Análisis de Clustering Jerárquico - Adult Dataset

Autor

2025-05-20

```
setwd("~/Descargas")
dd <- read.csv("credscoClean.csv", sep=";", stringsAsFactors = TRUE);
names(dd)
```

```
##  [1] "Dictamen"                  "Antiguedad.Trabajo"
##  [3] "Vivienda"                  "Plazo"
##  [5] "Edad"                      "Estado.civil"
##  [7] "Registros"                 "Tipo.trabajo"
##  [9] "Gastos"                    "Ingresos"
## [11] "Patrimonio"                "Cargas.patrimoniales"
## [13] "Importe.solicitado"        "Precio.del.bien.financiado"
## [15] "Estalvi"                   "RatiFin"
```

```
dim(dd)
```

```
## [1] 4455   16
```

```
summary(dd)
```

```
##     Dictamen    Antiguedad.Trabajo         Vivienda        Plazo
##  negatiu:1254   Min.   : 0.000     altres viv  : 319   Min.   : 6.00
##  positiu:3200   1st Qu.: 2.000     contr_privat: 247   1st Qu.:36.00
##  NA's   :   1   Median : 5.000     escriptura  :2107   Median :48.00
##                 Mean   : 7.987     ignora_cont :  20   Mean   :46.44
##                 3rd Qu.:12.000     lloguer     : 973   3rd Qu.:60.00
##                 Max.   :48.000     pares       : 783   Max.   :72.00
##                                    VivUnkown   :   6
##       Edad          Estado.civil   Registros                Tipo.trabajo
##  Min.   :18.00   casat    :3241   reg_no:3682   altres sit        : 171
##  1st Qu.:28.00   divorciat:  38   reg_si: 773   autonom           :1024
##  Median :36.00   ECUnknown:   1                 fixe              :2806
##  Mean   :37.08   separat  : 130                 temporal          : 452
##  3rd Qu.:45.00   solter   : 978                 WorkingTypeUnknown:   2
##  Max.   :68.00   vidu     :  67
##
##      Gastos          Ingresos       Patrimonio     Cargas.patrimoniales
##  Min.   : 35.00   Min.   :  0.0   Min.   :     0   Min.   :    0.0
##  1st Qu.: 35.00   1st Qu.: 80.0   1st Qu.:     0   1st Qu.:    0.0
##  Median : 51.00   Median :120.0   Median :  3000   Median :    0.0
##  Mean   : 55.57   Mean   :130.5   Mean   :  5347   Mean   :  341.6
```

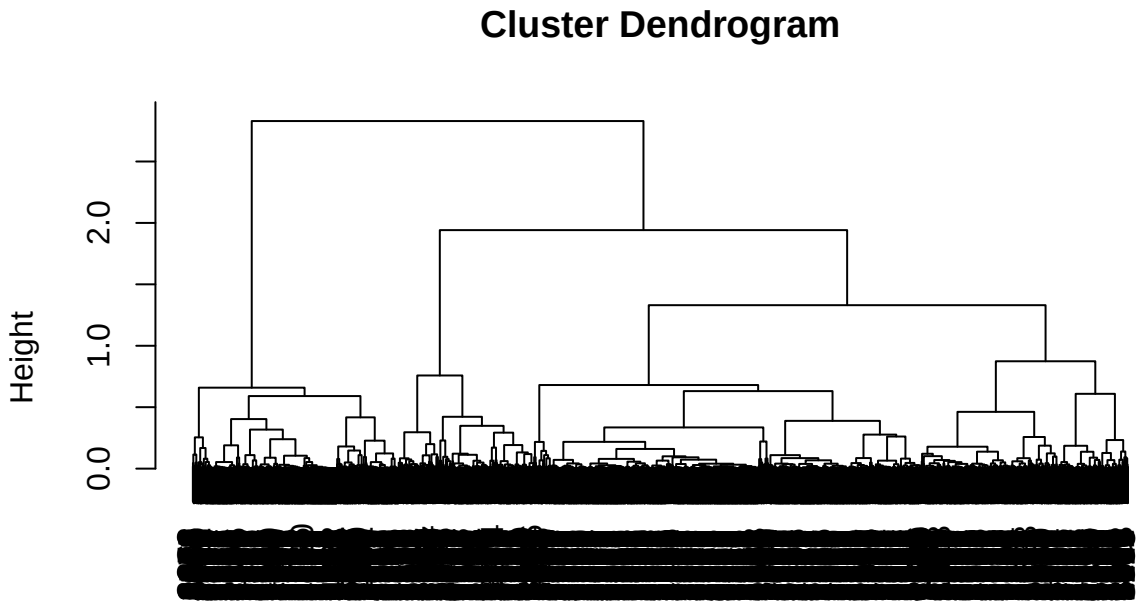
```
##  3rd Qu.: 72.00   3rd Qu.:165.0   3rd Qu.:  6000   3rd Qu.:    0.0
##  Max.   :180.00   Max.   :959.0   Max.   :300000   Max.   :30000.0
##
##  Importe.solicitado Precio.del.bien.financiado    Estalvi
##  Min.   : 100       Min.   :  105              Min.   :-16.080
##  1st Qu.: 700       1st Qu.: 1118              1st Qu.:  1.227
##  Median :1000       Median : 1400              Median :  2.916
##  Mean   :1039       Mean   : 1463              Mean   :  3.415
##  3rd Qu.:1300       3rd Qu.: 1692              3rd Qu.:  4.996
##  Max.   :5000       Max.   :11140              Max.   : 33.250
##
##     RatiFin
##  Min.   :  6.702
##  1st Qu.: 60.040
##  Median : 77.101
##  Mean   : 72.610
##  3rd Qu.: 88.422
##  Max.   :100.000
##
```

```
attach(dd)

#set a list of numerical variables
names(dd)
```

```
##  [1] "Dictamen"                  "Antiguedad.Trabajo"
##  [3] "Vivienda"                  "Plazo"
##  [5] "Edad"                      "Estado.civil"
##  [7] "Registros"                 "Tipo.trabajo"
##  [9] "Gastos"                    "Ingresos"
## [11] "Patrimonio"                "Cargas.patrimoniales"
## [13] "Importe.solicitado"        "Precio.del.bien.financiado"
## [15] "Estalvi"                   "RatiFin"
```

```
#hierarchical clustering

#euclidean distance si totes son numeriques
dcon<-data.frame (Antiguedad.Trabajo,Plazo,Edad,Gastos,Ingresos,Patrimonio,Cargas.patrimoniales,Importe

d  <- dist(dcon[1:10,])

#move to Gower mixed distance to deal
#simoultaneously with numerical and qualitative data

library(cluster)

#dissimilarity matrix
#do not include in actives the identifier variables nor the potential response variable

actives<-c(2:16)
dissimMatrix <- daisy(dd[,actives], metric = "gower", stand=TRUE)

distMatrix<-dissimMatrix^2
```

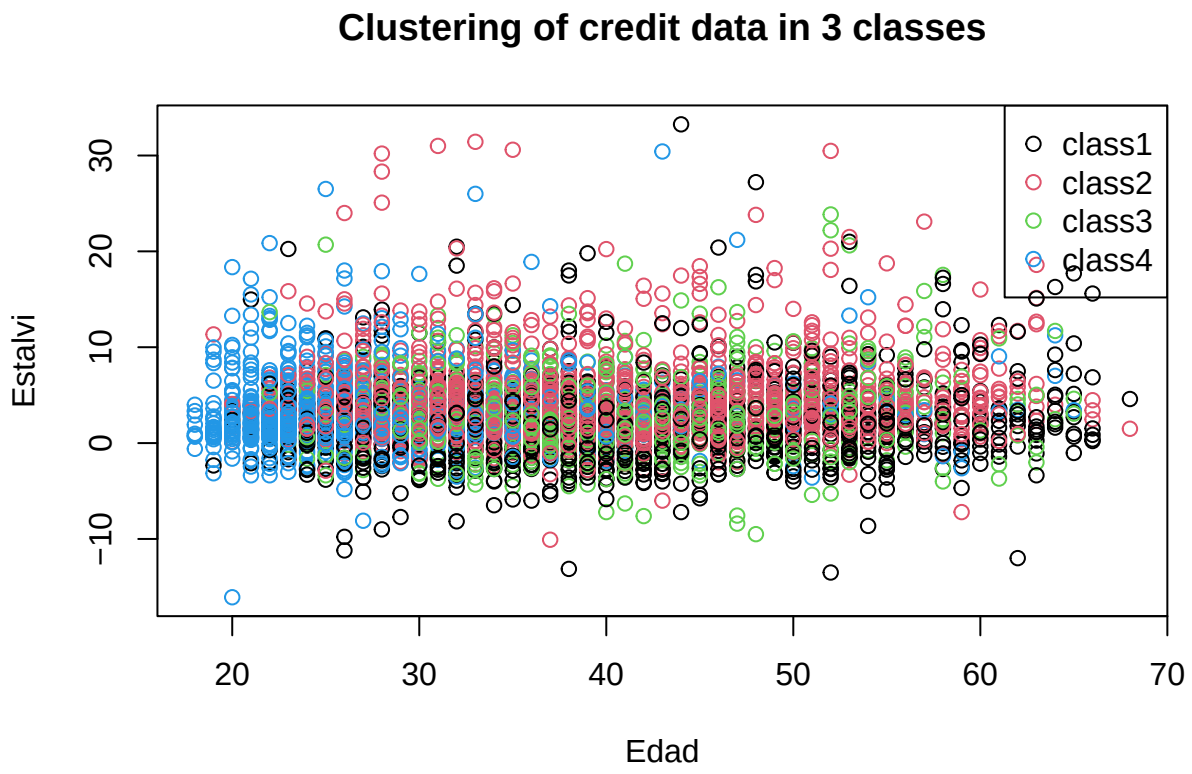
```r
h1 <- hclust(distMatrix,method="ward.D2")  # NOTICE THE COST
#versions noves "ward.D" i abans de plot: par(mar=rep(2,4)) si se quejara de los margenes del plot

plot(h1)
```



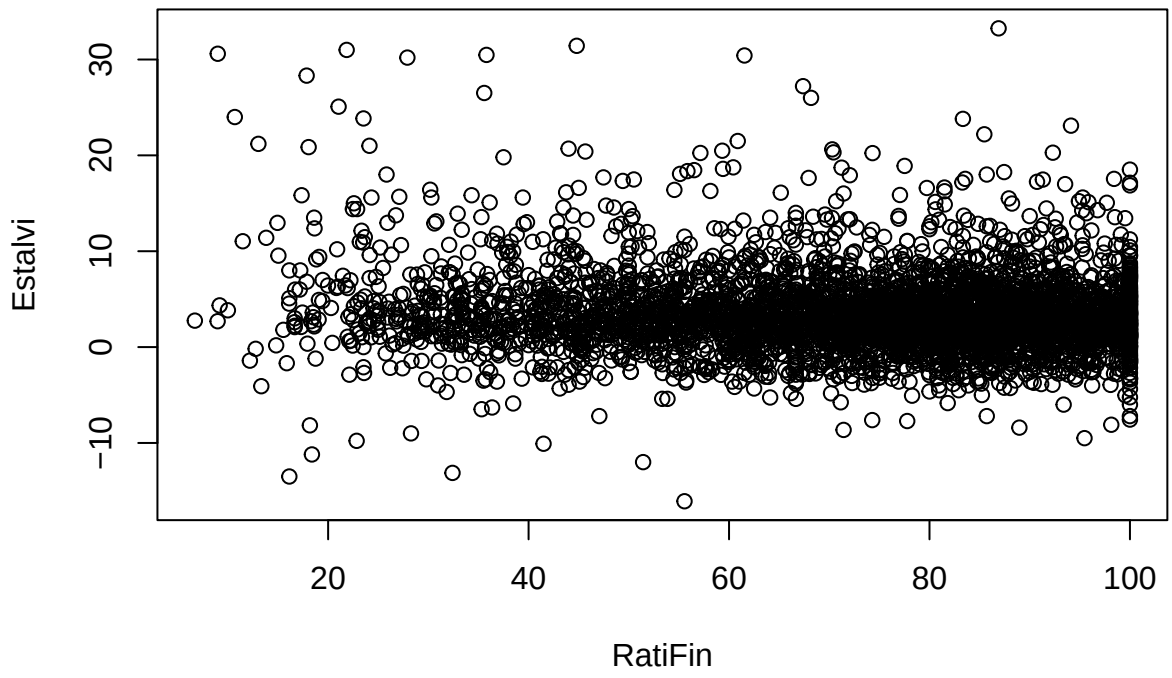

```r
k<-4

c2 <- cutree(h1,k)

#class sizes
table(c2)
```

```
## c2
##    1    2    3    4
##  984 1847  644  980
```

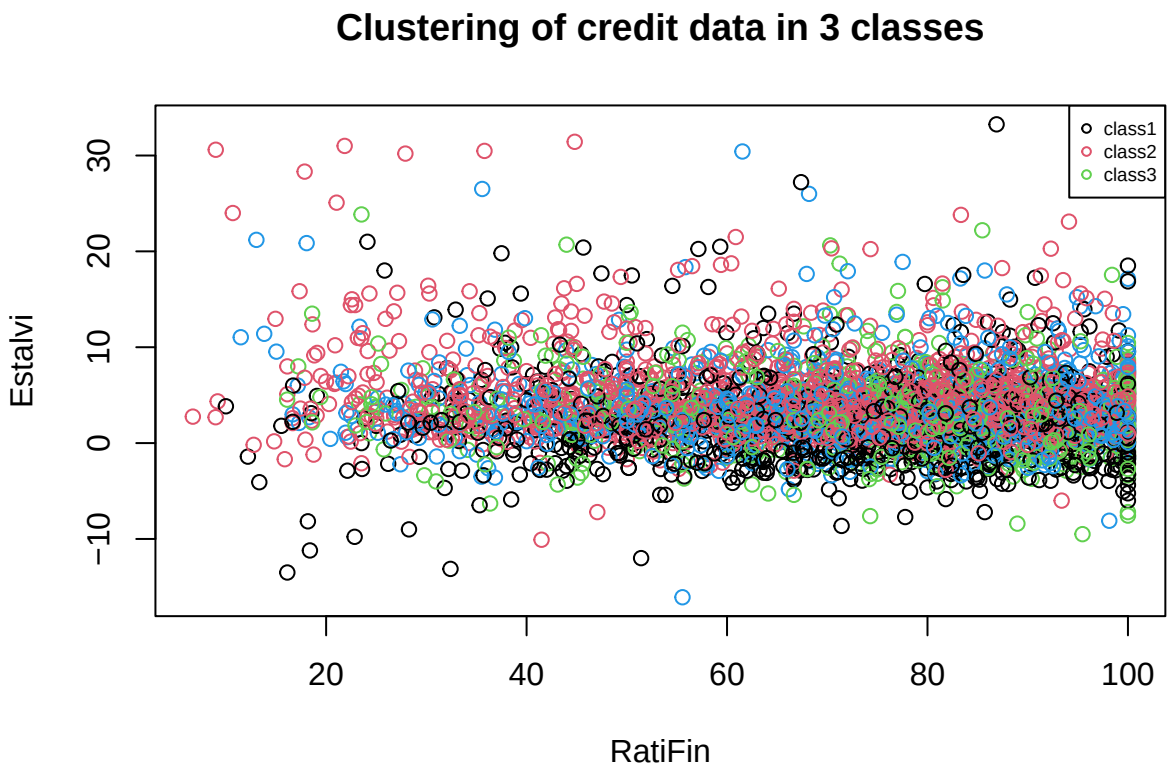
```r
#comparing with other partitions
#table(c1,c2)

# LETS SEE THE PARTITION VISUALLY

c1<-c2
plot(Edad,Estalvi,col=c1,main="Clustering of credit data in 3 classes")
legend("topright",c("class1","class2","class3","class4"),pch=1,col=c(1:k))
```



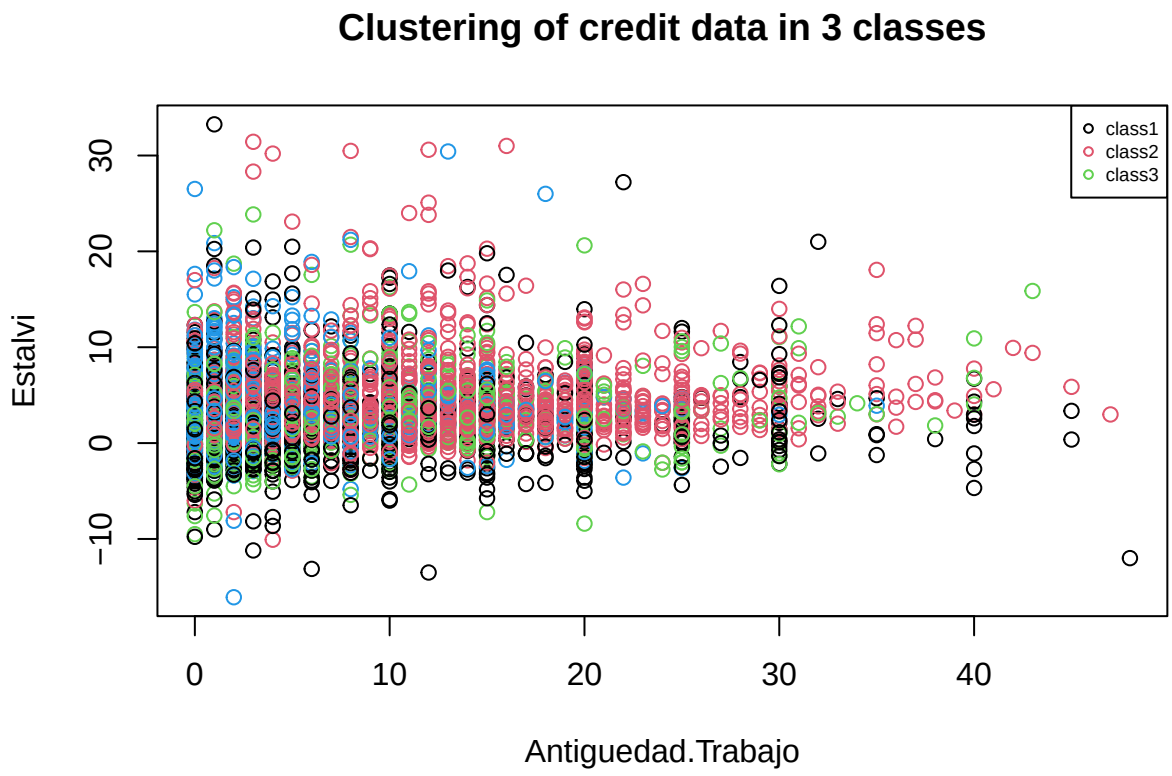

```
plot(RatiFin,Estalvi)
```



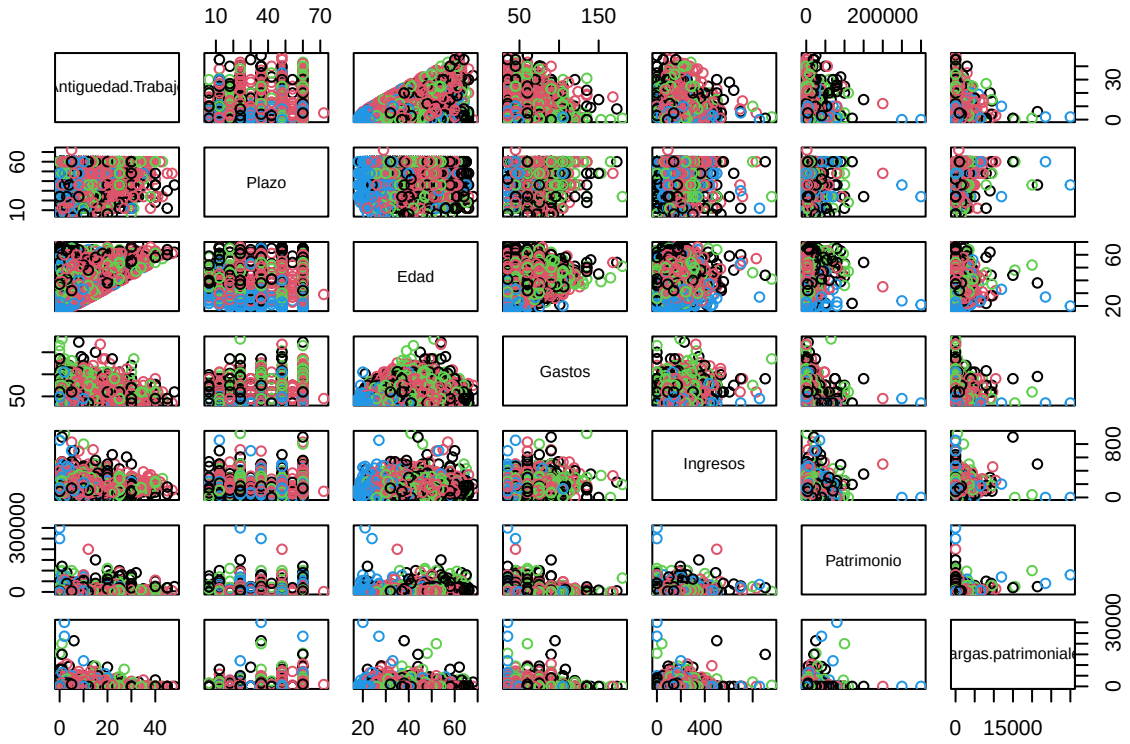

```
plot(RatiFin,Estalvi,col=c1,main="Clustering of credit data in 3 classes")
legend("topright",c("class1","class2","class3"),pch=1,col=c(1:3), cex=0.6)
```




```
plot(Antiguedad.Trabajo,Estalvi,col=c1,main="Clustering of credit data in 3 classes")
legend("topright",c("class1","class2","class3"),pch=1,col=c(1:3), cex=0.6)
```



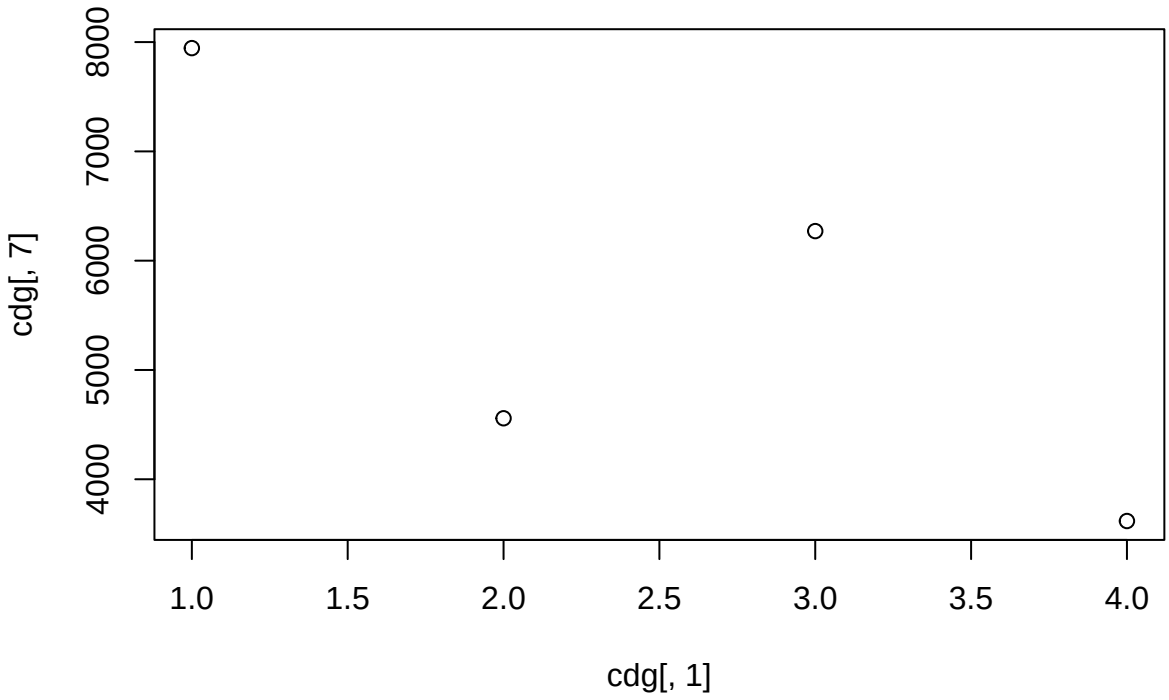

```
pairs(dcon[,1:7], col=c1)
```




```
#plot(FI[,1],FI[,2],col=c1,main="Clustering of credit data in 3 classes")
#legend("topleft",c("c1","c2","c3"),pch=1,col=c(1:3))

# LETS SEE THE QUALITY OF THE HIERARCHICAL PARTITION

cdg <- aggregate(as.data.frame(dcon),list(c1),mean)
cdg
```

```
##   Group.1 Antiguedad.Trabajo    Plazo     Edad   Gastos Ingresos Patrimonio
## 1       1           7.280488 44.29268 41.50000 61.15244 118.0295   7944.875
## 2       2          10.391987 47.87656 38.58148 59.55712 145.4824   4557.923
## 3       3           8.232919 47.61801 40.10093 61.72205 138.0357   6270.693
## 4       4           4.003061 45.12245 27.81633 38.40204 109.7235   3617.989
##   Cargas.patrimoniales Importe.solicitado Precio.del.bien.financiado  Estalvi
## 1              363.9106           1054.9939                   1496.905 2.189703
## 2              363.3844           1017.4781                   1451.941 4.141474
## 3              435.1429           1175.0839                   1618.376 3.205064
## 4              216.7378            974.1755                   1347.130 3.415744
##    RatiFin
## 1 71.98972
## 2 71.74981
## 3 74.44884
## 4 73.64464
```

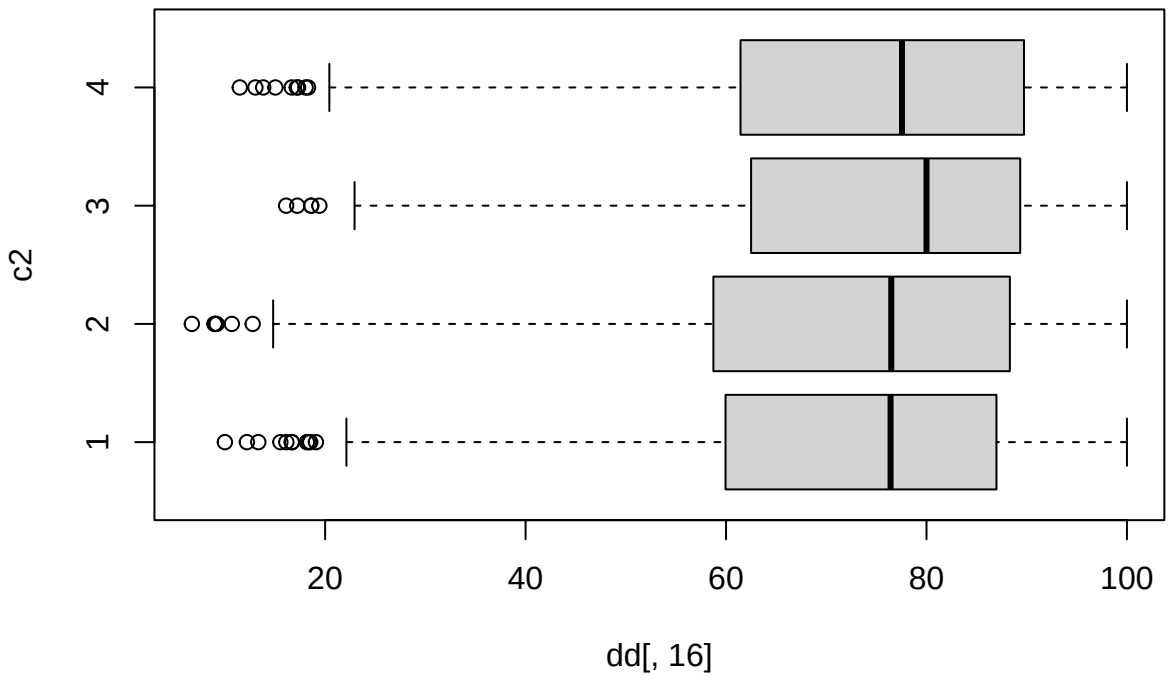
```r
plot(cdg[,1], cdg[,7])
```



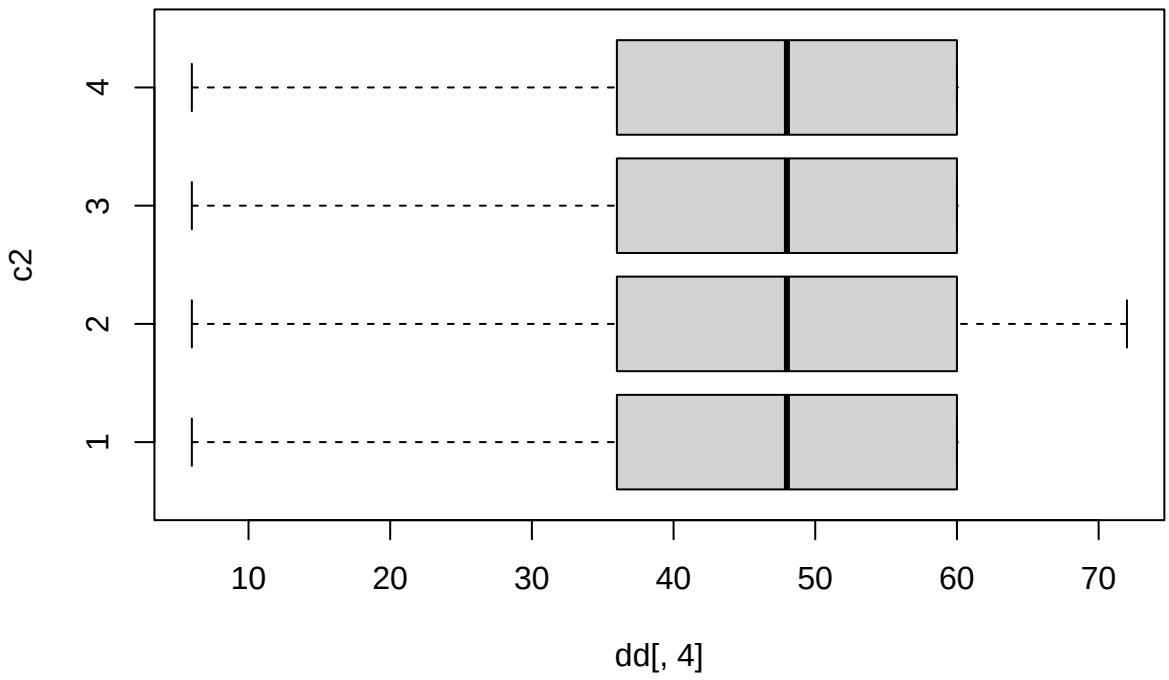

```r
#Tss <- h1$totss
#Bss <- sum(rowSums(cdg^2)*as.numeric(table(c1)))

#Ib4 <- 100*Bss/Tss
#Ib4



names(dd)
```

```
##  [1] "Dictamen"                "Antiguedad.Trabajo"
##  [3] "Vivienda"                "Plazo"
##  [5] "Edad"                    "Estado.civil"
##  [7] "Registros"               "Tipo.trabajo"
##  [9] "Gastos"                  "Ingresos"
## [11] "Patrimonio"              "Cargas.patrimoniales"
## [13] "Importe.solicitado"      "Precio.del.bien.financiado"
## [15] "Estalvi"                 "RatiFin"
```

