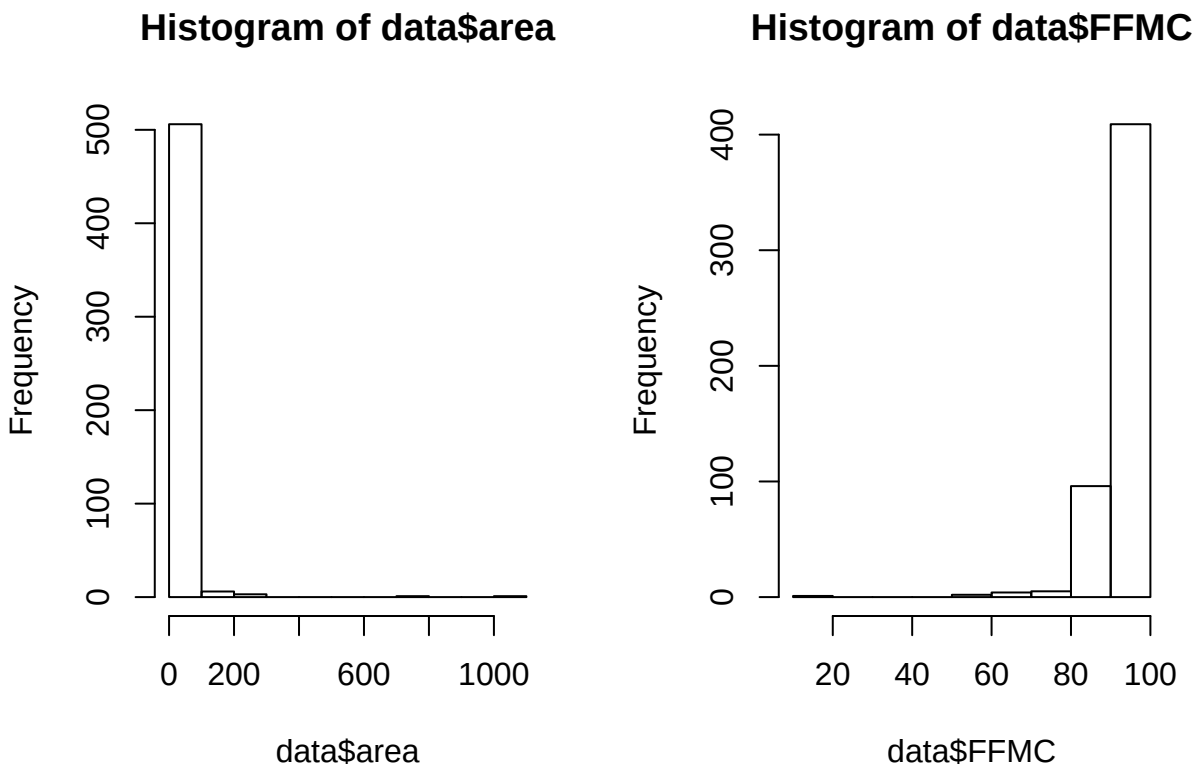
Forest Fire Project Part 1

Shi Fan Jin, Esther Law, Changning Liu

November 23, 2018

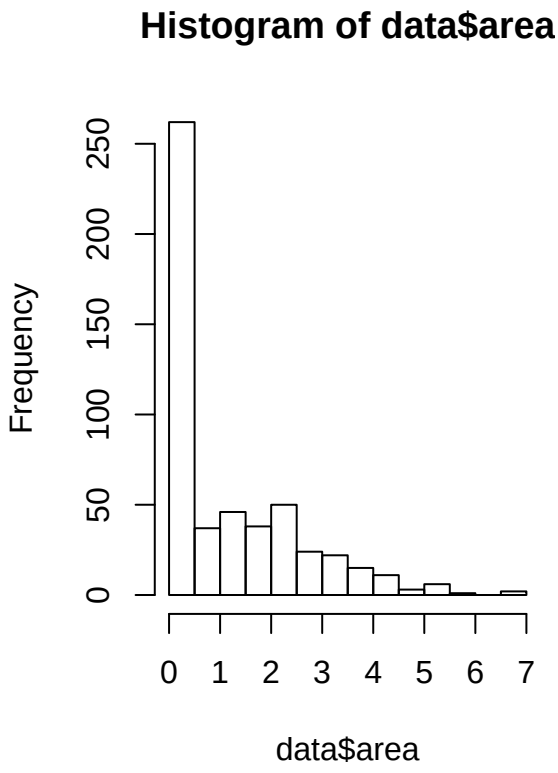
```
#clean dataset
par(mfrow=c(1,2))
hist(data$area)
hist(data$FFMC)
```




```
data$area=log(data$area+1)
hist(data$area)
data$FFMC=data$FFMC^13
data=data[-which.max(data$ISI),]
data=data[,-12]
summary(data)
```

```
##        X               Y          month              day
##  Min.   :1.000   Min.   :2.0   Length:516         Length:516
##  1st Qu.:3.000   1st Qu.:4.0   Class :character   Class :character
##  Median :4.000   Median :4.0   Mode  :character   Mode  :character
##  Mean   :4.665   Mean   :4.3
##  3rd Qu.:7.000   3rd Qu.:5.0
##  Max.   :9.000   Max.   :9.0
##       FFMC                DMC               DC             ISI
##  Min.   :3.419e+16   Min.   :  1.10   Min.   :  7.9   Min.   : 0.000
##  1st Qu.:2.616e+25   1st Qu.: 67.03   1st Qu.:440.1   1st Qu.: 6.475
##  Median :3.196e+25   Median :108.30   Median :664.2   Median : 8.400
```

```
##  Mean   :3.178e+25   Mean   :110.90   Mean   :548.6   Mean   : 8.930
##  3rd Qu.:3.839e+25   3rd Qu.:142.40   3rd Qu.:713.9   3rd Qu.:10.725
##  Max.   :6.043e+25   Max.   :291.30   Max.   :860.6   Max.   :22.700
##       temp              RH              wind            area
##  Min.   : 2.20   Min.   : 15.00   Min.   :0.400   Min.   :0.0000
##  1st Qu.:15.50   1st Qu.: 32.75   1st Qu.:2.700   1st Qu.:0.0000
##  Median :19.30   Median : 41.50   Median :4.000   Median :0.4252
##  Mean   :18.89   Mean   : 44.29   Mean   :4.017   Mean   :1.1132
##  3rd Qu.:22.80   3rd Qu.: 53.00   3rd Qu.:4.900   3rd Qu.:2.0245
##  Max.   :33.30   Max.   :100.00   Max.   :9.400   Max.   :6.9956
```

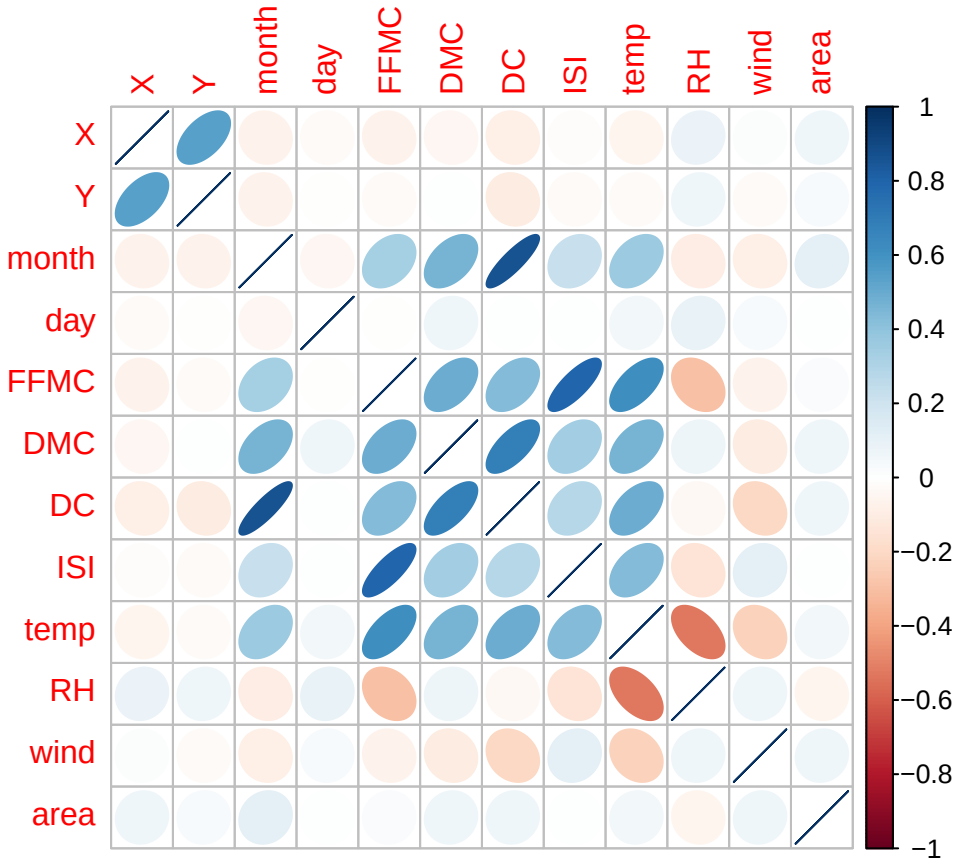



```r
#give numbers to categorical variables
num.data=data
num.data$month=gsub('jan', 1, num.data$month)
num.data$month=gsub('feb', 2, num.data$month)
num.data$month=gsub('mar', 3, num.data$month)
num.data$month=gsub('apr', 4, num.data$month)
num.data$month=gsub('may', 5, num.data$month)
num.data$month=gsub('jun', 6, num.data$month)
num.data$month=gsub('jul', 7, num.data$month)
num.data$month=gsub('aug', 8, num.data$month)
num.data$month=gsub('sep', 9, num.data$month)
num.data$month=gsub('oct', 10, num.data$month)
num.data$month=gsub('nov', 11, num.data$month)
num.data$month=gsub('dec', 12, num.data$month)
num.data$day=gsub('mon', 1, num.data$day)
num.data$day=gsub('tue', 2, num.data$day)
num.data$day=gsub('wed', 3, num.data$day)
num.data$day=gsub('thu', 4, num.data$day)
num.data$day=gsub('fri', 5, num.data$day)
```

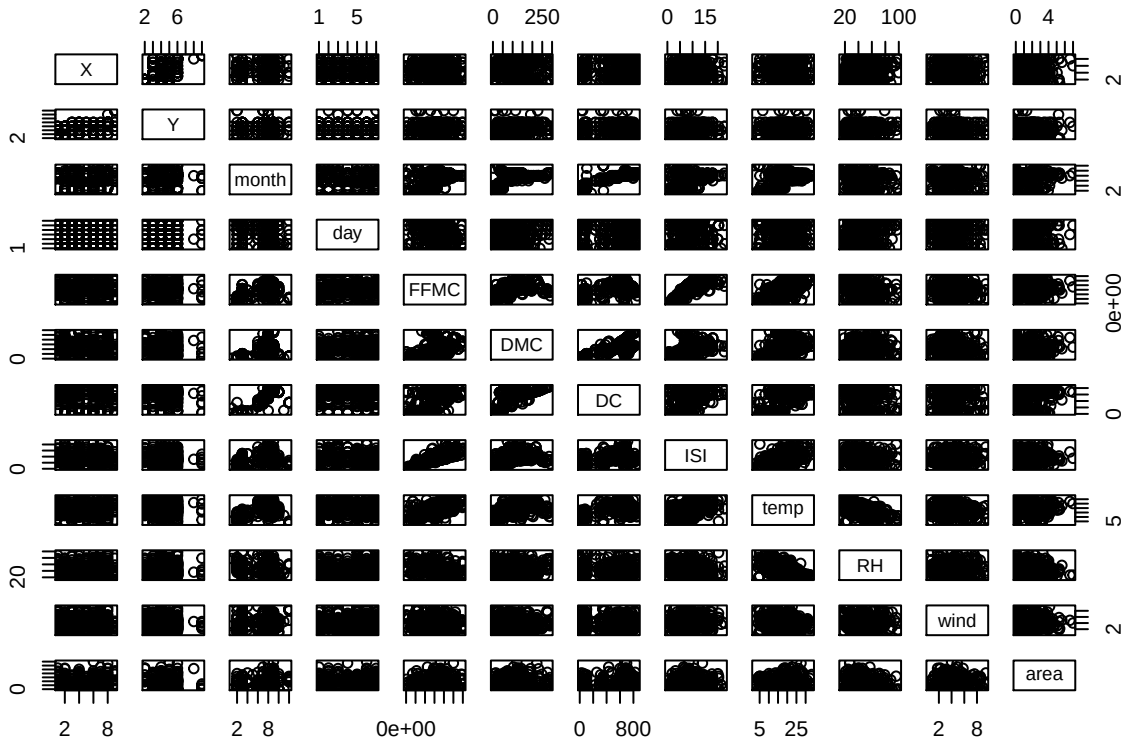
```
num.data$day=gsub('sat', 6, num.data$day)
num.data$day=gsub('sun', 7, num.data$day)

num.data$day=as.numeric(num.data$day)
num.data$month=as.numeric(num.data$month)

num.cor=cor(num.data)
corrplot(num.cor, method = 'ellipse')
```




```
pairs(num.data)
```




```
#make binary variables for categorical data
jan=rep(0,516)
feb=rep(0,516)
mar=rep(0,516)
apr=rep(0,516)
may=rep(0,516)
jun=rep(0,516)
jul=rep(0,516)
aug=rep(0,516)
sep=rep(0,516)
oct=rep(0,516)
nov=rep(0,516)
dec=rep(0,516)
for(i in 1:516){
  if(data$month[i]=='jan'){jan[i]=1}
  else if(data$month[i]=='feb'){feb[i]=1}
  else if(data$month[i]=='mar'){mar[i]=1}
  else if(data$month[i]=='apr'){apr[i]=1}
  else if(data$month[i]=='may'){may[i]=1}
  else if(data$month[i]=='jun'){jun[i]=1}
  else if(data$month[i]=='jul'){jul[i]=1}
  else if(data$month[i]=='aug'){aug[i]=1}
  else if(data$month[i]=='sep'){sep[i]=1}
  else if(data$month[i]=='oct'){oct[i]=1}
  else if(data$month[i]=='nov'){nov[i]=1}
  else if(data$month[i]=='dec'){dec[i]=1}
}

mon=rep(0,516)
tue=rep(0,516)
```

```
wed=rep(0,516)
thu=rep(0,516)
fri=rep(0,516)
sat=rep(0,516)
sun=rep(0,516)

for(i in 1:516){
  if(data$day[i]=='mon'){mon[i]=1}
  else if(data$day[i]=='tue'){tue[i]=1}
  else if(data$day[i]=='wed'){wed[i]=1}
  else if(data$day[i]=='thu'){thu[i]=1}
  else if(data$day[i]=='fri'){fri[i]=1}
  else if(data$day[i]=='sat'){sat[i]=1}
  else if(data$day[i]=='sun'){sun[i]=1}
}

data=data[,-(3:4)]
data=data.frame(data[,1:10], jan,feb,mar,apr,may,jun,jul,aug,sep,oct,nov,dec,mon,tue,wed,thu,fri,sat,su
```

```
train=sample(1:516,344)

#try model selection methods
min.model=glm(area~1, data=data[train,])
biggest=formula(glm(area~., data=data[train,]))
fwd.model=step(min.model, direction = 'forward', scope=biggest)
```

```
## Start:  AIC=1194.97
## area ~ 1
##
##        Df Deviance    AIC
## + dec   1   629.17 1189.9
## + DC    1   636.23 1193.8
## + DMC   1   637.21 1194.3
## + X     1   637.88 1194.7
## + sep   1   637.97 1194.7
## <none>      642.20 1195.0
## + jun   1   638.55 1195.0
## + mar   1   638.95 1195.2
## + jan   1   639.89 1195.7
## + Y     1   640.06 1195.8
## + temp  1   640.43 1196.0
## + wind  1   640.67 1196.2
## + may   1   641.01 1196.3
## + FFMC  1   641.31 1196.5
## + apr   1   641.38 1196.5
## + oct   1   641.62 1196.7
## + thu   1   641.66 1196.7
## + jul   1   641.71 1196.7
## + RH    1   641.85 1196.8
## + wed   1   641.86 1196.8
## + tue   1   641.90 1196.8
## + feb   1   641.91 1196.8
## + sat   1   641.98 1196.8
## + aug   1   642.01 1196.9
```

```
