

PRACTICA 09, SEMANA 3

Nelson de Jesus Magaña Godinez

16/8/2022

REALICE UN ANÁLISIS ESTADÍSTICO DE LOS DATOS.

- 1) Activa tu directorio de trabajo.

```
getwd()
```

```
## [1] "C:/Users/pc 1/Desktop/PAQUETE R/PRACTICAS_S3"
```

```
setwd("C:/Users/pc 1/Desktop/PAQUETE R/PRACTICAS_S3")
```

- 2) Limpia de objetos el área de trabajo (Workspace).

```
ls()
```

```
## character(0)
```

```
#rm(list=ls(all=TRUE))
```

```
ls()
```

```
## character(0)
```

- 3) Crea un nuevo Script y llámale “Script09-DatosBivariados1”.
- 4) Crea en Excel una hoja de datos con dos columnas o variables

```
# Recuerda que al guardar la hoja, el tipo de archivo es de extensión .csv(delimitado por comas).
```

```
# Llámale al archivo: HojaCat
```

```
# Otra forma de crear la hoja de datos es la siguiente (Vea la Práctica 04):
```

```
# Primero crear las dos variables categóricas en un editor de texto como NotePad o WordPad, colocando nombre a cada columna, y llamándole “HojaCat.txt”.
```

```
# Luego puede leer o recuperar este archivo con la función read.table()
```

```

Estado <- c("casado", "soltero", "soltero", "casado", "acompañado",
            "soltero", "casado", "casado", "acompañado", "acompañado",
            "casado", "soltero", "acompañado", "casado", "soltero",
            "soltero", "casado", "soltero")
Ocupacion <- c("desocupado", "estudia", "trabaja", "estudia", "trabaja",
               "estudia", "desocupado", "estudia", "trabaja", "estudia",
               "desocupado", "desocupado", "estudia", "desocupado",
               "trabaja", "trabaja", "desocupado", "trabaja")
HojaCat.txt<- data.frame(Estado=Estado, Ocupacion=Ocupacion)
HojaCat.txt

```

```

##      Estado  Ocupacion
## 1    casado desocupado
## 2    soltero  estudia
## 3    soltero  trabaja
## 4    casado  estudia
## 5 acompañado trabaja
## 6    soltero  estudia
## 7    casado desocupado
## 8    casado  estudia
## 9 acompañado trabaja
## 10 acompañado estudia
## 11   casado desocupado
## 12   soltero desocupado
## 13 acompañado estudia
## 14   casado desocupado
## 15   soltero  trabaja
## 16   soltero  trabaja
## 17   casado desocupado
## 18   soltero  trabaja

```

5) Recupera desde el entorno de R la hoja de datos de Excel.

```

HojaCat <- read.csv("HojaCat.csv", strip.white = TRUE);
HojaCat

```

```

##      Estado.Ocupación
## 1    casado;desocupado
## 2      soltero;estudia
## 3      soltero;trabaja
## 4      casado;estudia
## 5 acompañado;trabaja
## 6      soltero;desocupado
## 7      casado;trabaja
## 8      casado;estudia
## 9 acompañado;desocupado
## 10 acompañado;estudia
## 11      casado;trabaja
## 12      soltero;estudia
## 13 acompañado;desocupado
## 14      casado;desocupado
## 15      soltero;estudia

```

```
## 16      soltero;trabaja
## 17      casado;desocupado
## 18      soltero;trabaja
```

6) Conecta la hoja de datos a la segunda ruta o lista de búsqueda.

```
attach(HojaCat, pos=2) # pos especifica la posición donde buscar la conexión
search()
```

```
## [1] ".GlobalEnv"      "HojaCat"           "package:stats"
## [4] "package:graphics"  "package:grDevices" "package:utils"
## [7] "package:datasets"  "package:methods"   "Autoloads"
## [10] "package:base"
```

7) Crea una tabla de contingencia o de doble entrada

```
tablaCont <- table(HojaCat);
tablaCont
```

```
## Estado.Ocupación
## acompañado;desocupado    acompañado;estudia    acompañado;trabaja
##                2                1                1
##      casado;desocupado      casado;estudia      casado;trabaja
##                3                2                2
##      soltero;desocupado      soltero;estudia      soltero;trabaja
##                1                3                3
```

```
length(HojaCat)
```

