.Radiation.daily.sum..sfc... +
X..Wind.Speed.daily.mean..10.m.above.gnd... +
##     X..Wind.Direction.daily.mean..10.m.above.gnd... +
X..Wind.Speed.daily.mean..80.m.above.gnd... +
##     X..Wind.Direction.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..900.mb... +
##     X..Wind.Gust.daily.mean..sfc...
##
##                                                    Df Deviance    AIC
## - X..Low.Cloud.Cover.daily.mean..low.cld.lay...     1   1314.2 1342.2
## - X..High.Cloud.Cover.daily.mean..high.cld.lay...   1   1314.4 1342.4
## - X..Wind.Direction.daily.mean..10.m.above.gnd...   1   1314.6 1342.6
## <none>                                                  1313.6 1343.6
## - X..Snowfall.amount.raw.daily.sum..sfc...          1   1315.9 1343.9
## - X..Shortwave.Radiation.daily.sum..sfc...          1   1316.1 1344.1
## - X..Wind.Direction.daily.mean..80.m.above.gnd...   1   1316.3 1344.3
## - X..Wind.Speed.daily.mean..10.m.above.gnd...       1   1316.9 1344.9
## - X..Total.Cloud.Cover.daily.mean..sfc...           1   1317.4 1345.4
## - X..Temperature.daily.mean..2.m.above.gnd...       1   1317.5 1345.5
## - X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...  1   1317.6 1345.6
## - X..Wind.Gust.daily.mean..sfc...                   1   1318.1 1346.1
## - X..Wind.Speed.daily.mean..80.m.above.gnd...       1   1318.3 1346.3
## - X..Wind.Direction.daily.mean..900.mb...           1   1331.4 1359.4
## - X..Mean.Sea.Level.Pressure.daily.mean..MSL...     1   1348.5 1376.5
##
## Step:  AIC=1342.23
## X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##     X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
X..Snowfall.amount.raw.daily.sum..sfc... +
##     X..Total.Cloud.Cover.daily.mean..sfc... +
X..High.Cloud.Cover.daily.mean..high.cld.lay... +
##     X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
X..Shortwave.Radiation.daily.sum..sfc... +
##     X..Wind.Speed.daily.mean..10.m.above.gnd... +
X..Wind.Direction.daily.mean..10.m.above.gnd... +
##     X..Wind.Speed.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..80.m.above.gnd... +
##     X..Wind.Direction.daily.mean..900.mb... +
X..Wind.Gust.daily.mean..sfc...
##
##                                                    Df Deviance    AIC
## - X..Wind.Direction.daily.mean..10.m.above.gnd...   1   1315.1 1341.1
## - X..High.Cloud.Cover.daily.mean..high.cld.lay...   1   1315.9 1341.9
## <none>                                                  1314.2 1342.2
## - X..Snowfall.amount.raw.daily.sum..sfc...          1   1316.7 1342.7
## - X..Wind.Direction.daily.mean..80.m.above.gnd...   1   1316.9 1342.9
## - X..Wind.Speed.daily.mean..10.m.above.gnd...       1   1317.3 1343.3
```