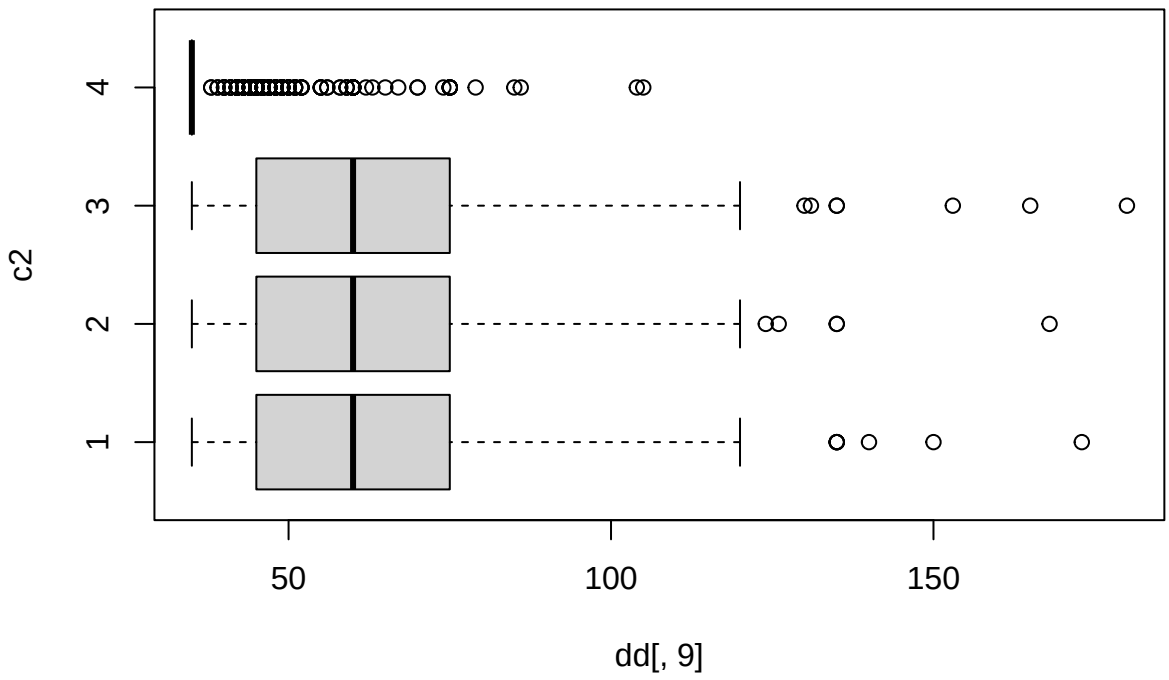
```r
#ratiFin
boxplot(dd[,16]~c2, horizontal=TRUE)
```



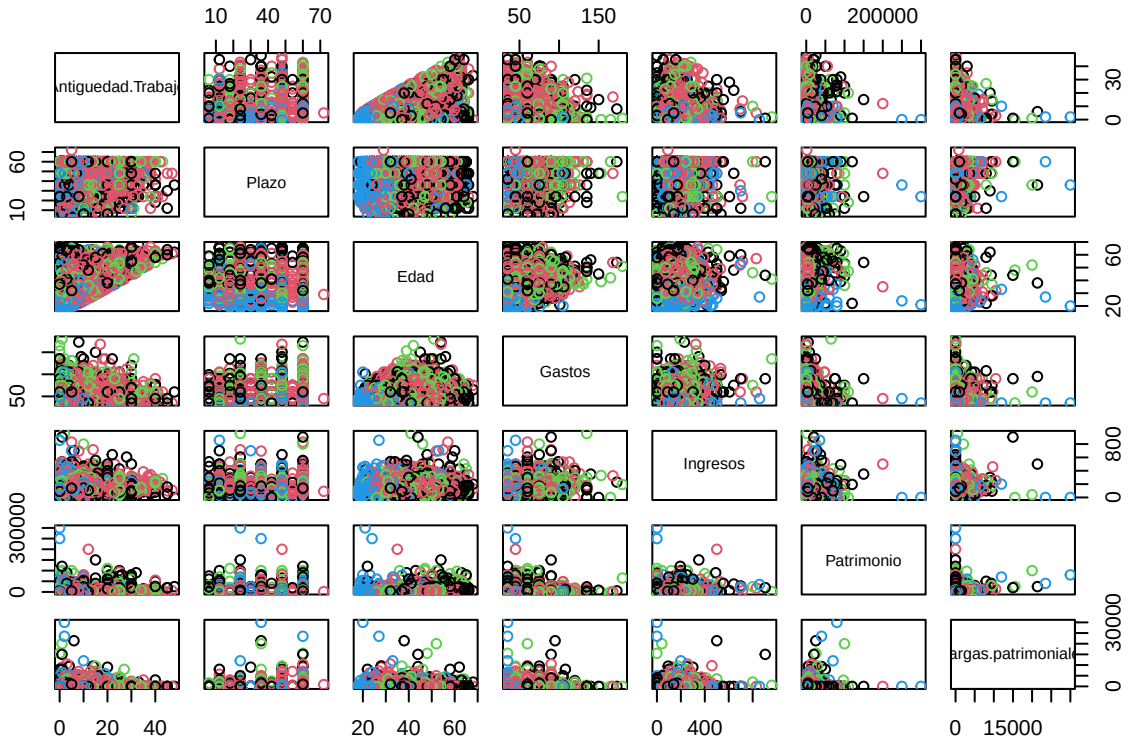

```
#plazo
boxplot(dd[,4]~c2, horizontal=TRUE)
```



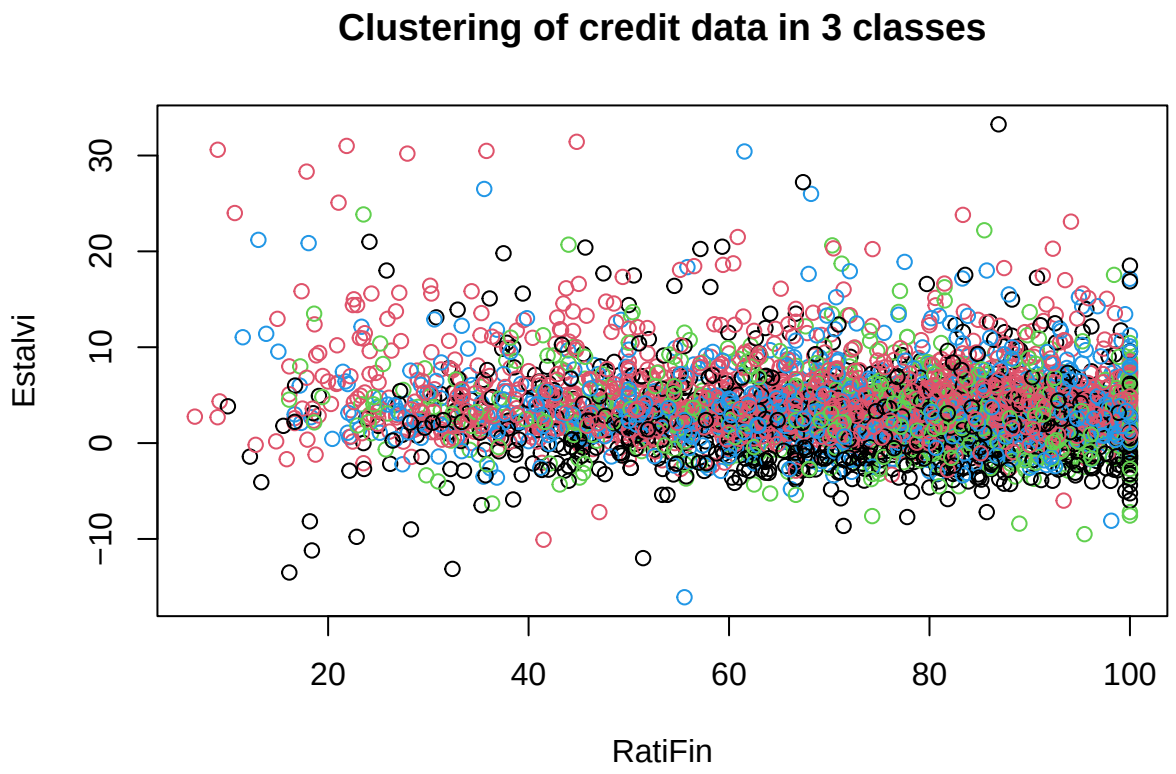

```
#gastos
boxplot(dd[,9]~c2, horizontal=TRUE)
```




```
pairs(dcon[,1:7], col=c2)
```



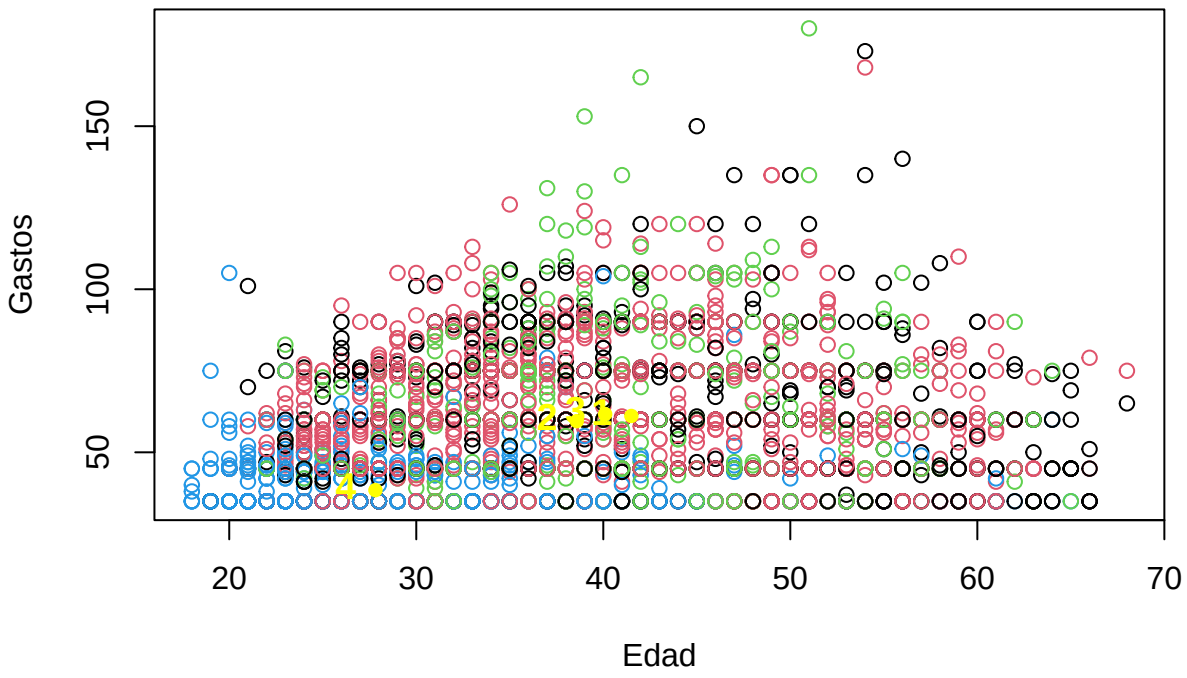

```
plot(RatiFin,Estalvi,col=c2,main="Clustering of credit data in 3 classes")
```



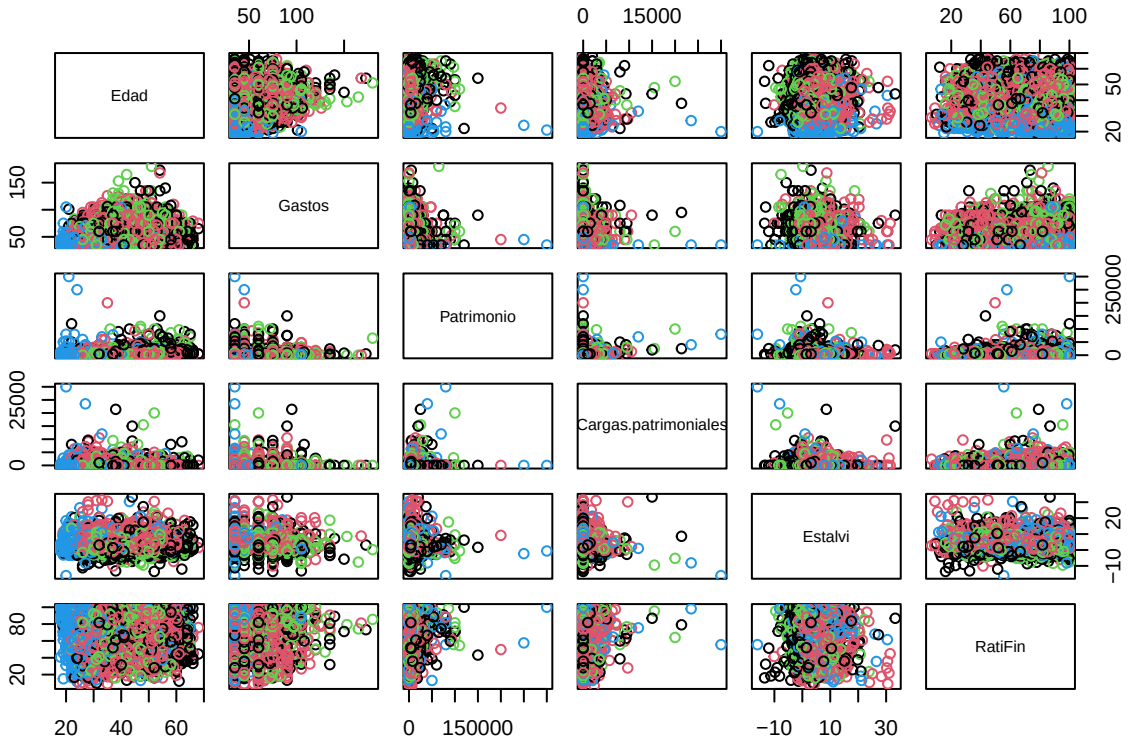

```
#legend("topright",levels(c2),pch=1,col=c(1:4), cex=0.6)

cdg <- aggregate(as.data.frame(dcon),list(c2),mean)
cdg
```

```
##   Group.1 Antiguedad.Trabajo    Plazo     Edad   Gastos Ingresos Patrimonio
## 1       1           7.280488 44.29268 41.50000 61.15244 118.0295   7944.875
## 2       2          10.391987 47.87656 38.58148 59.55712 145.4824   4557.923
## 3       3           8.232919 47.61801 40.10093 61.72205 138.0357   6270.693
## 4       4           4.003061 45.12245 27.81633 38.40204 109.7235   3617.989
##   Cargas.patrimoniales Importe.solicitado Precio.del.bien.financiado  Estalvi
## 1              363.9106           1054.9939                   1496.905 2.189703
## 2              363.3844           1017.4781                   1451.941 4.141474
## 3              435.1429           1175.0839                   1618.376 3.205064
## 4              216.7378            974.1755                   1347.130 3.415744
##    RatiFin
## 1 71.98972
## 2 71.74981
## 3 74.44884
## 4 73.64464
```

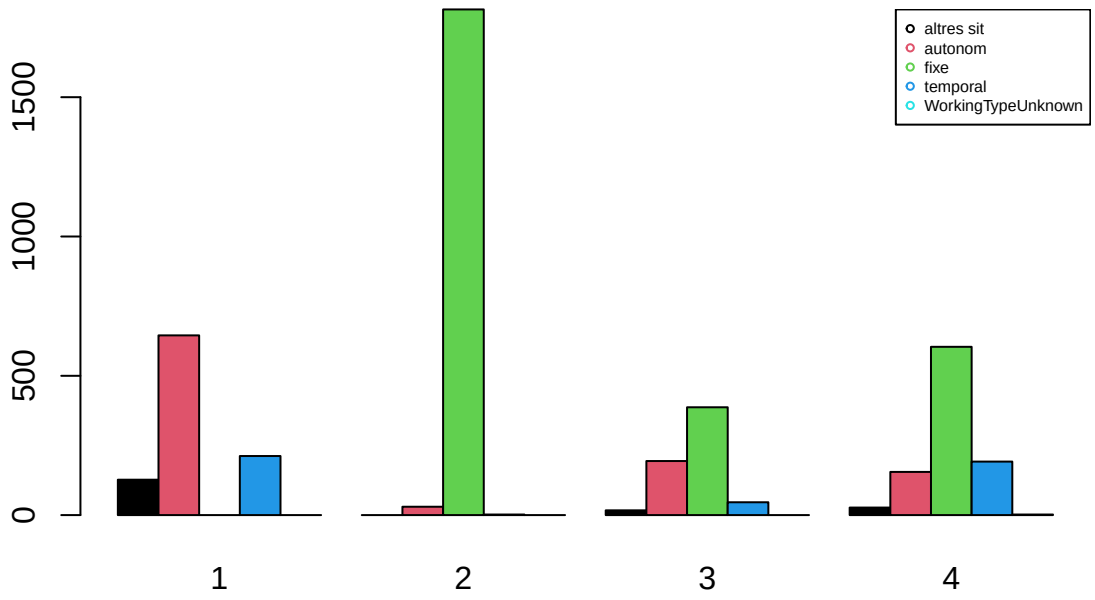
```
plot(Edad, Gastos, col= c2)
points(cdg[,4],cdg[,5],pch=16,col="yellow")
text(cdg[,4],cdg[,5], labels=cdg[,1], pos=2, font=2, cex=1.2, col="yellow")
```



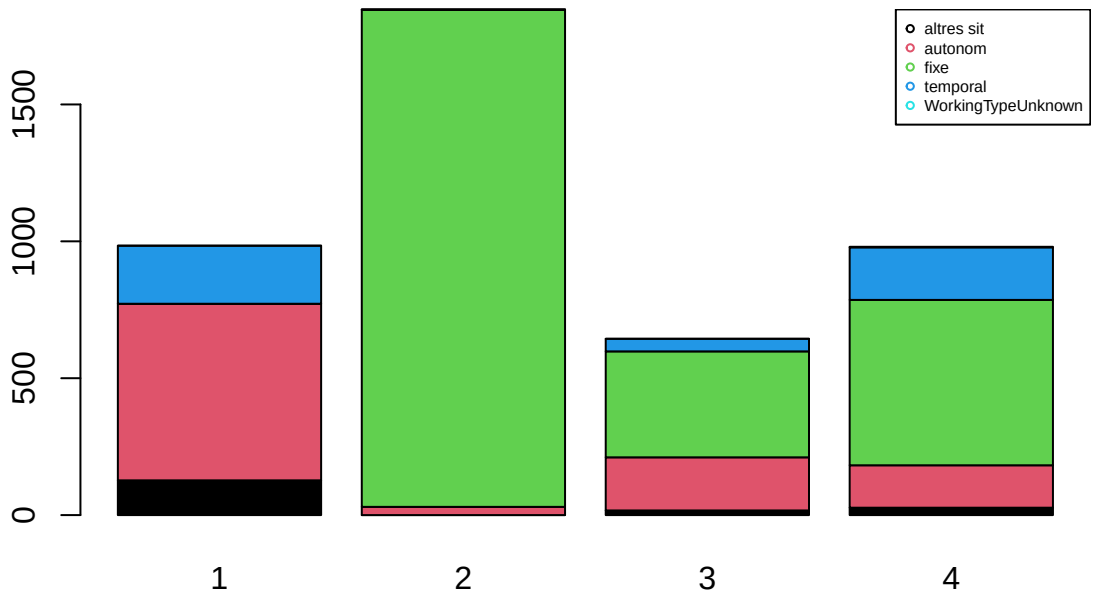

```
potencials<-c(3,4,6,7,10,11)
pairs(dcon[,potencials],col=c2)
```



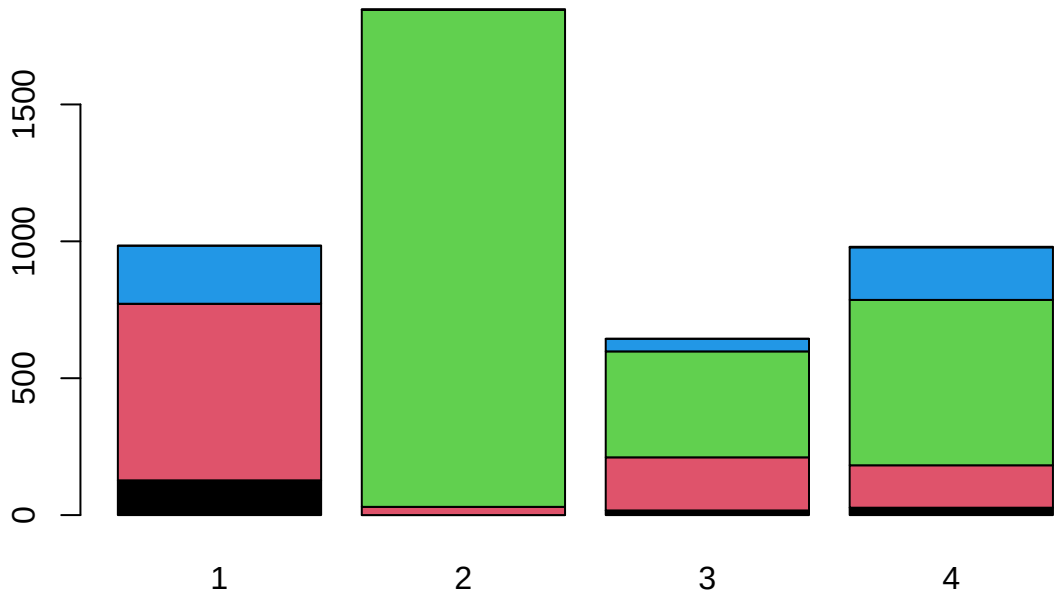

```
tt<-table(Tipo.trabajo,c2)

barplot(table(Tipo.trabajo, c2), beside=TRUE,col=c(1:length(levels(Tipo.trabajo))) )
legend("topright",levels(Tipo.trabajo),pch=1,cex=0.5, col=c(1:length(levels(Tipo.trabajo))))
```




```
barplot(table(Tipo.trabajo, c2), beside=FALSE,col=c(1:length(levels(Tipo.trabajo))) )
legend("topright",levels(Tipo.trabajo),pch=1,cex=0.5, col=c(1:length(levels(Tipo.trabajo))))
```




```
barplot(table(Tipo.trabajo, c2), beside=FALSE,col=c(1:length(levels(Tipo.trabajo))) )
```




```
#Profiling plots
names(dd)
```

```
##  [1] "Dictamen"                 "Antiguedad.Trabajo"
##  [3] "Vivienda"                 "Plazo"
##  [5] "Edad"                     "Estado.civil"
##  [7] "Registros"                "Tipo.trabajo"
##  [9] "Gastos"                   "Ingresos"
## [11] "Patrimonio"               "Cargas.patrimoniales"
## [13] "Importe.solicitado"       "Precio.del.bien.financiado"
## [15] "Estalvi"                  "RatiFin"
```

```
attach(dd)
```

```
## The following objects are masked from dd (pos = 4):
##
##     Antiguedad.Trabajo, Cargas.patrimoniales, Dictamen, Edad,
##     Estado.civil, Estalvi, Gastos, Importe.solicitado, Ingresos,
##     Patrimonio, Plazo, Precio.del.bien.financiado, RatiFin, Registros,
##     Tipo.trabajo, Vivienda
```

```
#Dictamen    <- as.factor(Dictamen)
#levels(Dictamen) <- c(NA, "positiu","negatiu")
```

```r
#Calcula els valor test de la variable Xnum per totes les modalitats del factor P
ValorTestXnum <- function(Xnum,P){
  #freq dis of fac
  nk <- as.vector(table(P));
  n <- sum(nk);
  #mitjanes x grups
  xk <- tapply(Xnum,P,mean);
  #valors test
  txk <- (xk-mean(Xnum))/(sd(Xnum)*sqrt((n-nk)/(n*nk)));
  #p-values
  pxk <- pt(txk,n-1,lower.tail=F);
  for(c in 1:length(levels(as.factor(P)))){if (pxk[c]>0.5){pxk[c]<-1-pxk[c]}}
  return (pxk)
}




ValorTestXquali <- function(P,Xquali){
  taula <- table(P,Xquali);
  n <- sum(taula);
  pk <- apply(taula,1,sum)/n;
  pj <- apply(taula,2,sum)/n;
  pf <- taula/(n*pk);
  pjm <- matrix(data=pj,nrow=dim(pf)[1],ncol=dim(pf)[2], byrow=TRUE);
  dpf <- pf - pjm;
  dvt <- sqrt(((1-pk)/(n*pk))%*%t(pj*(1-pj)));
  #i hi ha divisions iguals a 0 dona NA i no funciona
  zkj <- dpf
  zkj[dpf!=0]<-dpf[dpf!=0]/dvt[dpf!=0];
  pzkj <- pnorm(zkj,lower.tail=F);
  for(c in 1:length(levels(as.factor(P)))){for (s in 1:length(levels(Xquali))){if (pzkj[c,s]> 0.5){pzkj
  return (list(rowpf=pf,vtest=zkj,pval=pzkj))
}


#source("file")
#dades contain the dataset
dades<-dd
#dades<-dd[filtro,]
#dades<-df
K<-dim(dades)[2]
par(ask=TRUE)


#P must contain the class variable
#P<-dd[,3]
P<-c2
#P<-dd[,18]
nameP<-"classe"
#P<-df[,33]

nc<-length(levels(factor(P)))
```

```r
nc
```