## + sun   1   642.13 1196.9
## + ISI   1   642.14 1196.9
## + fri   1   642.19 1197.0
## + mon   1   642.20 1197.0
##
## Step:  AIC=1189.92
## area ~ dec
##
##        Df Deviance    AIC
## + DMC   1   620.98 1187.4
## + DC    1   621.27 1187.6
## + temp  1   622.75 1188.4
## + sep   1   623.55 1188.8
## + X     1   625.27 1189.8
## <none>      629.17 1189.9
## + jun   1   625.83 1190.1
## + FFMC  1   626.11 1190.2
## + mar   1   626.45 1190.4
## + jan   1   626.97 1190.7
## + Y     1   627.70 1191.1
## + may   1   627.91 1191.2
## + thu   1   628.35 1191.5
## + apr   1   628.43 1191.5
## + oct   1   628.47 1191.5
## + wed   1   628.65 1191.6
## + jul   1   628.82 1191.7
## + feb   1   628.97 1191.8
## + tue   1   628.98 1191.8
## + RH    1   628.99 1191.8
## + ISI   1   629.07 1191.9
## + mon   1   629.07 1191.9
## + sat   1   629.09 1191.9
## + sun   1   629.10 1191.9
## + fri   1   629.10 1191.9
## + wind  1   629.15 1191.9
## + aug   1   629.16 1191.9
##
## Step:  AIC=1187.42
## area ~ dec + DMC
##
##        Df Deviance    AIC
## + X     1   616.32 1186.8
## + sep   1   616.94 1187.2
## <none>      620.98 1187.4
## + jun   1   618.09 1187.8
## + aug   1   618.41 1188.0
## + temp  1   618.69 1188.1
## + may   1   618.93 1188.3
## + oct   1   619.07 1188.3
## + Y     1   619.20 1188.4
## + DC    1   619.60 1188.7
## + jan   1   619.75 1188.7
## + wed   1   620.24 1189.0
## + thu   1   620.43 1189.1
```

```
## + RH    1   620.57 1189.2
## + ISI   1   620.64 1189.2
## + tue   1   620.70 1189.3
## + FFMC  1   620.71 1189.3
## + mar   1   620.80 1189.3
## + feb   1   620.80 1189.3
## + jul   1   620.83 1189.3
## + sun   1   620.84 1189.3
## + apr   1   620.84 1189.3
## + wind  1   620.92 1189.4
## + fri   1   620.94 1189.4
## + sat   1   620.95 1189.4
## + mon   1   620.98 1189.4
##
## Step:  AIC=1186.83
## area ~ dec + DMC + X
##
##        Df Deviance    AIC
## + sep   1   612.00 1186.4
## + jun   1   612.47 1186.7
## <none>      616.32 1186.8
## + temp  1   613.65 1187.3
## + aug   1   614.08 1187.6
## + may   1   614.31 1187.7
## + oct   1   614.67 1187.9
## + DC    1   614.75 1188.0
## + wed   1   615.28 1188.2
## + jan   1   615.37 1188.3
## + RH    1   615.60 1188.4
## + FFMC  1   615.78 1188.5
## + thu   1   615.86 1188.6
## + tue   1   615.92 1188.6
## + ISI   1   616.08 1188.7
## + jul   1   616.11 1188.7
## + apr   1   616.14 1188.7
## + fri   1   616.20 1188.8
## + mar   1   616.24 1188.8
## + wind  1   616.25 1188.8
## + sun   1   616.25 1188.8
## + mon   1   616.26 1188.8
## + feb   1   616.28 1188.8
## + sat   1   616.29 1188.8
## + Y     1   616.29 1188.8
##
## Step:  AIC=1186.4
## area ~ dec + DMC + X + sep
##
##        Df Deviance    AIC
## <none>      612.00 1186.4
## + jun   1   609.07 1186.8
## + temp  1   609.35 1186.9
## + may   1   609.73 1187.1
## + oct   1   609.79 1187.2
## + jan   1   611.19 1188.0
```

```
## + wed   1   611.23 1188.0
## + tue   1   611.40 1188.1
## + RH    1   611.44 1188.1
## + FFMC  1   611.46 1188.1
## + thu   1   611.55 1188.2
## + wind  1   611.70 1188.2
## + Y     1   611.82 1188.3
## + feb   1   611.83 1188.3
## + mon   1   611.87 1188.3
## + apr   1   611.89 1188.3
## + ISI   1   611.91 1188.4
## + sun   1   611.96 1188.4
## + mar   1   611.97 1188.4
## + jul   1   611.99 1188.4
## + aug   1   611.99 1188.4
## + fri   1   611.99 1188.4
## + sat   1   612.00 1188.4
## + DC    1   612.00 1188.4
```

```
summary(fwd.model)
```

```
##
## Call:
## glm(formula = area ~ dec + DMC + X + sep, data = data[train,
##     ])
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.7275  -1.0296  -0.6187   0.7914   5.2285
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.463222   0.217028   2.134   0.0335 *
## dec         1.880864   0.614836   3.059   0.0024 **
## DMC         0.002302   0.001147   2.007   0.0456 *
## X           0.051585   0.031172   1.655   0.0989 .
## sep         0.241941   0.156250   1.548   0.1225
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.805297)
##
##     Null deviance: 642.2  on 343  degrees of freedom
## Residual deviance: 612.0  on 339  degrees of freedom
## AIC: 1186.4
##
## Number of Fisher Scoring iterations: 2
```

```
full.model=glm(area~.,data=data[train,])
back.model=step(full.model, direction='backward', trace=T)
```

```
## Start:  AIC=1216.56
## area ~ X + Y + FFMC + DMC + DC + ISI + temp + RH + wind + jan +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + nov +
##     dec + mon + tue + wed + thu + fri + sat + sun
```

```
##
##
## Step:  AIC=1216.56
## area ~ X + Y + FFMC + DMC + DC + ISI + temp + RH + wind + jan +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + nov +
##     dec + mon + tue + wed + thu + fri + sat
##
##
## Step:  AIC=1216.56
## area ~ X + Y + FFMC + DMC + DC + ISI + temp + RH + wind + jan +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + nov +
##     mon + tue + wed + thu + fri + sat
##
##
## Step:  AIC=1216.56
## area ~ X + Y + FFMC + DMC + DC + ISI + temp + RH + wind + jan +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + mon +
##     tue + wed + thu + fri + sat
##
##        Df Deviance    AIC
## - sat   1   591.29 1214.6
## - mon   1   591.30 1214.6
## - fri   1   591.36 1214.6
## - Y     1   591.41 1214.6
## - FFMC  1   591.47 1214.7
## - wed   1   591.51 1214.7
## - wind  1   591.85 1214.9
## - tue   1   591.99 1215.0
## - thu   1   592.02 1215.0
## - ISI   1   592.04 1215.0
## - RH    1   593.73 1216.0
## - DC    1   593.76 1216.0
## - oct   1   593.95 1216.1
## - X     1   594.41 1216.4
## <none>      591.29 1216.6
## - may   1   595.39 1216.9
## - DMC   1   596.60 1217.6
## - sep   1   596.72 1217.7
## - temp  1   597.28 1218.0
## - jan   1   602.41 1221.0
## - aug   1   602.48 1221.0
## - apr   1   604.46 1222.1
## - feb   1   605.28 1222.6
## - jul   1   608.04 1224.2
## - mar   1   609.67 1225.1
## - jun   1   614.50 1227.8
##
## Step:  AIC=1214.56
## area ~ X + Y + FFMC + DMC + DC + ISI + temp + RH + wind + jan +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + mon +
##     tue + wed + thu + fri
##
##        Df Deviance    AIC
## - mon   1   591.31 1212.6
```