```
## - X..Shortwave.Radiation.daily.sum..sfc...              1    1317.4 1343.4
## - X..Temperature.daily.mean..2.m.above.gnd...           1    1317.8 1343.8
## - X..Wind.Gust.daily.mean..sfc...                       1    1318.6 1344.6
## - X..Wind.Speed.daily.mean..80.m.above.gnd...           1    1318.7 1344.7
## - X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...      1    1320.0 1346.0
## - X..Total.Cloud.Cover.daily.mean..sfc...               1    1320.3 1346.3
## - X..Wind.Direction.daily.mean..900.mb...               1    1332.4 1358.4
## - X..Mean.Sea.Level.Pressure.daily.mean..MSL...         1    1348.8 1374.8
##
## Step:  AIC=1341.13
## X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##      X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
X..Snowfall.amount.raw.daily.sum..sfc... +
##      X..Total.Cloud.Cover.daily.mean..sfc... +
X..High.Cloud.Cover.daily.mean..high.cld.lay... +
##      X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
X..Shortwave.Radiation.daily.sum..sfc... +
##      X..Wind.Speed.daily.mean..10.m.above.gnd... +
X..Wind.Speed.daily.mean..80.m.above.gnd... +
##      X..Wind.Direction.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..900.mb... +
##      X..Wind.Gust.daily.mean..sfc...
##
##                                                        Df Deviance    AIC
## - X..High.Cloud.Cover.daily.mean..high.cld.lay...      1    1316.7 1340.7
## <none>                                                      1315.1 1341.1
## - X..Snowfall.amount.raw.daily.sum..sfc...             1    1317.8 1341.8
## - X..Wind.Speed.daily.mean..10.m.above.gnd...          1    1318.3 1342.3
## - X..Shortwave.Radiation.daily.sum..sfc...             1    1318.5 1342.5
## - X..Temperature.daily.mean..2.m.above.gnd...          1    1318.7 1342.7
## - X..Wind.Gust.daily.mean..sfc...                      1    1319.2 1343.2
## - X..Wind.Speed.daily.mean..80.m.above.gnd...          1    1319.4 1343.4
## - X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...     1    1320.6 1344.6
## - X..Total.Cloud.Cover.daily.mean..sfc...              1    1322.0 1346.0
## - X..Wind.Direction.daily.mean..80.m.above.gnd...      1    1322.1 1346.1
## - X..Wind.Direction.daily.mean..900.mb...              1    1332.6 1356.6
## - X..Mean.Sea.Level.Pressure.daily.mean..MSL...        1    1349.5 1373.5
##
## Step:  AIC=1340.71
## X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##      X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
X..Snowfall.amount.raw.daily.sum..sfc... +
##      X..Total.Cloud.Cover.daily.mean..sfc... +
X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
##      X..Shortwave.Radiation.daily.sum..sfc... +
X..Wind.Speed.daily.mean..10.m.above.gnd... +
##      X..Wind.Speed.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..80.m.above.gnd... +
##      X..Wind.Direction.daily.mean..900.mb... +
X..Wind.Gust.daily.mean..sfc...
```

```
## 
##                                                    Df Deviance    AIC
## <none>                                                 1316.7 1340.7
## - X..Wind.Speed.daily.mean..10.m.above.gnd...       1   1319.4 1341.4
## - X..Snowfall.amount.raw.daily.sum..sfc...          1   1319.6 1341.6
## - X..Shortwave.Radiation.daily.sum..sfc...          1   1319.7 1341.7
## - X..Wind.Speed.daily.mean..80.m.above.gnd...       1   1320.5 1342.5
## - X..Wind.Gust.daily.mean..sfc...                   1   1321.0 1343.0
## - X..Temperature.daily.mean..2.m.above.gnd...       1   1321.5 1343.5
## - X..Total.Cloud.Cover.daily.mean..sfc...           1   1323.2 1345.2
## - X..Wind.Direction.daily.mean..80.m.above.gnd...   1   1324.3 1346.3
## - X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...  1   1329.4 1351.4
## - X..Wind.Direction.daily.mean..900.mb...           1   1334.3 1356.3
## - X..Mean.Sea.Level.Pressure.daily.mean..MSL...     1   1350.7 1372.7
```

summary(g2)