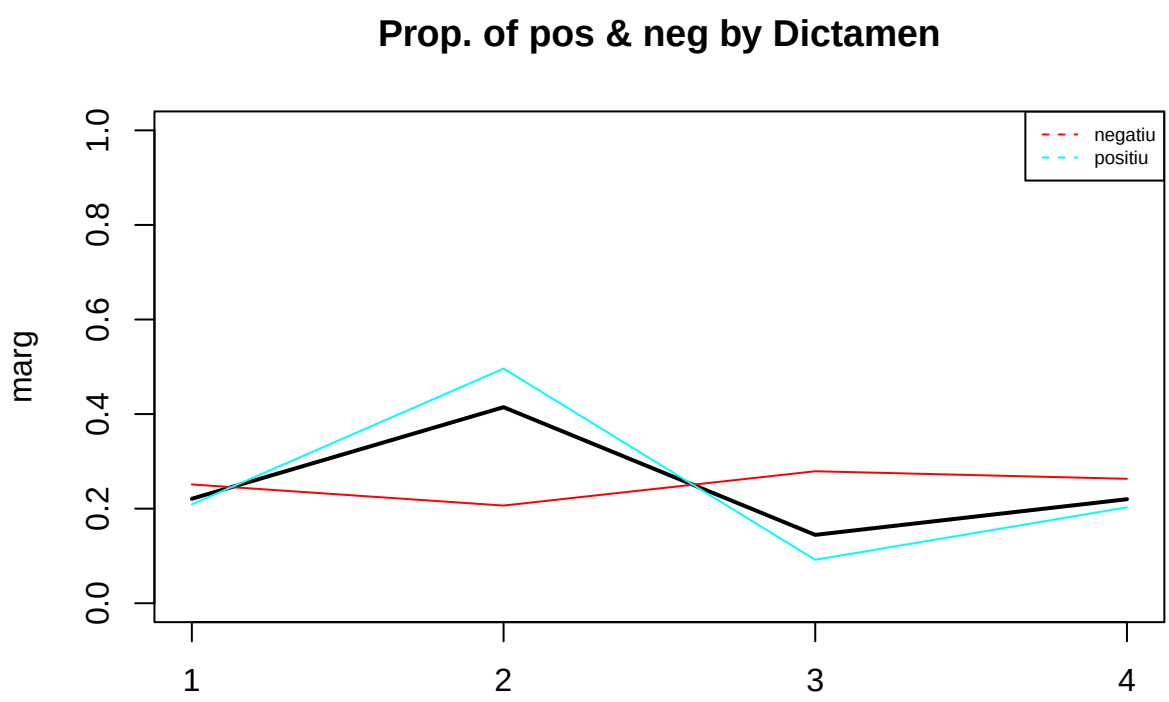
```
## [1] 4
```

```r
pvalk <- matrix(data=0,nrow=nc,ncol=K, dimnames=list(levels(P),names(dades)))
nameP<-"Class"
n<-dim(dades)[1]

for(k in 1:K){
  if (is.numeric(dades[,k])){
    print(paste("Anàlisi per classes de la Variable:", names(dades)[k]))

    boxplot(dades[,k]~P, main=paste("Boxplot of", names(dades)[k], "vs", nameP ), horizontal=TRUE)

    barplot(tapply(dades[[k]], P, mean),main=paste("Means of", names(dades)[k], "by", nameP ))
    abline(h=mean(dades[[k]]))
    legend(0,mean(dades[[k]]),"global mean",bty="n")
    print("Estadístics per groups:")
    for(s in levels(as.factor(P))) {print(summary(dades[P==s,k]))}
    o<-oneway.test(dades[,k]~P)
    print(paste("p-valueANOVA:", o$p.value))
    kw<-kruskal.test(dades[,k]~P)
    print(paste("p-value Kruskal-Wallis:", kw$p.value))
    pvalk[,k]<-ValorTestXnum(dades[,k], P)
    print("p-values ValorsTest: ")
    print(pvalk[,k])
  }else{
    if(class(dd[,k])=="Date"){
      print(summary(dd[,k]))
      print(sd(dd[,k]))
      #decide breaks: weeks, months, quarters...
      hist(dd[,k],breaks="weeks")
    }else{
      #qualitatives
      print(paste("Variable", names(dades)[k]))
      table<-table(P,dades[,k])
      #   print("Cross-table")
      #   print(table)
      rowperc<-prop.table(table,1)

      colperc<-prop.table(table,2)
      #  print("Distribucions condicionades a files")
      # print(rowperc)

      #ojo porque si la variable es true o false la identifica amb el tipus Logical i
      #aquest no te levels, por tanto, coertion preventiva

      dades[,k]<-as.factor(dades[,k])


      marg <- table(as.factor(P))/n
      print(append("Categories=",levels(as.factor(dades[,k]))))
```

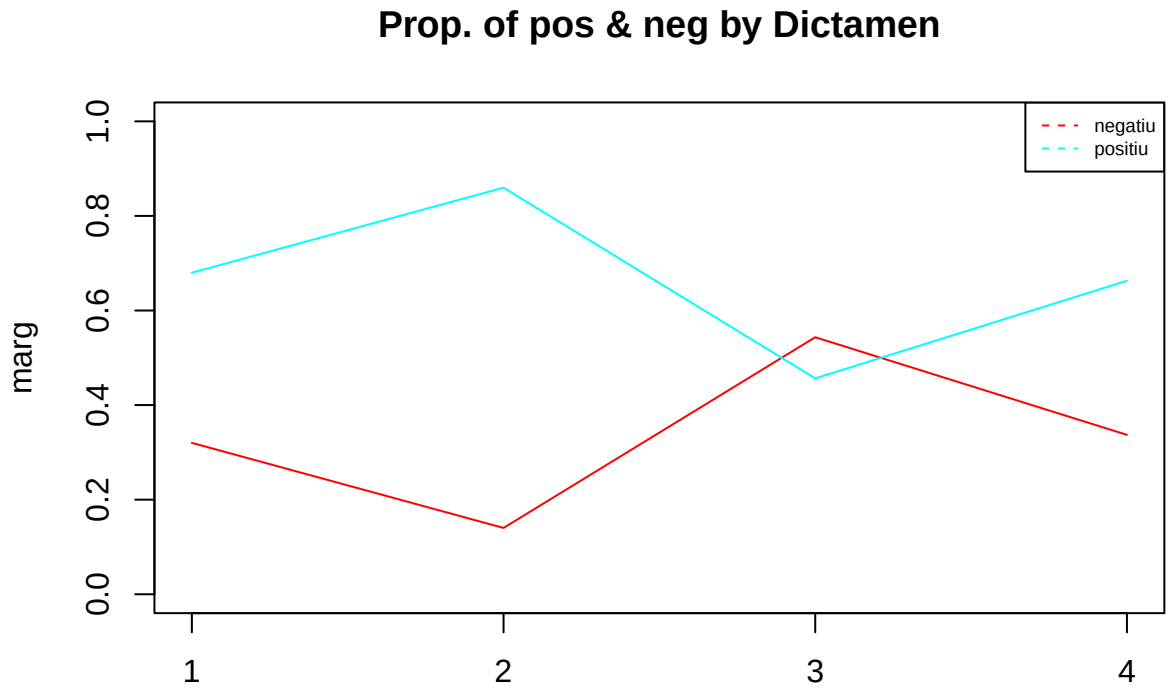
```r
#from next plots, select one of them according to your practical case
plot(marg,type="l",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]))
paleta<-rainbow(length(levels(dades[,k])))
for(c in 1:length(levels(dades[,k]))){lines(colperc[,c],col=paleta[c]) }

#with legend
plot(marg,type="l",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]))
paleta<-rainbow(length(levels(dades[,k])))
for(c in 1:length(levels(dades[,k]))){lines(colperc[,c],col=paleta[c]) }
legend("topright", levels(dades[,k]), col=paleta, lty=2, cex=0.6)

#condicionades a classes
print(append("Categories=",levels(dades[,k])))
plot(marg,type="n",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]))
paleta<-rainbow(length(levels(dades[,k])))
for(c in 1:length(levels(dades[,k]))){lines(rowperc[,c],col=paleta[c]) }

#with legend
plot(marg,type="n",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]))
paleta<-rainbow(length(levels(dades[,k])))
for(c in 1:length(levels(dades[,k]))){lines(rowperc[,c],col=paleta[c]) }
legend("topright", levels(dades[,k]), col=paleta, lty=2, cex=0.6)

#amb variable en eix d'abcisses
marg <-table(dades[,k])/n
print(append("Categories=",levels(dades[,k])))
plot(marg,type="l",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]), las=3)
#x<-plot(marg,type="l",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]), xaxt="n")
#text(x=x+.25, y=-1, adj=1, levels(CountryName), xpd=TRUE, srt=25, cex=0.7)
paleta<-rainbow(length(levels(as.factor(P))))
for(c in 1:length(levels(as.factor(P)))){lines(rowperc[c,],col=paleta[c]) }

#with legend
plot(marg,type="l",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]), las=3)
for(c in 1:length(levels(as.factor(P)))){lines(rowperc[c,],col=paleta[c])}
legend("topright", levels(as.factor(P)), col=paleta, lty=2, cex=0.6)

#condicionades a columna
plot(marg,type="n",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]), las=3)
paleta<-rainbow(length(levels(as.factor(P))))
for(c in 1:length(levels(as.factor(P)))){lines(colperc[c,],col=paleta[c]) }

#with legend
plot(marg,type="n",ylim=c(0,1),main=paste("Prop. of pos & neg by",names(dades)[k]), las=3)
for(c in 1:length(levels(as.factor(P)))){lines(colperc[c,],col=paleta[c])}
legend("topright", levels(as.factor(P)), col=paleta, lty=2, cex=0.6)

table<-table(dades[,k],P)
print("Cross Table:")
print(table)
print("Distribucions condicionades a columnes:")
print(colperc)
```

```
      #diagrames de barres apilades

      paleta<-rainbow(length(levels(dades[,k])))
      barplot(table(dades[,k], as.factor(P)), beside=FALSE,col=paleta )

      barplot(table(dades[,k], as.factor(P)), beside=FALSE,col=paleta )
      legend("topright",levels(as.factor(dades[,k])),pch=1,cex=0.5, col=paleta)

      #diagrames de barres adosades
      barplot(table(dades[,k], as.factor(P)), beside=TRUE,col=paleta )

      barplot(table(dades[,k], as.factor(P)), beside=TRUE,col=paleta)
      legend("topright",levels(as.factor(dades[,k])),pch=1,cex=0.5, col=paleta)

      print("Test Chi quadrat: ")
      print(chisq.test(dades[,k], as.factor(P)))

      print("valorsTest:")
      print( ValorTestXquali(P,dades[,k]))
      #calcular els pvalues de les quali
    }
  }
}#endfor
```

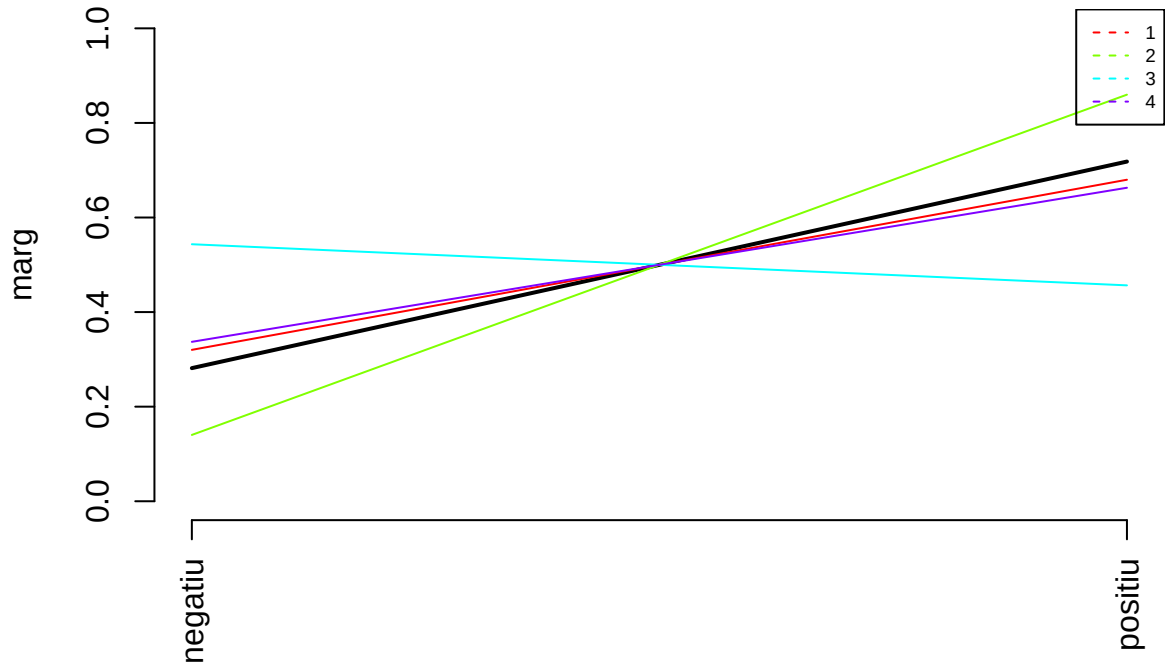
```
## [1] "Variable Dictamen"
## [1] "Categories=" "negatiu"     "positiu"
```

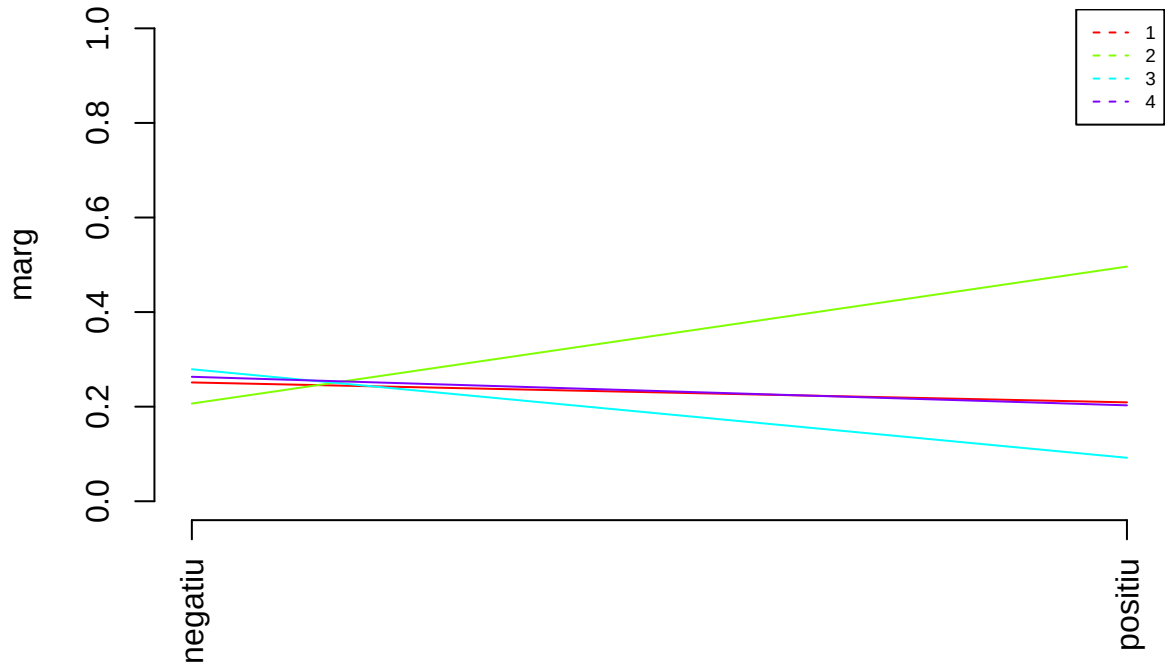



```
## [1] "Categories=" "negatiu"     "positiu"
```

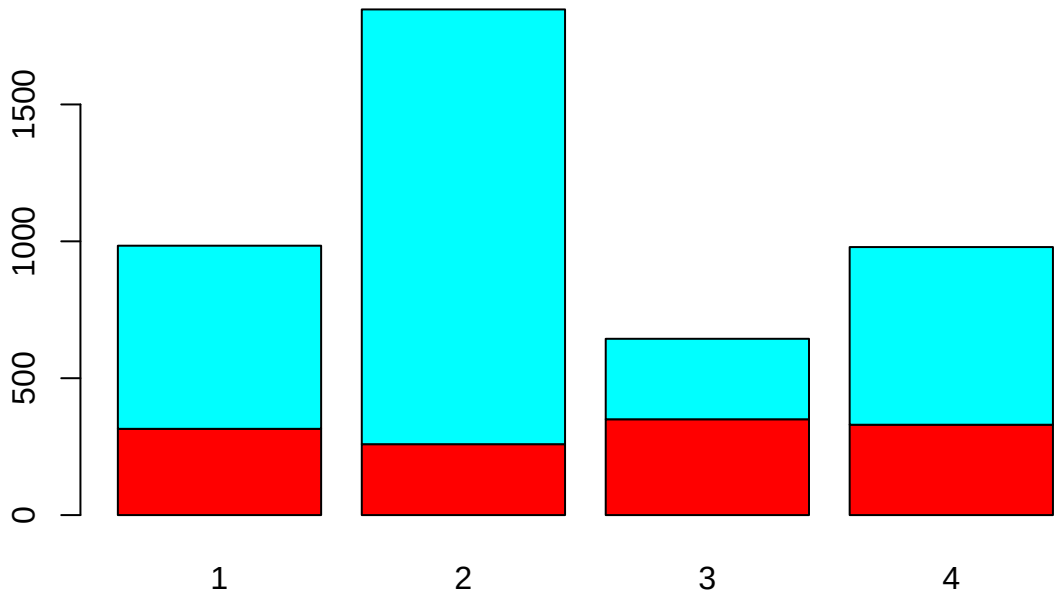


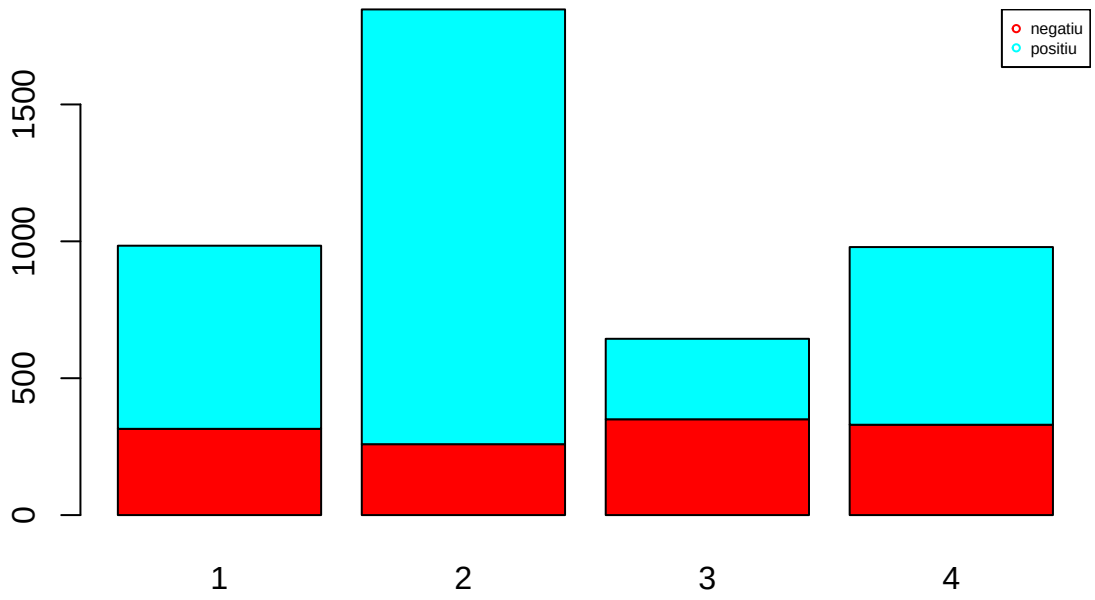

```
## [1] "Categories=" "negatiu"     "positiu"
```

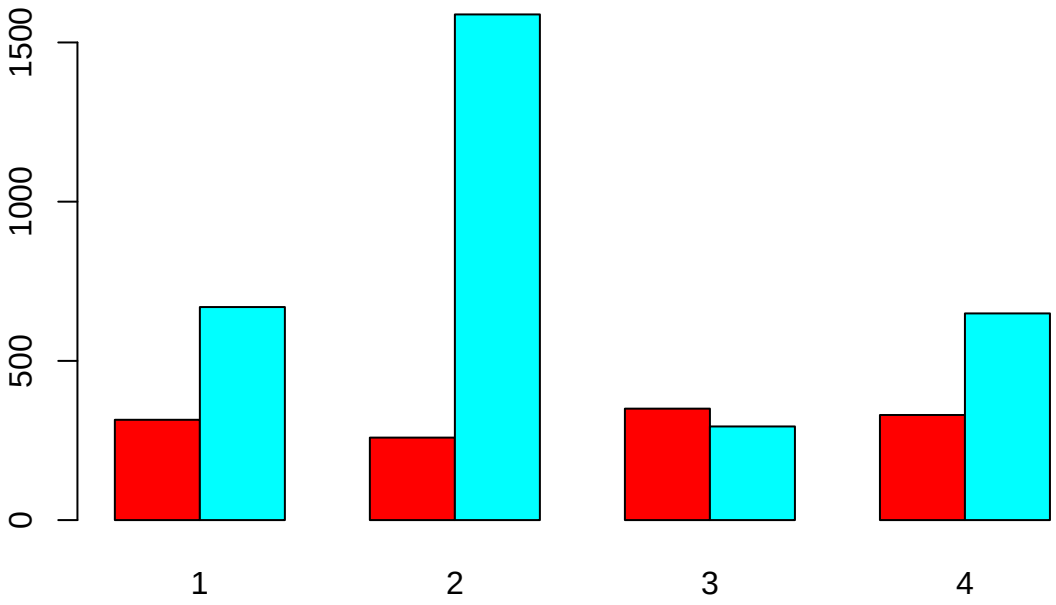

Prop. of pos & neg by Dictamen
marg
0.0
0.2
0.4
0.6
0.8
1.0
negatiu
positiu
1
2
3
4

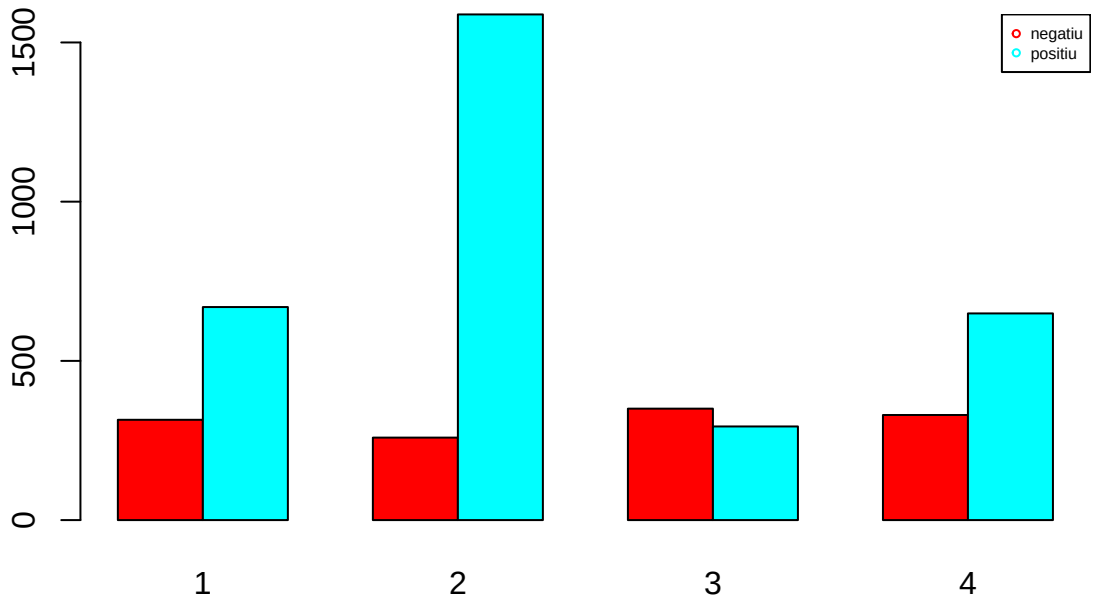



```
## [1] "Cross Table:"
##          P
##              1    2    3    4
##   negatiu  315  259  350  330
##   positiu  669 1588  294  649
## [1] "Distribucions condicionades a columnes:"
##
## P     negatiu   positiu
##   1 0.2511962 0.2090625
##   2 0.2065391 0.4962500
##   3 0.2791069 0.0918750
##   4 0.2631579 0.2028125
```


0
500
1000
1500
1
2
3
4

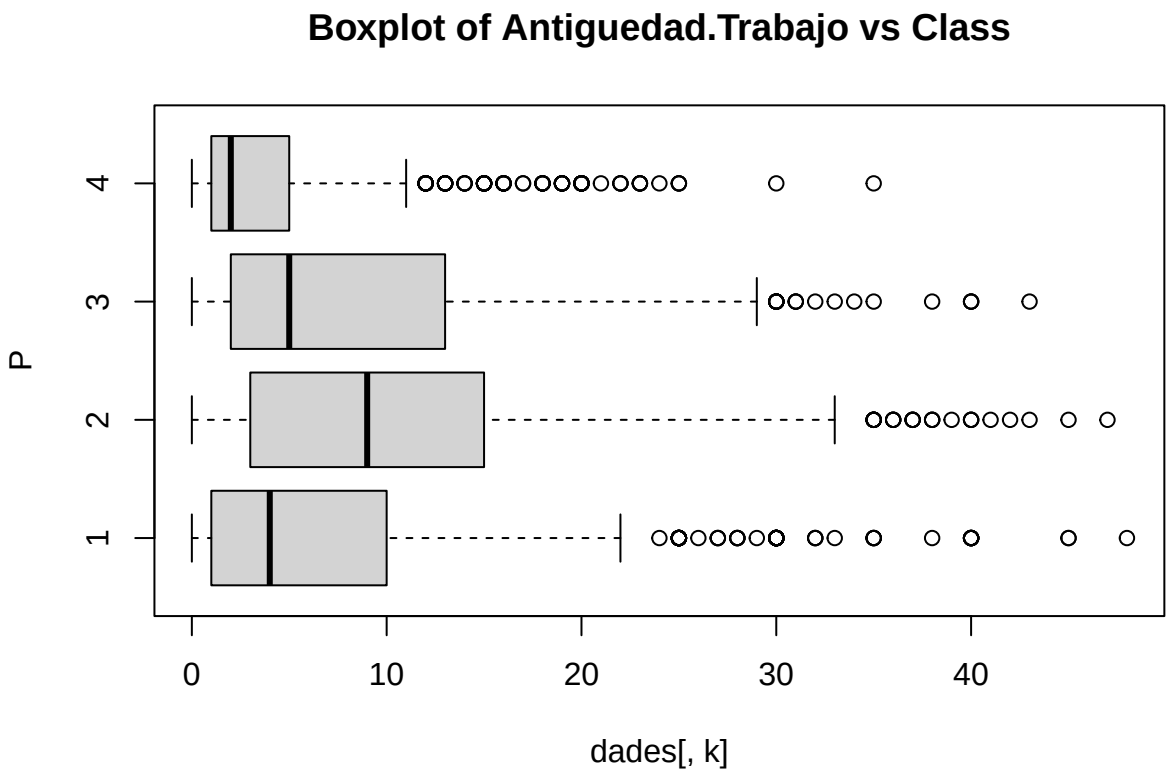

negatiu
positiu
0
500
1000
1500
1
2
3
4

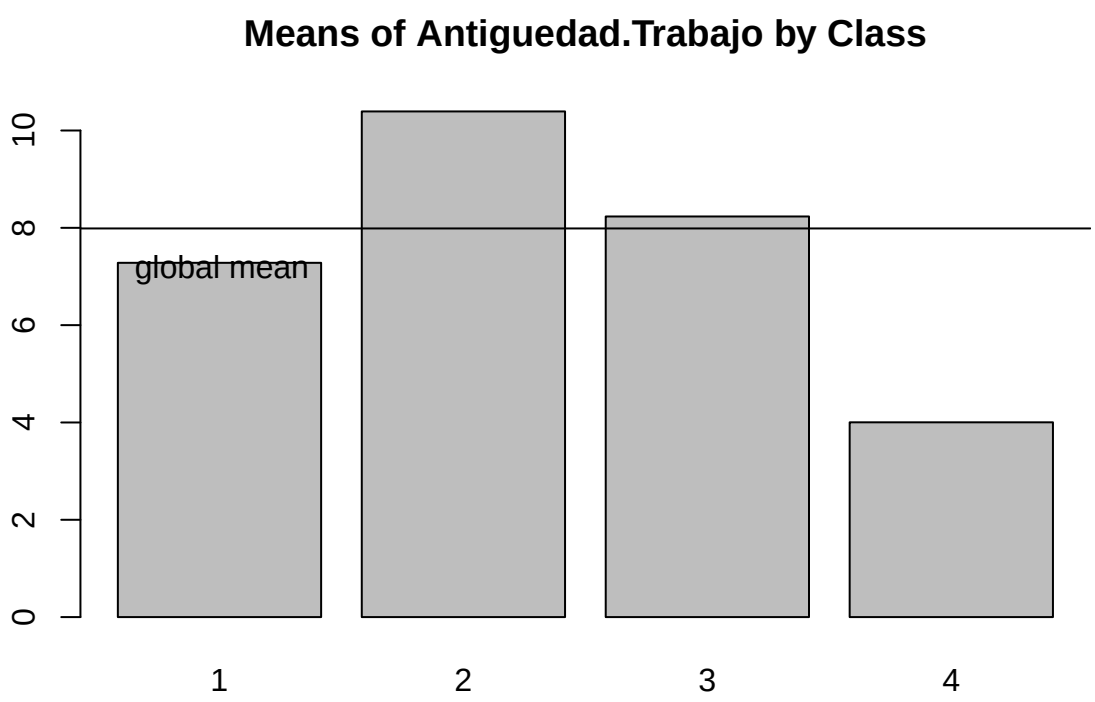

0
500
1000
1500
1
2
3
4



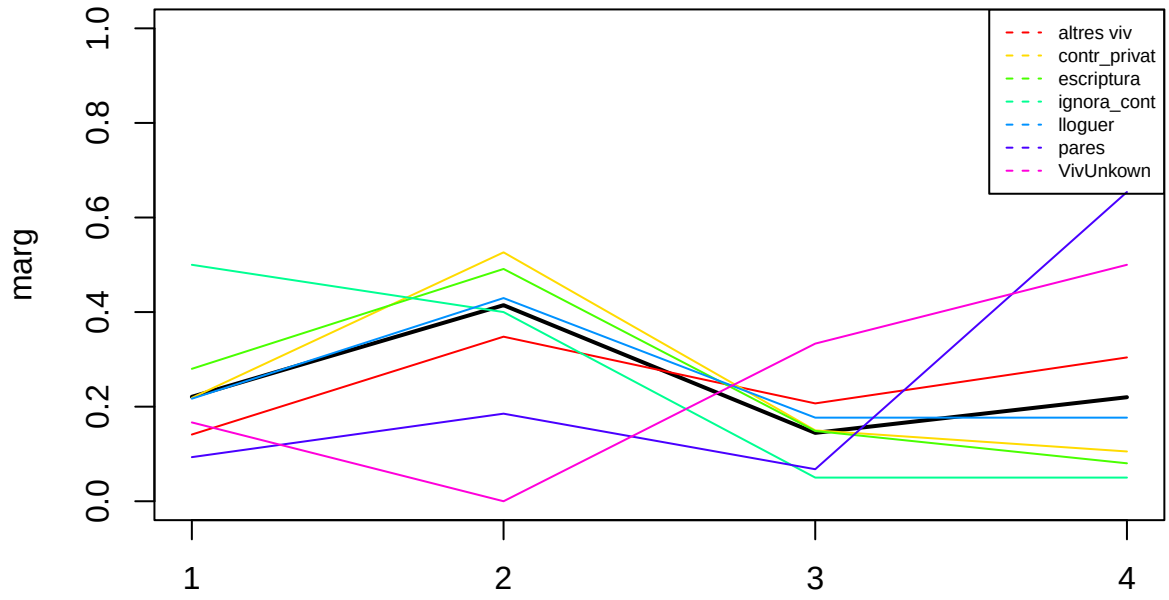

```
## [1] "Test Chi quadrat: "
##
##  Pearson's Chi-squared test
##
## data:  dades[, k] and as.factor(P)
## X-squared = 422.95, df = 3, p-value < 2.2e-16
##
## [1] "valorsTest:"
## $rowpf
##    Xquali
## P     negatiu   positiu
##   1 0.3201220 0.6798780
##   2 0.1402274 0.8597726
##   3 0.5434783 0.4565217
##   4 0.3370787 0.6629213
##
## $vtest
##    Xquali
## P      negatiu    positiu
##   1   3.048357  -3.048357
##   2 -17.650612  17.650612
##   3  15.979868 -15.979868
##   4   4.373954  -4.373954
##
## $pval
##    Xquali
```

```
## P        negatiu      positiu
##   1 1.150480e-03 1.150480e-03
##   2 0.000000e+00 5.032956e-70
##   3 8.825956e-58 0.000000e+00
##   4 6.100795e-06 6.100795e-06
##
## [1] "Anàlisi per classes de la Variable: Antiguedad.Trabajo"
```

