

# Práctica 1: R Markdown

Escribiendo informes reproducibles

6 de Junio, 2017

## Enunciado

La **privación material** es un concepto sociológico que se corresponde con la *'falta de bienes, servicios, recursos y comodidades que son habituales, o están ampliamente extendidos, en una sociedad determinada'*. Un problema habitual a la hora de estudiar este concepto es que la privación no puede ser medida directamente y, por tanto, tiene que ser medida de forma indirecta a partir de otras variables que sean reflejo de la presencia o ausencia de este concepto en la población.

En el archivo de datos `datos_practicas.RData` encontrarás el banco de datos `privacion_valencia` y el objeto espacial `carto_valencia` con la cartografía de la ciudad de Valencia. El banco de datos `privacion_valencia` contiene una serie de indicadores sociales para las 598 secciones censales de la ciudad de Valencia. Estos indicadores son en este orden: desempleo, instrucción insuficiente, instrucción insuficiente en jóvenes (16-29 años), trabajadores manuales, envejecimiento, trabajadores eventuales, residentes extranjeros, hogares mono-parentales, problemas de ruido, problemas de contaminación, problemas de limpieza, problemas de comunicación, problemas de zonas verdes, problemas de delincuencia y problemas de aseo. Todas las variables han sido medidas como el porcentaje de población (en cada sección) que pertenece a estos colectivos o que dice observar dichos problemas. Todos estos indicadores podrían ser reflejo de la privación económica de cada una de las secciones censales de la ciudad de Valencia.

## Tareas

1. Crea un proyecto de RStudio en un nuevo directorio con el nombre `practica_01_rmarkdown`. Este directorio debe tener una estructura de carpetas adecuada para la realización de esta práctica.
2. Descarga el archivo de datos `datos_practicas.RData` del repositorio *figshare* cuya url es <https://ndownloader.figshare.com/files/8455721>, guárdalo en la carpeta adecuada del directorio de tu proyecto y carga los datos en R.
3. Construye a partir de la información disponible en el banco de datos `privacion_valencia` un índice de privación que cuantifique este factor en cada sección censal de la ciudad de Valencia. Para ello, debes realizar un análisis de componentes principales utilizando la función `princomp` de R. Resume los resultados obtenidos mediante:
  - Una tabla con la desviación estándar, la proporción de varianza explicada y la varianza acumulada de las componentes principales.
  - Otra tabla con los pesos de las variables originales en cada una de las componentes.
4. Representa gráficamente las 4 primeras componentes principales del análisis en el mapa de la ciudad de Valencia. La primera componente principal resumirá la privación material de cada sección censal. Considera 5 grupos para categorizar las variables y utiliza la siguiente función para realizar la representación gráfica:

```
pinta_mapas <- function(cartografia, variable, n_grupos) {  
  if (!require(RColorBrewer)){  
    install.packages("RColorBrewer")  
    library(RColorBrewer)  
  }  
  if (!require(sp)){
```

```

install.packages("sp")
library(sp)
}
paleta <- brewer.pal(n_grupos, "BrBG")
grupos <- quantile(variable, probs = seq(0, 1, 1 / n_grupos))
pcorte <- c(grupos[1] - 0.5, grupos[2:n_grupos], grupos[n_grupos + 1] + 0.5)
leyenda <- c()
for (j in 2:length(pcorte)){
  leyenda[j] <- paste0(round(pcorte[j - 1], 2), " - ", round(pcorte[j], 2))
}
plot(cartografia, col = paleta[findInterval(variable[match(cartografia@data$CUSEC,
                                                             privacion_valencia[, 1])], pcorte)])
legend("bottomright", leyenda[-1], title = "Índice de privación",
       border = NULL, fill = paleta, bty = "n")
}

```

5. Genera un informe PDF que contenga el código de R y los resultados obtenidos utilizando R Markdown. No olvides tener en cuenta las normas de sintaxis estudiadas en la sesión 2 de este curso.

