

# Anomalías del volumen de agua embalsada en España

## Actividad Guiada 3

### Introducción

El objetivo de esta actividad es analizar las anomalías del volumen de agua embalsada en España utilizando datos oficiales del MITECO. Se calcula el llenado de embalses como porcentaje de su capacidad, y se categorizan las observaciones según su posición en los percentiles históricos para el mismo mes. Se visualiza la evolución tanto nacional como por cuenca hidrográfica y se analiza un año hidrológico específico.

### Carga de datos

```
if (!require("RODBC")) install.packages("RODBC")
```

Cargando paquete requerido: RODBC

```
url <- "https://www.miteco.gob.es/content/dam/miteco/es/agua/temas/evaluacion-de-los-recursos"
tempf <- tempfile()
download.file(url, tempf, mode = "wb")
unzip(tempf, exdir = tempdir())

library(RODBC)
mdb_path <- list.files(tempdir(), pattern = "\\*.mdb$", full.names = TRUE)
conn <- odbcConnectAccess2007(mdb_path)
embalses <- sqlFetch(conn, "T_Datos Embalses 1988-2025")
close(conn)
```

## Tratamiento y cálculo de percentiles

```
library(dplyr)
```

Warning: package 'dplyr' was built under R version 4.4.3

Adjuntando el paquete: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(lubridate)
```

Warning: package 'lubridate' was built under R version 4.4.3

Adjuntando el paquete: 'lubridate'

The following objects are masked from 'package:base':

date, intersect, setdiff, union

```
library(Hmisc)
```

Warning: package 'Hmisc' was built under R version 4.4.3

Adjuntando el paquete: 'Hmisc'

The following objects are masked from 'package:dplyr':

src, summarize

The following objects are masked from 'package:base':

format.pval, units

```
library(ggplot2)
```

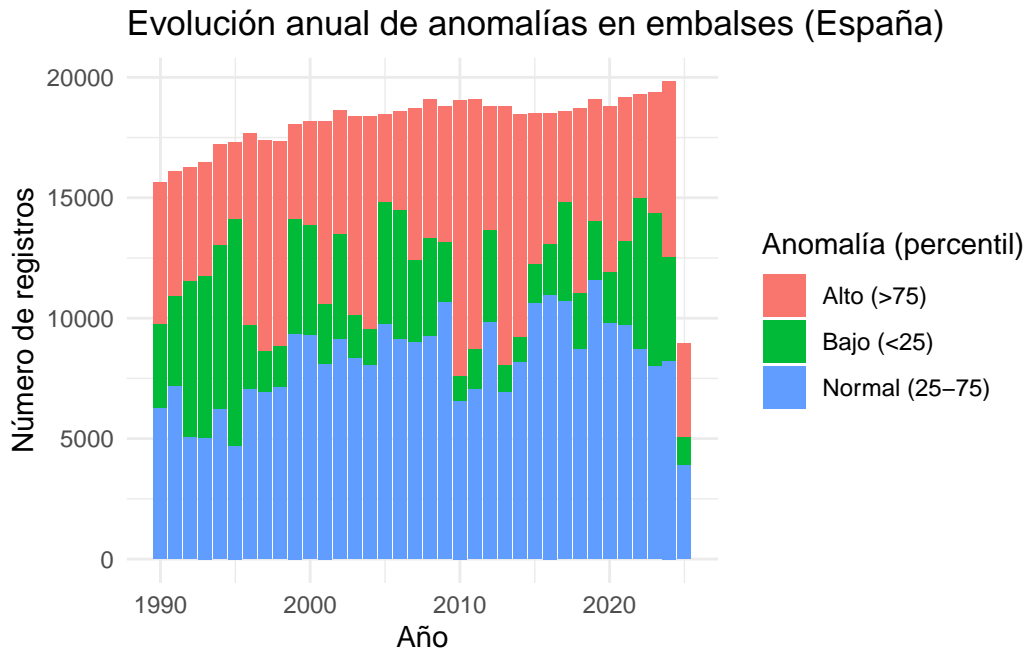
Warning: package 'ggplot2' was built under R version 4.4.3

```
embalses <- embalses %>%
  mutate(
    AGUA_ACTUAL = as.numeric(gsub(",", ".", AGUA_ACTUAL)),
    AGUA_TOTAL = as.numeric(gsub(",", ".", AGUA_TOTAL)),
    FECHA = as.Date(FECHA),
    año = year(FECHA),
    mes = month(FECHA),
    porcentaje = round((AGUA_ACTUAL / AGUA_TOTAL) * 100, 2)
  ) %>%
  filter(año >= 1990, AGUA_TOTAL > 0, !is.na(porcentaje), porcentaje <= 100) %>%
  group_by(EMBALSE_NOMBRE, mes) %>%
  mutate(percentil = round(ecdf(porcentaje)(porcentaje) * 100, 1)) %>%
  ungroup() %>%
  mutate(categoria_anomalia = case_when(
    percentil < 25 ~ "Bajo (<25)",
    percentil <= 75 ~ "Normal (25-75)",
    percentil > 75 ~ "Alto (>75)"
  ))
```

## Visualización nacional