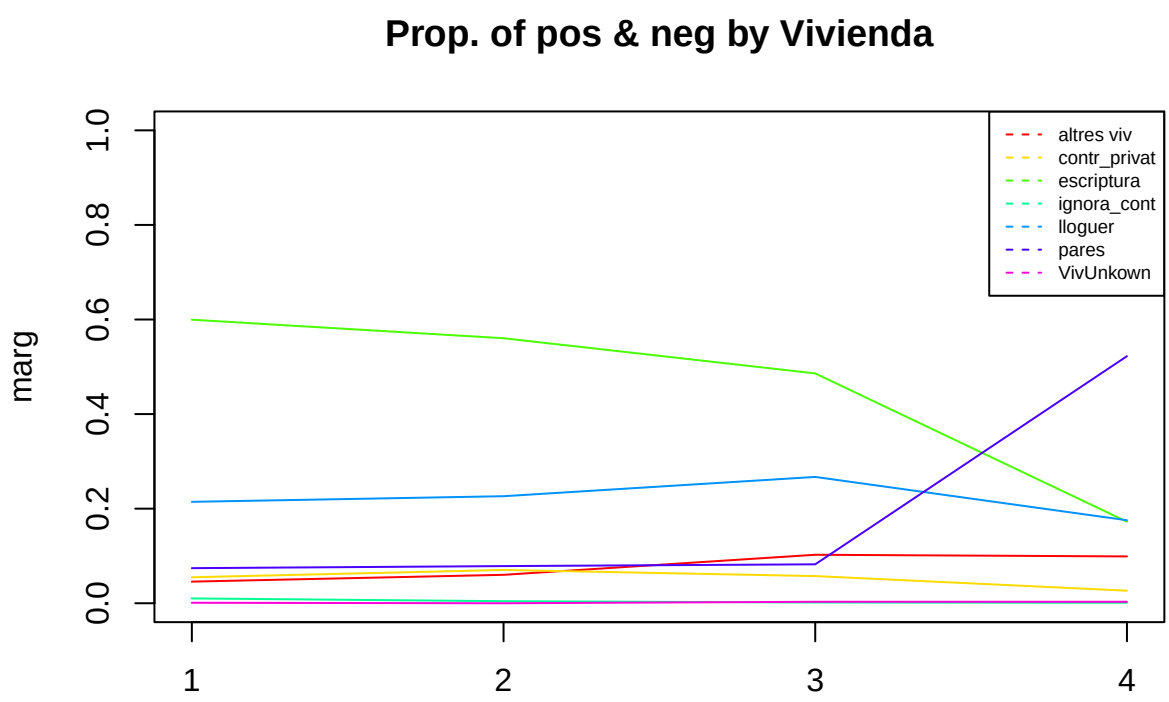




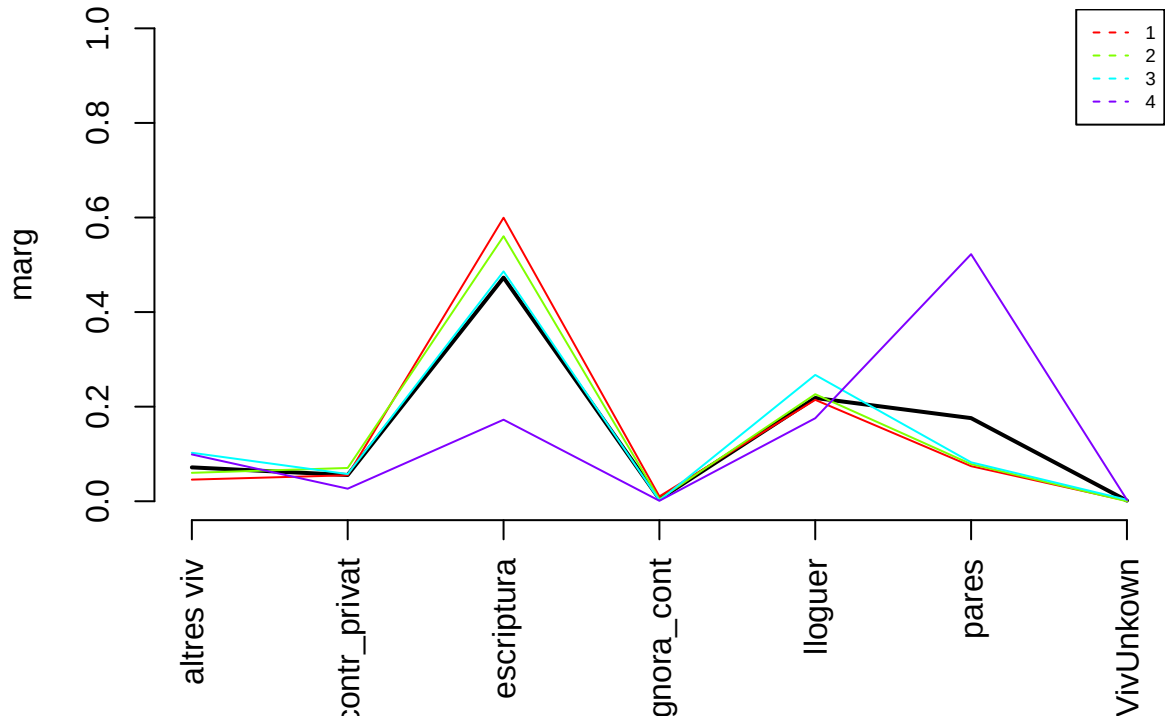

```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    1.00    4.00    7.28   10.00   48.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    3.00    9.00   10.39   15.00   47.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   2.000   5.000   8.233  13.000  43.000
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   1.000   2.000   4.003   5.000  35.000
## [1] "p-valueANOVA: 3.51883960322038e-125"
## [1] "p-value Kruskal-Wallis: 3.21800259830286e-118"
## [1] "p-values ValorsTest: "
## [1] 1.066696e-03 6.835670e-60 2.047521e-01 0.000000e+00
## [1] "Variable Vivienda"
## [1] "Categories="  "altres viv"   "contr_privat" "escriptura"   "ignora_cont"
## [6] "lloguer"      "pares"        "VivUnkown"
```

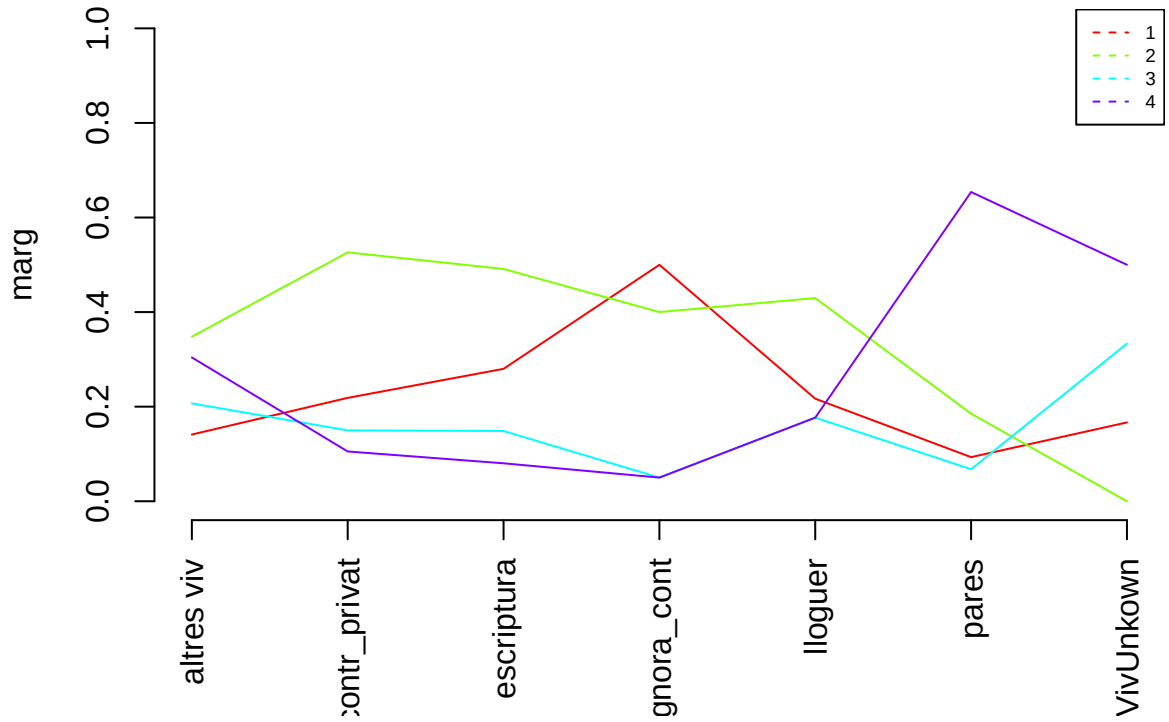



```
## [1] "Categories="  "altres viv"   "contr_privat" "escriptura"   "ignora_cont"
## [6] "lloguer"      "pares"        "VivUnkown"
```

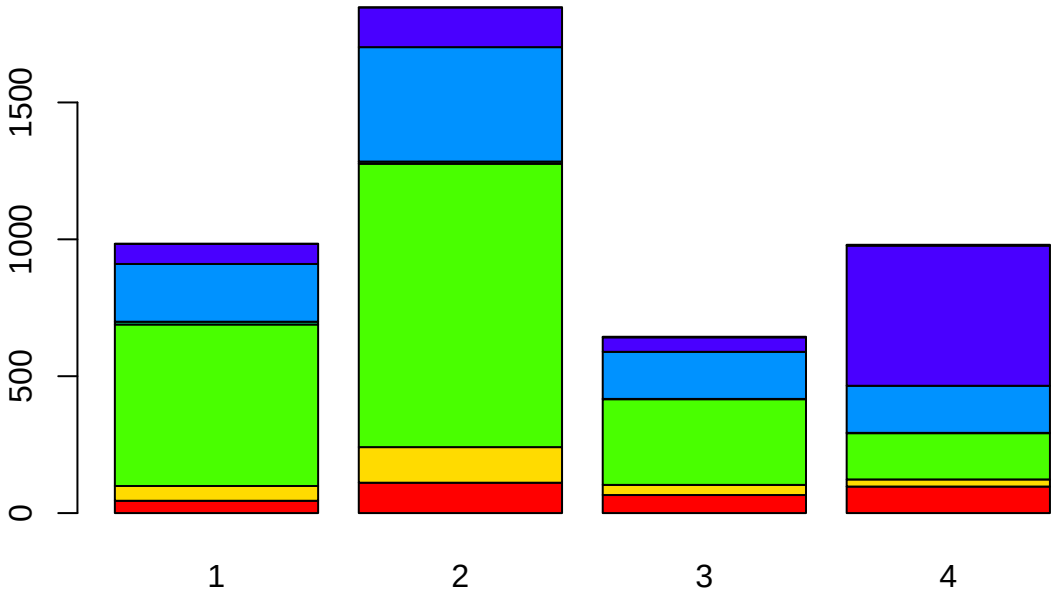


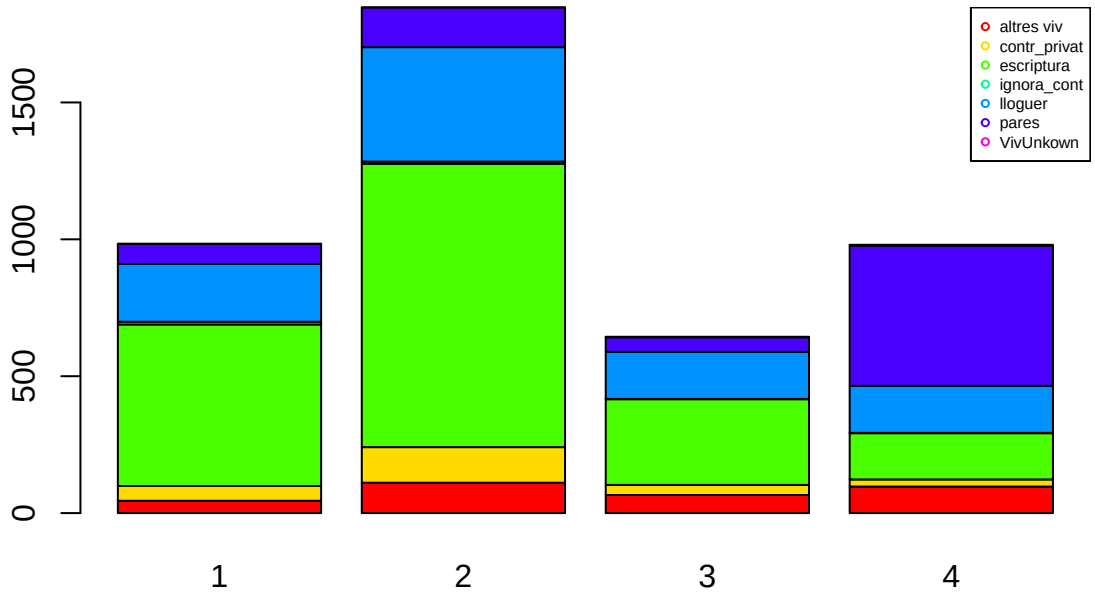

```
## [1] "Categories="  "altres viv"   "contr_privat" "escriptura"   "ignora_cont"
## [6] "lloguer"      "pares"        "VivUnkown"
```

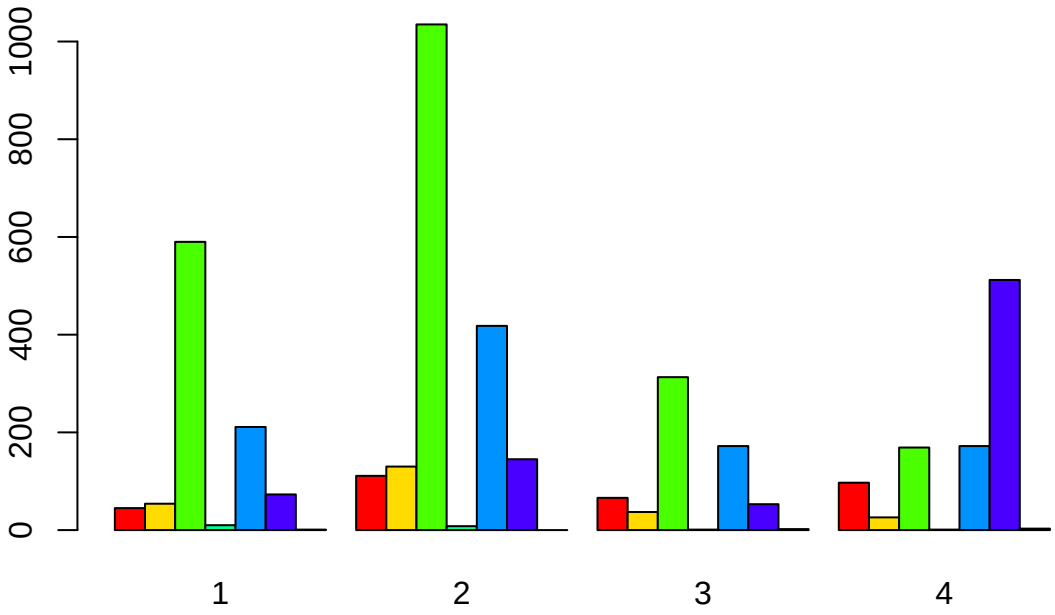

Prop. of pos & neg by Vivienda
marg
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
altres viv
contr_privat
escriptura
gnora_cont
lloguer
pares
VivUnkown

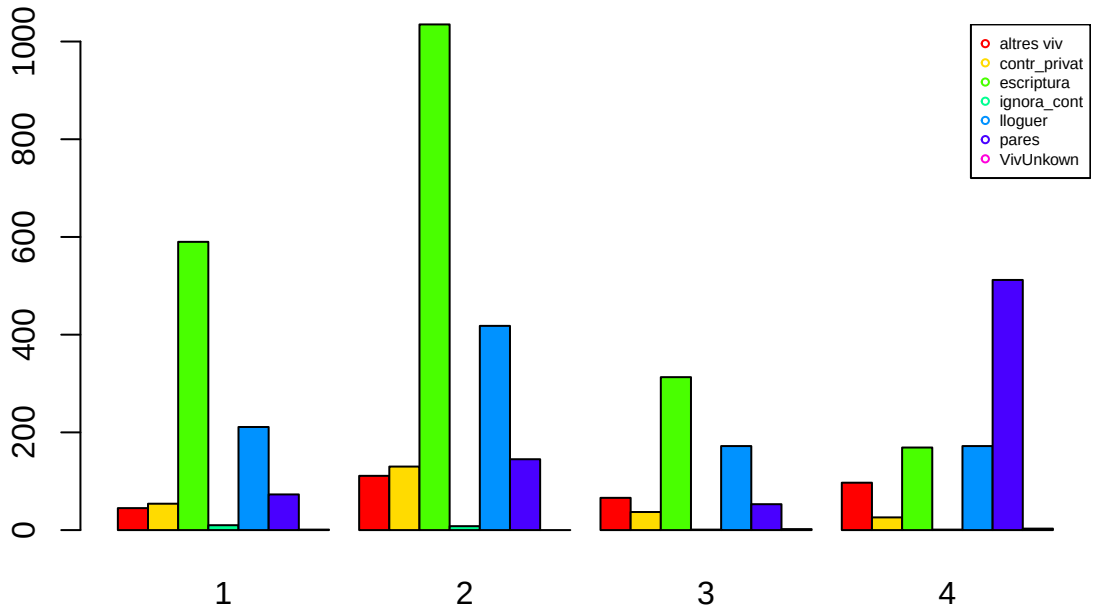



```
## [1] "Cross Table:"
##                P
##                    1    2    3    4
##   altres viv      45  111   66   97
##   contr_privat    54  130   37   26
##   escriptura     590 1035  313  169
##   ignora_cont     10    8    1    1
##   lloguer        211  418  172  172
##   pares           73  145   53  512
##   VivUnkown        1    0    2    3
## [1] "Distribucions condicionades a columnes:"
##
## P   altres viv contr_privat escriptura ignora_cont    lloguer      pares
##   1 0.14106583   0.21862348 0.28001898  0.50000000 0.21685509 0.09323116
##   2 0.34796238   0.52631579 0.49121974  0.40000000 0.42959918 0.18518519
##   3 0.20689655   0.14979757 0.14855244  0.05000000 0.17677287 0.06768838
##   4 0.30407524   0.10526316 0.08020883  0.05000000 0.17677287 0.65389527
##
## P    VivUnkown
##   1 0.16666667
##   2 0.00000000
##   3 0.33333333
##   4 0.50000000
```


0
500
1000
1500
1
2
3
4


altres viv
contr_privat
escriptura
ignora_cont
lloguer
pares
VivUnkown
0
500
1000
1500
1
2
3
4

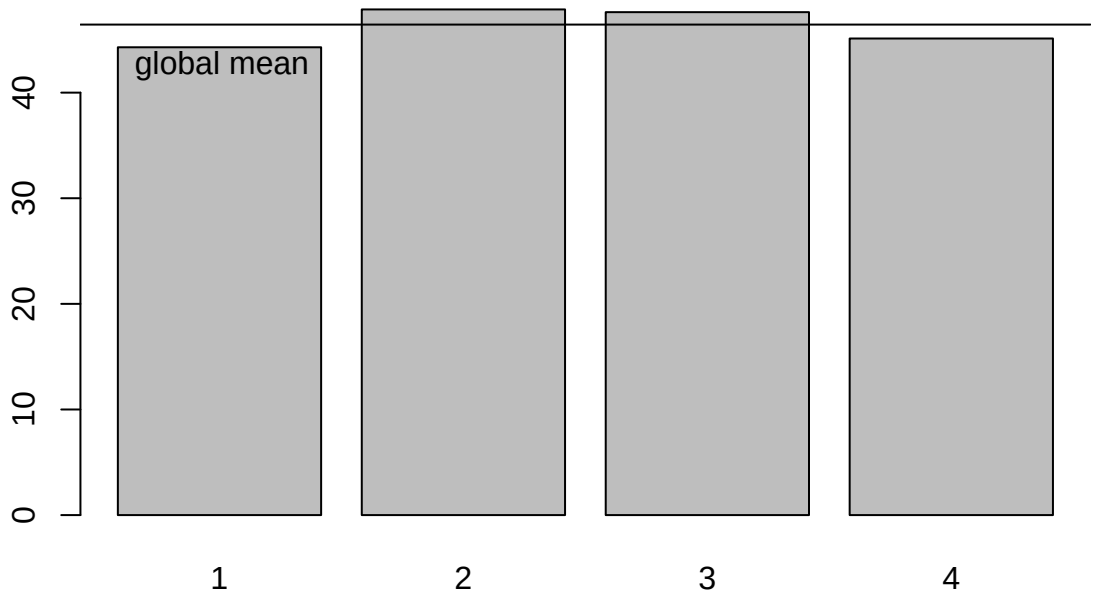

0
200
400
600
800
1000
1
2
3
4



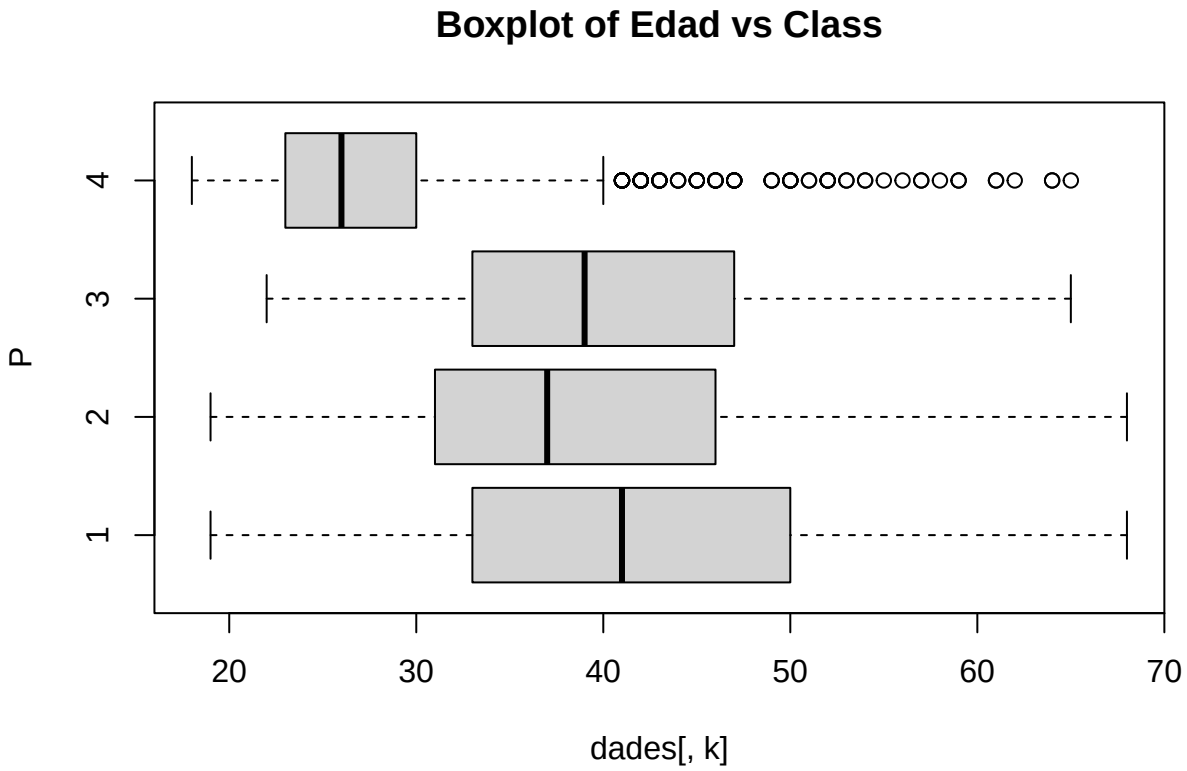

```
## [1] "Test Chi quadrat: "

## Warning in chisq.test(dades[, k], as.factor(P)): Chi-squared approximation may
## be incorrect

##
##  Pearson's Chi-squared test
##
## data:  dades[, k] and as.factor(P)
## X-squared = 1196.6, df = 18, p-value < 2.2e-16
##
## [1] "valorsTest:"
## $rowpf
##    Xquali
## P    altres viv contr_privat  escriptura ignora_cont     lloguer       pares
##   1 0.045731707  0.054878049 0.599593496 0.010162602 0.214430894 0.074186992
##   2 0.060097455  0.070384407 0.560368165 0.004331348 0.226312940 0.078505685
##   3 0.102484472  0.057453416 0.486024845 0.001552795 0.267080745 0.082298137
##   4 0.098979592  0.026530612 0.172448980 0.001020408 0.175510204 0.522448980
##    Xquali
## P     VivUnkown
##   1 0.001016260
##   2 0.000000000
##   3 0.003105590
##   4 0.003061224
```

