

Reto — Sesión 14: panel de ventas de una cadena de librerías

Curso de R — DataFrames y visualización (Bloques 7 y 8)

Contexto

Trabajas como analista de datos para una cadena de librerías con tiendas en cuatro regiones. Cada mes se registran las ventas de cada tienda. Tu reto es analizar estos datos y construir un pequeño panel de visualización, usando únicamente lo visto en esta sesión: creación y manipulación de data frames, `aggregate()`, `order()`, `write.csv()` y los gráficos base (`hist`, `boxplot`, `barplot`, `pie`) organizados con `par(mfrow)`.

Punto de partida: genera los datos

Copia y ejecuta este código para generar el dataset con el que vas a trabajar (la semilla fija asegura que todos obtengáis los mismos resultados):

```
set.seed(123)
tiendas <- c("Norte", "Sur", "Este", "Oeste")
meses   <- c("Enero", "Febrero", "Marzo", "Abril", "Mayo", "Junio")

ventas <- data.frame(
  tienda = sample(tiendas, 150, replace = TRUE),
  mes     = sample(meses, 150, replace = TRUE),
  importe = round(rnorm(150, mean = 3000, sd = 800), 2)
)

# Factor con el orden logico de los meses
ventas$mes <- factor(ventas$mes, levels = meses)
```

Tareas a resolver

1. Exploración inicial

Inspecciona el data frame `ventas` con las funciones que conoces (`nrow`, `dim`, `str`, `summary`) y responde: ¿cuántos registros hay en total? ¿cuántas tiendas y meses distintos aparecen?

2. Filtrado

Filtra únicamente las ventas de las tiendas Norte o Sur que superen los 2500 euros de importe. ¿Cuántos registros cumplen esa condición?

3. Totales con `aggregate()`

Calcula el importe total de ventas agrupado por tienda y mes con `aggregate()`. Guarda el resultado en una variable llamada `resumen`.

4. Ordenar y exportar

Ordena `resumen` primero por tienda y después por mes (recuerda que `mes` ya es un factor con orden lógico). Exporta el resultado ordenado a un fichero `resumen_ventas.csv` con `write.csv()`, sin incluir los

nombres de fila.

5. Panel de 4 gráficos (2x2)

Usa `par(mfrow = c(2, 2))` para construir un panel con los siguientes cuatro gráficos, en este orden:

- Un histograma de la columna `importe`.
- Un boxplot del `importe` agrupado por `tienda` (`importe ~ tienda`).
- Un gráfico de barras con la media de `importe` por `tienda` (apóyate en `tapply()`).
- Un gráfico de pastel (`pie`) con la proporción de registros por `tienda` (apóyate en `table()`).

No olvides restaurar el layout con `par(mfrow = c(1, 1))` al terminar, y añade títulos descriptivos (`main`) a cada gráfico.

Reto extra (opcional)

Añade colores distintos a cada `tienda` en el boxplot y en el `pie` usando `rainbow(4)`, y una leyenda (`legend()`) en el gráfico de barras que identifique cada `tienda`.