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

```
# Note que esta instrucción no devuelve el número de elementos, sino
# más bien el número de variables o columnas consideradas en el
# conjunto de datos
# Encuentra la suma de cada fila de la tabla de contingencia
# Distribución marginal de X=Estado civil
suma.filas <- apply(tablaCont, 1, sum);
suma.filas
```

```
## acompañado;desocupado    acompañado;estudia    acompañado;trabaja
##                2                1                1
##      casado;desocupado      casado;estudia      casado;trabaja
##                3                2                2
##      soltero;desocupado      soltero;estudia      soltero;trabaja
##                1                3                3
```

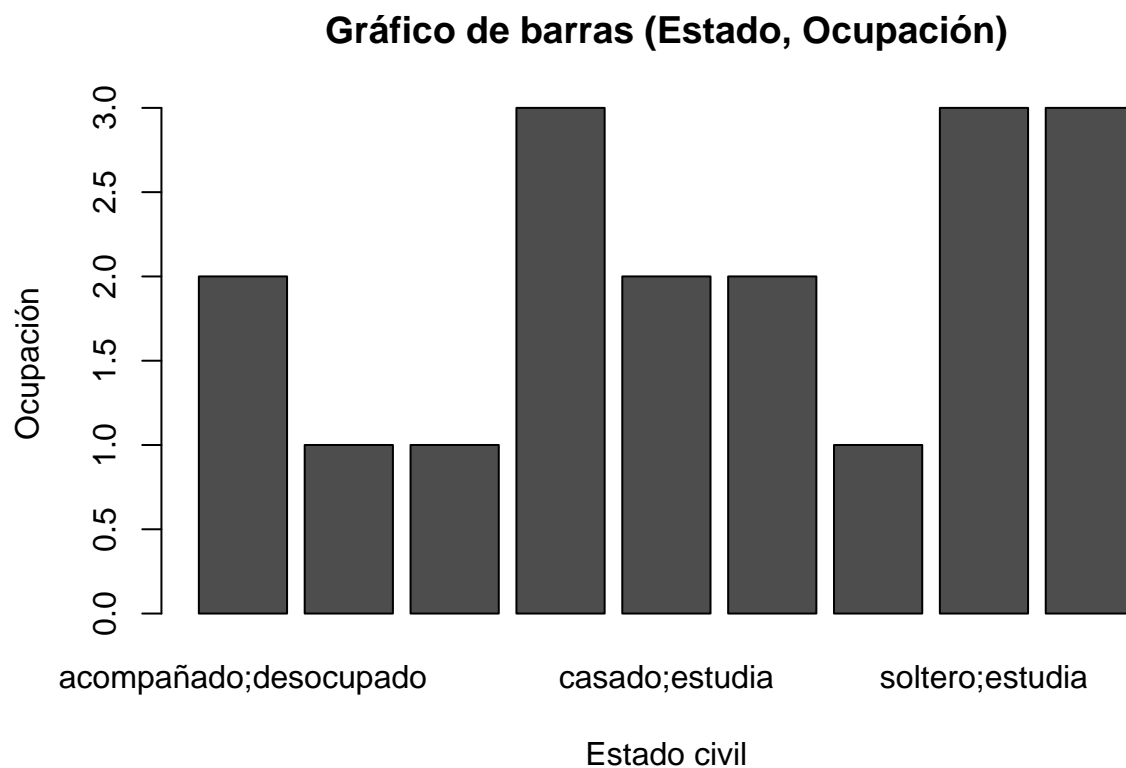
Encuentra la suma de cada fila de la tabla de contingencia # distribución marginal de Y=Ocupación

```
suma.columnas <- apply(tablaCont, 1, sum);
suma.columnas
```

```
## acompañado;desocupado    acompañado;estudia    acompañado;trabaja
##                2                1                1
##    casado;desocupado      casado;estudia      casado;trabaja
##                3                2                2
##    soltero;desocupado    soltero;estudia    soltero;trabaja
##                1                3                3
```

2 indica que son totales por columna # Gráficos de barras para tabla de contingencia. # Barras apiladas

```
barplot(t(tablaCont), main = "Gráfico de barras (Estado, Ocupación)",
        xlab = "Estado civil", ylab = "Ocupación",
        legend.text = TRUE)
```