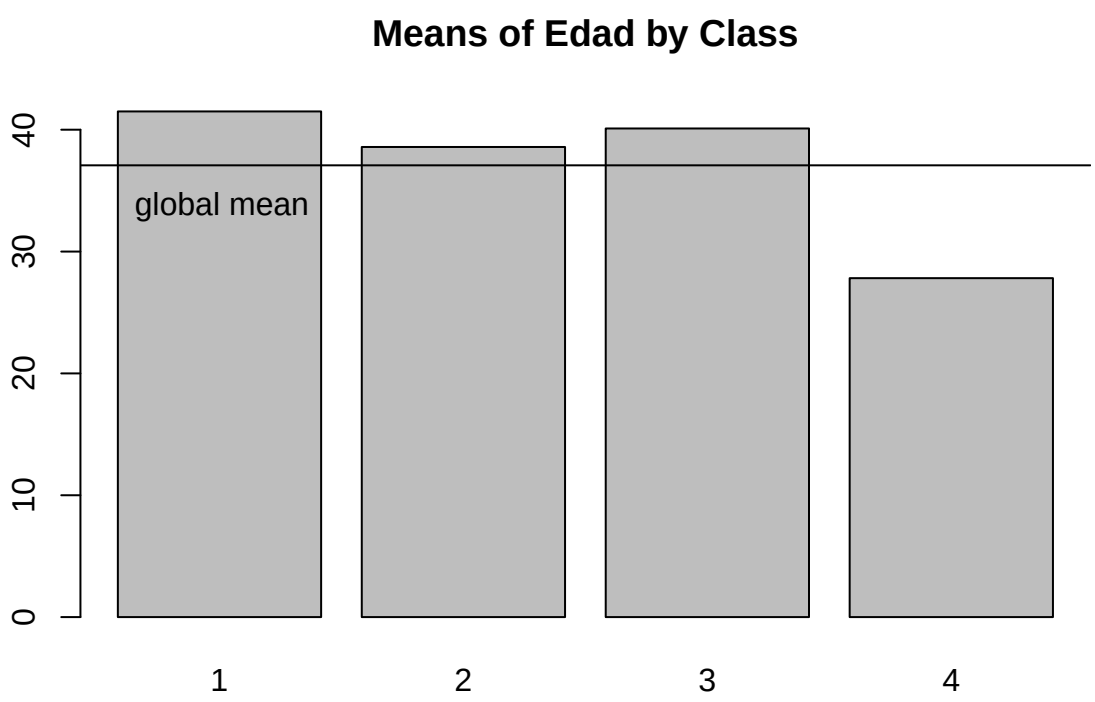
```
##
## $vtest
##    Xquali
## P     altres viv contr_privat   escriptura  ignora_cont      lloguer
##   1  -3.56620959  -0.08778372   9.01441225   3.01587247  -0.34194072
##   2  -2.50695160   3.66730484   9.83474138  -0.13274499   1.07491270
##   3   3.28609527   0.24100445   0.71844324  -1.20522892   3.23240197
##   4   3.76331196  -4.47826358 -21.33410101  -1.83925049  -3.68005245
##    Xquali
## P          pares    VivUnkown
##   1  -9.48369409  -0.32030267
##   2 -14.35215293  -2.06275634
##   3  -6.73730138   1.31583630
##   4  32.28628322   1.65698428
##
## $pval
##    Xquali
## P      altres viv  contr_privat    escriptura   ignora_cont       lloguer
##   1  1.810909e-04  4.650243e-01  9.896446e-20  1.281205e-03  3.661978e-01
##   2  6.088869e-03  1.225603e-04  3.988615e-23  4.471975e-01  1.412069e-01
##   3  5.079332e-04  4.047758e-01  2.362420e-01  1.140575e-01  6.137712e-04
##   4  8.383891e-05  3.762632e-06  0.000000e+00  3.293918e-02  1.165930e-04
##    Xquali
## P           pares     VivUnkown
##   1  0.000000e+00  3.743695e-01
##   2  0.000000e+00  1.956789e-02
##   3  8.067769e-12  9.411450e-02
##   4 5.448858e-229  4.876133e-02
##
## [1] "Anàlisi per classes de la Variable: Plazo"
```


Boxplot of Plazo vs Class
P
4
3
2
1
10
20
30
40
50
60
70
dades[, k]

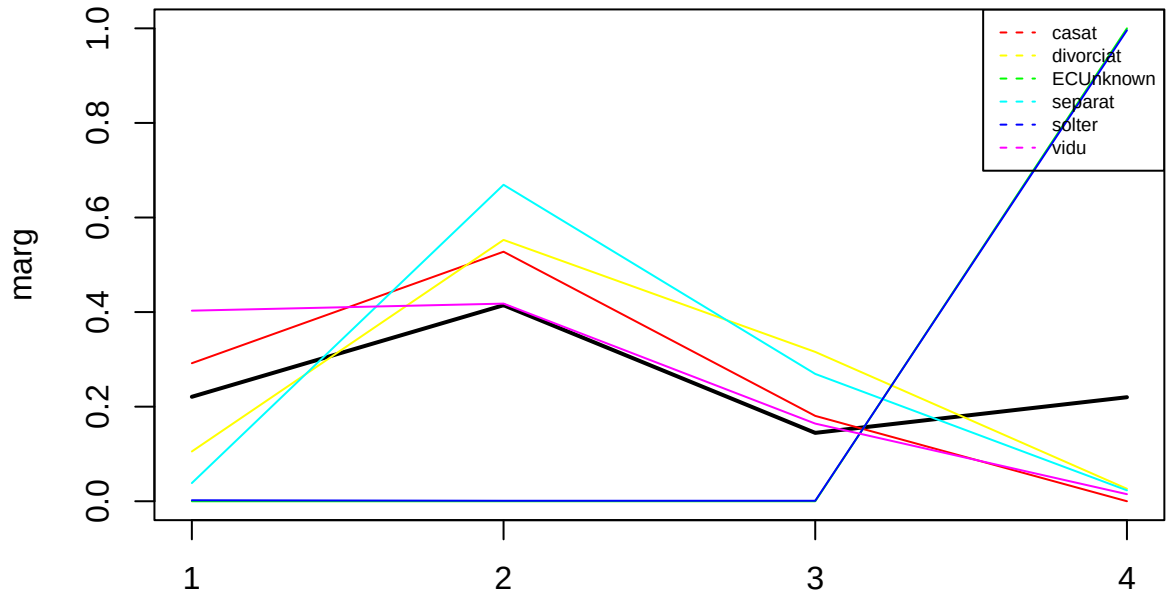



```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    6.00   36.00   48.00   44.29   60.00   60.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    6.00   36.00   48.00   47.88   60.00   72.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    6.00   36.00   48.00   47.62   60.00   60.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    6.00   36.00   48.00   45.12   60.00   60.00
## [1] "p-valueANOVA: 1.0026716586696e-10"
## [1] "p-value Kruskal-Wallis: 8.84196879467652e-11"
## [1] "p-values ValorsTest: "
## [1] 9.796502e-08 2.013658e-08 1.385077e-02 7.141138e-04
## [1] "Anàlisi per classes de la Variable: Edad"
```

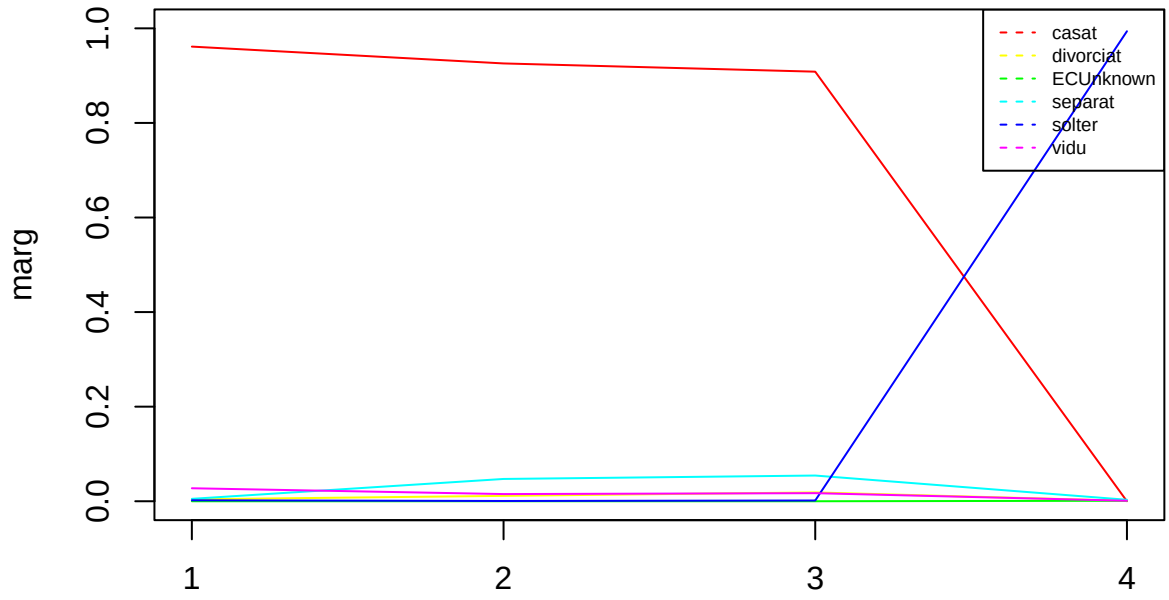

Boxplot of Edad vs Class
P
4
3
2
1
20
30
40
50
60
70
dades[, k]



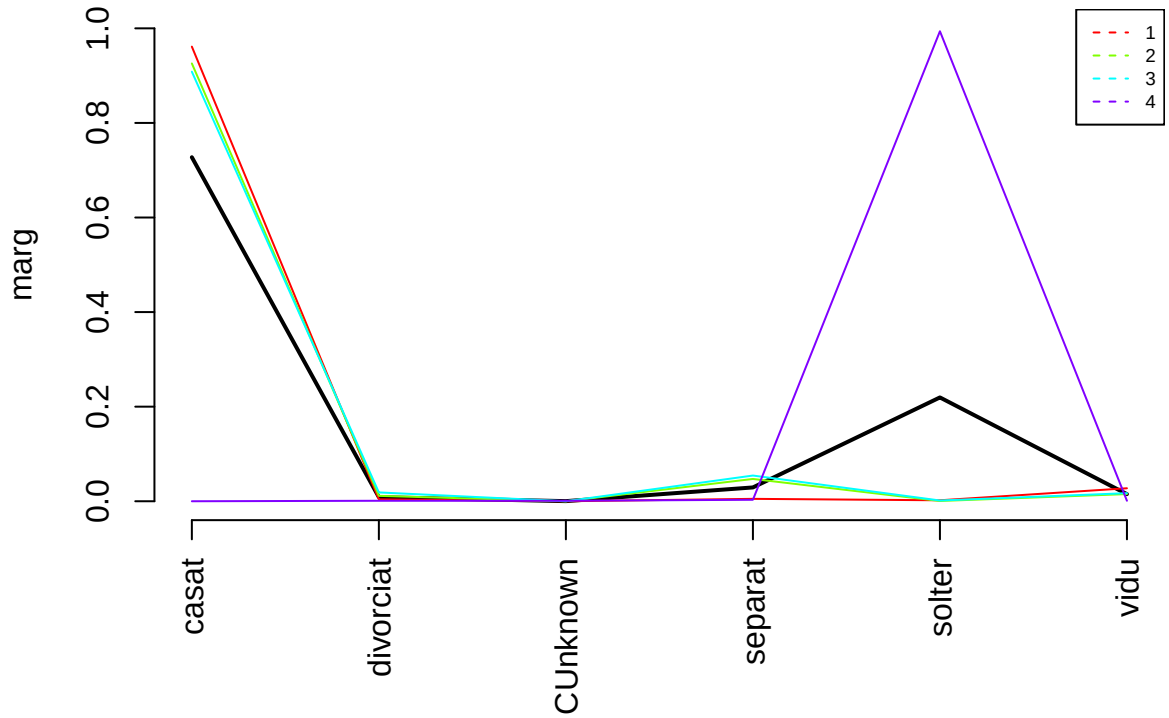

```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    19.0    33.0    41.0    41.5    50.0    68.0
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   19.00   31.00   37.00   38.58   46.00   68.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    22.0    33.0    39.0    40.1    47.0    65.0
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   18.00   23.00   26.00   27.82   30.00   65.00
## [1] "p-valueANOVA: 1.42367984171841e-248"
## [1] "p-value Kruskal-Wallis: 2.32658401793925e-236"
## [1] "p-values ValorsTest: "
## [1] 9.899767e-46 9.022580e-15 2.595850e-14 0.000000e+00
## [1] "Variable Estado.civil"
## [1] "Categories=" "casat"       "divorciat"   "ECUnknown"   "separat"
## [6] "solter"      "vidu"
```

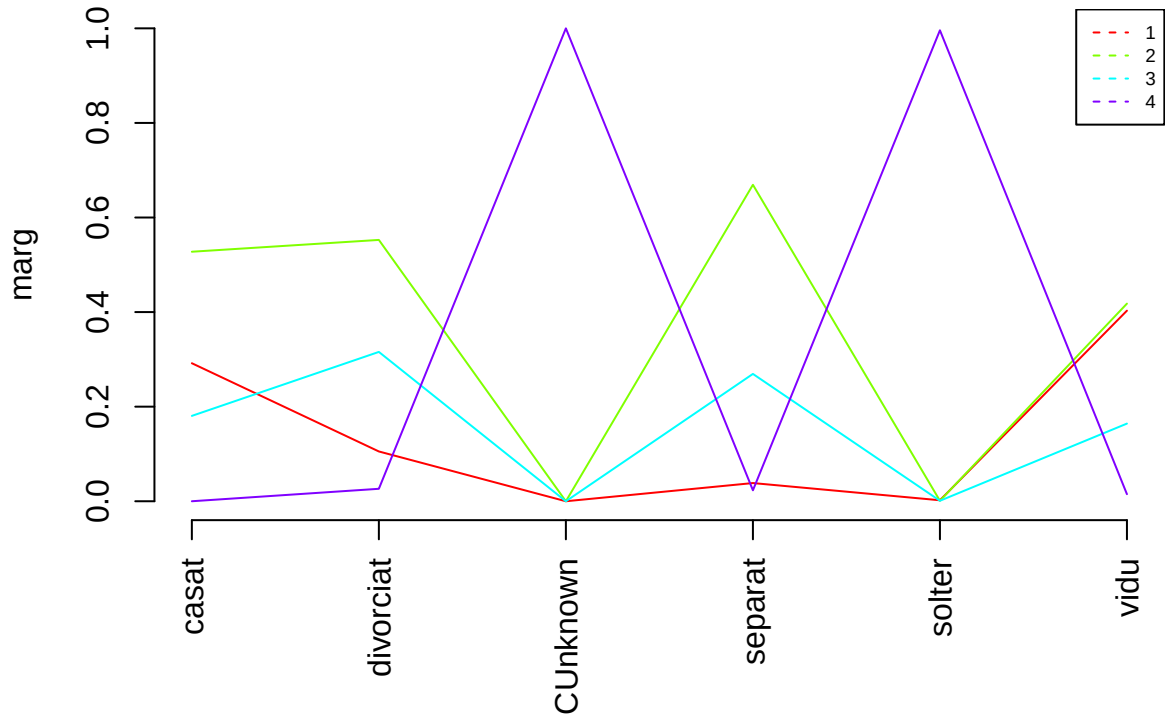



```
## [1] "Categories=" "casat"       "divorciat"   "ECUnknown"   "separat"
## [6] "solter"      "vidu"
```

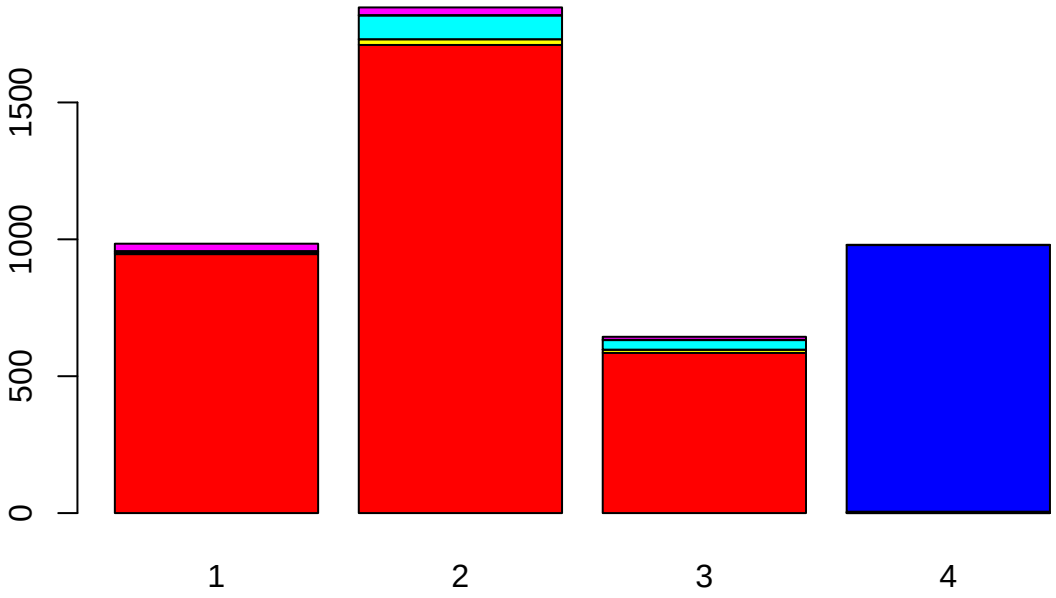


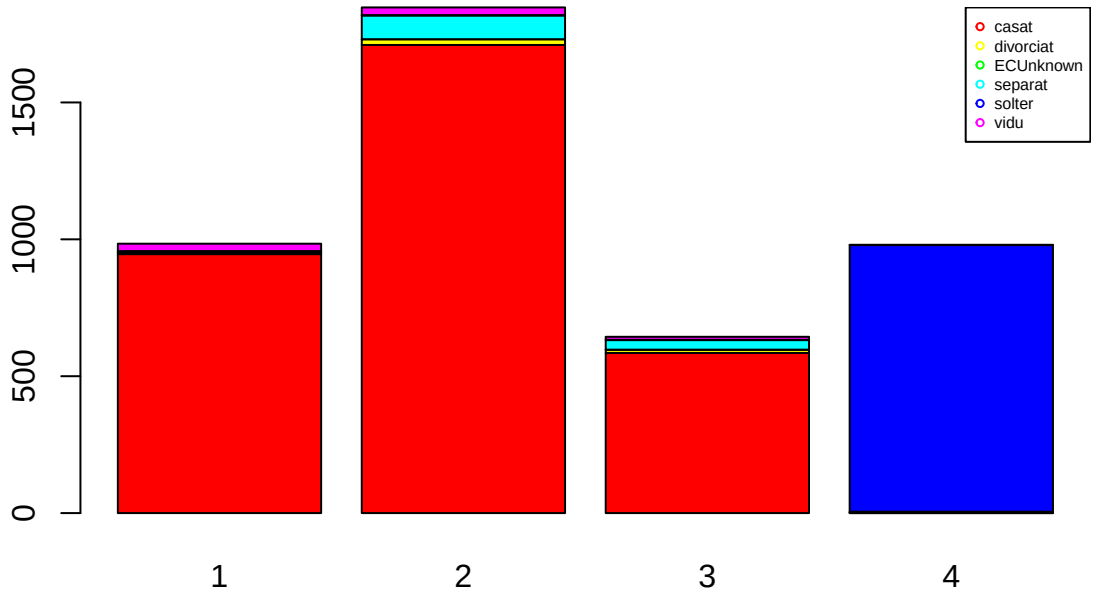

```
## [1] "Categories=" "casat"       "divorciat"   "ECUnknown"   "separat"
## [6] "solter"      "vidu"
```

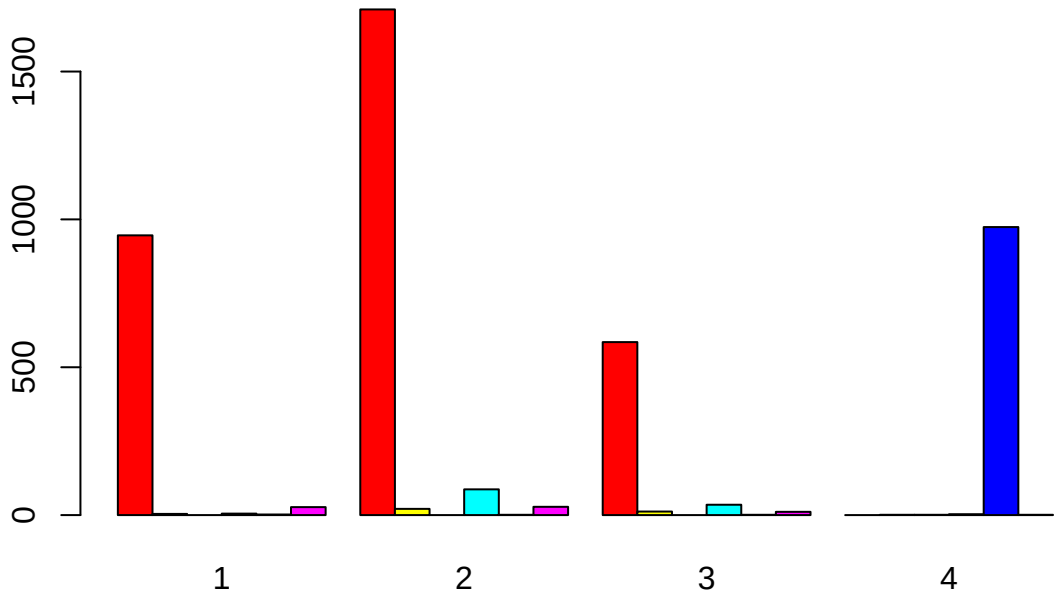

Prop. of pos & neg by Estado.civil
marg
0.0
0.2
0.4
0.6
0.8
1.0
casat
divorciat
CUnknown
separat
solter
vidu
1
2
3
4

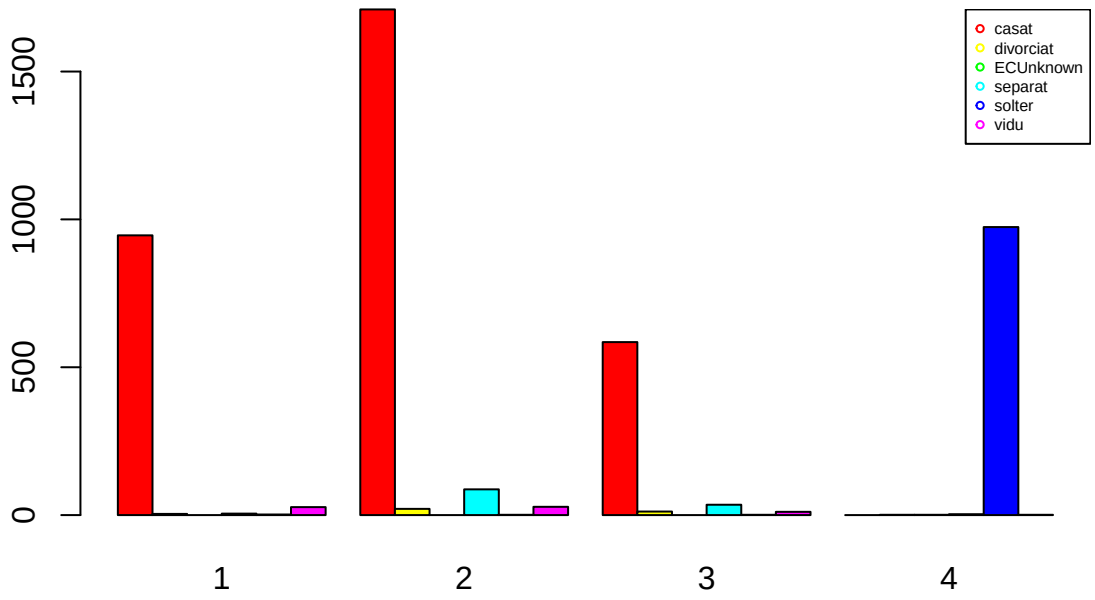



```
## [1] "Cross Table:"
##             P
##                1    2    3    4
##   casat      946 1710  585    0
##   divorciat    4   21   12    1
##   ECUnknown    0    0    0    1
##   separat      5   87   35    3
##   solter       2    1    1  974
##   vidu        27   28   11    1
## [1] "Distribucions condicionades a columnes:"
##
## P         casat   divorciat   ECUnknown     separat      solter        vidu
##   1 0.291885221 0.105263158 0.000000000 0.038461538 0.002044990 0.402985075
##   2 0.527614934 0.552631579 0.000000000 0.669230769 0.001022495 0.417910448
##   3 0.180499846 0.315789474 0.000000000 0.269230769 0.001022495 0.164179104
##   4 0.000000000 0.026315789 1.000000000 0.023076923 0.995910020 0.014925373
```


0
500
1000
1500
1
2
3
4

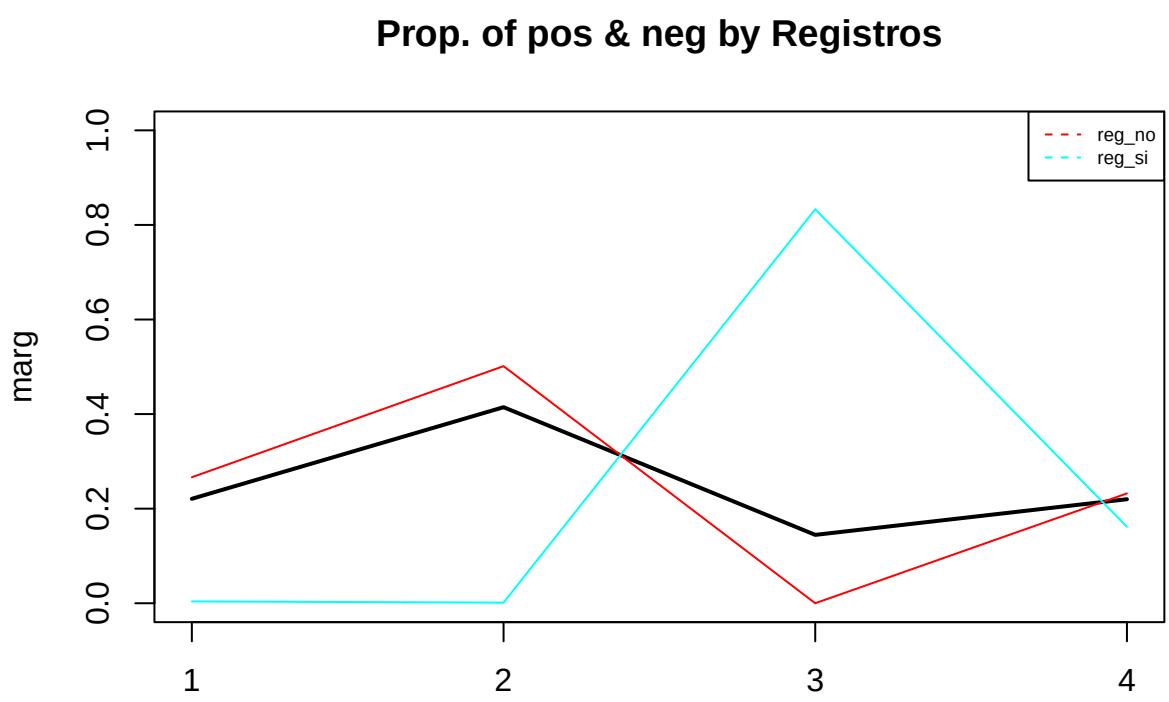

casat
divorciat
ECUnknown
separat
solter
vidu
0
500
1000
1500
1
2
3
4


0
500
1000
1500
1
2
3
4

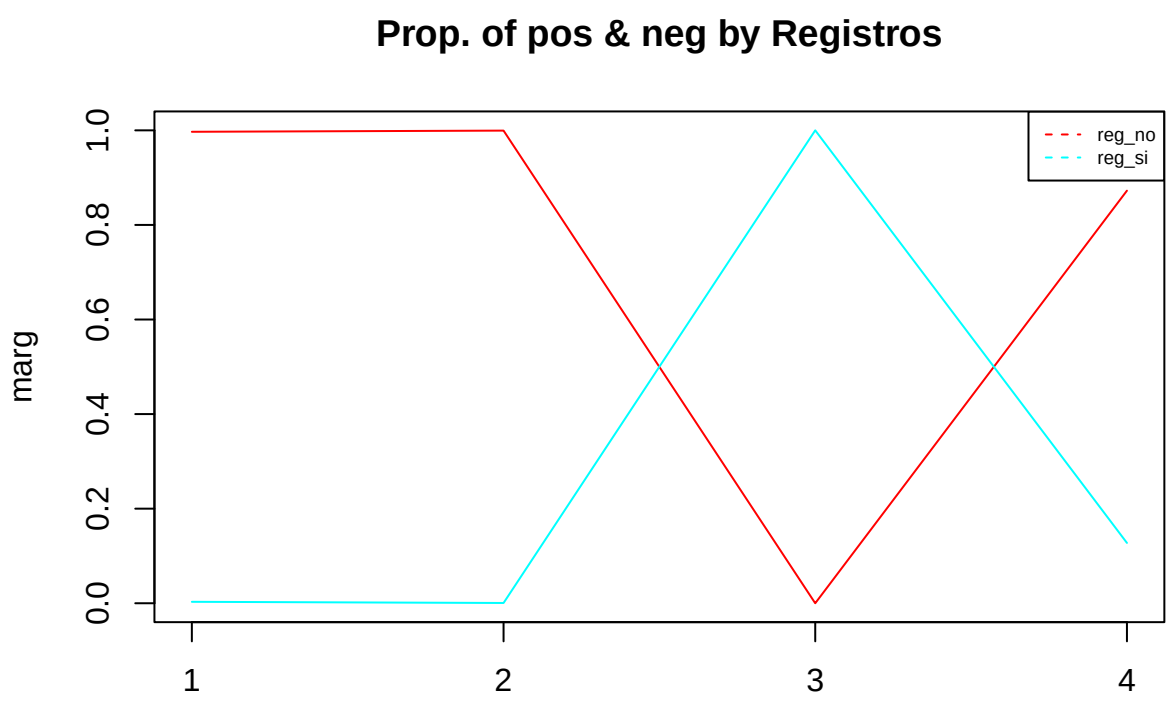



```
## [1] "Test Chi quadrat: "

## Warning in chisq.test(dades[, k], as.factor(P)): Chi-squared approximation may
## be incorrect

##
##  Pearson's Chi-squared test
##
## data:  dades[, k] and as.factor(P)
## X-squared = 4469, df = 15, p-value < 2.2e-16
##
## [1] "valorsTest:"
## $rowpf
##    Xquali
## P          casat    divorciat    ECUnknown      separat       solter
##   1 0.9613821138 0.0040650407 0.0000000000 0.0050813008 0.0020325203
##   2 0.9258256632 0.0113697888 0.0000000000 0.0471034109 0.0005414185
##   3 0.9083850932 0.0186335404 0.0000000000 0.0543478261 0.0015527950
##   4 0.0000000000 0.0010204082 0.0010204082 0.0030612245 0.9938775510
##    Xquali
## P           vidu
##   1 0.0274390244
##   2 0.0151597185
##   3 0.0170807453
##   4 0.0010204082
```