```
## 
## Call:
## glm(formula = X..pluie.demain.. ~ 
X..Temperature.daily.mean..2.m.above.gnd... + 
##      X..Mean.Sea.Level.Pressure.daily.mean..MSL... + 
X..Snowfall.amount.raw.daily.sum..sfc... + 
##      X..Total.Cloud.Cover.daily.mean..sfc... + 
X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... + 
##      X..Shortwave.Radiation.daily.sum..sfc... + 
X..Wind.Speed.daily.mean..10.m.above.gnd... + 
##      X..Wind.Speed.daily.mean..80.m.above.gnd... + 
X..Wind.Direction.daily.mean..80.m.above.gnd... + 
##      X..Wind.Direction.daily.mean..900.mb... + 
X..Wind.Gust.daily.mean..sfc..., 
##      family = binomial, data = MeteoTrain)
## 
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.3325  -0.9154   0.3609   0.9024   2.6031
## 
## Coefficients:
##                                                  Estimate Std. Error z
value
## (Intercept)                                     6.103e+01  1.127e+01
5.414
## X..Temperature.daily.mean..2.m.above.gnd...     3.457e-02  1.586e-02
2.180
## X..Mean.Sea.Level.Pressure.daily.mean..MSL...  -6.237e-02  1.098e-02  -
5.681
## X..Snowfall.amount.raw.daily.sum..sfc...       -3.127e-01  1.891e-01  -
1.654
## X..Total.Cloud.Cover.daily.mean..sfc...         1.082e-02  4.242e-03
2.551
```

```
## X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...   1.452e-02  4.104e-03
3.539
## X..Shortwave.Radiation.daily.sum..sfc...            1.028e-04  5.999e-05
1.714
## X..Wind.Speed.daily.mean..10.m.above.gnd...         1.131e-01  6.942e-02
1.628
## X..Wind.Speed.daily.mean..80.m.above.gnd...        -1.011e-01  5.204e-02  -
1.944
## X..Wind.Direction.daily.mean..80.m.above.gnd...    -4.165e-03  1.527e-03  -
2.727
## X..Wind.Direction.daily.mean..900.mb...             5.150e-03  1.246e-03
4.134
## X..Wind.Gust.daily.mean..sfc...                     3.878e-02  1.895e-02
2.047
##                                                    Pr(>|z|)
## (Intercept)                                        6.16e-08 ***
## X..Temperature.daily.mean..2.m.above.gnd...        0.029253 *
## X..Mean.Sea.Level.Pressure.daily.mean..MSL...      1.34e-08 ***
## X..Snowfall.amount.raw.daily.sum..sfc...           0.098149 .
## X..Total.Cloud.Cover.daily.mean..sfc...            0.010753 *
## X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...   0.000401 ***
## X..Shortwave.Radiation.daily.sum..sfc...           0.086496 .
## X..Wind.Speed.daily.mean..10.m.above.gnd...        0.103423
## X..Wind.Speed.daily.mean..80.m.above.gnd...        0.051926 .
## X..Wind.Direction.daily.mean..80.m.above.gnd...    0.006390 **
## X..Wind.Direction.daily.mean..900.mb...            3.57e-05 ***
## X..Wind.Gust.daily.mean..sfc...                    0.040703 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1635.4  on 1179  degrees of freedom
## Residual deviance: 1316.7  on 1168  degrees of freedom
## AIC: 1340.7
##
## Number of Fisher Scoring iterations: 4
```

Choix du meilleur modèle

Puis on détermine le meilleur modèle pour expliquer les données en comparant via
l'ANOVA puis via les critères BIC et AIC. C'est toujours le modèle g2 qui l'emporte…

```
anova(g1,g2,test="LRT")

## Analysis of Deviance Table
##
## Model 1: X..pluie.demain.. ~ X..Mean.Sea.Level.Pressure.daily.mean..MSL...
+
##     X..Wind.Direction.daily.mean..900.mb...
```

```
## Model 2: X..pluie.demain.. ~ X..Temperature.daily.mean..2.m.above.gnd... +
##      X..Mean.Sea.Level.Pressure.daily.mean..MSL... +
X..Snowfall.amount.raw.daily.sum..sfc... +
##      X..Total.Cloud.Cover.daily.mean..sfc... +
X..Medium.Cloud.Cover.daily.mean..mid.cld.lay... +
##      X..Shortwave.Radiation.daily.sum..sfc... +
X..Wind.Speed.daily.mean..10.m.above.gnd... +
##      X..Wind.Speed.daily.mean..80.m.above.gnd... +
X..Wind.Direction.daily.mean..80.m.above.gnd... +
##      X..Wind.Direction.daily.mean..900.mb... +
X..Wind.Gust.daily.mean..sfc...
##   Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
## 1     1177    1410.4
## 2     1168    1316.7  9   93.692 2.949e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
c(AIC(g1),AIC(g2))
```

```
## [1] 1416.398 1340.706
```

```
c(BIC(g1),BIC(g2))
```

```
## [1] 1431.618 1401.585
```

Prédiction

On commence par télécharger la base de test.