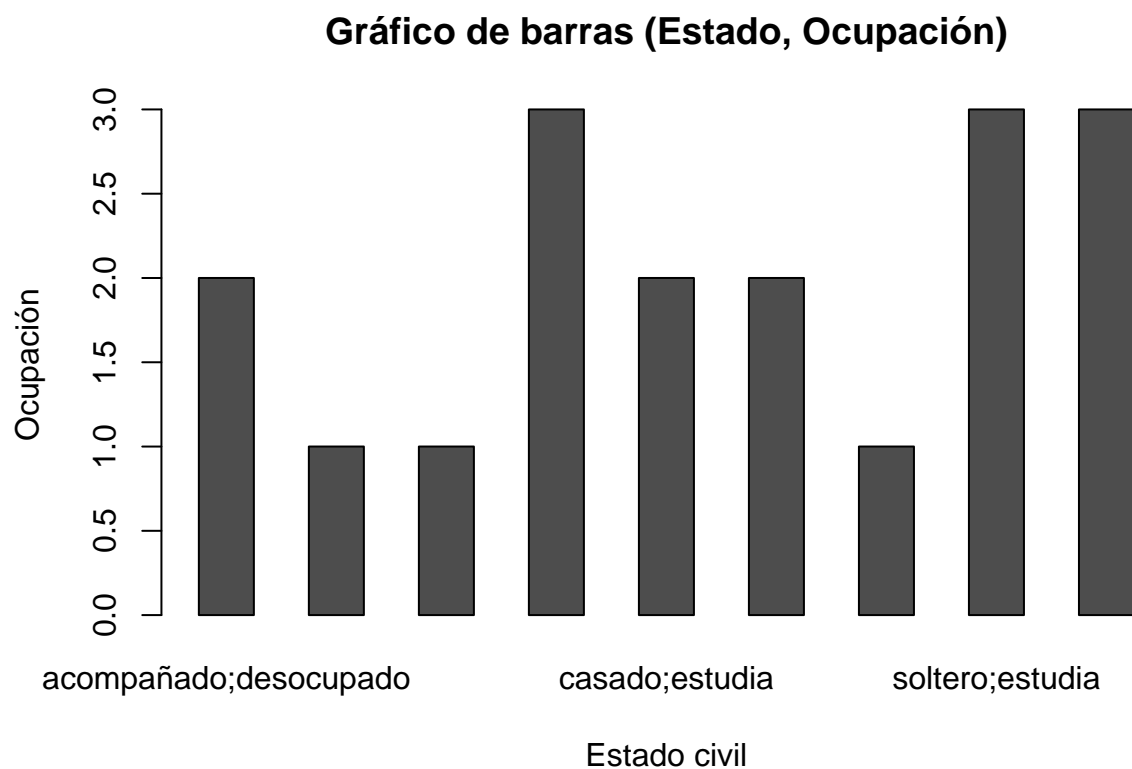


Note que `t(tablaCont)` indica que las barras representan el Estado civil de los encuestados y que éstas se subdividen en cada una de las diferentes ocupaciones consideradas.

En caso de usar únicamente `tablaCont`; las barras representarán las diferentes ocupaciones y éstas estarán subdivididas en cada uno de los estados civiles.

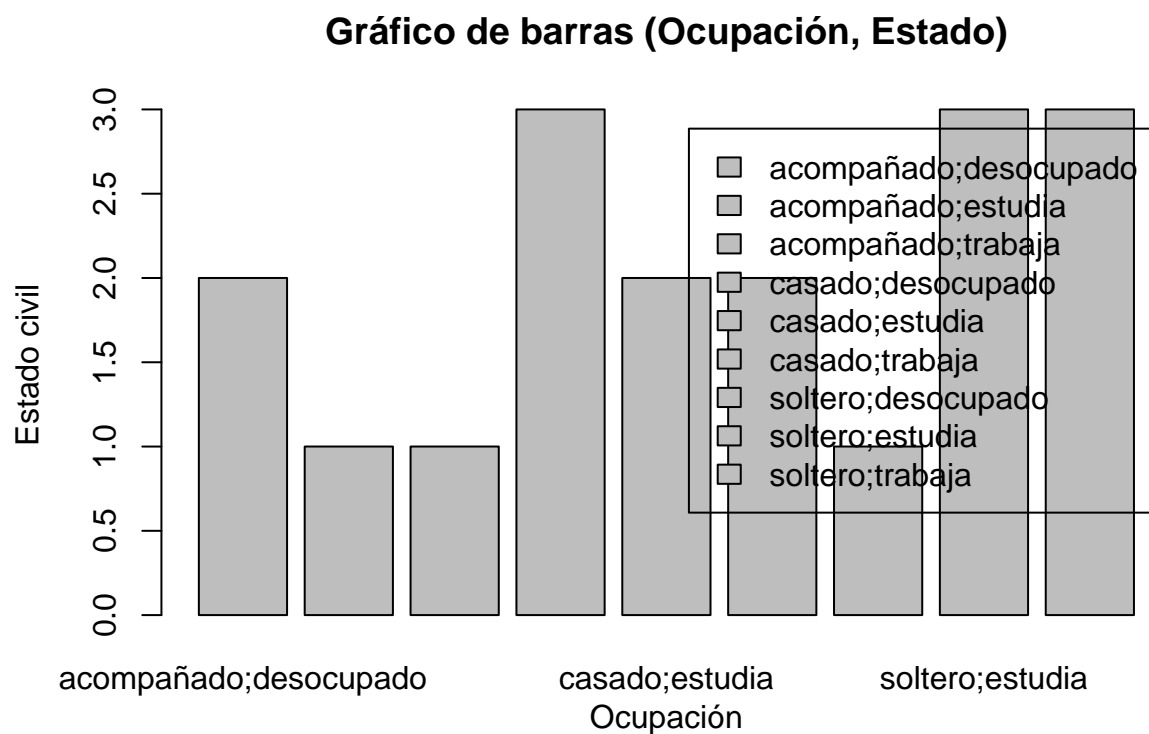
Barras agrupadas

```
barplot(t(tablaCont), main = "Gráfico de barras (Estado, Ocupación)",
        xlab = "Estado civil", ylab = "Ocupación", beside = TRUE,
        legend.text = TRUE)
```