## Resolución

```
# Establezco el directorio raíz del documento y las opciones globales del código
knitr::opts_knit$set(root.dir = normalizePath("../"))
knitr::opts_chunk$set(echo = TRUE, message = FALSE, warning = FALSE)

# Tarea 1
## Establezco la estructura de directorios del proyecto
if (!dir.exists("datos")) dir.create("datos")
if (!dir.exists("informes")) dir.create("informes")

# Tarea 2
## Descargo los datos de figshare
if (!file.exists("datos/datos_practicas.RData")) {
  file_url <- "https://ndownloader.figshare.com/files/8455721"
  download.file(file_url, destfile = "datos/datos_practicas.RData", mode = "wb")
}

## Cargo las librerías necesarias y los datos
library(pander)
load("datos/datos_practicas.RData")

# Tarea 3
## Realizo ACP
ACP <- princomp(privacion_valencia[, 2:16])

## Resumen del análisis
standard_dev <- ACP$sdev
prop_variance <- (ACP$sdev ^ 2) / sum(ACP$sdev ^ 2)
cumulative_prop <- cumsum(prop_variance)

tab_summary <- data.frame(standard_dev, prop_variance, cumulative_prop)
colnames(tab_summary) <- c("Desviación estandar", "Proporción de varianza explicada",
  "Varianza acumulada")

pandoc.table(tab_summary, justify = 'centre', round = 3,
  caption = "Resumen del análisis de componentes principales")
```

Tabla 1: Resumen del análisis de componentes principales

|         | Desviación estandar | Proporción de varianza<br>explicada | Varianza acumulada |
|---------|---------------------|-------------------------------------|--------------------|
| Comp.1  | 26.43               | 0.377                               | 0.377              |
| Comp.2  | 19.47               | 0.205                               | 0.581              |
| Comp.3  | 16.8                | 0.152                               | 0.734              |
| Comp.4  | 12.38               | 0.083                               | 0.816              |
| Comp.5  | 11.55               | 0.072                               | 0.888              |
| Comp.6  | 8.82                | 0.042                               | 0.93               |
| Comp.7  | 7.027               | 0.027                               | 0.957              |
| Comp.8  | 5.028               | 0.014                               | 0.971              |
| Comp.9  | 3.883               | 0.008                               | 0.979              |
| Comp.10 | 3.71                | 0.007                               | 0.986              |
| Comp.11 | 2.865               | 0.004                               | 0.991              |

|         | Desviación estandar | Proporción de varianza<br>explicada | Varianza acumulada |
|---------|---------------------|-------------------------------------|--------------------|
| Comp.12 | 2.512               | 0.003                               | 0.994              |
| Comp.13 | 2.337               | 0.003                               | 0.997              |
| Comp.14 | 2.109               | 0.002                               | 1                  |
| Comp.15 | 0.959               | 0                                   | 1                  |

```
## Peso de las variables originales en las componentes principales
panderOptions('table.continues', "Continuación.")
panderOptions('table.continues.affix', "(continúa en la página siguiente).")
pandoc.table(as.table(ACP$loadings), justify = 'centre', round = 3,
caption = "Peso de las variables originales en las componentes principales.")
```

Tabla 2: Peso de las variables originales en las componentes principales. (continúa en la página siguiente).

|             | Comp.1 | Comp.2 | Comp.3 | Comp.4 | Comp.5 | Comp.6 | Comp.7 |
|-------------|--------|--------|--------|--------|--------|--------|--------|
| desempt     | -0.073 | -0.088 | 0.049  | 0.076  | -0.038 | -0.001 | -0.053 |
| insufict16  | -0.345 | -0.379 | 0.034  | 0.198  | -0.121 | 0.131  | -0.272 |
| inst1629    | -0.156 | -0.159 | 0.042  | 0.088  | -0.035 | 0.048  | -0.027 |
| manualest   | -0.524 | -0.506 | -0.024 | -0.005 | -0.187 | 0.02   | 0.361  |
| envejet     | 0.002  | -0.03  | 0.069  | 0.316  | 0.046  | 0.209  | -0.669 |
| eventualt   | -0.139 | -0.151 | 0.046  | 0.113  | -0.057 | 0.006  | -0.051 |
| resexj      | -0.01  | -0.008 | 0.018  | 0.061  | -0.036 | -0.061 | -0.054 |
| monomart    | 0.024  | 0.01   | 0.076  | 0.136  | 0.021  | 0.022  | -0.23  |
| ruido       | 0.047  | 0.259  | 0.176  | -0.003 | -0.636 | 0.176  | 0.238  |
| contamina   | -0.05  | 0.246  | 0.113  | 0.039  | -0.611 | 0.137  | -0.208 |
| limpieza    | -0.258 | 0.12   | 0.144  | -0.117 | -0.151 | -0.883 | -0.246 |
| comunica    | -0.2   | -0.011 | -0.236 | -0.842 | -0.04  | 0.236  | -0.34  |
| zverdes     | -0.605 | 0.603  | -0.393 | 0.244  | 0.188  | 0.11   | 0.078  |
| delicuencia | -0.282 | 0.195  | 0.84   | -0.17  | 0.32   | 0.188  | 0.066  |
| aseo        | -0.01  | 0.004  | 0.009  | 0.009  | -0.005 | 0.001  | -0.028 |