```
## - fri   1   591.37 1212.6
## - Y     1   591.41 1212.6
## - FFMC  1   591.47 1212.7
## - wed   1   591.56 1212.7
## - wind  1   591.85 1212.9
## - ISI   1   592.04 1213.0
## - tue   1   592.13 1213.0
## - thu   1   592.18 1213.1
## - RH    1   593.73 1214.0
## - DC    1   593.76 1214.0
## - oct   1   593.95 1214.1
## - X     1   594.43 1214.4
## <none>      591.29 1214.6
## - may   1   595.41 1215.0
## - DMC   1   596.60 1215.6
## - sep   1   596.72 1215.7
## - temp  1   597.28 1216.0
## - jan   1   602.44 1219.0
## - aug   1   602.48 1219.0
## - apr   1   604.47 1220.1
## - feb   1   605.30 1220.6
## - jul   1   608.08 1222.2
## - mar   1   609.75 1223.1
## - jun   1   614.50 1225.8
##
## Step:  AIC=1212.57
## area ~ X + Y + FFMC + DMC + DC + ISI + temp + RH + wind + jan +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + tue +
##     wed + thu + fri
##
##        Df Deviance    AIC
## - fri   1   591.38 1210.6
## - Y     1   591.43 1210.6
## - FFMC  1   591.50 1210.7
## - wed   1   591.63 1210.8
## - wind  1   591.85 1210.9
## - ISI   1   592.07 1211.0
## - tue   1   592.15 1211.1
## - thu   1   592.20 1211.1
## - RH    1   593.73 1212.0
## - DC    1   593.77 1212.0
## - oct   1   594.02 1212.2
## - X     1   594.51 1212.4
## <none>      591.31 1212.6
## - may   1   595.45 1213.0
## - DMC   1   596.60 1213.6
## - sep   1   596.81 1213.8
## - temp  1   597.30 1214.0
## - jan   1   602.58 1217.1
## - aug   1   602.63 1217.1
## - apr   1   604.55 1218.2
## - feb   1   605.31 1218.6
## - jul   1   608.18 1220.2
## - mar   1   609.77 1221.2
```

```
## - jun   1   614.57 1223.8
##
## Step:  AIC=1210.61
## area ~ X + Y + FFMC + DMC + DC + ISI + temp + RH + wind + jan +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + tue +
##     wed + thu
##
##        Df Deviance    AIC
## - Y     1   591.50 1208.7
## - FFMC  1   591.57 1208.7
## - wed   1   591.78 1208.8
## - wind  1   591.95 1208.9
## - ISI   1   592.10 1209.0
## - tue   1   592.15 1209.1
## - thu   1   592.20 1209.1
## - RH    1   593.78 1210.0
## - DC    1   593.81 1210.0
## - oct   1   594.10 1210.2
## - X     1   594.54 1210.5
## <none>      591.38 1210.6
## - may   1   595.45 1211.0
## - DMC   1   596.69 1211.7
## - sep   1   596.84 1211.8
## - temp  1   597.33 1212.1
## - jan   1   602.62 1215.1
## - aug   1   602.68 1215.1
## - apr   1   604.55 1216.2
## - feb   1   605.32 1216.6
## - jul   1   608.18 1218.2
## - mar   1   609.77 1219.2
## - jun   1   614.57 1221.8
##
## Step:  AIC=1208.68
## area ~ X + FFMC + DMC + DC + ISI + temp + RH + wind + jan + feb +
##     mar + apr + may + jun + jul + aug + sep + oct + tue + wed +
##     thu
##
##        Df Deviance    AIC
## - FFMC  1   591.69 1206.8
## - wed   1   591.87 1206.9
## - wind  1   592.03 1207.0
## - ISI   1   592.21 1207.1
## - tue   1   592.26 1207.1
## - thu   1   592.32 1207.2
## - RH    1   593.98 1208.1
## - DC    1   594.07 1208.2
## - oct   1   594.30 1208.3
## <none>      591.50 1208.7
## - may   1   595.81 1209.2
## - DMC   1   597.02 1209.9
## - X     1   597.05 1209.9
## - sep   1   597.09 1209.9
## - temp  1   597.63 1210.2
## - jan   1   602.91 1213.3
```

```
## - aug   1   603.08 1213.3
## - apr   1   605.15 1214.5
## - feb   1   605.87 1214.9
## - jul   1   608.79 1216.6
## - mar   1   610.51 1217.6
## - jun   1   615.45 1220.3
##
## Step:  AIC=1206.8
## area ~ X + DMC + DC + ISI + temp + RH + wind + jan + feb + mar +
##     apr + may + jun + jul + aug + sep + oct + tue + wed + thu
##
##        Df Deviance    AIC
## - wed   1   592.07 1205.0
## - wind  1   592.16 1205.1
## - ISI   1   592.26 1205.1
## - tue   1   592.45 1205.2
## - thu   1   592.64 1205.3
## - RH    1   594.12 1206.2
## - DC    1   594.31 1206.3
## - oct   1   594.41 1206.4
## <none>      591.69 1206.8
## - may   1   596.04 1207.3
## - X     1   597.13 1207.9
## - sep   1   597.24 1208.0
## - DMC   1   597.80 1208.3
## - temp  1   598.77 1208.9
## - aug   1   603.28 1211.5
## - jan   1   603.32 1211.5
## - apr   1   605.80 1212.9
## - feb   1   606.45 1213.3
## - jul   1   609.14 1214.8
## - mar   1   610.70 1215.7
## - jun   1   615.84 1218.5
##
## Step:  AIC=1205.02
## area ~ X + DMC + DC + ISI + temp + RH + wind + jan + feb + mar +
##     apr + may + jun + jul + aug + sep + oct + tue + thu
##
##        Df Deviance    AIC
## - wind  1   592.58 1203.3
## - ISI   1   592.76 1203.4
## - tue   1   593.02 1203.6
## - thu   1   593.23 1203.7
## - RH    1   594.47 1204.4
## - oct   1   594.70 1204.5
## - DC    1   594.77 1204.6
## <none>      592.07 1205.0
## - may   1   596.32 1205.5
## - X     1   597.33 1206.1
## - sep   1   597.42 1206.1
## - DMC   1   598.33 1206.6
## - temp  1   599.08 1207.1
## - aug   1   603.45 1209.6
## - jan   1   603.56 1209.6
```

```