Note que la instrucción `beside = TRUE`, indica que por cada una de las diferentes ocupaciones se creará una barra para cada estado civil. Note que al usar `beside = FALSE` se obtiene el mismo gráfico de la instrucción anterior.

```
barplot(tablaCont, main = "Gráfico de barras (Ocupación, Estado)",
        xlab = "Ocupación\n", ylab = "Estado civil",
        beside = TRUE, legend.text = TRUE)
```



8) Calcula tablas de proporciones o de probabilidades.

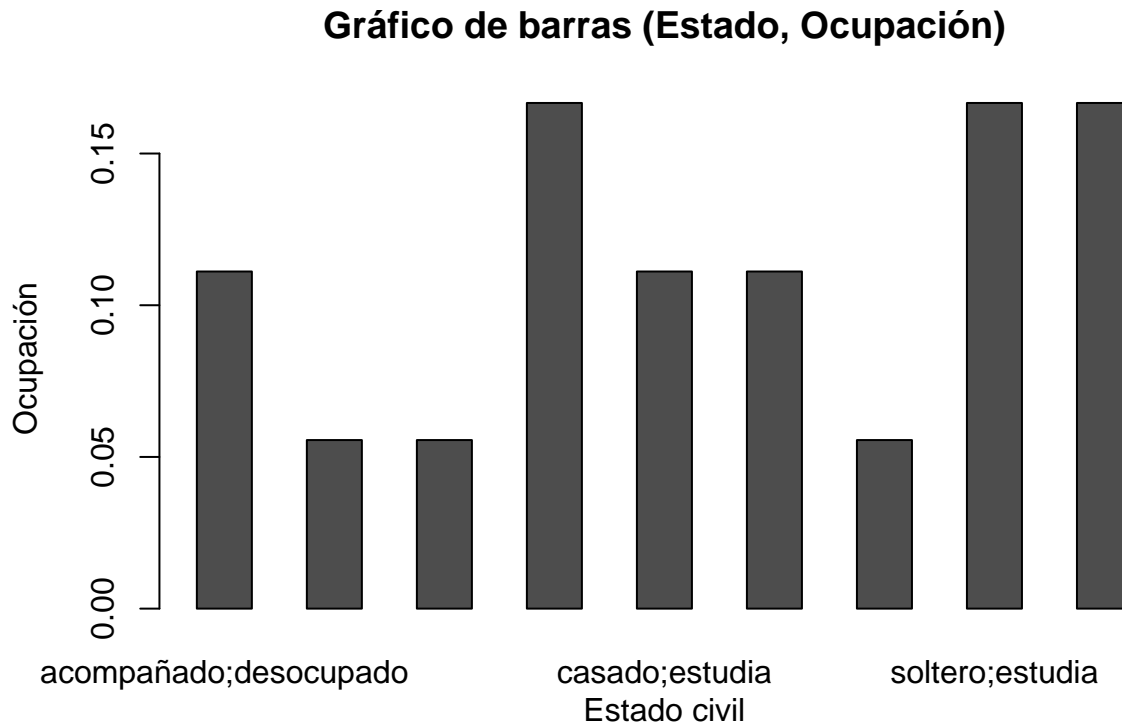
```
# Guardar las todas las opciones iniciales y modificar número de decimales
op <- options()
options(digits = 3) # sólo imprime 3 lugares decimales
options('digits')
```

```
## $digits
## [1] 3
```

```
# Proporciones basadas en el total de la muestra, la suma de filas y columnas suman 1.
propTotal <- prop.table(tablaCont);
propTotal
```

```
## Estado.Ocupación
## acompañado;desocupado    acompañado;estudia    acompañado;trabaja
##                0.1111                0.0556                0.0556
##    casado;desocupado      casado;estudia      casado;trabaja
##                0.1667                0.1111                0.1111
##    soltero;desocupado    soltero;estudia    soltero;trabaja
##                0.0556                0.1667                0.1667
```