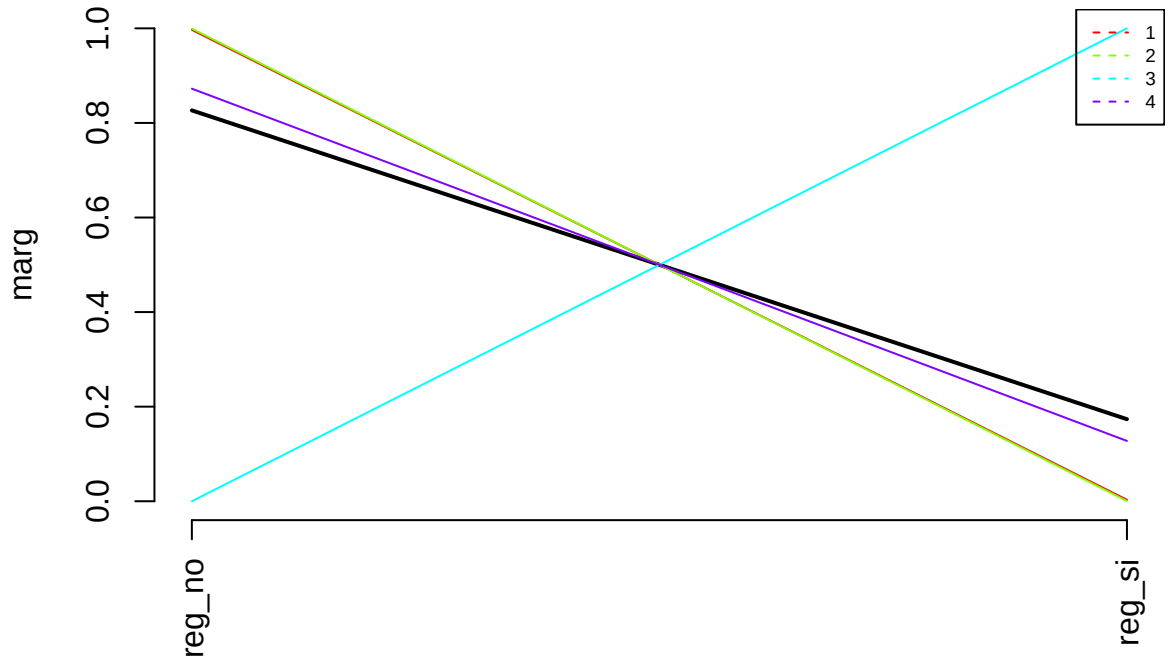
```
##
## $vtest
##    Xquali
## P          casat    divorciat    ECUnknown      separat       solter
##   1  18.66788713  -1.72535596  -0.53249924  -5.08842308 -18.67331849
##   2  25.02000645   1.73468748  -0.84164394   5.98122701 -29.71653556
##   3  11.14694036   3.01456186  -0.41112341   4.10264544 -14.44879330
##   4 -57.91481318  -2.89436369   1.88327228  -5.50055925  66.30894629
##    Xquali
## P           vidu
##   1   3.62062684
##   2   0.05558269
##   3   0.46022050
##   4  -4.08272007
##
## $pval
##    Xquali
## P           casat     divorciat     ECUnknown       separat        solter
##   1  4.517866e-78  4.223167e-02  2.971901e-01  1.805266e-07  0.000000e+00
##   2 1.851833e-138  4.139809e-02  1.999936e-01  1.107315e-09  0.000000e+00
##   3  3.705552e-29  1.286754e-03  3.404910e-01  2.042265e-05  0.000000e+00
##   4  0.000000e+00  1.899639e-03  2.983174e-02  1.892943e-08  0.000000e+00
##    Xquali
## P            vidu
##   1  1.469451e-04
##   2  4.778371e-01
##   3  3.226790e-01
##   4  2.225581e-05
##
## [1] "Variable Registros"
## [1] "Categories=" "reg_no"      "reg_si"
```

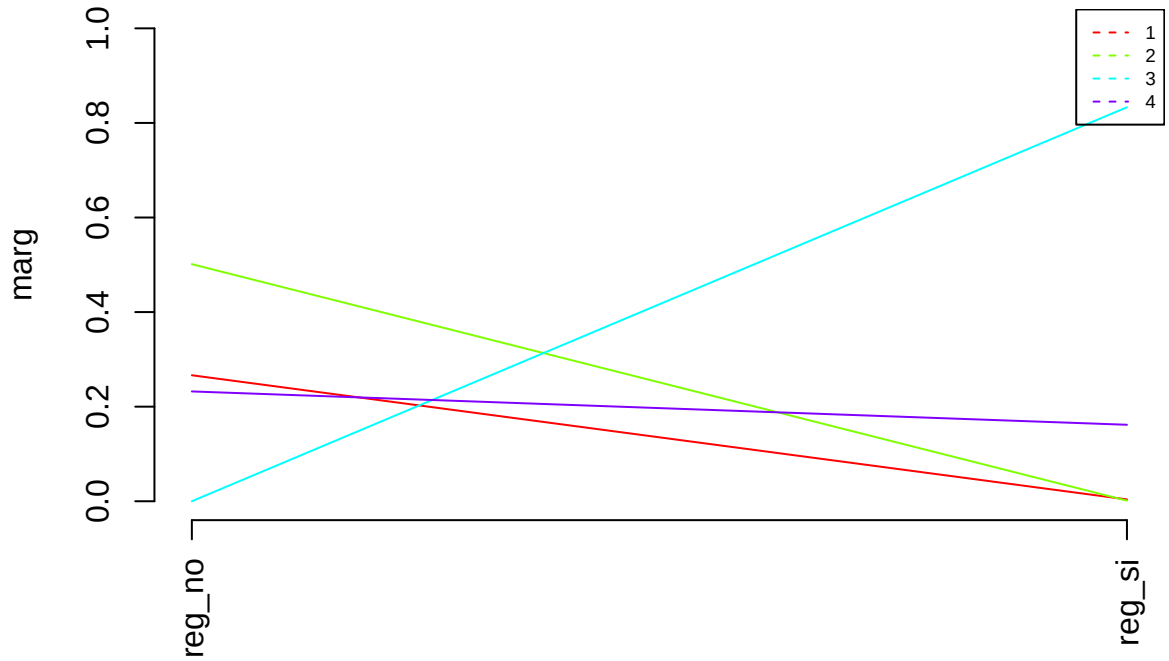



```
## [1] "Categories=" "reg_no"      "reg_si"
```

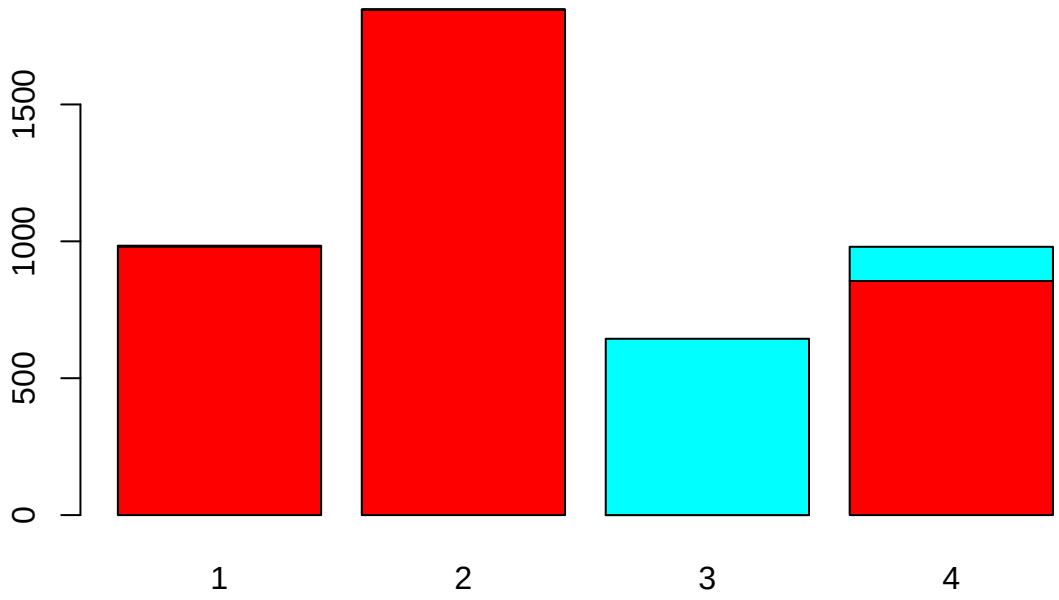


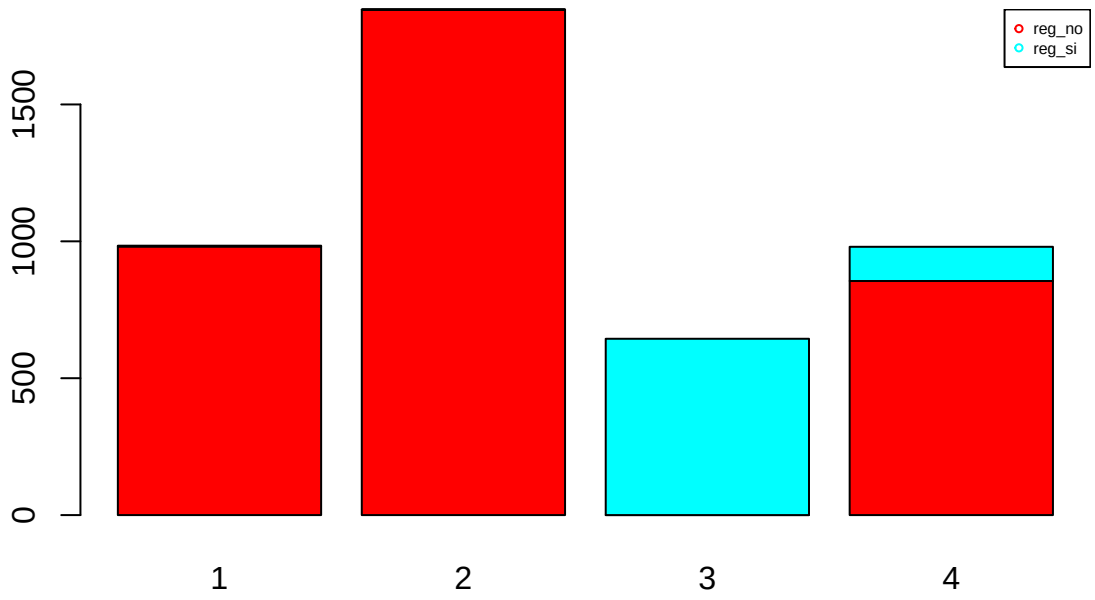

```
## [1] "Categories=" "reg_no"      "reg_si"
```

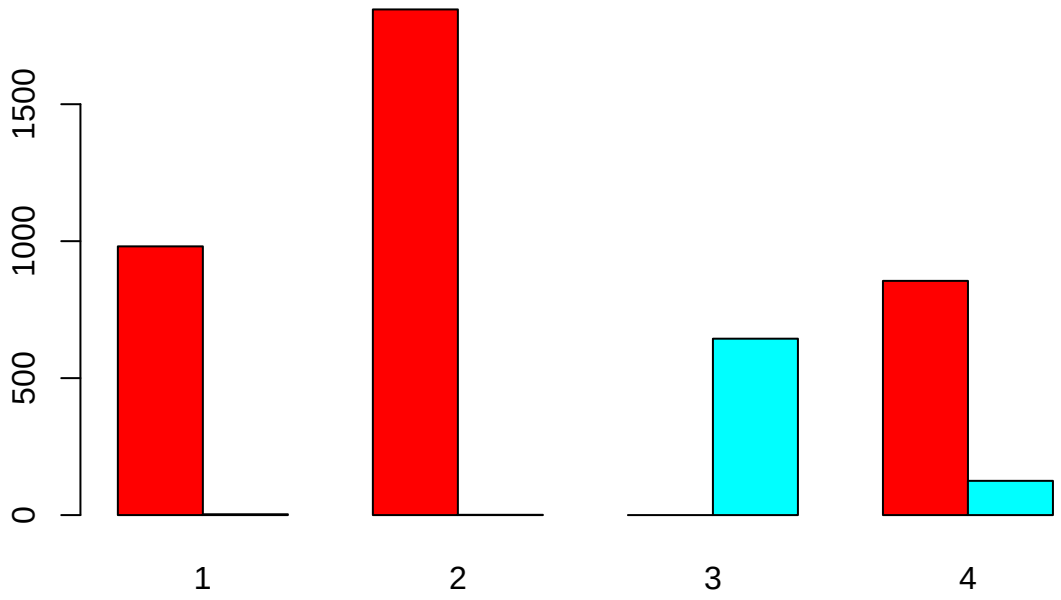

Prop. of pos & neg by Registros
marg
0.0
0.2
0.4
0.6
0.8
1.0
reg_no
reg_si
1
2
3
4

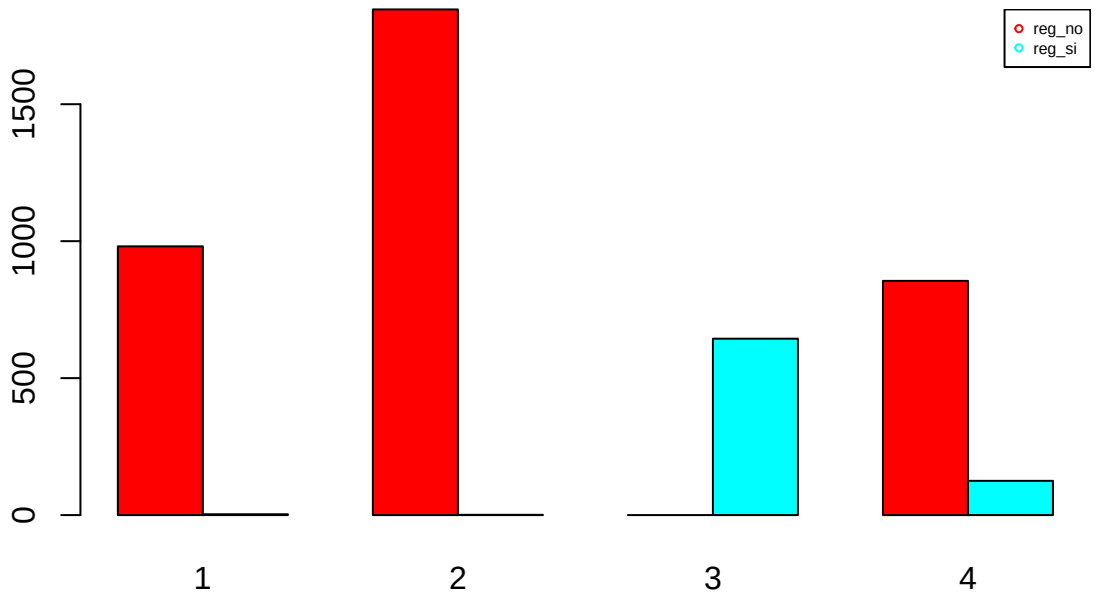



```
## [1] "Cross Table:"
##         P
##             1    2    3    4
##   reg_no  981 1846    0  855
##   reg_si    3    1  644  125
## [1] "Distribucions condicionades a columnes:"
##
## P        reg_no      reg_si
##   1 0.266431287 0.003880983
##   2 0.501357958 0.001293661
##   3 0.000000000 0.833117723
##   4 0.232210755 0.161707633
```


0
500
1000
1500
1
2
3
4

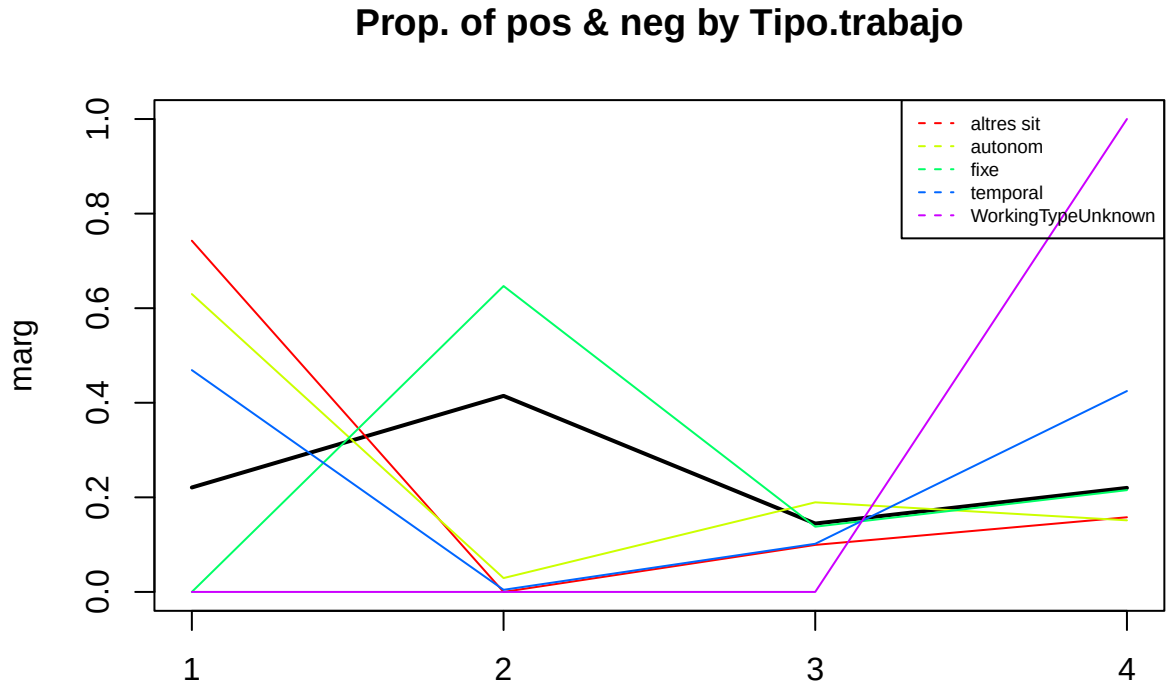

reg_no
reg_si
0
500
1000
1500
1
2
3
4


0
500
1000
1500
1
2
3
4

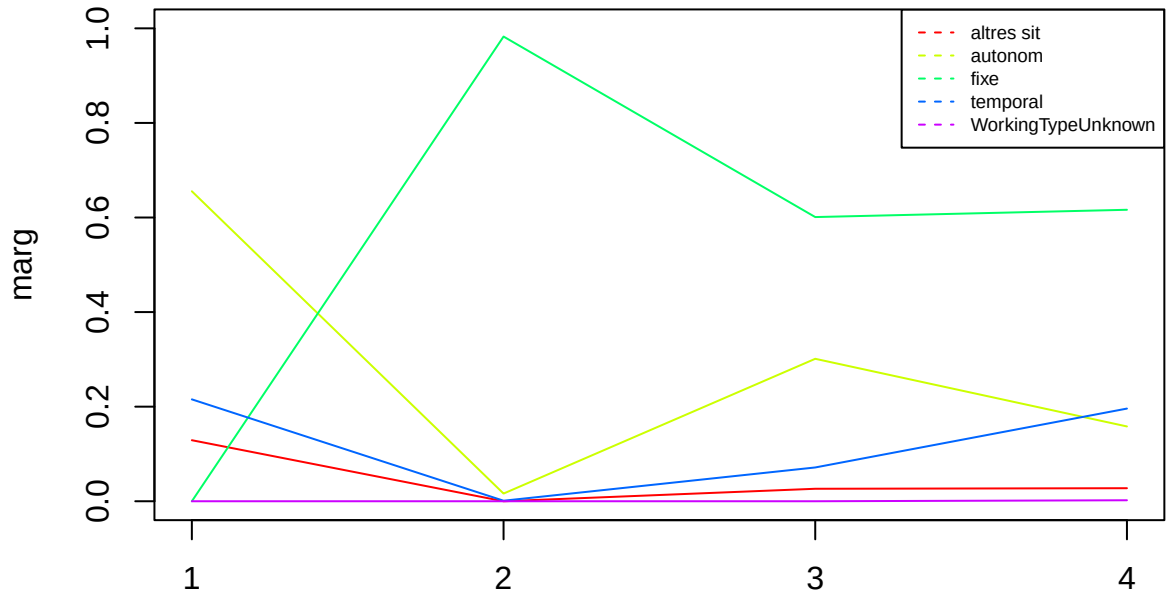



```
## [1] "Test Chi quadrat: "
##
##  Pearson's Chi-squared test
##
## data:  dades[, k] and as.factor(P)
## X-squared = 3666.7, df = 3, p-value < 2.2e-16
##
## [1] "valorsTest:"
## $rowpf
##    Xquali
## P         reg_no       reg_si
##   1 0.9969512195 0.0030487805
##   2 0.9994585815 0.0005414185
##   3 0.0000000000 1.0000000000
##   4 0.8724489796 0.1275510204
##
## $vtest
##    Xquali
## P      reg_no    reg_si
##   1  15.99717 -15.99717
##   2  25.65630 -25.65630
##   3 -59.88244  59.88244
##   4   4.30203  -4.30203
##
## $pval
##    Xquali
```

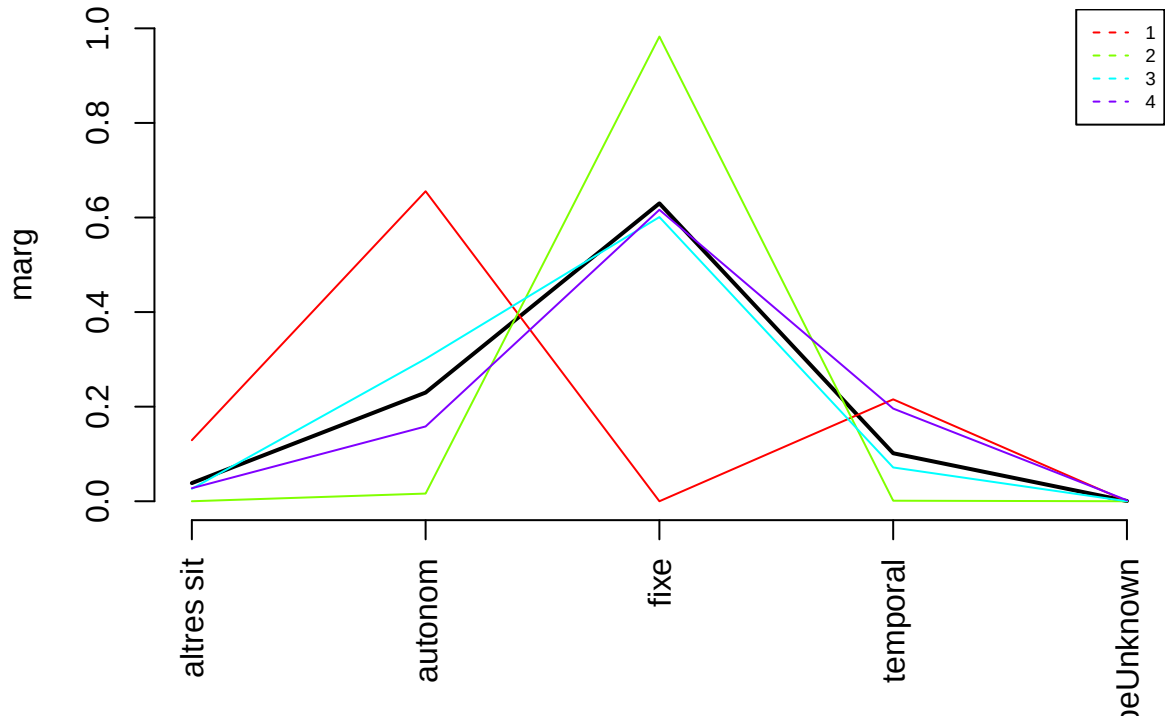
```
## P          reg_no        reg_si
##   1  6.685822e-58  0.000000e+00
##   2 1.797899e-145  0.000000e+00
##   3  0.000000e+00  0.000000e+00
##   4  8.462032e-06  8.462032e-06
##
## [1] "Variable Tipo.trabajo"
## [1] "Categories="        "altres sit"         "autonom"
## [4] "fixe"               "temporal"           "WorkingTypeUnknown"
```

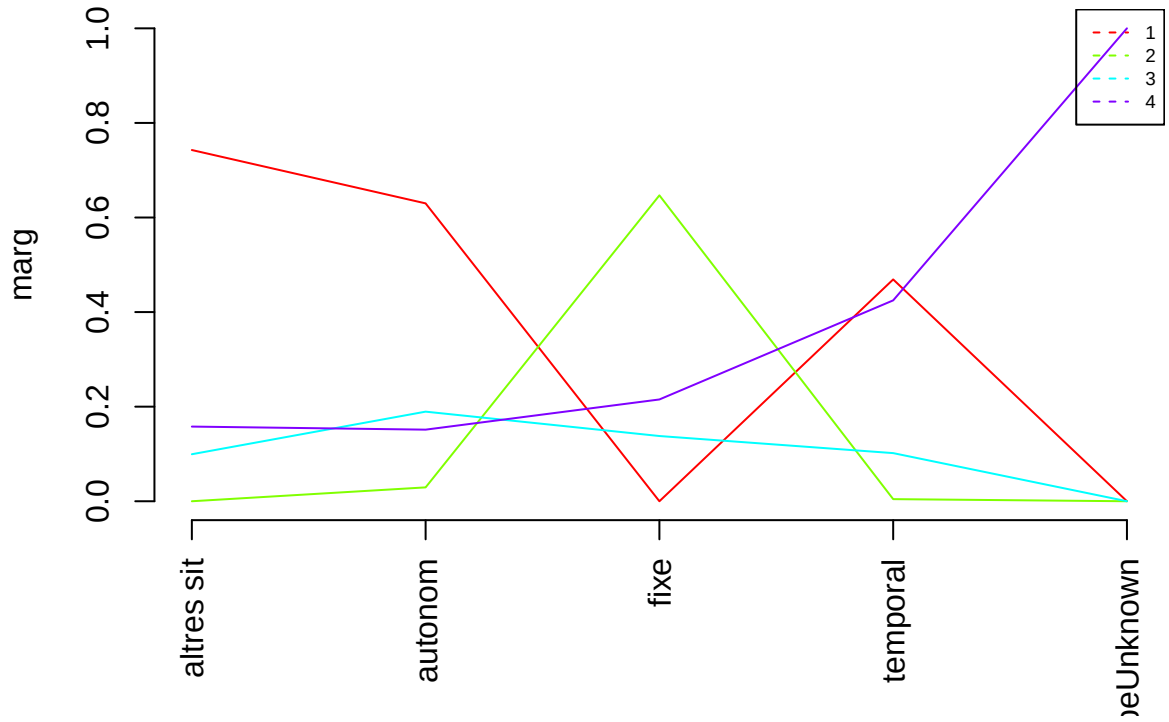



```
## [1] "Categories="        "altres sit"         "autonom"
## [4] "fixe"               "temporal"           "WorkingTypeUnknown"
```