```
MeteoTest <- read.csv("meteo.test.csv", header = T, quote="")
summary(MeteoTest)
```

```
##        X.            X..Year..        X..Month..        X..Day..
X..Hour..
##  Min.   :  36   Min.   :2010   Min.   : 1.000   Min.   : 1.00   Min.    :0
##  1st Qu.: 825   1st Qu.:2012   1st Qu.: 4.000   1st Qu.: 8.00   1st Qu.:0
##  Median :1554   Median :2014   Median : 7.000   Median :16.00   Median :0
##  Mean   :1517   Mean   :2014   Mean   : 6.859   Mean   :15.26   Mean    :0
##  3rd Qu.:2252   3rd Qu.:2016   3rd Qu.:10.000   3rd Qu.:22.00   3rd Qu.:0
##  Max.   :2936   Max.   :2018   Max.   :12.000   Max.   :31.00   Max.    :0
##   X..Minute.. X..Temperature.daily.mean..2.m.above.gnd...
##  Min.   :0    Min.   :-5.040
##  1st Qu.:0    1st Qu.: 5.625
##  Median :0    Median :13.085
##  Mean   :0    Mean   :12.062
##  3rd Qu.:0    3rd Qu.:18.170
##  Max.   :0    Max.   :26.600
##  X..Relative.Humidity.daily.mean..2.m.above.gnd...
##  Min.   :36.42
##  1st Qu.:64.35
##  Median :71.64
```

```
##   Mean    :71.08
##   3rd Qu.:78.06
##   Max.    :91.46
##   X..Mean.Sea.Level.Pressure.daily.mean..MSL...
##   Min.    : 983
##   1st Qu.:1012
##   Median :1017
##   Mean    :1017
##   3rd Qu.:1022
##   Max.    :1046
##   X..Total.Precipitation.daily.sum..sfc...
##   Min.    : 0.000
##   1st Qu.: 0.000
##   Median : 0.100
##   Mean    : 2.202
##   3rd Qu.: 2.475
##   Max.    :39.300
##   X..Snowfall.amount.raw.daily.sum..sfc...
##   Min.    :0.00000
##   1st Qu.:0.00000
##   Median :0.00000
##   Mean    :0.08255
##   3rd Qu.:0.00000
##   Max.    :6.65000
##   X..Total.Cloud.Cover.daily.mean..sfc...
##   Min.    :  0.00
##   1st Qu.: 30.20
##   Median : 52.30
##   Mean    : 53.62
##   3rd Qu.: 79.60
##   Max.    :100.00
##   X..High.Cloud.Cover.daily.mean..high.cld.lay...
##   Min.    : 0.000
##   1st Qu.: 1.603
##   Median :12.210
##   Mean    :20.451
##   3rd Qu.:34.358
##   Max.    :99.330
##   X..Medium.Cloud.Cover.daily.mean..mid.cld.lay...
##   Min.    :  0.00
##   1st Qu.:  3.23
##   Median : 26.09
##   Mean    : 32.66
##   3rd Qu.: 51.34
##   Max.    :100.00
##   X..Low.Cloud.Cover.daily.mean..low.cld.lay...
##   Min.    :  0.00
##   1st Qu.: 12.55
##   Median : 39.65
##   Mean    : 41.75
```