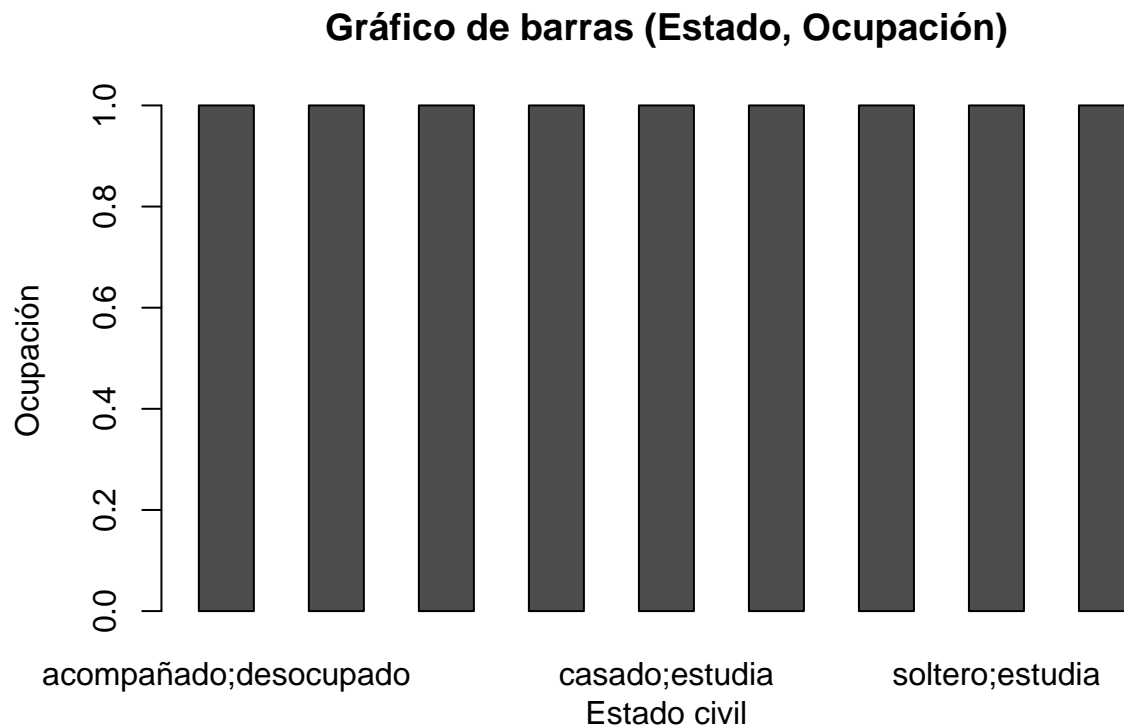
```
barplot(t(propTotal), main = "Gráfico de barras (Estado, Ocupación)",
        xlab = "Estado civil\n", ylab = "Ocupación",
        beside = TRUE, legend.text = TRUE)
```



```
# Proporciones basadas en el total por fila, cada fila suma 1.
propFila <- prop.table(tablaCont, 1);
propFila
```

```
## Estado.Ocupación
## acompañado;desocupado    acompañado;estudia    acompañado;trabaja
##                1                1                1
##    casado;desocupado      casado;estudia      casado;trabaja
##                1                1                1
##    soltero;desocupado    soltero;estudia    soltero;trabaja
##                1                1                1
```