Tabla 3: Continuación.

|             | Comp.8 | Comp.9 | Comp.10 | Comp.11 | Comp.12 | Comp.13 |
|-------------|--------|--------|---------|---------|---------|---------|
| desempt     | -0.047 | -0.155 | -0.05   | 0.666   | 0.703   | 0.082   |
| insufict16  | -0.157 | 0.088  | 0.506   | -0.328  | 0.248   | -0.066  |
| inst1629    | -0.141 | -0.716 | 0.271   | 0.01    | -0.261  | 0.229   |
| manualest   | 0.166  | 0.227  | -0.252  | 0.256   | -0.304  | 0.056   |
| envejet     | -0.099 | 0.385  | -0.172  | 0.172   | -0.224  | 0.155   |
| eventualt   | -0.109 | -0.186 | -0.584  | -0.435  | 0.302   | -0.432  |
| resexj      | -0.01  | -0.14  | -0.401  | -0.277  | 0.113   | 0.782   |
| monomart    | -0.185 | -0.389 | -0.253  | 0.29    | -0.354  | -0.322  |
| ruido       | -0.618 | 0.143  | -0.006  | 0.028   | -0.034  | 0.043   |
| contamina   | 0.678  | -0.148 | 0.024   | -0.025  | 0.012   | -0.061  |
| limpieza    | -0.098 | 0.066  | 0.044   | 0.021   | -0.048  | -0.018  |
| comunica    | -0.116 | -0.048 | -0.076  | 0.017   | 0.015   | 0.016   |
| zverdes     | -0.044 | -0.012 | -0.003  | 0.002   | 0.023   | -0.002  |
| delicuencia | 0.054  | 0.024  | 0.007   | -0.03   | 0.014   | 0.017   |

|             | Comp.8 | Comp.9 | Comp.10 | Comp.11 | Comp.12 | Comp.13 |
|-------------|--------|--------|---------|---------|---------|---------|
| <b>aseo</b> | -0.017 | 0.003  | -0.012  | 0.022   | 0.012   | 0.023   |

|                    | Comp.14 | Comp.15 |
|--------------------|---------|---------|
| <b>desempt</b>     | -0.038  | 0.029   |
| <b>insufict16</b>  | 0.359   | 0.003   |
| <b>inst1629</b>    | -0.457  | 0.003   |
| <b>manualest</b>   | -0.003  | -0.002  |
| <b>envejet</b>     | -0.331  | 0.032   |
| <b>eventualt</b>   | -0.285  | -0.003  |
| <b>resexj</b>      | 0.325   | 0.019   |
| <b>monomart</b>    | 0.599   | 0.008   |
| <b>ruido</b>       | -0.033  | 0.011   |
| <b>contamina</b>   | 0.021   | -0.003  |
| <b>limpieza</b>    | -0.057  | 0.011   |
| <b>comunica</b>    | 0.002   | 0.006   |
| <b>zverdes</b>     | 0.016   | 0.005   |
| <b>delicuencia</b> | 0.014   | 0.005   |
| <b>aseo</b>        | -0.001  | -0.999  |

```
## Construcción de las componentes principales
Ind_privacion <- list()

for (i in 1:15) {
  Ind_privacion[[i]] <- as.matrix(privacion_valencia[, 2:16]) %*% ACP$loadings[, i]
}
```