```
##   3rd Qu.: 66.50
##   Max.    :100.00
##   X..Sunshine.Duration.daily.sum..sfc...
##   Min.    :    0.0
##   1st Qu.: 116.7
##   Median : 347.6
##   Mean    : 353.0
##   3rd Qu.: 560.1
##   Max.    :1021.1
##   X..Shortwave.Radiation.daily.sum..sfc...
##   Min.    : 121.9
##   1st Qu.:1946.0
##   Median :3391.8
##   Mean    :3813.1
##   3rd Qu.:5509.5
##   Max.    :8259.2
##   X..Wind.Speed.daily.mean..10.m.above.gnd...
##   Min.    : 3.03
##   1st Qu.: 6.35
##   Median : 9.35
##   Mean    :11.23
##   3rd Qu.:13.47
##   Max.    :49.14
##   X..Wind.Direction.daily.mean..10.m.above.gnd...
##   Min.    : 45.08
##   1st Qu.:151.67
##   Median :206.38
##   Mean    :201.27
##   3rd Qu.:250.45
##   Max.    :326.87
##   X..Wind.Speed.daily.mean..80.m.above.gnd...
##   Min.    : 2.990
##   1st Qu.: 8.082
##   Median :12.565
##   Mean    :14.725
##   3rd Qu.:18.340
##   Max.    :63.100
##   X..Wind.Direction.daily.mean..80.m.above.gnd...
##   Min.    : 37.18
##   1st Qu.:159.47
##   Median :217.01
##   Mean    :207.55
##   3rd Qu.:256.15
##   Max.    :330.47
##   X..Wind.Speed.daily.mean..900.mb...
X..Wind.Direction.daily.mean..900.mb...
##   Min.    :  2.16                    Min.    : 32.09
##   1st Qu.: 12.58                     1st Qu.:165.43
##   Median : 21.76                     Median :239.06
##   Mean    : 25.57                    Mean    :211.65
```

```
##   3rd Qu.: 34.66                        3rd Qu.:265.90
##   Max.   :104.89                        Max.   :333.05
##   X..Wind.Gust.daily.mean..sfc...
X..Temperature.daily.max..2.m.above.gnd...
##   Min.   : 3.290                        Min.   :-2.130
##   1st Qu.: 9.293                        1st Qu.: 9.295
##   Median :14.445                        Median :17.145
##   Mean   :17.591                        Mean   :16.231
##   3rd Qu.:22.065                        3rd Qu.:22.823
##   Max.   :90.750                        Max.   :33.120
##   X..Temperature.daily.min..2.m.above.gnd...
##   Min.   :-8.870
##   1st Qu.: 2.458
##   Median : 8.595
##   Mean   : 8.022
##   3rd Qu.:13.367
##   Max.   :21.890
##   X..Relative.Humidity.daily.max..2.m.above.gnd...
##   Min.   : 63.00
##   1st Qu.: 83.00
##   Median : 88.00
##   Mean   : 87.47
##   3rd Qu.: 94.00
##   Max.   :100.00
##   X..Relative.Humidity.daily.min..2.m.above.gnd...
##   Min.   :17.00
##   1st Qu.:45.00
##   Median :54.00
##   Mean   :53.73
##   3rd Qu.:62.00
##   Max.   :88.00
##   X..Mean.Sea.Level.Pressure.daily.max..MSL...
##   Min.   : 987.3
##   1st Qu.:1015.1
##   Median :1019.3
##   Mean   :1019.7
##   3rd Qu.:1024.5
##   Max.   :1049.0
##   X..Mean.Sea.Level.Pressure.daily.min..MSL...
##   Min.   : 980.6
##   1st Qu.:1009.1
##   Median :1014.9
##   Mean   :1014.1
##   3rd Qu.:1019.7
##   Max.   :1040.5
##   X..Total.Cloud.Cover.daily.max..sfc...
X..Total.Cloud.Cover.daily.min..sfc...
##   Min.   :  0.00                        Min.   : 0.000
##   1st Qu.:100.00                        1st Qu.: 0.000
##   Median :100.00                        Median : 0.000
```