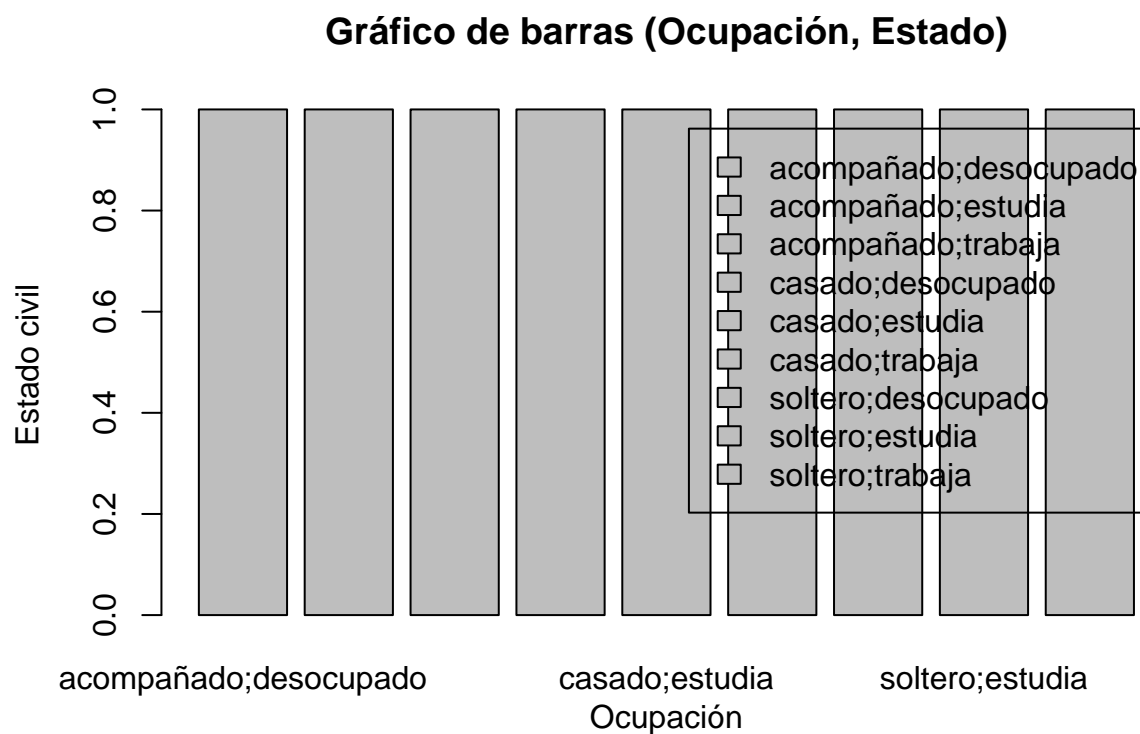
```
# Total por fila se indica en 1
barplot(t(propFila), main = "Gráfico de barras (Estado, Ocupación)",
        xlab = "Estado civil\n", ylab = "Ocupación", beside = TRUE,
        legend.text = TRUE)
```



```
# Proporciones basadas en el total por columna, cada columna suma 1.
propColum <- prop.table(tablaCont, 1);
propColum
```

```
## Estado.Ocupación
## acompañado;desocupado    acompañado;estudia    acompañado;trabaja
##                1                1                1
##    casado;desocupado      casado;estudia      casado;trabaja
##                1                1                1
##    soltero;desocupado    soltero;estudia    soltero;trabaja
##                1                1                1
```