Representación gráfica de las 4 primeras componentes principales del análisis.

```
# Tarea 4
## Función para realizar la representación gráfica de una variable en un mapa
pinta_mapas <- function(cartografia, variable, n_grupos) {
  if (!require(RColorBrewer)) {
    install.packages("RColorBrewer")
    library(RColorBrewer)
  }
  if (!require(sp)) {
    install.packages("sp")
    library(sp)
  }
  paleta <- brewer.pal(n_grupos, "BrBG")
  grupos <- quantile(variable, probs = seq(0, 1, 1 / n_grupos))
  pcorte <- c(grupos[1] - 0.5, grupos[2:n_grupos], grupos[n_grupos + 1] + 0.5)
  colores <- paleta[
    findInterval(variable[match(cartografia@data$CUSEC, privacion_valencia[, 1])], pcorte)
  ]
  leyenda <- c()
  for (j in 2:length(pcorte)){
    leyenda[j] <- paste0(round(pcorte[j - 1], 2), " - ", round(pcorte[j], 2))
  }
}
```

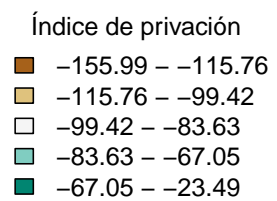
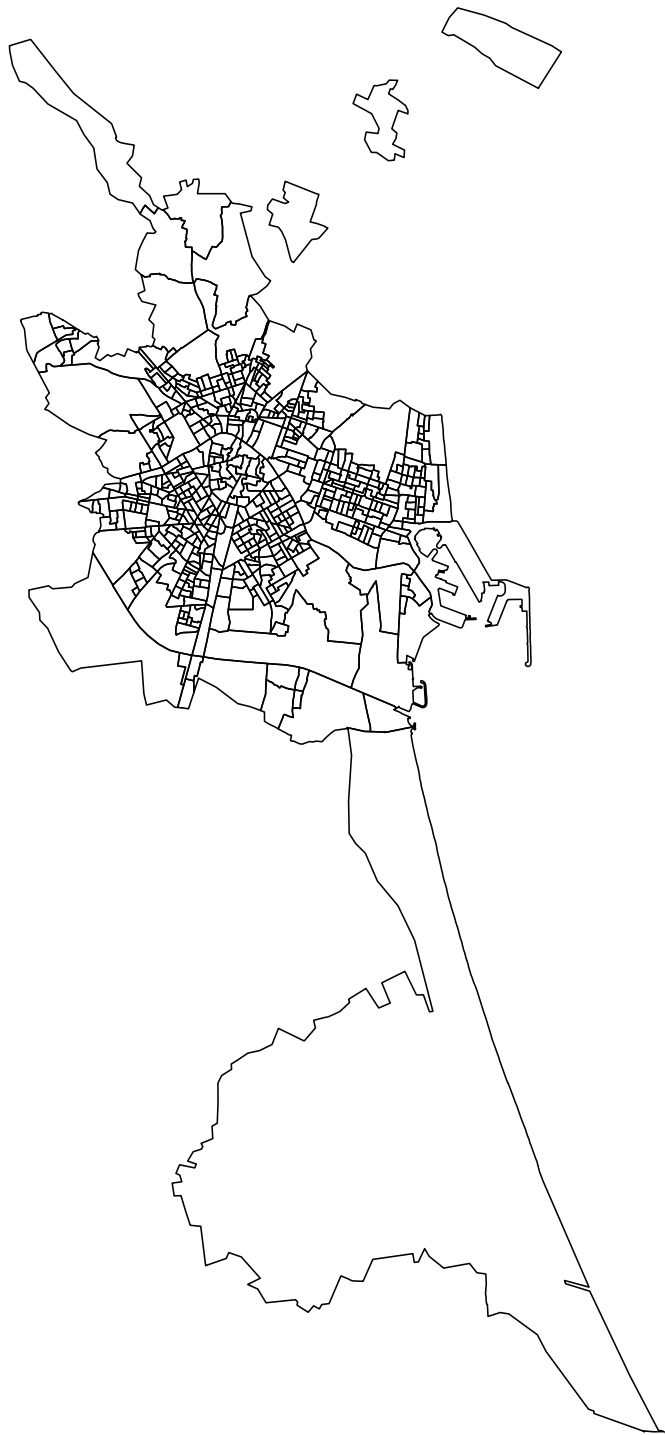
```

    }
    plot(cartografia, col = colores)
    legend("bottomright", leyenda[-1], title = "Índice de privación",
          border = NULL, fill = paleta, bty = "n")
}

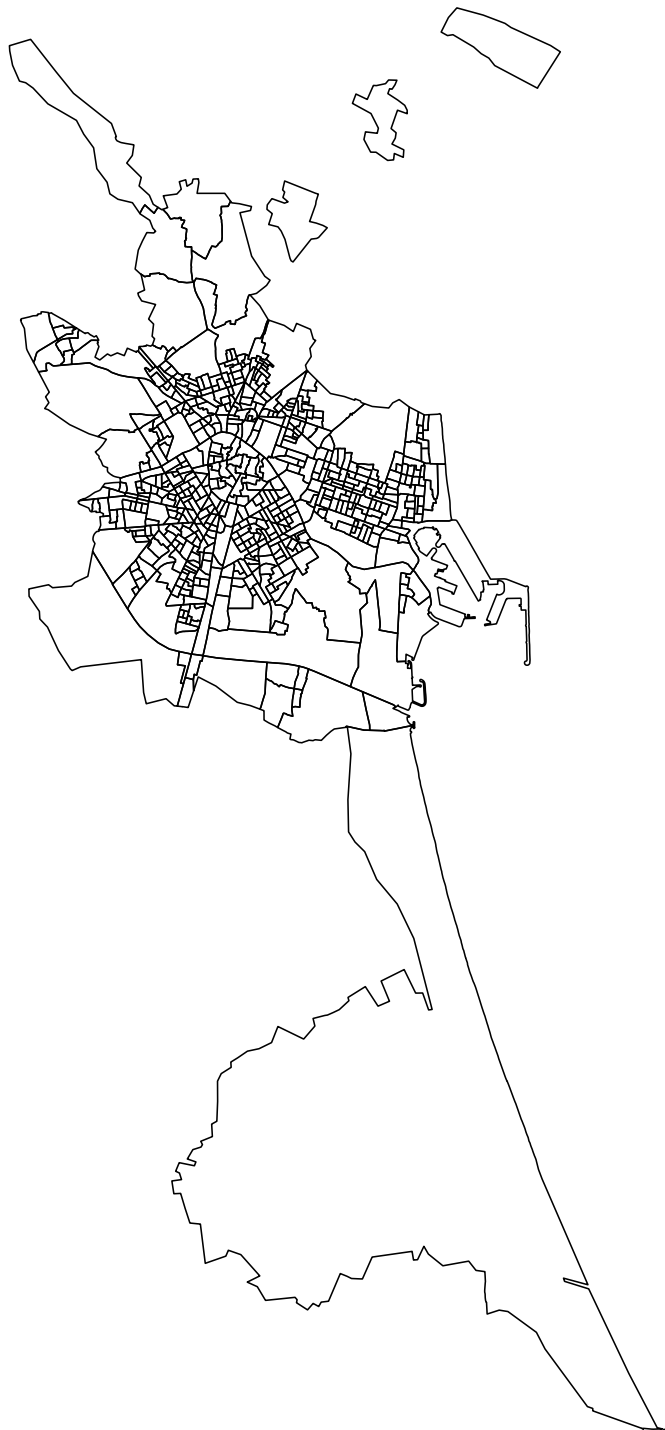
for (i in 1:4){
  cat("## Componente principal", i)
  pinta_mapas(carto_valencia, Ind_privacion[[i]], 5)
}

```

## Componente principal 1



## Componente principal 2



Índice de privación

- -39.35 – 1.32
- 1.32 – 13.55
- 13.55 – 22.51
- 22.51 – 31.86
- 31.86 – 72.97



### Componente principal 3



Índice de privación

- -30.13 – 28.01
- 28.01 – 35.84
- 35.84 – 44.49
- 44.49 – 55.58
- 55.58 – 86.27

## Componente principal 4