```
##   Mean   : 91.36                          Mean    : 10.103
##   3rd Qu.:100.00                          3rd Qu.:  2.925
##   Max.   :100.00                          Max.    :100.000
##   X..High.Cloud.Cover.daily.max..high.cld.lay...
##   Min.   :  0.00
##   1st Qu.: 16.25
##   Median : 71.00
##   Mean   : 58.67
##   3rd Qu.:100.00
##   Max.   :100.00
##   X..High.Cloud.Cover.daily.min..high.cld.lay...
##   Min.   : 0.000
##   1st Qu.: 0.000
##   Median : 0.000
##   Mean   : 1.169
##   3rd Qu.: 0.000
##   Max.   :84.000
##   X..Medium.Cloud.Cover.daily.max..mid.cld.lay...
##   Min.   :  0.00
##   1st Qu.: 41.25
##   Median :100.00
##   Mean   : 74.00
##   3rd Qu.:100.00
##   Max.   :100.00
##   X..Medium.Cloud.Cover.daily.min..mid.cld.lay...
##   Min.   :  0.000
##   1st Qu.:  0.000
##   Median :  0.000
##   Mean   :  2.779
##   3rd Qu.:  0.000
##   Max.   :100.000
##   X..Low.Cloud.Cover.daily.max..low.cld.lay...
##   Min.   :  0.00
##   1st Qu.:100.00
##   Median :100.00
##   Mean   : 84.48
##   3rd Qu.:100.00
##   Max.   :100.00
##   X..Low.Cloud.Cover.daily.min..low.cld.lay...
##   Min.   :  0.000
##   1st Qu.:  0.000
##   Median :  0.000
##   Mean   :  4.745
##   3rd Qu.:  0.000
##   Max.   :100.000
##   X..Wind.Speed.daily.max..10.m.above.gnd...
##   Min.   : 4.35
##   1st Qu.:12.44
##   Median :17.31
##   Mean   :19.67
```

```
##   3rd Qu.:23.90
##   Max.   :74.34
##   X..Wind.Speed.daily.min..10.m.above.gnd...
##   Min.   : 0.000
##   1st Qu.: 1.080
##   Median : 2.465
##   Mean   : 3.906
##   3rd Qu.: 5.008
##   Max.   :33.840
##   X..Wind.Speed.daily.max..80.m.above.gnd...
##   Min.   : 5.82
##   1st Qu.:17.93
##   Median :24.14
##   Mean   :26.08
##   3rd Qu.:30.95
##   Max.   :94.46
##   X..Wind.Speed.daily.min..80.m.above.gnd...
X..Wind.Speed.daily.max..900.mb...
##   Min.   : 0.000                              Min.   :  6.92
##   1st Qu.: 1.095                              1st Qu.: 25.36
##   Median : 2.520                              Median : 39.55
##   Mean   : 5.068                              Mean   : 43.20
##   3rd Qu.: 6.145                              3rd Qu.: 55.47
##   Max.   :43.210                              Max.   :148.95
##   X..Wind.Speed.daily.min..900.mb... X..Wind.Gust.daily.max..sfc...
##   Min.   : 0.00                      Min.   :  5.76
##   1st Qu.: 3.34                      1st Qu.: 18.00
##   Median : 6.92                      Median : 26.64
##   Mean   :11.66                      Mean   : 30.30
##   3rd Qu.:15.58                      3rd Qu.: 38.43
##   Max.   :81.12                      Max.   :133.20
##   X..Wind.Gust.daily.min..sfc... X..pluie.demain.. X..temp.demain...
##   Min.   : 0.000                  Mode :logical     Min.   :-5.250
##   1st Qu.: 1.800                  FALSE:133         1st Qu.: 6.303
##   Median : 3.960                  TRUE :157         Median :12.470
##   Mean   : 6.825                                    Mean   :12.026
##   3rd Qu.: 8.190                                    3rd Qu.:18.000
##   Max.   :54.720                                    Max.   :27.900
```

Table sur les 2 modèles

Nous allons faire maintenant une prédiction avec les 2 modèles, puis on affiche les 2 tables.

```
pred1 = predict(g1, MeteoTest, type="response")
pre1 = pred1 > 0.5
table(pre1, MeteoTest$X..pluie.demain..)