```
anomalías_nacional <- embalses %>%
  group_by(año, categoria_anomalia) %>%
  summarise(n = n(), .groups = "drop")

ggplot(anomalías_nacional, aes(x = año, y = n, fill = categoria_anomalia)) +
  geom_col() +
  labs(
    title = "Evolución anual de anomalías en embalses (España)",
    x = "Año", y = "Número de registros", fill = "Anomalia (percentil)"
  ) +
  theme_minimal()
```

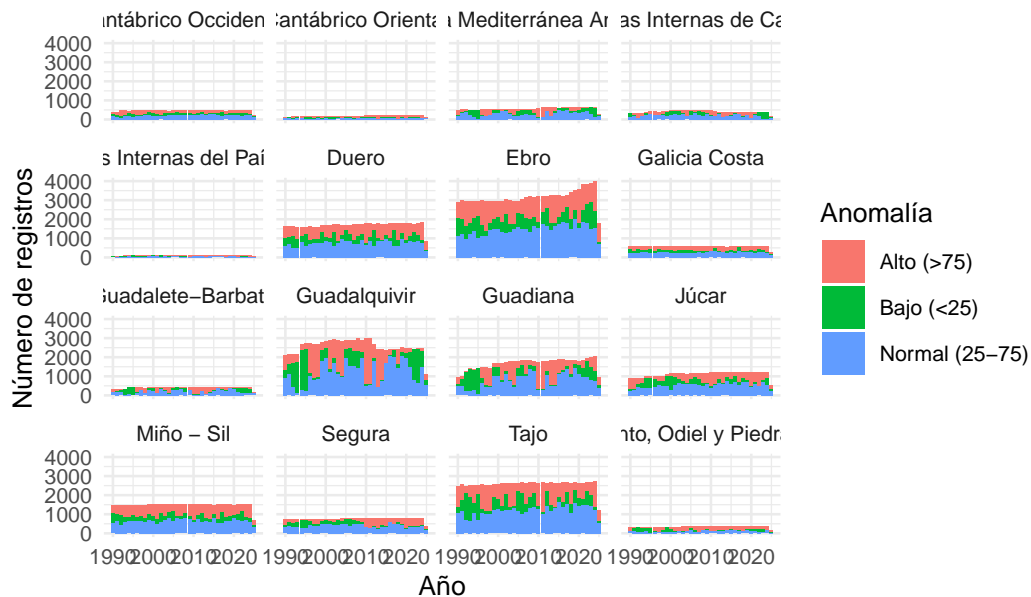


### Visualización por cuenca hidrográfica

```
anomalías_cuenca <- embalses %>%
  group_by(año, AMBITO_NOMBRE, categoria_anomalia) %>%
  summarise(n = n(), .groups = "drop")

ggplot(anomalías_cuenca, aes(x = año, y = n, fill = categoria_anomalia)) +
  geom_col() +
  facet_wrap(~ AMBITO_NOMBRE) +
  labs(
    title = "Anomalías por cuenca hidrográfica en España",
    x = "Año", y = "Número de registros", fill = "Anomalía"
  ) +
  theme_minimal(base_size = 10)
```

## Anomalías por cuenca hidrográfica en España



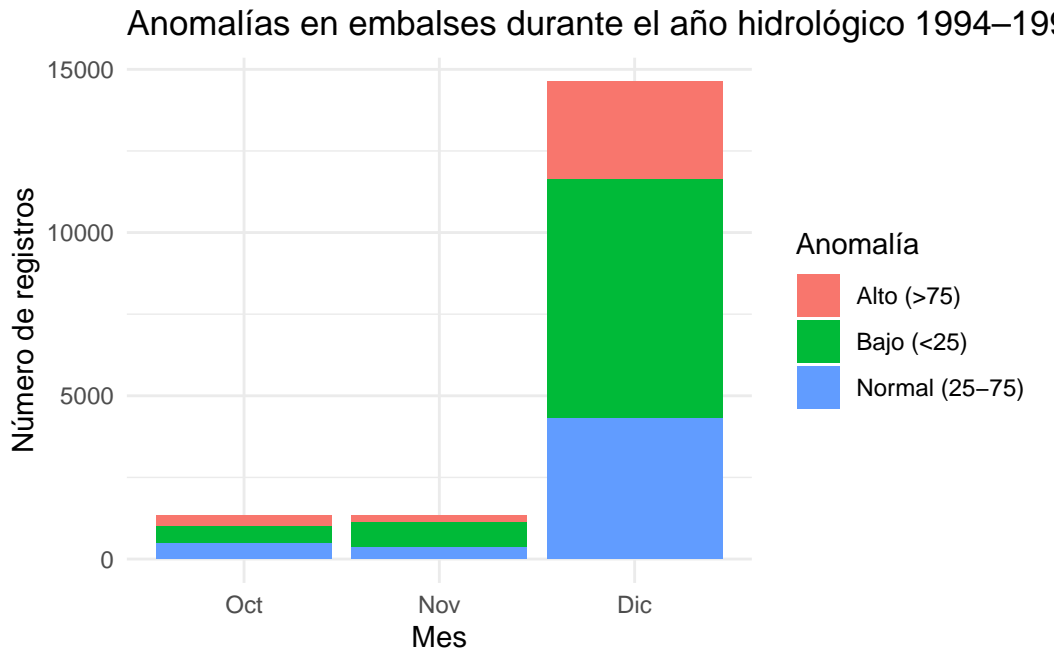
## Año hidrológico 1994–1995

