```
data <- read.csv("/home/xavi/Documents/models/treballs/train.csv", header=TRUE)
#data_2 es fara servir per identificar els passatger que assignem a la particio cross
data_2 <- read.csv("/home/xavi/Documents/models/treballs/train.csv", header=TRUE)
#treiem NA
data = data[!is.na(data[,3]),]
data = data[!is.na(data[,5]),]
data = data[!is.na(data[,6]),]
data = data[!is.na(data[,7]),]
data = data[!is.na(data[,8]),]
data = data[!is.na(data[,10]),]
data = data[!is.na(data[,11]),]
data = data[!is.na(data[,12]),]

y = factor(data[,2])
y_num = data[,2]
id = data[,1]
class = factor(data[,3]) #Pclass
sex = factor(data[,5]) #Sex
age = data[,6] #Age
sib = factor(data[,7]) #sib sp
parch = factor(data[,8]) #par ch
#tiket = factor(data[,9]) #tiket
fare = data[,10] #fare
cabin = factor(data[,11]) #cabin
port = factor(data[,12]) #port of embarqued

y_2 = factor(data_2[,2])
y_num_2 = data_2[,2]
id_2 = data_2[,1]
class_2 = factor(data_2[,3]) #Pclass
sex_2 = factor(data_2[,5]) #Sex
age_2 = data_2[,6] #Age
sib_2 = factor(data_2[,7]) #sib sp
parch_2 = factor(data_2[,8]) #par ch
#tiket = factor(data[,9]) #tiket
fare_2 = data_2[,10] #fare
cabin_2 = factor(data_2[,11]) #cabin
port_2 = factor(data_2[,12]) #port of embarqued


#class
summary(class)
plot(class)
plot(class,y)
plot(class,y, xlab = "pclass", ylab="% supervivents", main = "% de supervivencia per clase", col = c
("blue","red"), sub = "1 sobreviu, 0 mort")

#sex
summary(sex)
plot(sex)
plot(sex,y)

#age
summary(age)
hist(age)
#test de normalitat, shapiro.test
shapiro.test(age)
boxplot(age)
qqnorm(age)
qqline(age)


ample_edat = 10
age_mod = array(dim = length(age))
for (i in 1:length(age)){
  age_mod[i] = as.integer(age[i]/ample_edat)
  if (is.na(age_mod[i])==FALSE){
    if(age_mod[i]>6){
      age_mod[i] = 6
    }
  }
}
```

```
age_mod = factor(age_mod)
plot(age_mod)
plot(age_mod,y)

#sib
summary(sib)
plot(sib)
plot(sib,y)
agrupem_sib = 2
sib_mod = array(dim = length(sib))
for (i in 1:length(sib)){
  if (as.integer(sib[i]) > (agrupem_sib)){
    sib_mod[i] = agrupem_sib
  }
  else{
    sib_mod[i] = as.integer(sib[i])-1
  }
}
sib_mod = factor(sib_mod)
summary(sib_mod)
plot(sib_mod)
plot(sib_mod,y)

sib_mod_plot = array(dim = length(sib))
for (i in 1:length(sib)){
  if (as.integer(sib[i]) > 1){
    sib_mod_plot[i] = "2 o +"
  }
  else{
    sib_mod_plot[i] = "0 o 1"
  }
}
sib_mod_plot = factor(sib_mod_plot)
summary(sib_mod_plot)
plot(sib_mod_plot)
plot(sib_mod_plot,y)

sib_mod_2 = sib_2
sib_mod_2[sib_2==1]=0
sib_mod_2[sib_2==3]=2
sib_mod_2[sib_2==4]=2
sib_mod_2[sib_2==5]=2
sib_mod_2[sib_2==8]=2

#parch
summary(parch)
plot(parch)
plot(parch,y)
agrupem_parch = 2
parch_mod = array(dim = length(parch))
for (i in 1:length(parch)){
  if (as.integer(parch[i]) > (agrupem_parch)){
    parch_mod[i] = agrupem_parch
  }
  else{
    parch_mod[i] = as.integer(parch[i])-1
  }
}
parch_mod = factor(parch_mod)
summary(parch_mod)
plot(parch_mod)
plot(parch_mod,y)
plot(parch_mod,sib_mod)

#fare
summary(fare)
hist(fare)
ample_fare = 20
fare_mod = array(dim = length(fare))
plot(fare)
for (i in 1:length(fare)){
  if (fare[i]<25){
```

```
    fare_mod[i] = 0
  }
  else{
    if (fare[i]<100){
      fare_mod[i] = 1
    }
    else{
      fare_mod[i] = 2
    }
  }
}
fare_mod = factor(fare_mod)
summary(fare_mod)
plot(fare_mod)
plot(fare_mod,y)
plot(fare_mod,class)
#cabin
cabin_mod = array(dim = length(cabin))
for (i in 1:length(cabin)){
  if (cabin[i]==''){
    cabin_mod[i]=0
  }
  else{
    cabin_mod[i]=1
  }
}
summary(cabin_mod)
cabin_mod = factor(cabin_mod)
plot(cabin_mod)
plot(cabin_mod,y)

cabin_mod_2 = array(dim = length(cabin_2))
for (i in 1:length(cabin_2)){
  if (cabin_2[i]==''){
    cabin_mod_2[i]=0
  }
  else{
    cabin_mod_2[i]=1
  }
}
cabin_mod_2 = factor(cabin_mod_2)

#Port
summary(port)
plot(port)
plot(port,y)


#particio
prop = 0.75
reord = sample(1:length(y), size=length(y), replace=F)
train = reord[1:length(y)*prop]
cross = reord[(length(y)*prop+1):length(y)]

id_train = id[train]
y_train = y[train]
class_train = class[train]
sex_train = sex[train]
age_train = age[train]
sib_mod_train = sib_mod[train]
parch_mod_train = parch_mod[train]
cabin_mod_train = cabin_mod[train]
fare_mod_train = fare_mod[train]
port_train = port[train]

logit<-glm(y_train~class_train+sex_train+age_train+sib_mod_train+parch_mod_train+cabin_mod_train
+fare_mod_train+port_train,family=binomial)
summary(logit)

logit<-glm(y_train~class_train+sex_train+age_train+sib_mod_train+parch_mod_train+cabin_mod_train
+fare_mod_train,family=binomial)
summary(logit)
```

```
logit<-glm(y_train~class_train+sex_train+age_train+sib_mod_train+parch_mod_train
+cabin_mod_train,family=binomial)
summary(logit)