## - apr   1   606.01 1211.0
## - feb   1   606.64 1211.4
## - jul   1   609.26 1212.9
## - mar   1   610.89 1213.8
## - jun   1   616.02 1216.7
##
## Step:  AIC=1203.31
## area ~ X + DMC + DC + ISI + temp + RH + jan + feb + mar + apr +
##     may + jun + jul + aug + sep + oct + tue + thu
##
##        Df Deviance    AIC
## - ISI   1   593.01 1201.6
## - tue   1   593.59 1201.9
## - thu   1   593.71 1202.0
## - RH    1   595.12 1202.8
## - DC    1   595.46 1203.0
## - oct   1   595.72 1203.1
## <none>      592.58 1203.3
## - may   1   597.44 1204.1
## - X     1   597.79 1204.3
## - sep   1   598.88 1205.0
## - DMC   1   599.15 1205.1
## - temp  1   599.33 1205.2
## - aug   1   605.72 1208.9
## - jan   1   606.20 1209.1
## - apr   1   608.30 1210.3
## - feb   1   609.66 1211.1
## - jul   1   612.82 1212.9
## - mar   1   613.99 1213.5
## - jun   1   620.03 1216.9
##
## Step:  AIC=1201.56
## area ~ X + DMC + DC + temp + RH + jan + feb + mar + apr + may +
##     jun + jul + aug + sep + oct + tue + thu
##
##        Df Deviance    AIC
## - tue   1   593.97 1200.1
## - thu   1   594.19 1200.2
## - DC    1   595.64 1201.1
## - RH    1   595.69 1201.1
## <none>      593.01 1201.6
## - oct   1   596.54 1201.6
## - may   1   597.77 1202.3
## - X     1   598.20 1202.6
## - DMC   1   599.31 1203.2
## - temp  1   599.56 1203.3
## - sep   1   600.06 1203.6
## - jan   1   606.37 1207.2
## - aug   1   607.64 1208.0
## - apr   1   608.51 1208.4
## - feb   1   609.74 1209.1
## - jul   1   613.98 1211.5
## - mar   1   614.34 1211.7
## - jun   1   621.10 1215.5
```

```
##
## Step:  AIC=1200.12
## area ~ X + DMC + DC + temp + RH + jan + feb + mar + apr + may +
##     jun + jul + aug + sep + oct + thu
##
##        Df Deviance    AIC
## - thu   1   594.89 1198.7
## - RH    1   596.49 1199.6
## - DC    1   596.55 1199.6
## <none>      593.97 1200.1
## - oct   1   597.53 1200.2
## - may   1   598.76 1200.9
## - X     1   598.99 1201.0
## - DMC   1   600.01 1201.6
## - temp  1   600.26 1201.7
## - sep   1   601.01 1202.2
## - jan   1   607.45 1205.8
## - aug   1   608.39 1206.4
## - apr   1   609.58 1207.0
## - feb   1   610.50 1207.6
## - jul   1   614.60 1209.9
## - mar   1   615.26 1210.2
## - jun   1   622.10 1214.0
##
## Step:  AIC=1198.65
## area ~ X + DMC + DC + temp + RH + jan + feb + mar + apr + may +
##     jun + jul + aug + sep + oct
##
##        Df Deviance    AIC
## - RH    1   597.04 1197.9
## - DC    1   597.41 1198.1
## <none>      594.89 1198.7
## - oct   1   598.40 1198.7
## - may   1   599.50 1199.3
## - X     1   600.10 1199.7
## - temp  1   600.83 1200.1
## - DMC   1   601.19 1200.3
## - sep   1   601.69 1200.6
## - jan   1   608.02 1204.2
## - aug   1   608.90 1204.7
## - apr   1   609.96 1205.3
## - feb   1   610.95 1205.8
## - jul   1   615.07 1208.1
## - mar   1   615.68 1208.5
## - jun   1   622.36 1212.2
##
## Step:  AIC=1197.89
## area ~ X + DMC + DC + temp + jan + feb + mar + apr + may + jun +
##     jul + aug + sep + oct
##
##        Df Deviance    AIC
## - DC    1   599.40 1197.2
## - oct   1   599.50 1197.3
## - may   1   600.08 1197.6
```

```
## <none>      597.04 1197.9
## - temp  1   601.31 1198.3
## - sep   1   602.15 1198.8
## - X     1   602.67 1199.1
## - DMC   1   605.16 1200.5
## - jan   1   608.07 1202.2
## - aug   1   608.90 1202.7
## - apr   1   610.27 1203.4
## - feb   1   610.97 1203.8
## - jul   1   615.60 1206.4
## - mar   1   615.88 1206.6
## - jun   1   622.95 1210.5
##
## Step:  AIC=1197.25
## area ~ X + DMC + temp + jan + feb + mar + apr + may + jun + jul +
##     aug + sep + oct
##
##        Df Deviance    AIC
## - may   1   601.02 1196.2
## <none>      599.40 1197.2
## - temp  1   603.69 1197.7
## - X     1   604.72 1198.3
## - DMC   1   605.27 1198.6
## - jan   1   608.24 1200.3
## - oct   1   608.45 1200.4
## - apr   1   610.35 1201.5
## - feb   1   611.51 1202.1
## - sep   1   616.19 1204.8
## - mar   1   617.58 1205.5
## - jul   1   618.25 1205.9
## - aug   1   620.20 1207.0
## - jun   1   623.22 1208.7
##
## Step:  AIC=1196.18
## area ~ X + DMC + temp + jan + feb + mar + apr + jun + jul + aug +
##     sep + oct
##
##        Df Deviance    AIC
## <none>      601.02 1196.2
## - temp  1   604.67 1196.3
## - X     1   606.33 1197.2
## - DMC   1   606.88 1197.5
## - jan   1   608.45 1198.4
## - oct   1   608.46 1198.4
## - apr   1   610.35 1199.5
## - feb   1   611.62 1200.2
## - sep   1   617.18 1203.3
## - mar   1   618.75 1204.2
## - jul   1   619.20 1204.4
## - aug   1   621.67 1205.8
## - jun   1   624.43 1207.3
```

```
summary(back.model)
```

```
##
```

```
## Call:
## glm(formula = area ~ X + DMC + temp + jan + feb + mar + apr +
##     jun + jul + aug + sep + oct, data = data[train, ])
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5710  -1.0110  -0.5401   0.8511   5.0487
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.899020   0.553211   3.433 0.000673 ***
## X            0.054504   0.031870   1.710 0.088166 .
## DMC          0.002904   0.001616   1.797 0.073318 .
## temp         0.023329   0.016461   1.417 0.157368
## jan         -2.191979   1.084102  -2.022 0.043988 *
## feb         -1.580064   0.654024  -2.416 0.016237 *
## mar         -1.757156   0.562279  -3.125 0.001935 **
## apr         -1.927354   0.850269  -2.267 0.024050 *
## jun         -2.450821   0.682641  -3.590 0.000380 ***
## jul         -2.018529   0.637968  -3.164 0.001700 **
## aug         -2.054782   0.609288  -3.372 0.000833 ***
## sep         -1.745756   0.585178  -2.983 0.003064 **
## oct         -1.417793   0.700605  -2.024 0.043807 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.815776)
##
##     Null deviance: 642.20  on 343  degrees of freedom
## Residual deviance: 601.02  on 331  degrees of freedom
## AIC: 1196.2
##
## Number of Fisher Scoring iterations: 2
```

```
#original MSE
test=data[-train,]
fwd.pred=predict(fwd.model, newdata=test)
fwd.mse=mean((fwd.pred-test$area)^2)
fwd.mse
```

```
## [1] 2.067513
```

```
back.pred=predict(back.model, newdata=test)
back.mse=mean((back.pred-test$area)^2)
back.mse
```

```
## [1] 2.134971
```

```
#k-fold cross validation
glm.fit.1=glm(area~dec+DMC+X+sep, data=data)
cv.err.1=cv.glm(data,glm.fit.1)
cv.err.1
```

```
## $call
## cv.glm(data = data, glmfit = glm.fit.1)
##
## $K
```