```
embalses_94_95 <- embalses %>%
  filter(FECHA >= as.Date("1994-10-01") & FECHA <= as.Date("1995-09-30"))

graf_94_95 <- embalses_94_95 %>%
  group_by(mes, categoria_anomalia) %>%
  summarise(n = n(), .groups = "drop")

meses <- c("Oct", "Nov", "Dic", "Ene", "Feb", "Mar", "Abr", "May", "Jun", "Jul", "Ago", "Sep")

ggplot(graf_94_95, aes(x = factor(mes, levels = 10:9), y = n, fill = categoria_anomalia)) +
  geom_col() +
  scale_x_discrete(labels = meses) +
  labs(
    title = "Anomalías en embalses durante el año hidrológico 1994-1995",
    x = "Mes", y = "Número de registros", fill = "Anomalía"
  ) +
  theme_minimal()
```

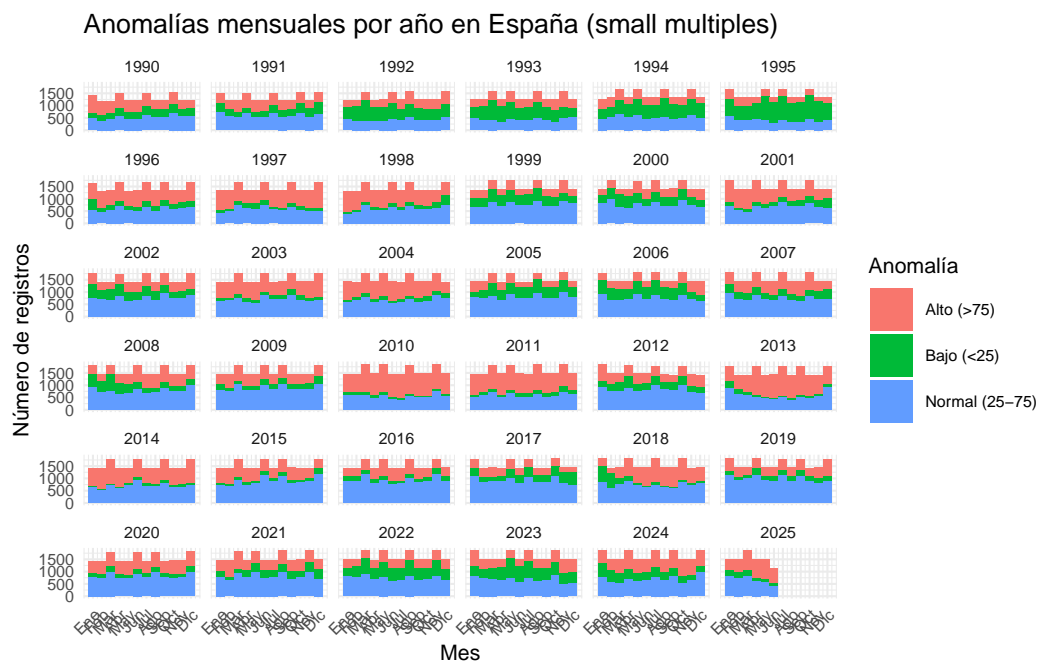


### Small multiples nacionales

```
anomalías_mes_nacional <- embalses %>%
  group_by(año, mes, categoria_anomalia) %>%
  summarise(n = n(), .groups = "drop")

g <- ggplot(anomalías_mes_nacional, aes(x = mes, y = n, fill = categoria_anomalia)) +
  geom_col() +
  facet_wrap(~ año, ncol = 6) +
  scale_x_continuous(
    breaks = 1:12,
    labels = c("Ene", "Feb", "Mar", "Abr", "May", "Jun",
               "Jul", "Ago", "Sep", "Oct", "Nov", "Dic")
  ) +
  labs(
    title = "Anomalías mensuales por año en España (small multiples)",
    x = "Mes", y = "Número de registros", fill = "Anomalía"
  ) +
  theme_minimal(base_size = 8) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
g # <- Esto asegura que el gráfico se muestre en el HTML
```



```
ggsave("anomalías_small_multiples.png", plot = g, width = 16, height = 12, dpi = 300)
```