logit<-glm(y_train~class_train+sex_train+age_train+sib_mod_train+cabin_mod_train,family=binomial)
summary(logit)

sib_mod[sib_mod==1]=0
logit<-glm(y_train~class_train+sex_train+age_train+sib_mod_train+cabin_mod_train,family=binomial)
summary(logit)

#mirem l'encert del cross
newdata = data.frame(class_train=class[cross],sex_train = sex[cross],age_train = age
[cross],sib_mod_train=sib_mod[cross], cabin_mod_train = cabin_mod[cross])
pre = predict(logit,newdata, type =c('response'),se.fit =T)
pre1 = pre[1]
output <- matrix(unlist(pre1), ncol = 1, byrow = TRUE)
output = as.integer(output+0.5)
sum(y[cross] == output)/length(output)

#interacció
logit<-glm(y_train~class_train+sex_train+(class_train):(sex_train)+age_train+sib_mod_train
+cabin_mod_train,family=binomial)
summary(logit)

#mirem l'encert del cross
newdata = data.frame(class_train=class[cross],sex_train = sex[cross],age_train = age
[cross],sib_mod_train=sib_mod[cross], cabin_mod_train = cabin_mod[cross])
pre = predict(logit,newdata, type =c('response'),se.fit =T)
pre1 = pre[1]
output <- matrix(unlist(pre1), ncol = 1, byrow = TRUE)
output = as.integer(output+0.5)
sum(y[cross] == output)/length(output)


# recomptes XVLE TAULA 1-----------------------------------------------
# 1 ?s viu

sum(age>16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age>16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age>16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==2 )

sum(age>16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age>16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age>16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==0 )

sum(age>16 & sex=='male' & class==1 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age>16 & sex=='male' & class==1 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age>16 & sex=='male' & class==1 & cabin_mod==0  & sib_mod ==2 )

sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 )


sum(age>16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age>16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age>16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==2 )

sum(age>16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age>16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age>16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==0 )

sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==2 )

sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age>16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 )
```

```
sum(age>16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age>16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age>16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==2 )

sum(age>16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age>16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age>16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==0 )

sum(age>16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age>16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age>16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==2 )

sum(age>16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age>16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age>16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==0 )


sum(age>16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age>16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age>16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==2 )

sum(age>16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age>16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age>16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==0 )

sum(age>16 & sex=='female' & class==1 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age>16 & sex=='female' & class==1 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age>16 & sex=='female' & class==1 & cabin_mod==0  & sib_mod ==2 )

sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 )


sum(age>16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age>16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age>16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==2 )

sum(age>16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age>16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age>16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==0 )

sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==2 )

sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age>16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 )


sum(age>16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age>16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age>16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==2 )