##
## pre1    FALSE TRUE
```

```
##    FALSE    85    44
##    TRUE     48   113

pred2 = predict(g2, MeteoTest, type="response")
pre2 = pred2 > 0.5
table(pre2, MeteoTest$X..pluie.demain..)

##
## pre2     FALSE TRUE
##    FALSE    87   48
##    TRUE     46  109
```

Performance des modèles

sur le modèle avec 2 variables, on a 44 faux négatifs et 48 faux positifs soit un taux d'erreur de 31,72%. Sur le modèle choisi automatiquement avec STEP, on a 48 faux négatifs et 46 faux positifs soit un taux d'erreur de 32,41%.
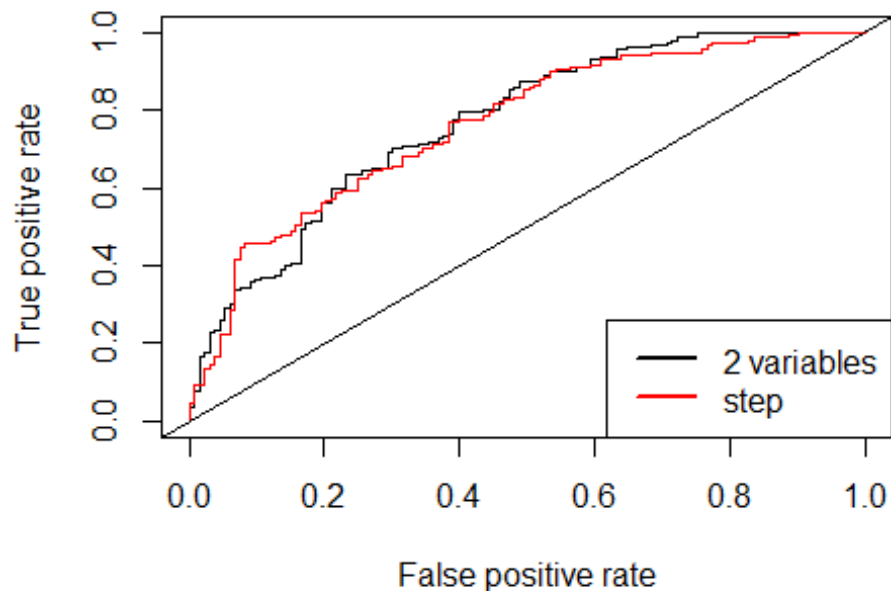
On trace la courbe ROC des différents modèles. Ils sont similaires mais on choisit tout de même le modèle avec 2 variables car l'erreur de prédiction est plus faible.

```
library(ROCR)

## Warning: package 'ROCR' was built under R version 3.6.3

p1 = prediction(pred1, MeteoTest$X..pluie.demain..)
p2 = prediction(pred2, MeteoTest$X..pluie.demain..)
perf1 = performance(p1, "tpr", "fpr")
perf2 = performance(p2, "tpr", "fpr")
plot(perf1)
plot(perf2,add=TRUE,col="red")
legend("bottomright",legend=c("2 variables","step"),
col=c("black","red"),lty=1,lwd=2)
title("comparaison modéle")
abline(0,1)
```

comparaison modéle

construction du fichier de prediction

On ajoute la colonne de prédiction au fichier MeteoTest.

```
class (MeteoTest)

## [1] "data.frame"

MeteoTest$prev<-pre1
```

Export du fichier de prediction

On "tronque" le fichier pour ne conserver que les jours. On exporte le fichier.

```
Prevision <- subset(MeteoTest, select = c(1,2,3,4,47,49) )
head(Prevision)

##   X. X..Year.. X..Month.. X..Day.. X..pluie.demain..  prev
## 1 36      2010          7        6             FALSE FALSE
## 2 44      2010          7       14             FALSE  TRUE
## 3 54      2010          7       24             FALSE  TRUE
## 4 58      2010          7       28              TRUE  TRUE
## 5 74      2010          8       13              TRUE FALSE
## 6 86      2010          8       25              TRUE  TRUE

write.csv2(Prevision, file = "PrevisionPluie.csv")
```