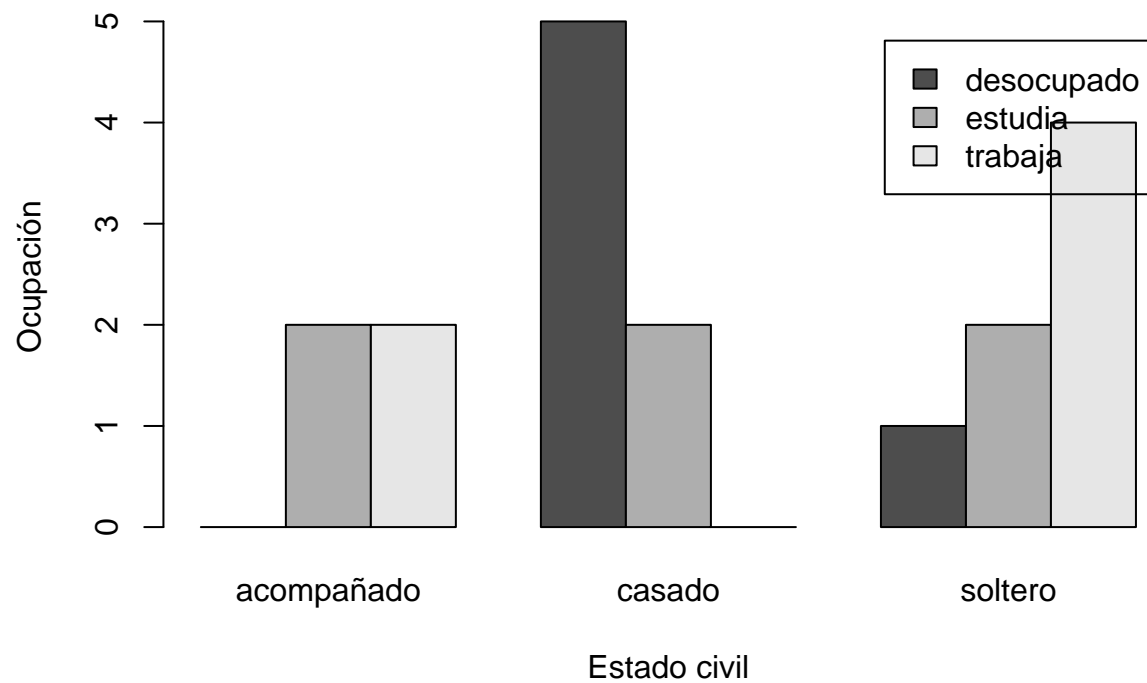
```
# Total por columna se indica en 2
barplot(propColum, main = "Gráfico de barras (Ocupación, Estado)",
        xlab = "Ocupación\n", ylab = "Estado civil", beside = TRUE,
        legend.text = TRUE)
```

9) Otra forma de elaborar los gráficos de barras para el vector bidimensional categórico.

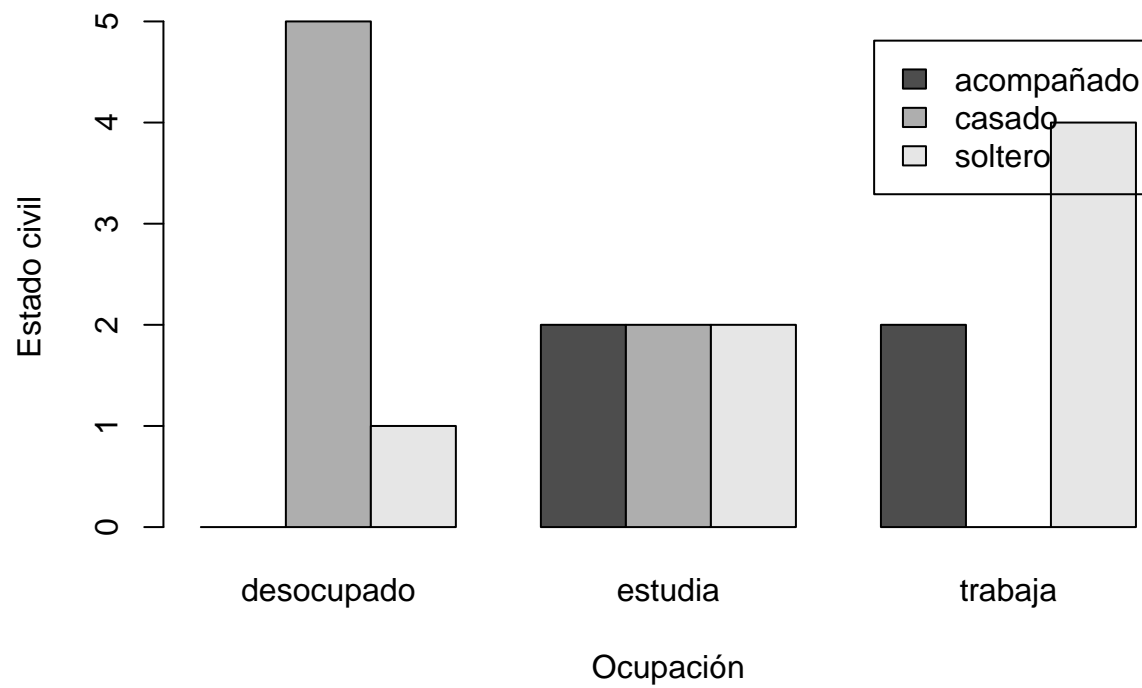
```
# Gráfico de barras no apiladas y colocación de leyenda
barplot(table(Ocupacion, Estado),
  main="Gráfico de barras (Estado, Ocupación)",
  xlab = "Estado civil", ylab = "Ocupación", beside = TRUE,
  legend.text = T)
```