sum(age>16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age>16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age>16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==0 )

sum(age>16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age>16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age>16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==2 )

sum(age>16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age>16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age>16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==0 )

-------------
  sum(age<=16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==2 )
```

```
sum(age<=16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='male' & class==1 & cabin_mod==1  & sib_mod ==0 )

sum(age<=16 & sex=='male' & class==1 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='male' & class==1 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='male' & class==1 & cabin_mod==0  & sib_mod ==2 )

sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 )


sum(age<=16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==2 )

sum(age<=16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==1  & sib_mod ==0 )

sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==2 )

sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='male' & class==2 & cabin_mod==0  & sib_mod ==0 )


sum(age<=16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==2 )

sum(age<=16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='male' & class==3 & cabin_mod==1  & sib_mod ==0 )

sum(age<=16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==2 )

sum(age<=16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='male' & class==3 & cabin_mod==0  & sib_mod ==0 )


sum(age<=16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==2 )

sum(age<=16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='female' & class==1 & cabin_mod==1  & sib_mod ==0 )

sum(age<=16 & sex=='female' & class==1 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='female' & class==1 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='female' & class==1 & cabin_mod==0  & sib_mod ==2 )

sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 )


sum(age<=16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==2 )

sum(age<=16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==1  & sib_mod ==0 )
```

```
sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==2 )

sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='female' & class==2 & cabin_mod==0  & sib_mod ==0 )


sum(age<=16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==2 )

sum(age<=16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='female' & class==3 & cabin_mod==1  & sib_mod ==0 )

sum(age<=16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==2 & y==0)
sum(age<=16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==2 & y==1)
sum(age<=16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==2 )

sum(age<=16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==0 & y==0)
sum(age<=16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==0 & y==1)
sum(age<=16 & sex=='female' & class==3 & cabin_mod==0  & sib_mod ==0 )




#analitzem els errors
op = output
yc = y[cross]

mv = op==0 & yc ==1
vm = op==1 & yc ==0

idc_mv = idc[mv]
idc_vm = idc[vm]


recmv = c(sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==0 &
```

```
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]>16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='male' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0 &
sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==1
& sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==1
& sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==0
& sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==1 & cabin_mod_2[idc_mv]==0
& sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1
& sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1
& sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0
& sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0
& sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1
& sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==1
& sib_mod_2[idc_mv] ==1),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0
& sib_mod_2[idc_mv] ==0),
        sum(age_2[idc_mv]<=16 & sex_2[idc_mv]=='female' & class_2[idc_mv]==2 & cabin_mod_2[idc_mv]==0
& sib_mod_2[idc_mv] ==1))

recvm = c(sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
```

```
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]>16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='male' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0 &
sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==1
& sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==1
& sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==0
& sib_mod_2[idc_vm] ==0),
```

```
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==1 & cabin_mod_2[idc_vm]==0
& sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1
& sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1
& sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0
& sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0
& sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1
& sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==1
& sib_mod_2[idc_vm] ==1),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0
& sib_mod_2[idc_vm] ==0),
        sum(age_2[idc_vm]<=16 & sex_2[idc_vm]=='female' & class_2[idc_vm]==2 & cabin_mod_2[idc_vm]==0
& sib_mod_2[idc_vm] ==1))

#Dones i nens
class_temp=1
age_d_1 = mean(age[class==class_temp & age>=16 & sex=='female'])
sib_d_1 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='female'])-1)
cab_d_1 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='female'])-1)

class_temp=2
age_d_2 = mean(age[class==class_temp & age>=16 & sex=='female'])
sib_d_2 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='female'])-1)
cab_d_2 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='female'])-1)

class_temp=3
age_d_3 = mean(age[class==class_temp & age>=16 & sex=='female'])
sib_d_3 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='female'])-1)
cab_d_3 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='female'])-1)

class_temp=1
age_nn_1 = mean(age[class==class_temp & age<16 & sex=='female'])
sib_nn_1 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='female'])-1)
cab_nn_1 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='female'])-1)

class_temp=2
age_nn_2 = mean(age[class==class_temp & age<16 & sex=='female'])
sib_nn_2 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='female'])-1)
cab_nn_2 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='female'])-1)

class_temp=3
age_nn_3 = mean(age[class==class_temp & age<16 & sex=='female'])
sib_nn_3 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='female'])-1)
cab_nn_3 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='female'])-1)

class_temp=1
age_h_1 = mean(age[class==class_temp & age>=16 & sex=='male'])
sib_h_1 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='male'])-1)
cab_h_1 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='male'])-1)

class_temp=2
age_h_2 = mean(age[class==class_temp & age>=16 & sex=='male'])
sib_h_2 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='male'])-1)
cab_h_2 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='male'])-1)

class_temp=3
age_h_3 = mean(age[class==class_temp & age>=16 & sex=='male'])
sib_h_3 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='male'])-1)
cab_h_3 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='male'])-1)

class_temp=1
age_n_1 = mean(age[class==class_temp & age<16 & sex=='male'])
sib_n_1 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='male'])-1)
cab_n_1 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='male'])-1)

class_temp=2
age_n_2 = mean(age[class==class_temp & age<16 & sex=='male'])
```

```
sib_n_2 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='male'])-1)
cab_n_2 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='male'])-1)

class_temp=3
age_n_3 = mean(age[class==class_temp & age<16 & sex=='male'])
sib_n_3 = median(as.numeric(sib_mod[class==class_temp & age>=16 & sex=='male'])-1)
cab_n_3 = median(as.numeric(cabin_mod[class==class_temp & age>=16 & sex=='male'])-1)


newdata = data.frame(class_train=factor(rep(c(1,2,3),4)),sex_train = c(rep('female',6),rep
('male',6)),age_train = c
(age_d_1,age_d_2,age_d_3,age_nn_1,age_nn_2,age_nn_3,age_h_1,age_h_2,age_h_3,age_n_1,age_n_2,age_n_3),sib_mo
(c
(sib_d_1,sib_d_2,sib_d_3,sib_nn_1,sib_nn_2,sib_nn_3,sib_h_1,sib_h_2,sib_h_3,sib_n_1,sib_n_2,sib_n_3)),
cabin_mod_train = factor(c
(cab_d_1,cab_d_2,cab_d_3,cab_nn_1,cab_nn_2,cab_nn_3,cab_h_1,cab_h_2,cab_h_3,cab_n_1,cab_n_2,cab_n_3)))
pre = predict(logit,newdata, type =c('response'),se.fit =T)
pre <- matrix(unlist(pre[1]), ncol = 3, byrow = TRUE)

x <- matrix(c(rep(c(1,2,3),4)),ncol=3,byrow =TRUE)
plot(c(x[1,],x[2,],x[3,],x[4,]),c(pre[1,],pre[2,],pre[3,],pre[4,]),xlab='clase',ylab ='prob')
lines(x[1,],pre[1,])
lines(x[2,],pre[2,])
lines(x[3,],pre[3,])
lines(x[4,],pre[4,])

#Probabilitat sobreviure noi titanic
newdata=data.frame(class_train=factor(3),sex_train='male',age_train=20,sib_mod_train=factor
(0),cabin_mod_train=factor(0))
pre = predict(logit,newdata, type =c('response'),se.fit =T)
pre <- matrix(unlist(pre[1]), ncol = 3, byrow = TRUE)
pre
```