```
## [1] 516
##
## $delta
## [1] 1.904671 1.904640
##
## $seed
##   [1]         403         344  1654269195 -1877109783  -961256264
##   [6]  1403523942   124639233   261424787  1836448066  1034917620
##  [11]   -13630729   468718317  1694379396  1559298986  1935866133
##  [16] -1450855505  2105396150  1802260960  1077391651   539731521
##  [21]   122505520   230898510 -1940184647  1223031755 -1597886342
##  [26] -1854140036 -1783225921  1484611221  1365746860  -346485118
##  [31]  1206044253  1201793367   956757054   350214264 -1324711077
##  [36] -1071776071 -1831283960  1862871478  1826141937   268853539
##  [41]   607122290  1403468228 -1913257593  -346768547  -463801004
##  [46]  1625143162  1387501253  -280700897   422985638   732086576
##  [51]  -725197869  1128169969 -1024391392 -1032957506  -353082935
##  [56] -1139485317   445858634  1224049580  -796824913   247650405
##  [61]   542595100  -404965774  -423589651 -1999189721  1430762830
##  [66] -1439787608 -1821953621  1808288457 -1668269352 -1099348090
##  [71]    24992353  -140278797  1709299298   269218004  1210153431
##  [76]  -238785587  -383705244  -472446134   -15672651  -538858993
##  [81]  1771793174   -40319552 -1632234109 -1613456479  -109108336
##  [86]  2056208494   -22837735  -195336533   835590810    86083164
##  [91]   517854047  -397095627 -1681163060   985204258   613283965
##  [96]  1720828279   929066654  -293216936   348261563  1602155097
## [101]  1188944360  -526197738 -1056856175  -999184317 -1361965230
## [106] -1531694108  1775225383  1578975869  -305972364   499275610
## [111]  -297283163 -1400910401  -628019642  -995328816  -416901389
## [116]  -131985391   838014528   160971294   516863209 -1388986853
## [121] -1140425174   -14278324  1980074767  1968938565  -674453508
## [126]  1981972050 -1121442867  -907997817   494782574  1629589576
## [131]    57352779 -1057292375  2043531000 -1557378522 -2033982655
## [136] -1166473005 -1243272062  1507691188   519478839  -214579923
## [141]   -15421884  -257117334   917217237  1604258159 -1793932170
## [146]  1404112800   895408227 -1228178175 -1129918352  1726708750
## [151] -1514890119 -1503646709 -1385642694   530830396 -1346935937
## [156]  1552658645   -31128340 -1631936958  1895821725   665429015
## [161]  1197209982    25108280  1382701083 -1944495751  1482522568
## [166]  1344742390  2075751985   887378019  2132045874   507710852
## [171] -1239616569 -1239992675  1300848148 -2078006086  1822724357
## [176]  1779634911  -346210458    68638448  -651383277   408315825
## [181] -2048675232  1331662846 -1134018679 -1715119941 -1026157430
## [186]  -262084756 -1837320977  -648290523  -685012516  -699648078
## [191] -1959870803  -939002265  1368704398  1604498024  -270455189
## [196]   878774281 -1496090984  1651241414  -719838943  -170397261
## [201]  1455721762   967288596  1974706071 -1131741299  1944879780
## [206]   693669002   967713525   849937743  1265996374   507097344
## [211]  -355466173  2106394593   467134288  1892604334   231167833
## [216]   250371435  1139934426  -641921636  -858574433  -474934283
## [221] -2013707636 -1380890782   119355197   993660855 -2002084770
## [226] -1013984872  -379702021   925663385   -31002328  -442860704
## [231] -1075363102   641540264  1223098572 -1416033796  1656623026
## [236]  -797919760  1826465220   503635800   479382394  -180521552
```

```
## [241]   335149500  -762107820 -1027404862 -1827060960  1940004796
## [246]   -55285168   820174306 -1241009816   780750156  1441990908
## [251]   182952690   711513856 -1637546380  2029275144  1579137466
## [256] -1032808496  -740935188  -653815916   804057810 -1415082400
## [261]  1411310924     1021216  -350965118  1885104616   987875308
## [266]  -131369412  1998212722 -1838333072   848054564  1637511736
## [271] -1985453670   185338320  1063452092 -1501837164  -244441982
## [276]  -655956544  1724949308  2018636944   664617762   824380936
## [281]  -454972148  1579726236 -1666261486  -245134464   802374420
## [286]  1526755624   260631610 -1205064752  -361627028   937556532
## [291]  -494404974   -99312288 -1390302900  1614862304   774509986
## [296]  1360092584  -193173876 -1314447940 -1663111566 -1553441808
## [301]  1474524740 -1146038184 -2113923014  1375914096 -1694107588
## [306]  1286052116  -244301118 -1394274592 -1799518084  1647537872
## [311]  -826656350   363032616   887266060  1523945916 -1606634702
## [316] -1705936192  1284490292  -611903800   -82717510  1223915152
## [321]  1740148460 -1417786796   261071506  -232439520  -735589940
## [326] -1901903136    53278914 -1123929048  -739412052  1446503100
## [331] -1800705038  1762987568 -1902032988 -1903022664   152627034
## [336]   717533328   246260604  1132269844 -1594365630   343599360
## [341] -1585484356   763185488 -1647048414  -696819128   134752268
## [346] -1533188516 -1147008558  1720888832 -1992442540 -1205072664
## [351]  1506019578   603356880  1553058604   599599348  -568221294
## [356]   -26303136  1900303372  -891679520 -1327582622  -770317144
## [361] -2141666228  1110458236 -1984349902   163365488  1002681156
## [366] -1830044712  -482759814  1341773872 -1905328196 -2080945580
## [371]   830894146  1654662816   550937276  -155556912  -506636958
## [376]  1307590376  1514790092  1301985788 -2135581198  1363303552
## [381]  2121865716  1892557448  1059986234  1182360912 -1907645332
## [386] -1564824172  1223534674 -1003414048  1916359884  -428286304
## [391]  -449023614 -1081851672  1379461868   -76178116    39708786
## [396] -1623293328   506321444  2118223928 -1103890150  -418716464
## [401] -1680746180 -1496452204   625145730  1623387072  1943565884
## [406]  1450672528  1732028450  1481551112   487458956   872206108
## [411]  -748118894  2140900736  -935907308   548710184 -1376221638
## [416]  1162922448 -1448468116   593599412 -1498189294 -1658751008
## [421]   462232012   124672224   557945890  1357773608   781625868
## [426]  1816482620  -142301198  -510500112  1926556228   -13969832
## [431]  -233818310 -1136707600   460294844  -416275180  1083468994
## [436]  -781480864  1355603324  1851695056   578619682  -798888280
## [441]  1535516556  -172522308  1128772402  -958671296 -1018973004
## [446]  1397935176 -1571451462  1620720144  -804320020  -120377004
## [451]   539237138  1130743200  -645778740  1151912544  1355687874
## [456]  1031070504 -1306313446   142889871   589573337 -1228012690
## [461] -1810984196 -1041194899  -926225385 -1469796896  1488304574
## [466] -1061851253  -338333683 -2027587526  -665470856  1572082721
## [471]  -864508205  -357575180   -92260734  -727936233 -1218739743
## [476] -1034735834  1959295524 -2018968363   970828927  2088100120
## [481]  1582403638 -1341110589 -1107547387   214011554  -657752944
## [486] -1164080071  -795818773  -913533476   200285194   752745215
## [491]    63615817 -1335801634 -1217612148  -805518755 -1818050457
## [496] -1804608112   621505966   918960219   676442173  -402544982
## [501] -1789574328   224036721  -764759837  -670177308   261209874
## [506]  -893263129  1281983665   362571510  1179824404  1440671269
```

```
## [511]  -385713361  -963744536  2009648582 -1648018317  -719300907
## [516]  1092938930   163127424   830016617  -184674085   343641132
## [521]   154989626  -583638289   424254841   484102670 -1823956004
## [526]   550505549  -397533257  -309821120 -1833768930   947035883
## [531]  1101549165  1819580122  1254016664   600796417  -119647565
## [536]   507059668  1219230946   981013879   178359489 -1798344634
## [541]   519307140  1302534453   -26966241  -277288328 -1009389738
## [546]   145067491   844768165 -1768828350 -1545351120  -672619431
## [551] -1158866485    15817276  -881306518   -50051425 -1092056279
## [556]  1196083134   248280748   296336381   891261767   207606832
## [561] -1985825458 -2102088325  1662966493 -1304487158   873339432
## [566]   806441617  1094718083  1123275652   840202674   725952263
## [571]  -597236399 -1409356586  -114177612  1521819077   401542415
## [576]  1975999368   -10714458 -1167764909  1721318325 -2086876910
## [581]   742005472   969609033  -890084997 -1277850548  -971870758
## [586]  1665781711   -20106855  1492599214  1206353596   231206701
## [591]  1215581399  -910904032  -626793858   303180875   704691789
## [596]  1846420602 -1294941000 -2057982367   787944467 -1144544716
## [601]  -630963134  1707080791   636622753  -866277274   434270308
## [606]   154340245   835108031 -1921487784   511924342 -1737570173
## [611]   907225285   803644898  -625275824 -1045666439 -1073554773
## [616] -1362001380  -645341878   459588927 -1339341943   -24970338
## [621] -2003084724  -426558435 -1332570713  -956733232   355793518
## [626]  1450597184
```

```
glm.fit.2=glm(area~X + DMC + temp + jan + feb + mar + apr + jun + jul + aug + sep + oct, data=data)
cv.err.2=cv.glm(data,glm.fit.2)
cv.err.2
```

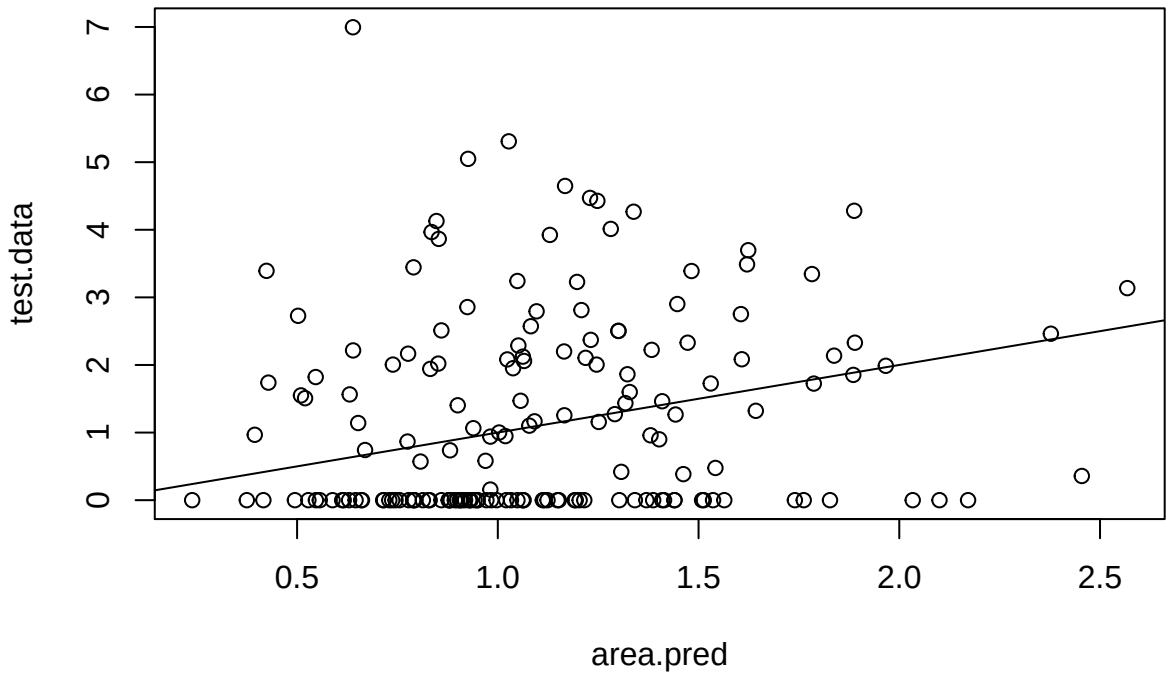
```
## $call
## cv.glm(data = data, glmfit = glm.fit.2)
##
## $K
## [1] 516
##
## $delta
## [1] 1.946604 1.946515
##
## $seed
##   [1]         403         236  1980538363  -125047968  -820381145
##   [6] -1073960099   -33499050  1164085917  1218464490 -1045685882
##  [11] -1933699935   795942478   495210185  1943097543  -229514104
##  [16] -1232480817   352489432   351006324  -867325393 -1442019852
##  [21]   629147155   966152721  1616591578   200016369  1339052126
##  [26]  -225299422 -1512315715   103118938  1508496509   226325507
##  [31]  1531586180  1851232347  -180012444  1939719752  1161394291
##  [36]  1375986488 -1867206193  -269449371  1080036398  1121658053
##  [41]  1200423730 -1114719938  1019413465   699098406 -2049351935
##  [46]  -216610673  1752238736 -1881625753  1183034576   289280140
##  [51] -1743926729  -243822724 -1713104501  -866739415 -2073616446
##  [56] -1798597783  1624378294  -511303094  -484622395 -1576496876
##  [61]   996804790  -917710736  1806554914   -22198181  -722998599
##  [66]   410108957  1149778011  1459463378 -1618359224   845376994
##  [71]   637274396 -1432005815  1437723723  1902160615  1262833869
##  [76] -1911506640 -1518922062  1373330532   883641086      380847
```

```
##  [81] -1184588235  -275900831 -1432937449   586023422  -849877260
##  [86]   463821014  -270234328 -2144570915 -1207087337  1190567323
##  [91]  1689142169  1175822956  1904310766 -2004549720   438376634
##  [96] -1394835005  -121754383  1087076069   584586339   107315546
## [101]   110581328  -774253782   322028116 -1036237599  1593876099
## [106]  1065739567  1846277637   963650216 -2078226102  -498934596
## [111]  -626315018   723134007   931715149    60391561   597152831
## [116]   791758566 -1727454852  1129299070  -826899616  1471815573
## [121]  1323965647   385117571 -1210125359   584755140 -1507369530
## [126] -1630192576  -517807086  -571062293   272680041  1894989389
## [131]  1425886795  -296480926   702856184 -1681846510 -1472159124
## [136]   273436889  1327605851  1856875895 -1096069219   924953536
## [141]  1110396098  -922946572   570967086   343580095   612081381
## [146]  2109955057  1830151207 -1652127890  2077242340 -1182212794
## [151]  1554609304  -178802003 -1741074873   876484779  -138661239
## [156]  -194324964 -1302787650  1604305720  -845765270   761367059
## [161]   286667873 -2061365355 -1822701485  -443807062 -1431880576
## [166]  -943009510   826639332   193208881  1060164691 -1018766401
## [171] -2131691499  1422972600   693581146   859577612   134206822
## [176]  -861972985   415151805 -1569592615 -1595615089   -38970474
## [181]  -573925972  -509242194  1450628496  1297158309  1107000575
## [186]   449088467  1181799169 -1383618252   141546518   575856144
## [191] -1510208958   791485563   571913753  -427005123   176980475
## [196]  -909485774  1663554728 -1719394814 -1913820036  -642612695
## [201] -1719239189 -1926862073  1015704877 -1550711536  1897118226
## [206] -1607617020  1275173406  1012447567   103797525 -1650061759
## [211]  1790336887   325626846 -1609980844 -1443351626 -1315423352
## [216]   -79557059  -311237705  1063514811  1174518585 -1703642100
## [221]  1304192654  1789749448  2048347290 -1932651421 -2024483631
## [226]  1563770629    73898051   707232506   616261360   704593034
## [231] -1416266956  -345343423 -1439682653 -1889542257  1834668133
## [236]  1053726728  -827427286 -2097712868   368026774  1417196055
## [241]  1519961325 -1825545879   -16419745  1825800198   716299100
## [246] -1807956770 -1324393152  1925488821   326096495  1697180579
## [251]   517499505  1088607076 -1897394138 -1965803424   656225202
## [256]   805488779  1644082121   213671917  2072270699 -1622521662
## [261]  1939475672  1890787634   707647436 -1916513735  -739951493
## [266]  1005398807  1273857405   516630432 -1924956126   941044500
## [271]  1814913486 -1704451617  -534041403   281412817  -475276921
## [276]   904658254   756476612 -1950127194  -370800520 -1416311027
## [281] -1555906073 -1976536885  -327526743   926657084  -872015138
## [286]  -455521832 -1105296132  1314064768 -1833163414  1525557867
## [291] -2086205717 -2106114529  1750183917  1402811018  2121599494
## [296] -1514668274  1789977404  1804828197   -82739535  1044404073
## [301]  1824048843   172317904  -418259704 -1460278612   931696878
## [306]   202599327  1450658015  -214279669 -1634230007 -1916141370
## [311]  1749006826  -685331598   919572488     -338767  -855323459
## [316]  -565531379  1798641407   421964228   577509668  1113407128
## [321] -1657021822  1575406067 -1781653165 -1834104809   127588245
## [326]  1666082818   417635150  1868688774 -1161682316   426014637
## [331] -1229536263  1481653585  2026697715   713524648   739408608
## [336] -2089412364 -1006873546  -892287913  1629967031   657944211
## [341]  1428628513   763220142  1141569474  -586336902  1359023952
## [346] -1821931639   619923605  1249165925  1256352759  -616761428
```

```
## [351]   857839788 -2005944240 -1895061446  -835570469 -1564404389
## [356]  1759702639  1436432445   210900858 -2023817930  1774770654
## [361]   -86437716 -1399046411  1418867169  1388343801  1742237851
## [366]  -818439648  1356679128  1159213852 -1996007714  2060155919
## [371]   623546095   873516987  1184786969   502484630   -75589606
## [376]  1458888066 -1622834760  1559476833 -1644782355 -1492794691
## [381]   207161839  -399818988  -682203980 -1250108536   247633394
## [386] -1444908733   259589347  1745526279 -1239523131  -507121518
## [391] -1169282338 -1011136010   -93315196  1886091229    68903689
## [396]  1881553217   -60416093 -1612118888   659924176  1690294724
## [401]  1616569542  1432268455   -50453529   559058339 -1991628271
## [406] -1820685154  -779481838  -749891094  -128191712  2143460313
## [411]   849139301  1709237493  1251924039  1351389212  -818950692
## [416]   482002080  1944196042 -1230654645  1639819147  1875917567
## [421]  -449099955   439110634  -507730778  1953268142 -1348028516
## [426] -1238417531 -1451510959  -521402167 -1296891861  1166177968
## [431]   -41050392  1017413452 -1816812402   743156863   351108991
## [436]  1850326443 -1433845335  2072566758   995646602  -842355822
## [441] -1826360856  -801381871  -667949795 -1932314131  1456116895
## [446]  1462468068  -365338364  1706811640   382005346   704987411
## [451] -1384319309 -1416245833  1063344693   738233250   882766702
## [456]  1034614909  -189422190   588242311  -267551754  -533240099
## [461]  1991196518 -1875437903  1794952952 -1561023755  1000699398
## [466]  -116544433   401214986   917641425  1692588966  1076991977
## [471]   428108540  1475461149  1736064394  1940532591   334313254
## [476]  1762388861     -776938  -616690543  2128094184 -1581197251
## [481]   198487006  -521431313  1986138074  1302007169  -923129554
## [486] -1966832191   -74710484   912392925    -2790254 -1296023881
## [491] -1779108266 -2007800003  -988022042 -1691069039  1425852776
## [496]  -540657931 -1166237722  -519239409 -1368853878 -1434350239
## [501]  -815532106  1402351913  1508708028  1792947533   933964298
## [506]   383417087 -1587872650 -2140432019   970437382  1106500577
## [511]  1417452616 -1209893827  1932748606  1300635465   784433113
## [516] -1772047062  -614851456   550550105  -934610935  1825560048
## [521]  -468798686 -1411812251 -1411738103  1662300882 -1561081636
## [526]  1446953397  1216616393   635548356  -814439558  1590512353
## [531] -1630559807   316142498 -2063831616  -217001639 -1138305823
## [536] -1161666696   167656098  -194523307  -582937775  2041999306
## [541]  1076095220  1799241709 -1443440591   182444372  -998449174
## [546] -1572227031  1226168249  -859164006 -1357239552  1370496361
## [551]  -757766903   572980640 -2067465326  -560435707   214208089
## [556] -1113319134 -2001200500   892451461  1585436489  2082318676
## [561]  -220896310  -996825263 -1704356495 -1711580894  1271979456
## [566]  -661656791   105809745  1278023928 -1579682830  -214878795
## [571] -1997389039   765062026  1731574196 -1240695091  1888972705
## [576]  1967925924   520547930  -647645783 -1985880871  -652827318
## [581]  1272906656   514889017   822618185   915869392  -465682366
## [586]   556848421    95648009  -651513870   863467964  -375347371
## [591]  1981426537   176783780   523648218  -589977567  1914642721
## [596] -1871760254 -1586947104  -346442695   703798945   732710968
## [601]  -165412766  1888428821  1893337521  -743130742 -1050505356
## [606]   471691885 -1581648751  1125688788  2098523018   358039561
## [611] -1567651463   597627866  1346302752  -160376471  1063411465
## [616]  -623243680   897524434   995310533 -1124453319 -1103252830
```

```
## [621] -1708500724  -983918331   324701481 -1224973676 -1261295304
## [626] -2057610510
```

```
bag.area=randomForest(area~., data=data[train,], mtry=6, importance=TRUE)
area.pred=predict(bag.area, newdata=data[-train,])
test.data=data[-train,]$area
plot(area.pred, test.data)
abline(0,1)
```




```
mean((area.pred-test.data)^2) #1.943381 in the same range as before
```

```
## [1] 2.116
```

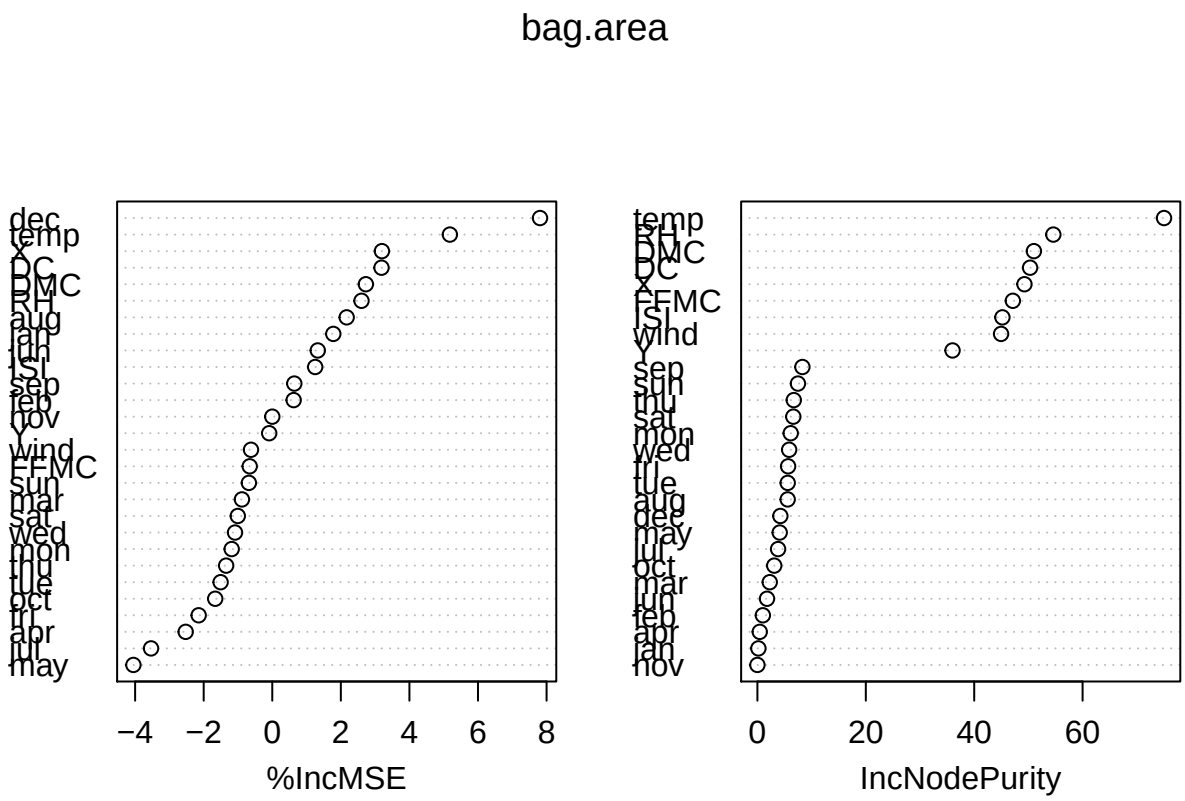
```
#my models so far are crap. again lol.

importance(bag.area)
```

```
##          %IncMSE IncNodePurity
## X     3.19806573    49.2712713
## Y    -0.08928522    36.0009894
## FFMC -0.65755976    47.1248016
## DMC   2.72976225    51.0142787
## DC    3.18589339    50.3125974
## ISI   1.25127371    45.2035645
## temp  5.18021014    75.0268518
## RH    2.60253154    54.6005124
## wind -0.62013637    44.9548646
## jan   1.77981456     0.1793389
## feb   0.62355874     1.0006808
## mar  -0.88516752     2.2787578
## apr  -2.52477342     0.4332846
## may  -4.05006674     4.1118277
## jun   1.32664757     1.7688555
## jul  -3.53558455     3.8122397
```

```
## aug   2.17042565     5.5823438
## sep   0.64267700     8.3010496
## oct  -1.66423243     3.1047610
## nov   0.00000000     0.0000000
## dec   7.80880209     4.2325123
## mon  -1.18319622     6.1466122
## tue  -1.50600633     5.5894805
## wed  -1.08631231     5.8528253
## thu  -1.34477981     6.7086144
## fri  -2.14661875     5.6637627
## sat  -1.00513486     6.6211796
## sun  -0.68218425     7.4799297
```

```
varImpPlot(bag.area)
```




```
#reclassify data
clus=data
clus$FFMC=exp(clus$FFMC)
clus$area=exp(clus$area)-1
#for(i in 1:517){
  #if(clus$area[i]==0){clus$area[i]='low'}
  #else if (clus$area[i]<=20){clus$area[i]='medium'}
  #else if (clus$area[i]<50){clus$area[i]='high'}
  #else (clus$area[i]='super high')
#}
```

```
#SVM
```

```
#KNN
test=clus[-train,]
```

```
folds=function(x){
  knn.fire=knn(clus[train,c(4,5)], test[,c(4,5)], clus[train,]$area, k=x)
  least=mean(knn.fire==test$area)

  knn.fire=knn(clus[train,c(4,7)], test[,c(4,7)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(4,8)], test[,c(4,8)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(4,19)], test[,c(4,19)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(4,22)], test[,c(4,22)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(5,7)], test[,c(5,7)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(5,8)], test[,c(5,8)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(5,19)], test[,c(5,19)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(5,22)], test[,c(5,22)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(7,8)], test[,c(7,8)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(7,19)], test[,c(7,19)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(7,22)], test[,c(7,22)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(8,19)], test[,c(8,19)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(8,22)], test[,c(8,22)], clus[train,]$area, k=x)
```

```
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  knn.fire=knn(clus[train,c(19,22)], test[,c(19,22)], clus[train,]$area, k=x)
  a=mean(knn.fire==test$area)
  if(a>least){least=a}

  return(least)
}

sapply(1:20, folds)
```

```
##  [1] 0.4651163 0.4651163 0.4651163 0.4651163 0.4651163 0.4651163 0.4651163
##  [8] 0.4651163 0.4651163 0.4651163 0.4651163 0.4651163 0.4651163 0.4651163
## [15] 0.4651163 0.4651163 0.4651163 0.4651163 0.4651163 0.4651163
```

```
#which combo of variables?
x=18

knn.fire=knn(clus[train,c(4,5)], test[,c(4,5)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(4,7)], test[,c(4,7)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(4,8)], test[,c(4,8)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(4,19)], test[,c(4,19)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(4,22)], test[,c(4,22)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(5,7)], test[,c(5,7)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(5,8)], test[,c(5,8)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(5,19)], test[,c(5,19)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(5,22)], test[,c(5,22)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(7,8)], test[,c(7,8)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(7,19)], test[,c(7,19)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(7,22)], test[,c(7,22)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(8,19)], test[,c(8,19)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(8,22)], test[,c(8,22)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```

```
knn.fire=knn(clus[train,c(19,22)], test[,c(19,22)], clus[train,]$area, k=x)
mean(knn.fire==test$area)
```

```
## [1] 0.4651163
```