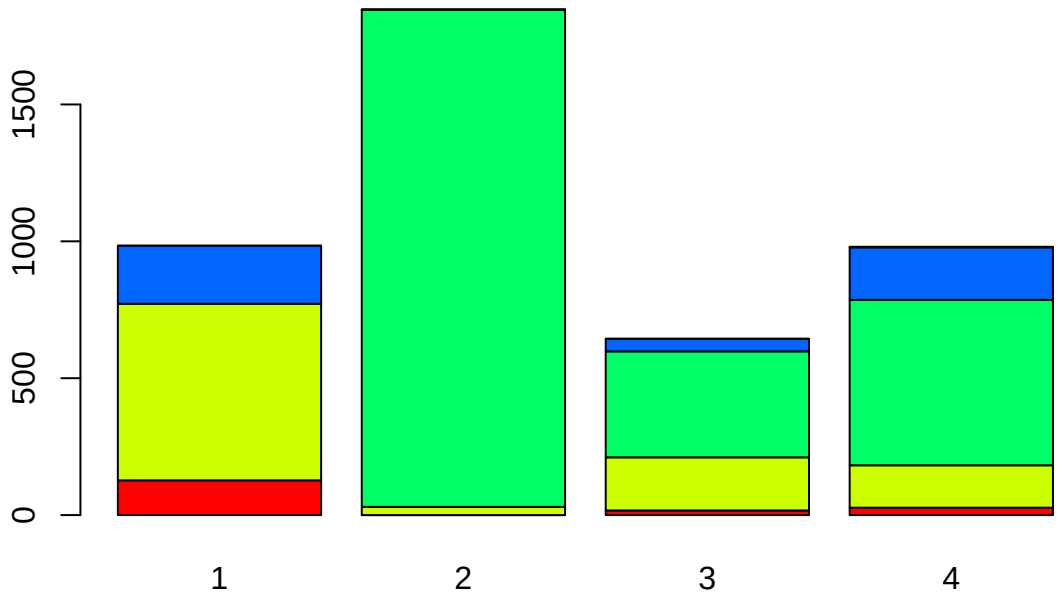


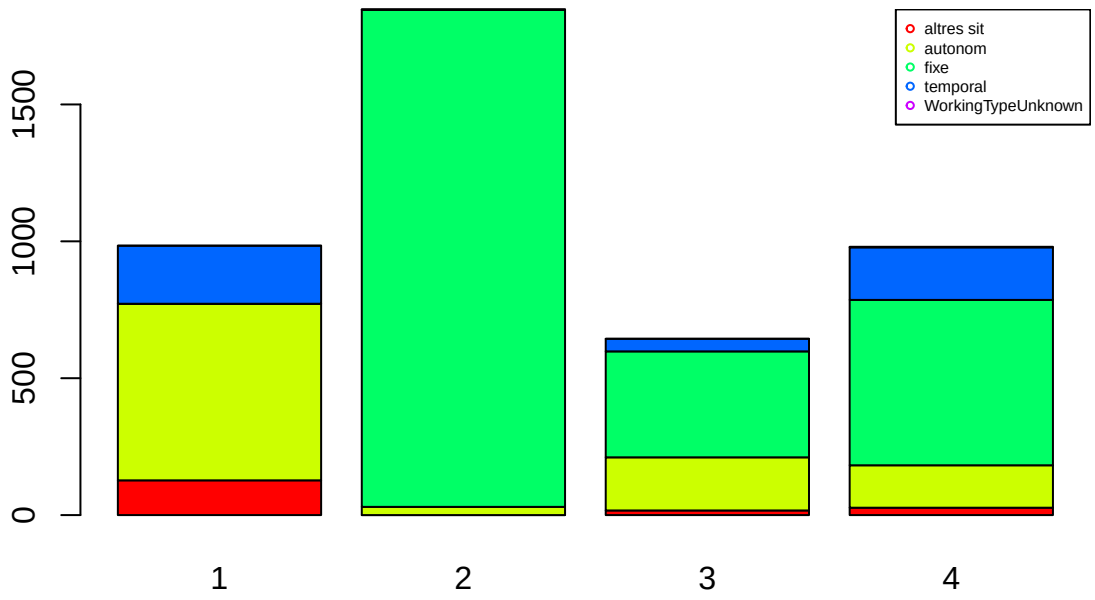

```
## [1] "Categories="        "altres sit"         "autonom"
## [4] "fixe"               "temporal"           "WorkingTypeUnknown"
```

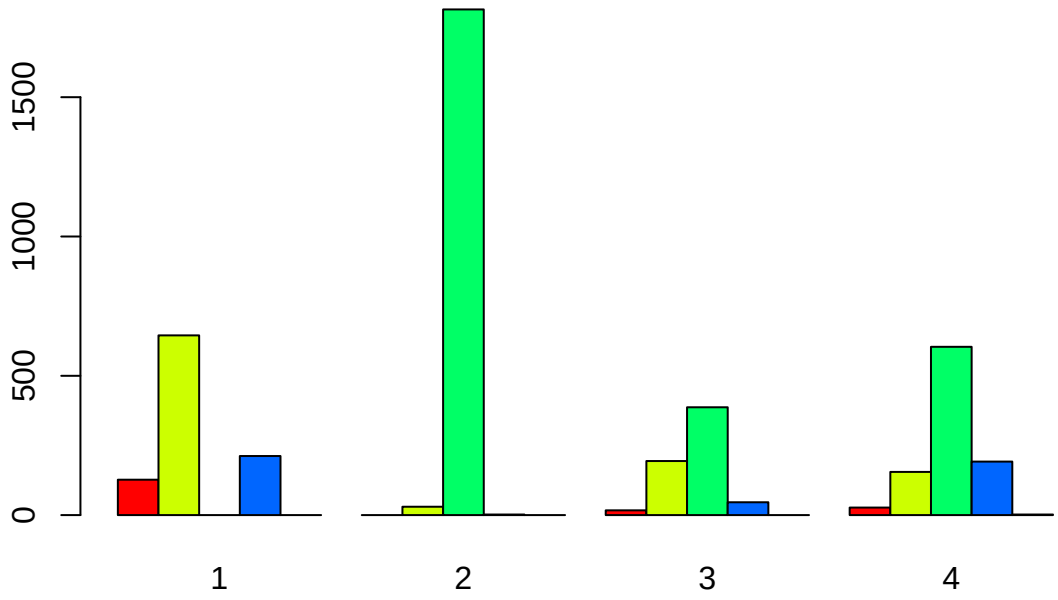

Prop. of pos & neg by Tipo.trabajo
marg
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
altres sit
autonom
fixe
temporal
eUnknown

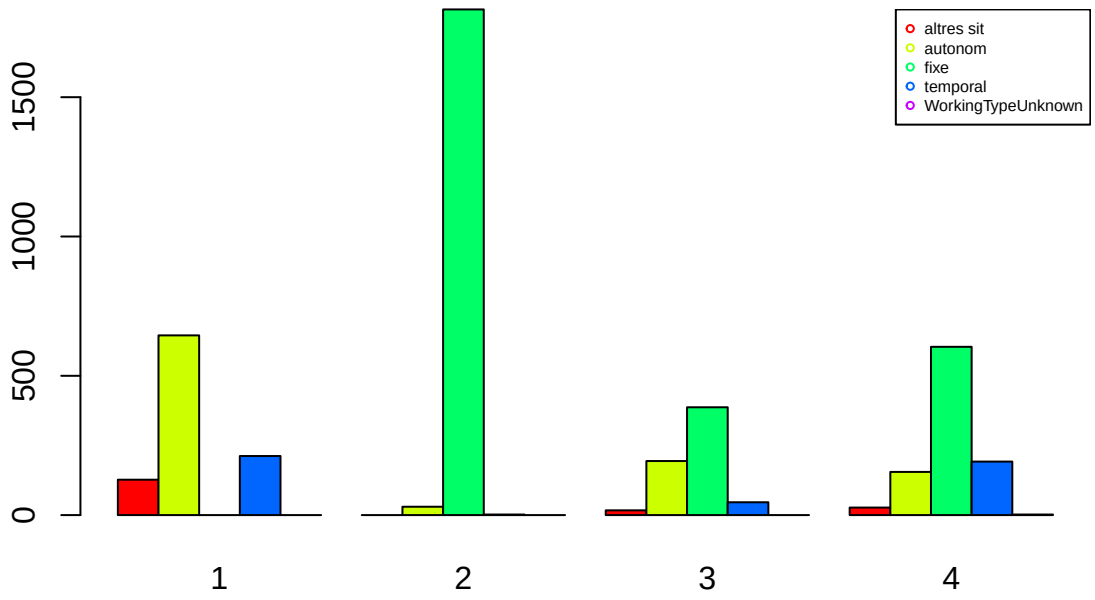



```
## [1] "Cross Table:"
##                       P
##                          1    2    3    4
##   altres sit           127    0   17   27
##   autonom              645   30  194  155
##   fixe                   0 1815  387  604
##   temporal             212    2   46  192
##   WorkingTypeUnknown     0    0    0    2
## [1] "Distribucions condicionades a columnes:"
##
## P    altres sit     autonom        fixe    temporal WorkingTypeUnknown
##   1 0.742690058 0.629882812 0.000000000 0.469026549        0.000000000
##   2 0.000000000 0.029296875 0.646828225 0.004424779        0.000000000
##   3 0.099415205 0.189453125 0.137918746 0.101769912        0.000000000
##   4 0.157894737 0.151367188 0.215253029 0.424778761        1.000000000
```


0
500
1000
1500
1
2
3
4

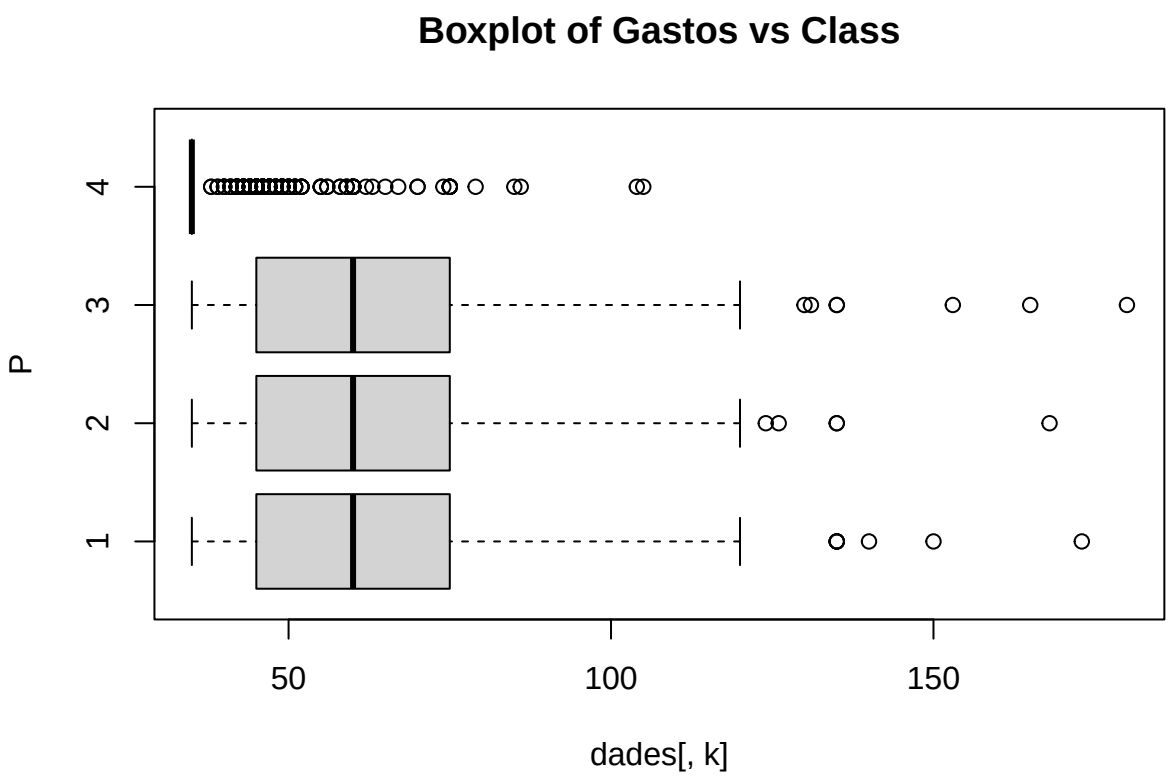

altres sit
autonom
fixe
temporal
WorkingTypeUnknown
0
500
1000
1500
1
2
3
4

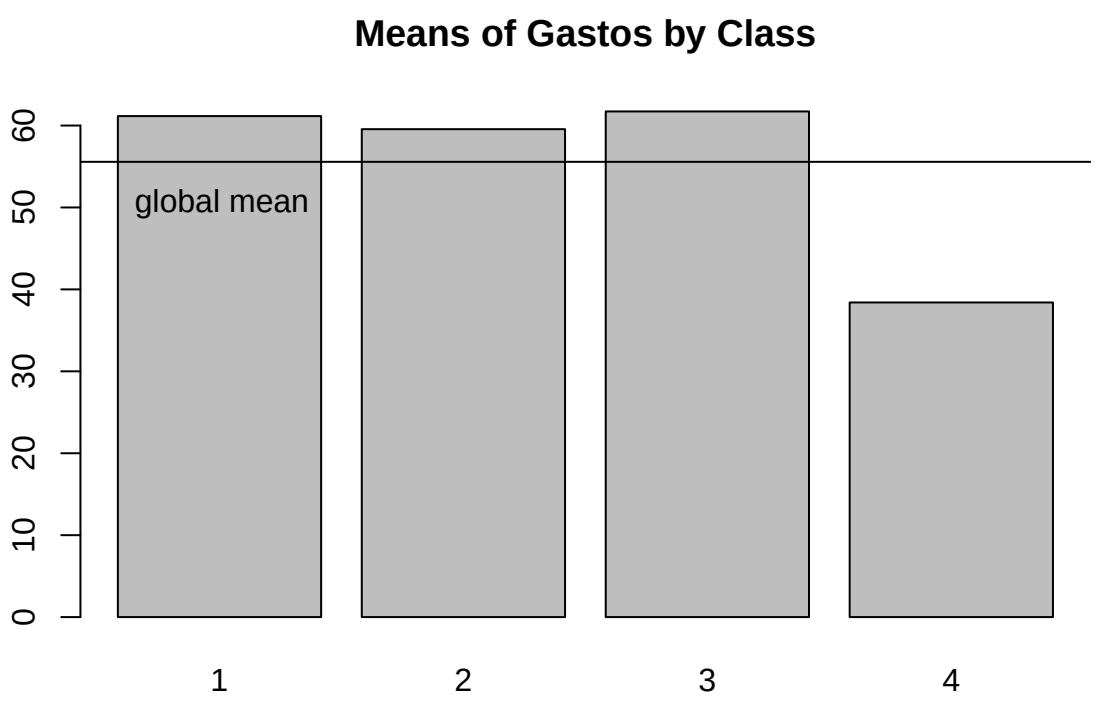

0
500
1000
1500
1
2
3
4



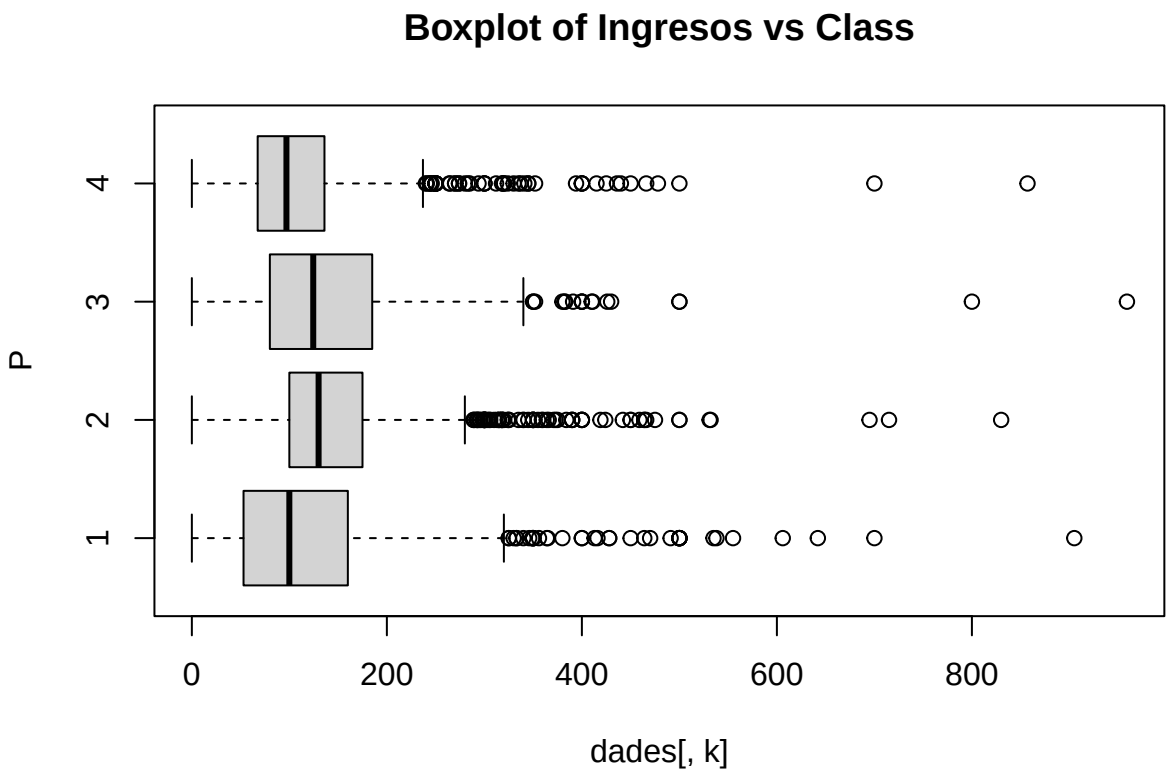

```
## [1] "Test Chi quadrat: "

## Warning in chisq.test(dades[, k], as.factor(P)): Chi-squared approximation may
## be incorrect

##
##  Pearson's Chi-squared test
##
## data:  dades[, k] and as.factor(P)
## X-squared = 2859.9, df = 12, p-value < 2.2e-16
##
## [1] "valorsTest:"
## $rowpf
##    Xquali
## P    altres sit     autonom        fixe    temporal WorkingTypeUnknown
##   1 0.129065041 0.655487805 0.000000000 0.215447154        0.000000000
##   2 0.000000000 0.016242555 0.982674607 0.001082837        0.000000000
##   3 0.026397516 0.301242236 0.600931677 0.071428571        0.000000000
##   4 0.027551020 0.158163265 0.616326531 0.195918367        0.002040816
##
## $vtest
##    Xquali
## P    altres sit     autonom        fixe    temporal WorkingTypeUnknown
##   1  16.7739865  35.9515387 -46.3583107  13.4165045         -0.7531522
##   2 -11.2221697 -28.5177720  41.0441135 -18.6732514         -1.1903979
```

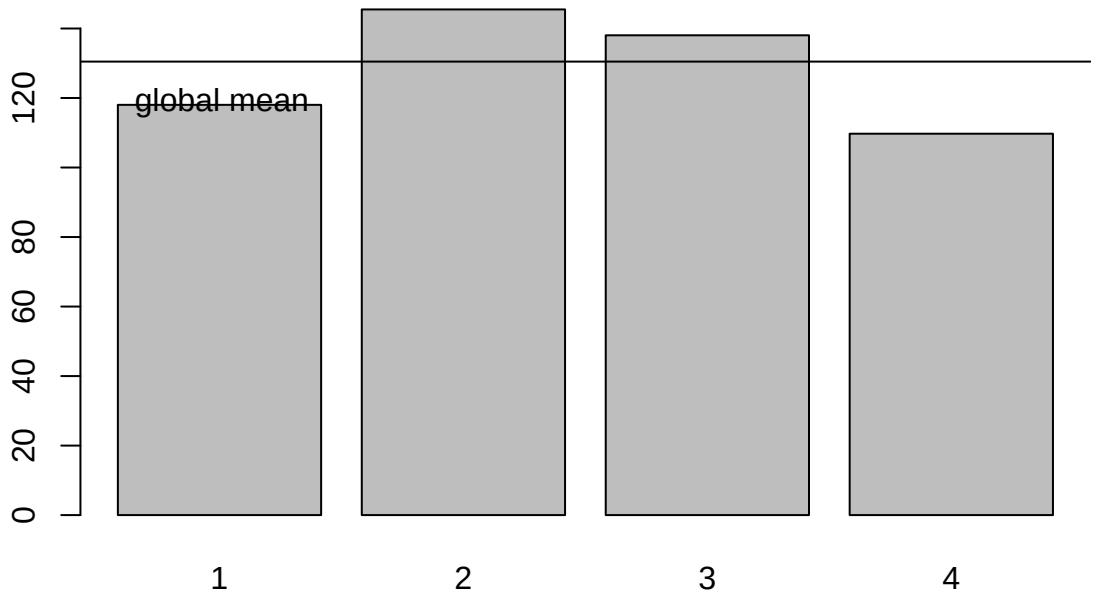
```
##   3  -1.7118204   4.6554402  -1.6435218  -2.7289466         -0.5814816
##   4  -1.9985999  -6.0396320  -0.9930536  11.0889406          2.6636482
##
## $pval
##    Xquali
## P      altres sit       autonom          fixe      temporal WorkingTypeUnknown
##   1  1.891290e-63 2.394397e-283  0.000000e+00  2.420047e-41       2.256792e-01
##   2  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00       1.169450e-01
##   3  4.346488e-02  1.616447e-06  5.013750e-02  3.176849e-03       2.804580e-01
##   4  2.282583e-02  7.723308e-10  1.603419e-01  7.097950e-29       3.864917e-03
##
## [1] "Anàlisi per classes de la Variable: Gastos"
```



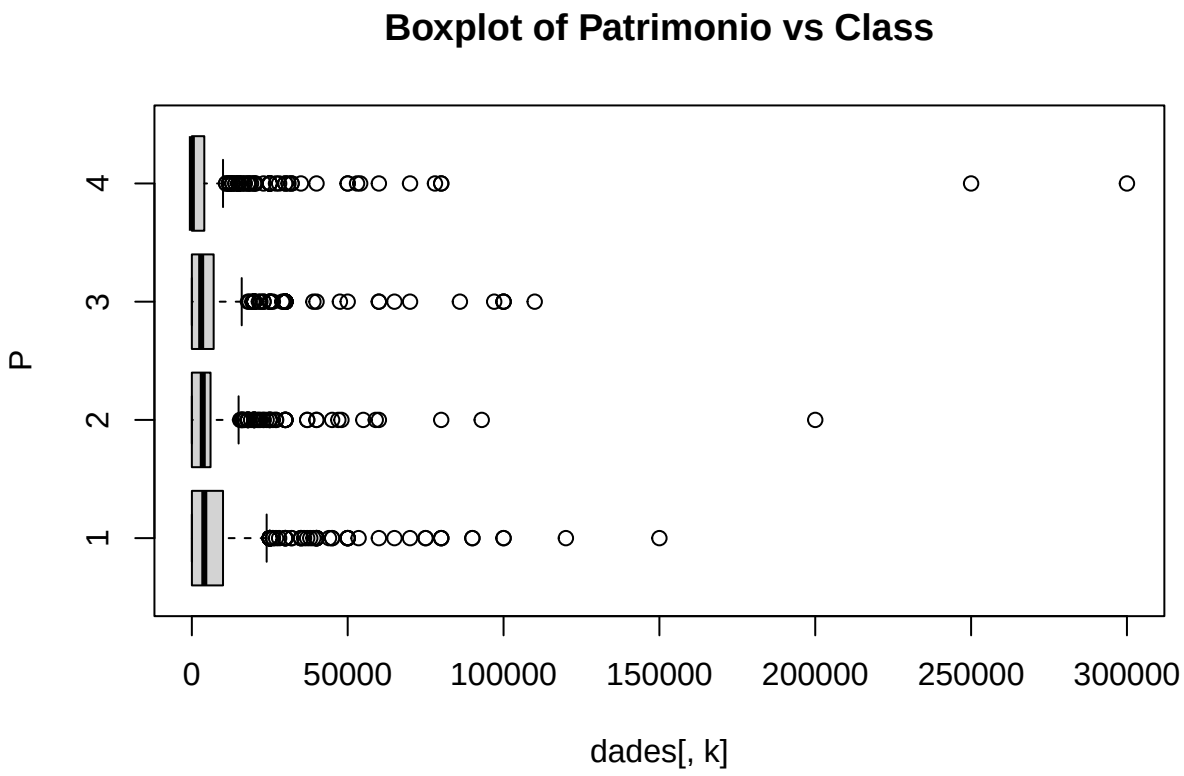


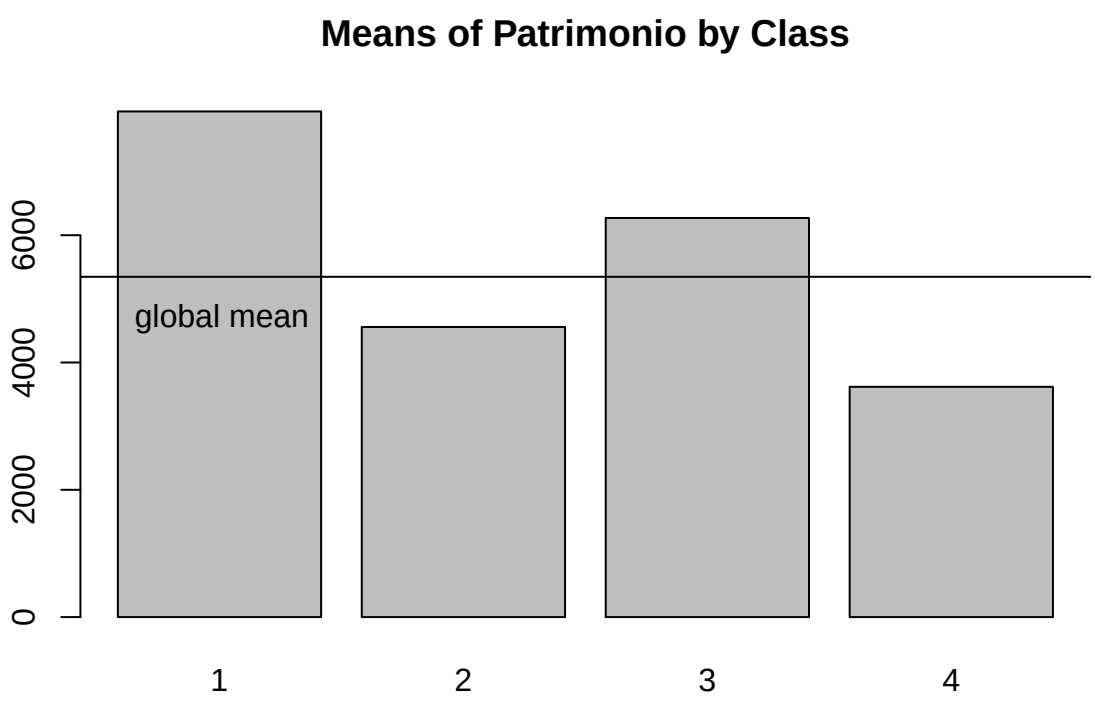

```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   35.00   45.00   60.00   61.15   75.00  173.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   35.00   45.00   60.00   59.56   75.00  168.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   35.00   45.00   60.00   61.72   75.00  180.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    35.0    35.0    35.0    38.4    35.0   105.0
## [1] "p-valueANOVA: 0"
## [1] "p-value Kruskal-Wallis: 7.20760650701279e-284"
## [1] "p-values ValorsTest: "
## [1] 2.522912e-24 2.218378e-30 3.513753e-18 0.000000e+00
## [1] "Anàlisi per classes de la Variable: Ingresos"
```


Boxplot of Ingresos vs Class
P
4
3
2
1
0
200
400
600
800
dades[, k]