Gráfico de barras (Estado, Ocupación)



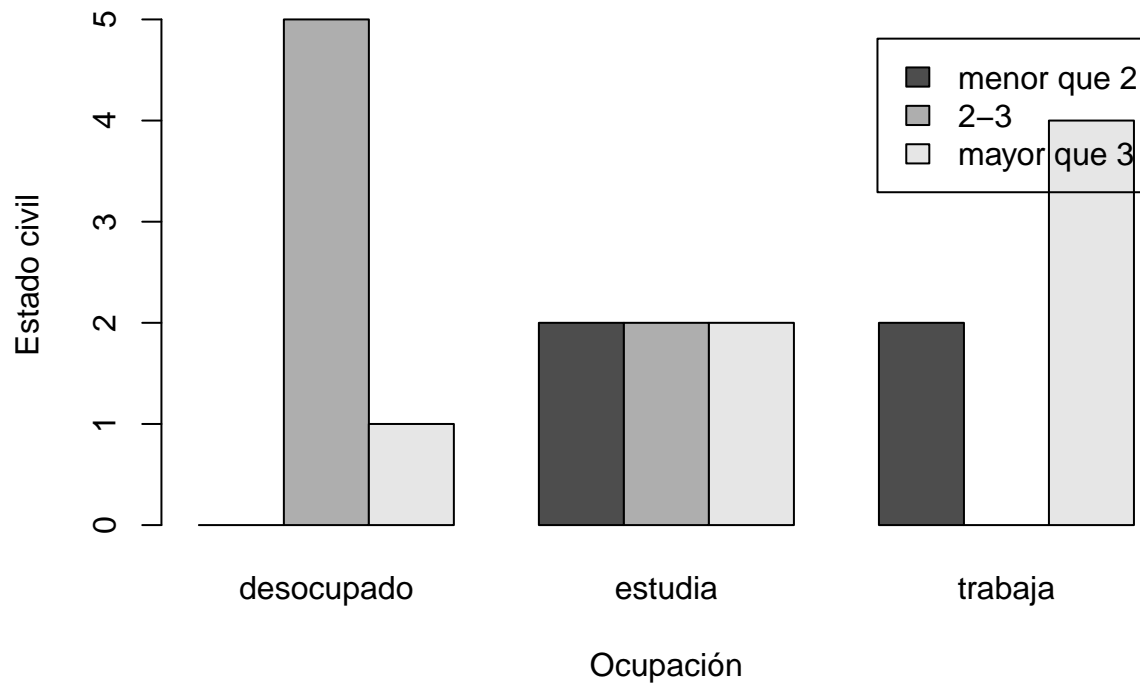
```
barplot(table(Estado, Ocupacion),  
        main = "Gráfico de barras (Ocupación, Estado)",  
        xlab = "Ocupación", ylab = "Estado civil", beside = TRUE,  
        legend.text = TRUE)
```

Gráfico de barras (Ocupación, Estado)



```
barplot(table(Estado, Ocupacion),  
  main = "Gráfico de barras (Ocupación, Estado)",  
  xlab = "Ocupación", ylab = "Estado civil", beside = TRUE,  
  legend.text = c("menor que 2", "2-3", "mayor que 3"))
```

Gráfico de barras (Ocupación, Estado)



*# Note que se puede definir a conveniencia
la leyenda que se desea incorporar en el gráfico con
la instrucción legend.text*

10) Realizar la prueba o contraste Chi-cuadrado de independencia

```
prueba <- chisq.test(tablaCont);
```

```
## Warning in chisq.test(tablaCont): Chi-squared approximation may be incorrect
```

```
prueba
```

```
##
## Chi-squared test for given probabilities
##
## data:  tablaCont
## X-squared = 3, df = 8, p-value = 0.9
```

*# Tenga en cuenta que las frecuencias esperadas deben ser todas mayores a 5
Frecuencias absolutas esperadas para la prueba Chi-cuadrada
prueba\$expected # fij = fi./No. column*

```
## acompañado;desocupado    acompañado;estudia    acompañado;trabaja
```

##	2	2	2
##	casado;desocupado	casado;estudia	casado;trabaja
##	2	2	2
##	soltero;desocupado	soltero;estudia	soltero;trabaja
##	2	2	2