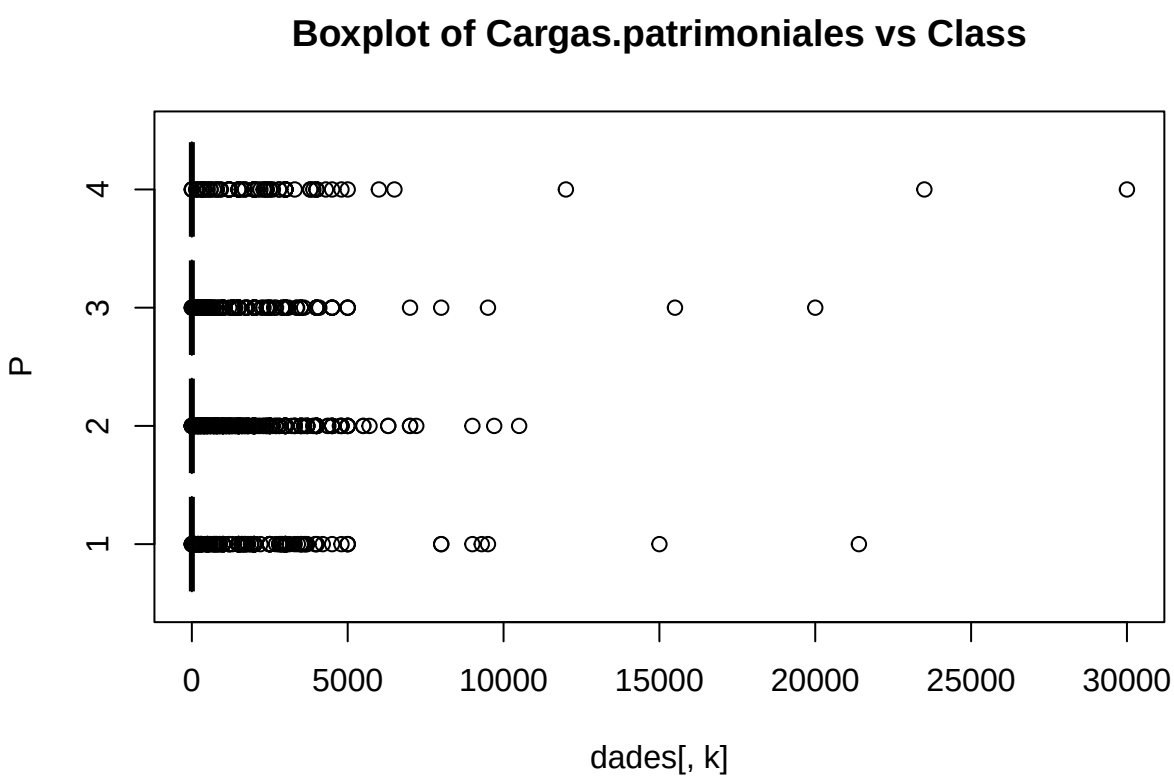


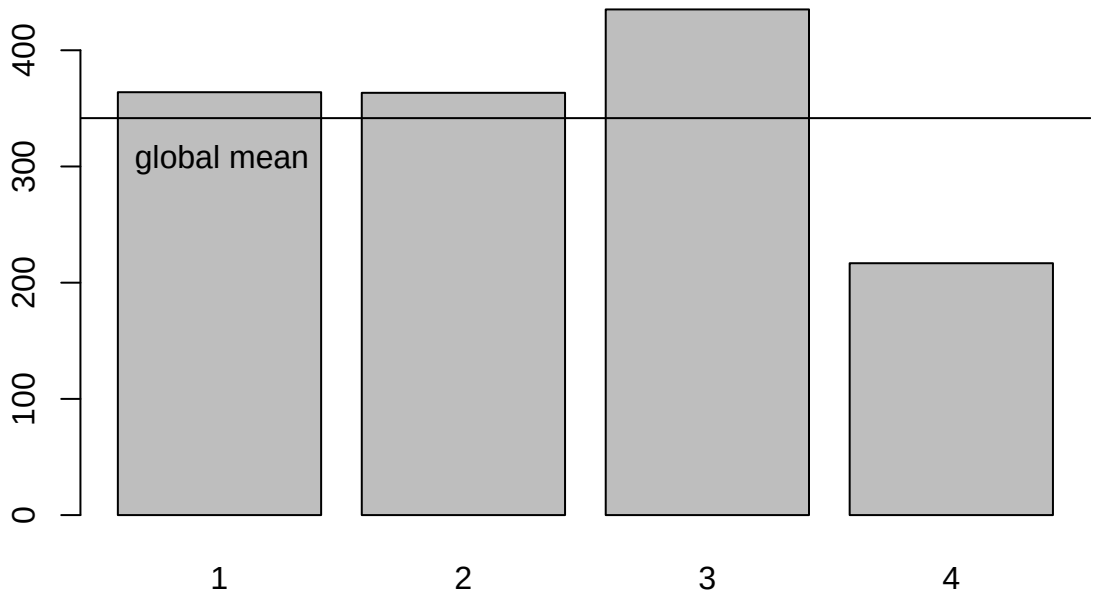

```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0      53     100     118     160     905
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0   100.0   130.0   145.5   175.0   830.0
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0    80.0   124.5   138.0   185.0   959.0
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   67.75   97.00  109.72  136.00  857.00
## [1] "p-valueANOVA: 3.19233193885177e-34"
## [1] "p-value Kruskal-Wallis: 4.53661056659362e-62"
## [1] "p-values ValorsTest: "
## [1] 1.570396e-07 1.373699e-22 8.170800e-03 0.000000e+00
## [1] "Anàlisi per classes de la Variable: Patrimonio"
```


Boxplot of Patrimonio vs Class
P
4
3
2
1
0
50000
100000
150000
200000
250000
300000
dades[, k]



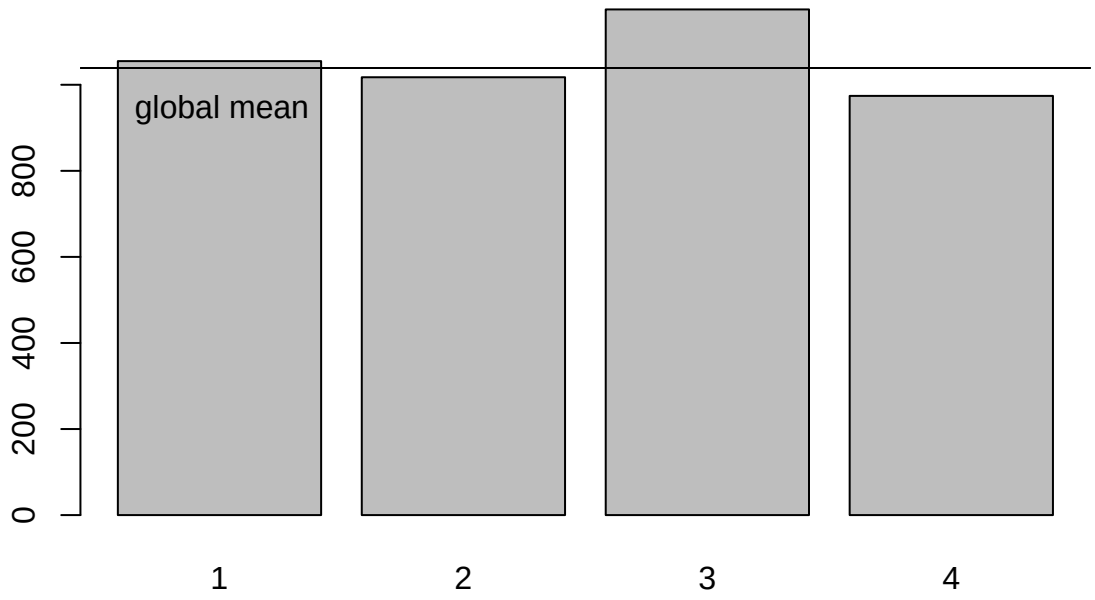

```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0      22    4000    7945   10000  150000
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0       0    3500    4558    6000  200000
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0       0    3000    6271    7000  110000
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0       0       0    3618    4000  300000
## [1] "p-valueANOVA: 3.54236009134637e-15"
## [1] "p-value Kruskal-Wallis: 6.33358739578153e-71"
## [1] "p-values ValorsTest: "
## [1] 7.191679e-16 6.111448e-05 1.394977e-02 5.528508e-08
## [1] "Anàlisi per classes de la Variable: Cargas.patrimoniales"
```


Boxplot of Cargas.patrimoniales vs Class
P
4
3
2
1
0
5000
10000
15000
20000
25000
30000
dades[, k]

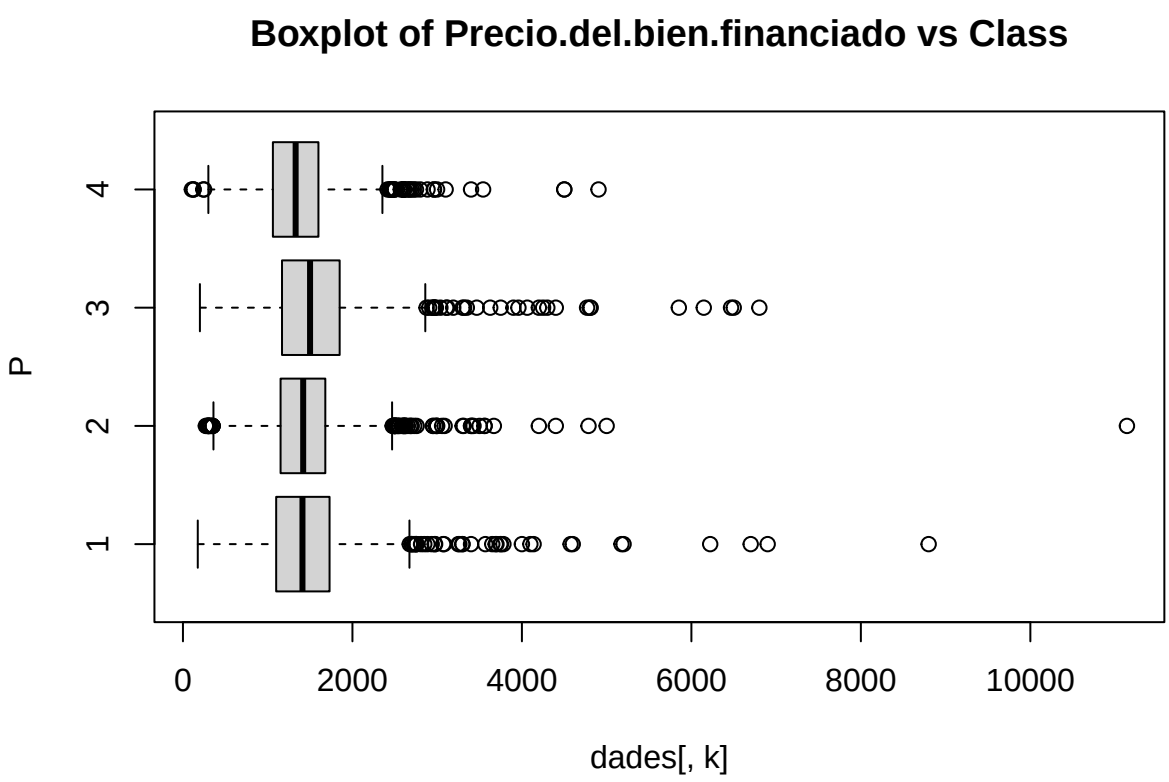


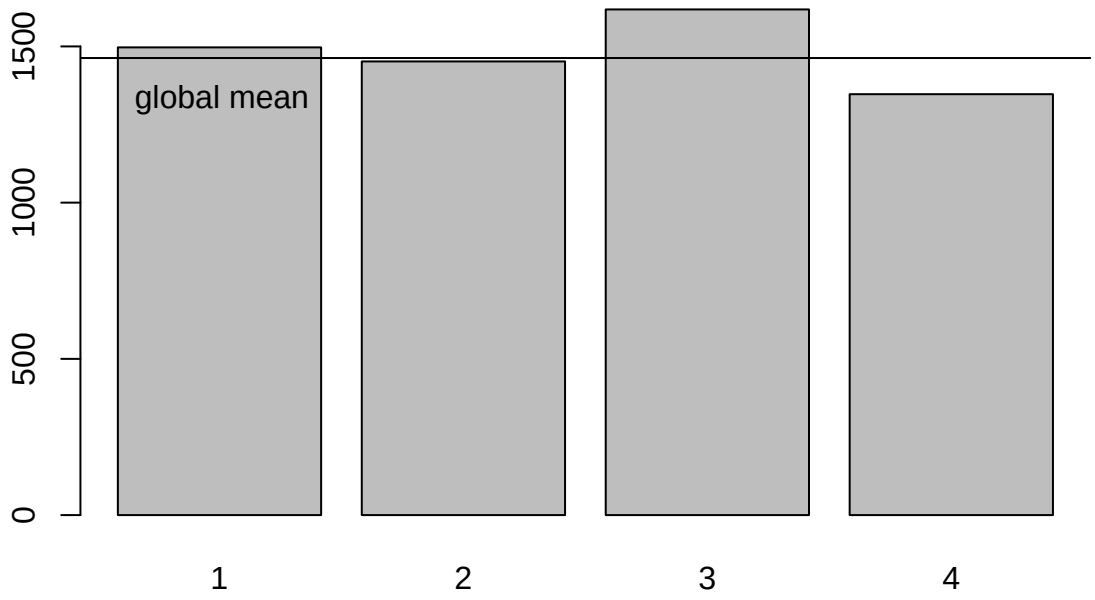

```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0     0.0     0.0   363.9     0.0 21400.0
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0     0.0     0.0   363.4     0.0 10500.0
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0     0.0     0.0   435.1     0.0 20000.0
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0     0.0     0.0   216.7     0.0 30000.0
## [1] "p-valueANOVA: 0.0110050156392372"
## [1] "p-value Kruskal-Wallis: 8.12120583371892e-19"
## [1] "p-values ValorsTest: "
## [1] 0.2620195482 0.1627482835 0.0195542903 0.0001877172
## [1] "Anàlisi per classes de la Variable: Importe.solicitado"
```


Boxplot of Importe.solicitado vs Class
P
4
3
2
1
0
1000
2000
3000
4000
5000
dades[, k]

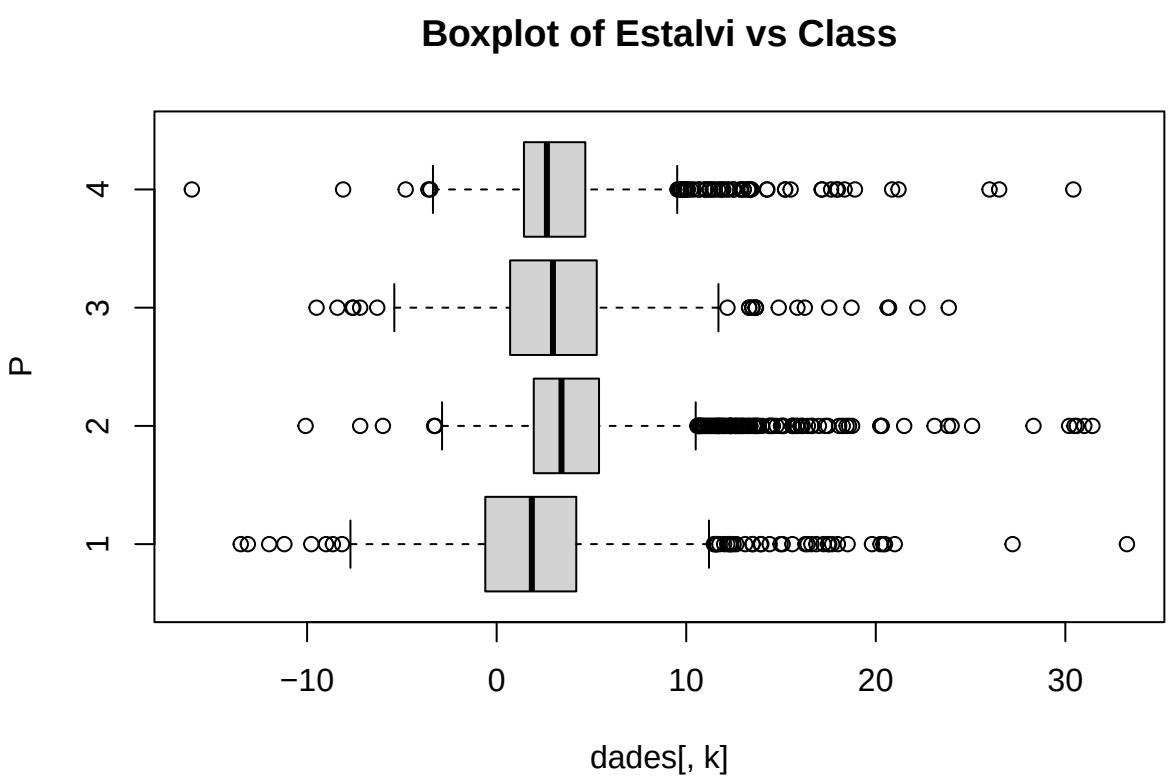


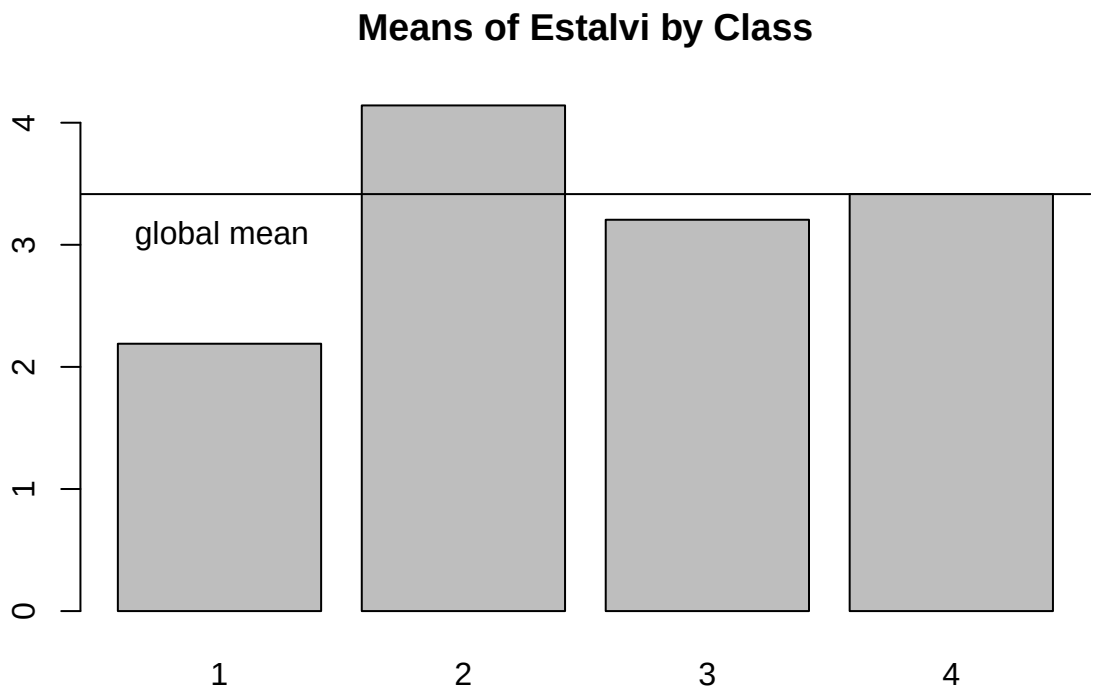

```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     100     700    1000    1055    1300    4500
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     100     700    1000    1017    1300    3000
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     200     800    1100    1175    1500    5000
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   100.0   697.5  1000.0   974.2  1200.0  4000.0
## [1] "p-valueANOVA: 9.21383384594724e-13"
## [1] "p-value Kruskal-Wallis: 5.79592496287302e-11"
## [1] "p-values ValorsTest: "
## [1] 1.158532e-01 5.402014e-03 2.262981e-15 6.591494e-07
## [1] "Anàlisi per classes de la Variable: Precio.del.bien.financiado"
```


Boxplot of Precio.del.bien.financiado vs Class
P
4
3
2
1
0
2000
4000
6000
8000
10000
dades[, k]

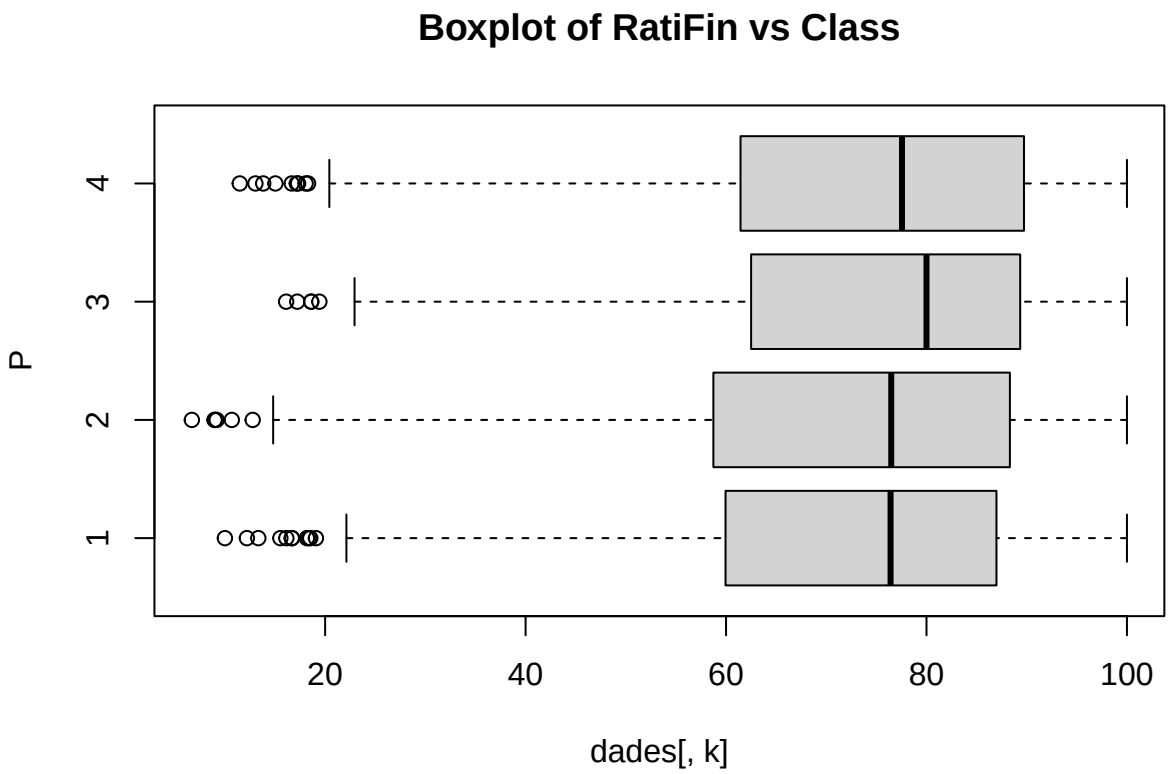


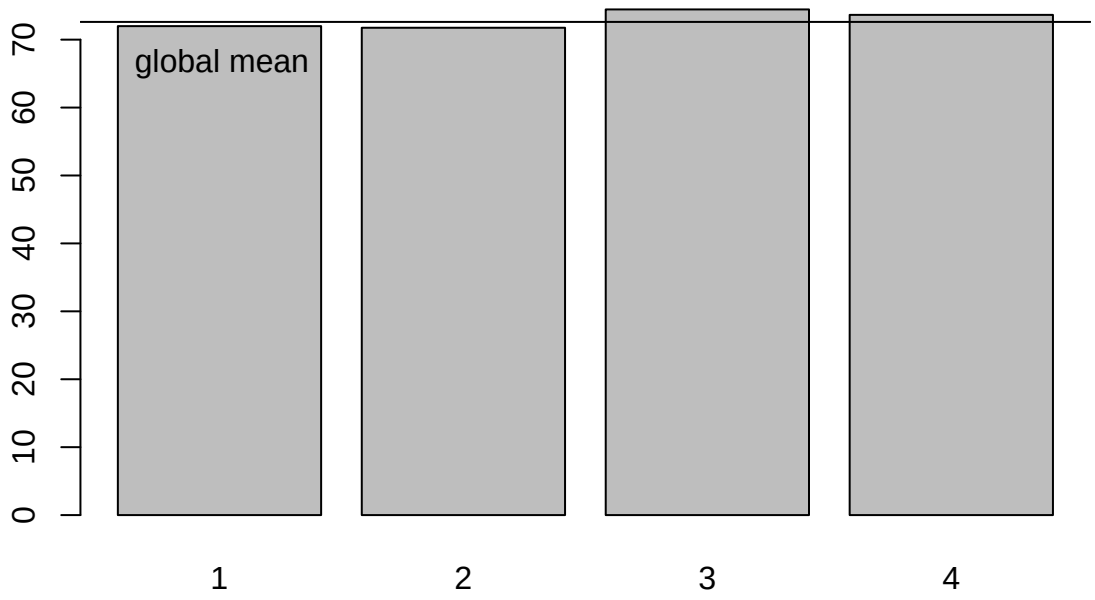

```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     175    1100    1410    1497    1730    8800
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     270    1152    1419    1452    1680   11140
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     200    1170    1500    1618    1849    6802
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     105    1062    1330    1347    1599    4904
## [1] "p-valueANOVA: 5.45702576708873e-15"
## [1] "p-value Kruskal-Wallis: 1.3145979256208e-13"
## [1] "p-values ValorsTest: "
## [1] 2.711835e-02 1.640926e-01 6.218571e-12 3.609446e-11
## [1] "Anàlisi per classes de la Variable: Estalvi"
```


Boxplot of Estalvi vs Class
P
4
3
2
1
−10
0
10
20
30
dades[, k]




```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -13.500  -0.600   1.855   2.190   4.200  33.250
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -10.080   1.956   3.420   4.141   5.400  31.428
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  -9.500   0.715   2.969   3.205   5.280  23.850
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -16.080   1.440   2.652   3.416   4.675  30.420
## [1] "p-valueANOVA: 3.11767704158316e-29"
## [1] "p-value Kruskal-Wallis: 2.84521379025959e-45"
## [1] "p-values ValorsTest: "
## [1] 0.000000e+00 7.411869e-24 7.736883e-02 4.986843e-01
## [1] "Anàlisi per classes de la Variable: RatiFin"
```


Boxplot of RatiFin vs Class
P
4
3
2
1
20
40
60
80
100
dades[, k]




```
## [1] "Estadístics per groups:"
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   10.00   59.95   76.42   71.99   86.97  100.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   6.702  58.737  76.482  71.750  88.319 100.000
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.11   62.50   80.00   74.45   89.34  100.00
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   11.49   61.47   77.56   73.64   89.71  100.00
## [1] "p-valueANOVA: 0.00663129783294964"
## [1] "p-value Kruskal-Wallis: 0.00617546381021875"
## [1] "p-values ValorsTest: "
## [1] 0.139845008 0.008916100 0.006667546 0.035989477
```

```
#descriptors de les classes més significatius. Afegir info qualits
for (c in 1:length(levels(as.factor(P)))) {
  if(!is.na(levels(as.factor(P))[c])){
    print(paste("P.values per class:",levels(as.factor(P))[c]));
    print(sort(pvalk[c,]), digits=3)
  }
}
```

```
## [1] "P.values per class: 1"
##                   Dictamen                   Vivienda
##                   0.00e+00                   0.00e+00
```

```
##               Estado.civil                  Registros
##                   0.00e+00                   0.00e+00
##              Tipo.trabajo                    Estalvi
##                   0.00e+00                   0.00e+00
##                       Edad                     Gastos
##                   9.90e-46                   2.52e-24
##                 Patrimonio                      Plazo
##                   7.19e-16                   9.80e-08
##                   Ingresos          Antiguedad.Trabajo
##                   1.57e-07                   1.07e-03
## Precio.del.bien.financiado          Importe.solicitado
##                   2.71e-02                   1.16e-01
##                    RatiFin       Cargas.patrimoniales
##                   1.40e-01                   2.62e-01
## [1] "P.values per class: 2"
##                   Dictamen                   Vivienda
##                   0.00e+00                   0.00e+00
##               Estado.civil                  Registros
##                   0.00e+00                   0.00e+00
##              Tipo.trabajo          Antiguedad.Trabajo
##                   0.00e+00                   6.84e-60
##                     Gastos                    Estalvi
##                   2.22e-30                   7.41e-24
##                   Ingresos                       Edad
##                   1.37e-22                   9.02e-15
##                      Plazo                  Patrimonio
##                   2.01e-08                   6.11e-05
##         Importe.solicitado                    RatiFin
##                   5.40e-03                   8.92e-03
##       Cargas.patrimoniales Precio.del.bien.financiado
##                   1.63e-01                   1.64e-01
## [1] "P.values per class: 3"
##                   Dictamen                   Vivienda
##                   0.00e+00                   0.00e+00
##               Estado.civil                  Registros
##                   0.00e+00                   0.00e+00
##              Tipo.trabajo                     Gastos
##                   0.00e+00                   3.51e-18
##         Importe.solicitado                       Edad
##                   2.26e-15                   2.60e-14
## Precio.del.bien.financiado                    RatiFin
##                   6.22e-12                   6.67e-03
##                   Ingresos                      Plazo
##                   8.17e-03                   1.39e-02
##                 Patrimonio       Cargas.patrimoniales
##                   1.39e-02                   1.96e-02
##                    Estalvi          Antiguedad.Trabajo
##                   7.74e-02                   2.05e-01
## [1] "P.values per class: 4"
##                   Dictamen          Antiguedad.Trabajo
##                   0.00e+00                   0.00e+00
##                   Vivienda                       Edad
##                   0.00e+00                   0.00e+00
##               Estado.civil                  Registros
```

```
##                   0.00e+00                   0.00e+00
##               Tipo.trabajo                     Gastos
##                   0.00e+00                   0.00e+00
##                   Ingresos Precio.del.bien.financiado
##                   0.00e+00                   3.61e-11
##                  Patrimonio         Importe.solicitado
##                   5.53e-08                   6.59e-07
##       Cargas.patrimoniales                      Plazo
##                   1.88e-04                   7.14e-04
##                    RatiFin                    Estalvi
##                   3.60e-02                   4.99e-01
```

```
#afegir la informacio de les modalitats de les qualitatives a la llista de pvalues i fer ordenacio glob

#saving the dataframe in an external file
#write.table(dd, file = "credscoClean.csv", sep = ";", na = "NA", dec = ".", row.names = FALSE, col.nam
```